Diplomarbeit

zur Erlangung des wissenschaftlichen Grades

Diplom-Physiker

Systematic and Experimental Uncertainties in the Search for Neutral MSSM Higgs Bosons Decaying to Two Tau Leptons

vorgelegt von

Lorenz Hauswald

geboren am 26.02.1989 in Meißen

Technische Universität Dresden

Institut für Kern- und Teilchenphysik

Fachrichtung Physik

Fakultät für Mathematik und Naturwissenschaften

2013

1. Gutachter: Jun.-Prof. Dr. Arno Straessner

2. Gutachter: Prof. Dr. Dominik Stöckinger

Abgabedatum: 16.09.2013

Abstract

For every physics measurement it is essential to precisely know its uncertainty. At the search for neutral MSSM Higgs bosons the decay channel to two tau leptons plays a major role. For this channel the influence of the uncertainty of Monte Carlo generator parameters on the acceptance of the analyses is studied in this thesis. The corresponding uncertainties are found to be in the range of 0 to 30 %. Physics with tau leptons is exceptionally challenging, because they can decay into hadrons, which, at the Large Hadron Collider, are hardly distinguishable from the overwhelming QCD background. Sophisticated multivariate algorithms aim to identify them correctly. The impact of experimental uncertainties on the outcome of the identification algortihms is subject of the second part of the thesis. The uncertainties obtained are below 7 %.

Kurzfassung

Für jede physikalische Messung ist es essentiell ihre Unsicherheit genau zu kennen. Bei der Suche nach neutralen MSSM Higgs-Bosonen spielt der Zerfallskanal in zwei Tau-Leptonen eine wichtige Rolle. Für diesen Kanal wurde der Einfluss der Unsicherheit von Parametern der Monte-Carlo-Generatoren auf die Akzeptanz der Analysen in dieser Diplomarbeit untersucht. Die entsprechenden Unsicherheiten befinden sich im Bereich von 0 bis 30 %. Physik mit Tau-Leptonen ist besonders anspruchsvoll, da sie in Hadronen zerfallen können, die am Large Hadron Collider kaum vom überwiegenden QCD-Untergrund zu unterscheiden sind. Komplexe multivariate Algorithmen sollen sie korrekt identifizieren. Die Auswirkungen experimenteller Unsicherheiten auf das Ergebnis der Identifizierungsalgorithmen ist Thema des zweiten Teils dieser Arbeit. Die ermittelten Unsicherheiten sind kleiner als 7 %.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Particle physics is one of the fastest developing fields in physics. Before the British physicist J.J. Thomson discovered the electron in 1897 [1], very little was known about the basic structure of matter. Since then many new particles have been discovered using diverse and increasingly complex experiments. In the late 1940s theoretical physicists, especially Sin-Itiro Tomonaga, Julian Schwinger and Richard Feynman, solved important problems of existing quantum theoretical descriptions of electromagnetic interactions between particles [2, 3, 4]. The resulting theory of Quantum Electrodynamics (QED) is a successful model for the following theoretical models of particle interactions. In the early 60s Sheldon Glashow played an important role in developing the model of electroweak unification and therefore realized that weak and electromagnetic interactions are manifestations of the same fundamental force [5]. A few years later Steven Weinberg and Abdus Salam included the Higgs mechanism [6, 7], which is often referred to as the birth of the Standard Model [8, 9]. With big improvements of Quantum Chromodynamics (QCD) [10, 11], the theory of strong interactions, in the early 1970s and experimental proofs of the existence of quarks, the Standard Model describes now all basic interactions except gravity. It is still the theoretical foundation of modern particle physics and was tested in many experiments. Its free parameters are measured with very high precision [12].

The success of particle physics in the last century, with many discoveries of new particles and properties, is driven by improvements of particle accelerators and detectors. Today, much effort is centered around experiments at the CERN laboratory in Geneva, Switzerland. The Large Hadron Collider (LHC) [13] is the biggest particle collider today. It is designed for proton beam energies of up to 7 TeV and has therefore the greatest collision energy ever achieved by man. Two of the seven particle detectors hosted at the LHC are so-called multipurpose detectors: ATLAS (A Toroidal LHC ApparatuS) [14] and CMS (Compact Muon Solenoid) [15]. They are more complex than any other particle detector so far, have diverse applications and are operated by big international collaborations. To process the immense amount of data taken by the detectors, new methods of data acquisition and data analysis have been developed. Improvements in grid computing enable the scientists to process the big number of captured events. The experimental setup of the ATLAS detector and the LHC is described in Chapter 3.

These are good conditions to study the properties of the in 2012 newly discovered Higgs boson, and to explore the borders of the Standard Model by finding indications for new physics beyond the Standard Model. The especially interesting theory of supersymmetry proposes a symmetry between bosons and fermions. Each particle is supposed to have a superpartner in the other group with a difference in spin of $^1/_2$. Due to symmetry breaking the superpartners do not have the same mass. Evidence of superpartners of the known elementary particles has yet to be found. One possible supersymmetric extension of the Standard Model implementing supersymmetry is the Minimal Supersymmetric Standard Model (MSSM) [16]. It predicts the existence of five Higgs bosons, three of them with neutral electric charge. To find them, the decay of the respective neutral Higgs bosons to two tau leptons is particularly important, as the branching ratio is comparatively high, especially for a high ratio of the vacuum expectation values of the two MSSM Higgs fields, $\tan\beta$ [17]. The tau lepton is the only lepton which can decay to a neutrino and hadrons, additionally to the decay to neutrinos and a lighter charged lepton. This makes reconstruction and analyses in that channel challenging. Further details of the theoretical basis of these processes are the subject of Chapter 2.

For understanding the captured events, it is essential to have good simulations of what is measured in the detector and why. This is done using Monte Carlo simulations which generating steps are discussed in Section 4.1. For understanding the significance of possible discoveries of Higgs bosons, it is essential to thoroughly study the influence of systematic uncertainties on the analyses. One source of systematic uncertainty is the dependence of simulated events on generator parameters. A goal of this study is to analyse how these uncertainties affect the resulting kinematic shapes and acceptances of two analyses in two subchannels of the neutral MSSM Higgs di-tau decay (Chapter 4). Hadronically decaying tau leptons form jets in the detector, which are difficult to distinguish from the QCD background. Therefore complex identification algorithms, including multivariate methods, are applied, which are outlined in Chapter 5. The second goal of this thesis is to understand the effect of experimental systematic uncertainties on the outcome of the identification algorithm for hadronic tau lepton decays.

Chapter 2

Theoretical Foundations

2.1 The Standard Model of Particle Physics

The Standard Model (SM) of particle physics is the theory of the 12 currently known elementary particles and their interaction [18]. Elementary particles do not have an observed substructure and therefore are considered to have no intrinsic geometric form. The Standard Model describes three of the four known fundamental forces: electromagnetic, weak and strong interaction. A unified theory including the gravitational force is yet to be established. The lack of description of gravitational interaction between particles does hardly limit the precision of Standard Model predictions, because the magnitude of gravitational forces is negligible on a subatomic scale. There are two classes of particles: bosons with integer spin quantum numbers and fermions having half-integer spin numbers.

Elementary fermions, the building blocks of matter in the universe, are divided in two categories: leptons and quarks. Contrary to quarks, leptons are not affected by the strong force. All elementary fermions can be organized in three generations, each being an almost exact copy of the others, only distinguished by different masses. Heavy particles tend to decay to lighter ones, if the decay is not forbidden by fundamental laws of conservation. This is the reason for the matter in the universe being almost exclusively constructed of fermions of the lightest generation.

Elementary bosons in the Standard Model are called gauge bosons, as they all arise theoretically as mediators of force in the respective gauge theory. Interactions of particles are mathematically described by quantum field theory with specific Lagrangian densities for each type of interaction. The invariance of these Lagrangians under certain gauge symmetries gives rise to gauge fields, which are represented by gauge bosons. Forces are mediated by the exchange of gauge bosons, the strength (and therefore probability) of an interaction is related to the respective charge of the interacting particles and the coupling constant. The coupling constants are parameters of the SM, i.e. they can not be predicted theoretically by the SM.

Electromagnetic interactions are described by the Lagrangian of the QED:

$$\mathcal{L} = i\bar{\psi}\gamma^{\mu}(\partial_{\mu} + ieA_{\mu})\psi - m\bar{\psi}\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \ .$$

This expression depends on the fermion field ψ, the electromagnetic field tensor $F_{\mu\nu}$ and the four-potential of the electromagnetic field A_{μ}. The absence of a mass term of the photon γ (the gauge boson of QED) is a consequence of $U(1)$ local gauge symmetry[1]. Because the photon carries no electric charge, it can not interact with itself.

The weak interaction is mediated by three bosons: W^+, W^- and Z^0, which are very massive. Weak and electromagnetic interactions are described in the SM as a unified theory of the electroweak interaction, a non-abelian gauge theory under the $SU(2)_L \times U(1)_Y$-group[2]. There are three gauge bosons arising from $SU(2)$ (W^0, W^1, W^2) and one from $U(1)$ (B^0), all have to be masseless because of the requirement of invariance of the Lagrangian under local gauge symmetry transformations. The observed bosons $\gamma, Z^0, W^{\pm}$ arise because of spontaneous symmetry breaking of $SU(2)_L \times U(1)_Y \to U(1)_{\text{QED}}$, which also accounts for the mass of the weak bosons. In the SM, this is generated by the Higgs mechanism [6, 7] and is the reason why electromagnetic and weak interactions are so different in interaction rate at commonly low energy, while at higher energy they are unified. The Higgs mechanism introduces a field ϕ with a nonzero vacuum expectation value - the Higgs field. The SM Higgs field is a complex doublet, so it has four components, three of them are Goldstone bosons, which are responsible for the masses of $W^{\pm}$ and Z. The remaining component is the field of a measurable massive neutral scalar boson, the SM Higgs boson. Scalar bosons have zero spin. The SM Higgs boson is the only elementary particle with spin 0, the other gauge bosons are vector bosons (spin 1). Fermions gain mass by interaction with the Higgs field (the coupling is proportional to the fermion mass). The photon does not interact with the Higgs field and therefore remains massless. It is not yet evident, whether probable small neutrino masses are also a result of the Higgs mechanism (Dirac neutrino theory) or if they acquire mass through a different mechanism (e.g. Majorana neutrino theory) [19]. As a result of the multiple decade

[1] $U(n)$ is the unitary group of n-th degree, the mathematical group of $n \times n$-matrices. Therefore $U(1)$ represents the set of complex numbers with the absolute value of 1 and the multiplication as group operation. The QED is an abelian gauge symmetry with the representing group $U(1)$. That means the Lagrangian is invariant under multiplication of any element of that group.

[2] $SU(n) = U(n) : det(e) = 1 \forall e \in U(n)$ (special unitary group of degree n). The index L of $SU(2)$ means the symmetry applies only to left chiral fields.

lasting search for the SM Higgs boson [20, 21], in 2012 a new boson with a mass[3] of 125 GeV was found by the ATLAS and CMS experiments at the LHC at CERN [22, 23]. So far its properties are compatible with the properties of the SM Higgs boson and evidence is growing that the remaining properties match, too [24, 25]. The neutral bosons Z^0 and γ arise from a rotational transformation of W^0 and B^0 by the Weinberg angle θ_W, which is an important parameter of the SM:

$$\begin{pmatrix} \gamma \\ Z^0 \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B^0 \\ W^0 \end{pmatrix} .$$

Quantum chromodynamics is the theory of strong interaction, with a $SU(3)$ gauge symmetry. Eight force carrying bosons exist, the so called gluons. They couple to particles, which have *color* (quarks and gluons). Color is a charge concept in QCD. There are three charges, called red, green and blue. Due to the fact that the potential energy in a quark-antiquark system rises when moving the quarks farer away from another, it is not possible to separate the quarks without a new quark-antiquark pair being created inbetween, each quark coupling to one of the original quarks. This property (the nonexistence of free quarks) is called confinement. Another unique property of QCD is the asymptotic freedom, which means strong interaction forces become negligible at small distances.

All elementary particles of the SM are listed in Table 2.1. Each particle has a corresponding antiparticle with the same mass and opposite quantum numbers. The neutral gauge bosons are their own antiparticles. The listed third component of the weak isospin T_3^W is only valid for left-handed particles. Right-handed anti-particles have the opposite weak isospin, while right-handed particles and left-handed antiparticles have no weak isospin at all. W bosons couple to the weak isospin, while the Z boson additionally couples to the electric charge.

In the SM neutrinos are supposed to be massless. However, experiments suggest the existence of neutrino oscillation (neutrino flavour change). This mechanism requires the neutrinos to have nonzero mass. Given in the table are experimental upper bounds of neutrino masses of the different flavours. Studies using measurements of redshifts of luminous galaxies and of cosmic microwave brackground radiation set limits on the sum of all neutrino masses, which are significantly lower than the directly measured values given in Table 2.1 (for instance $\sum m_\nu < 0.26\,\mathrm{eV}$ in [26]).

[3] In particle physics masses are usually given in units of electron volts divided by the square of the velocity of light in a vacuum c^2. The system of natural units is chosen, where $c = \hbar = k_\mathrm{B} = 1$ with the Boltzmann constant k_B and the reduced Planck constant $\hbar$.

	name	mass	electric charge	T_3^{W}	color
Fermions					
Quarks					
u	up	2.3 MeV	$\frac{2}{3}$	$+\frac{1}{2}$	r,g,b
c	charm	1.3 GeV			
t	top	173.5 GeV			
d	down	4.8 MeV	$-\frac{1}{3}$	$-\frac{1}{2}$	r,g,b
s	strange	95 MeV			
b	bottom	4.2 GeV			
Leptons					
ν_e	electron neutrino	< 2 eV	0	$+\frac{1}{2}$	-
ν_μ	muon neutrino	< 0.2 MeV			
ν_τ	tau neutrino	< 18.2 MeV			
e	electron	511 keV	-1	$-\frac{1}{2}$	-
μ	muon	105.7 MeV			
τ	tau	1.78 GeV			
Bosons					
γ	photon	0	0	0	-
$W^\pm$	W-Boson	80.4 GeV	±1	±1	-
Z^0	Z-Boson	91.2 GeV	0	0	-
g	gluons	0	0	0	8 color combinations
H	Higgs	125 GeV	0	—	-

Table 2.1: Elementary particles of the Standard Model [27]. The entry of the Higgs boson is preliminary, since it is not yet confirmed that the newfound boson is, in fact, the SM Higgs boson.

2.2 Minimal Supersymmetric Extension of the Standard Model

Despite the big success of the SM in predicting and describing particle physics phenomena, more and more open questions appear, which cannot be satisfactorily answered without changing important aspects of the model. Therefore physicists aim to find and experimentally confirm Beyond Standard Model (BSM) theories.

The *hierarchy problem* is one of the problems of the SM [28]. It basically is the believe that the big difference between electroweak and Planck scale is unnatural. To avoid divergences when calculating the fermionic loop contributions to the Higgs mass, a cut-off energy has to be introduced. This cut-off energy should resemble a scale where the model is believed to break down. The Planck scale is a sensible (and also maximum) guess, as it is the energy scale where gravity cannot be neglected anymore. On the other hand, the Higgs mass is very small compared to that scale and in fact should be that small for various reasons. For instance, the perturbative SM broke down, if the Higgs mass would be bigger than about 1 TeV

[29]. The exact value of this upper bound depends on the cut-off scale. This limit is only valid within the SM, as there could be new physics restoring the unitarity. To make up for this discrepancy, there has to be a correcting value, the bare mass of the Higgs, so that the difference is of the order of the electroweak scale. The precision of this fine-tuning could be as high as 32 decimal places, which is widely considered unphysical. In other words, there is currently no explaination why the weak force is so much stronger than gravity. Lower cut-offs would decrease the amount of fine-tuning, which hints at new physics at the TeV scale. Cut-off scales there would lead to fine-tuning of an extend which many physicists consider small enough.

The existence of *dark matter* is strongly supported by many observations, but it is not described by the SM. Dark matter could be consisting of WIMPs: weakly interacting massive particles (which do not participate in strong and electromagnetic interactions). There are, in fact, three particles in the SM which only interact by weak force and probably have mass: the neutrinos. However, the upper bounds on neutrino masses are already too low, therefore they are moving very fast and are candidates for a *hot dark matter* hypothesis. Hot dark matter alone can not constitute dark matter in the universe, as it can not explain the formation of astrophysical structures smaller than galaxy clusters [30]. Scientists are now in search for WIMPs beyond the SM.

Gauge coupling constants are, in fact, not constant, as they are functions of Q, the energy scale of the interaction. This phenomenon is called *running gauge couplings*. In the SM the coupling constants nearly meet at very high energies. The unification of the couplings at some energy would be a great hint for a more fundamental theory and is expected to be part of a possible Grand Unified Theory, describing weak, electromagnetic and strong interactions using only one gauge symmetry.

While not solving all problems of the SM, supersymmetry [16] could resolve at least the three important ones described above. In the theory of supersymmetry a new fundamental symmetry is introduced between bosons and fermions. Each fermion is suggested to have a bosonic superpartner and vice versa. The superpartners share the same quantum numbers (except their spin). Naturally one would expect their masses to be the same, too, but since there are no observed SUSY particles within the experimentally accessible mass range, SUSY has to be spontaneously broken. This means SUSY particles would have higher masses. It would be convenient for solving the hierarchy problem, if these masses were in the TeV range. One specific SUSY model, which is tailored to resolve problems with the SM with minimum additional assumptions, is the MSSM. To cancel out anomalies and give mass to all particles, there is the need for two Higgs doublets (instead of one in the SM) in the MSSM. This implies the existence of four additional Higgs bosons: two neutral and two charged. Since baryon and lepton numbers are no longer strictly conserved in the MSSM, the concept of R-parity is introduced [16]. Among other things, this suppresses non observed decay modes like the proton decay. The additional particles would contribute to the Higgs loop diagrams in a opposite way, so that a significantly lower amount of fine-tuning is necessary. Due to R-parity, the lightest supersymmetric particle (LSP) is stable and a WIMP and currently the most promising candidate to explain dark matter. In the MSSM the gauge couplings are naturally unified at the expected energy scale.

No SUSY particle has been experimentally observed yet, but with the LHC reaching energies where supersymmetric partners of SM particles are exptected at high luminosity there is hope to find first experimental evidence soon. While the mass of the newly observed boson is relatively high for a possible MSSM Higgs, it is still compatible [31].

2.3 Characteristics of Tau Lepton Physics

In comparison to the other leptons, the tau lepton is very massive particle ($1.77682 \pm 0.00016\,\text{GeV}$ [27]). Because of the big resulting phase space, its mean life time ($2.906 \pm 0.010 \times 10^{-13}\,\text{s}$ [27]) is small compared to other particles in its mass region. Hence, tau leptons can not be detected directly in collider experiments, instead they have to be reconstructed using their decay products. Its high mass allows the tau lepton to not only decay to the lighter charged leptons (and neutrinos), but also to light mesons (and a neutrino). The branching ratio for $\tau^- \to \mu^- \overline{\nu_\mu} \nu_\tau$ is 17.4 %, for $\tau^- \to e^- \overline{\nu_e} \nu_\tau$ it is 17.8 % [27]. Hadronic decay modes account for 64.8 %, they are categorized by the number of charged decay products (prongs), which has to be odd, because of charge conservation. While technically not being purely hadronic decays (because of the neutrino), they are widely called hadronic, since the neutrino is not detected in most of the experiments. The majority of hadronic decay products are pions, but in rare cases there are also kaons involved.

Of primary interest are 1-prong (one charged pion + neutrals) and 3-prong decay modes, as they have the largest branching fraction. The leptonic tau decay channels are theoretically well described and among others good processes to study lepton universality. However, in collider experiments they are hard to reconstruct, since the resulting charged leptons are hard to distinguish from prompt electrons and muons produced in other processes. The hadronic decay channel has to deal with difficult background discrimination, too. For the differentiation between hadronic tau decay products and quark and gluon jets (QCD jets), complex identification algorithms (covered in Chapter 5) are applied. A jet is a collection of hadrons, which are emitted in a cone. Other particles within the cone are often also associated with the jet. Jets originate from hadronizing quarks or gluons when their momentum is large enough.

Tau leptons are important for Higgs searches within the Standard Model, because the $\tau^+\tau^-$ decay channel has a high branching ratio compared to other decay channels. The decay channel with the highest BR[4] at a Higgs mass of 125 GeV is to $b\bar{b}$ with 57.7 % [32], but the channel is one of the more difficult ones, as background discrimination relies heavily on b-tagging algorithms[5] and has to deal with a very high QCD background. The second highest Higgs BR is that of the WW channel (21.5 % [32]). It is divided into several subchannels, the one with the biggest BR (both W bosons decay hadronically) is not feasible to study at the moment, because the large QCD background is hard to discriminate. The other subchannels (2 charged leptons+2 neutrinos ($2\ell 2\nu$) and $\ell\nu qq$, both $\ell \neq \tau$) are very promising, especially $2\ell 2\nu$, since it has a relatively small background, but its BR is only at 1.06 % [32]. The Higgs decay into a pair of gluons has a BR of 8.57 % at $m_H = 125$ GeV, but is not studied either, since there is no feasible way to reduce the background. With 6.32 % BR the $\tau\tau$ decay channel not only can contribute significantly to measuring properties of a SM Higgs boson, but also is especially interesting for decays of neutral Higgs bosons in SUSY models [33, 34, 35]. Furthermore, the channels $H \to ZZ$ with a BR of 2.64 % and $H \to \gamma\gamma$ with a BR of 0.23 % are important, because of their clean detector signatures and comparably easy background discrimination [32].

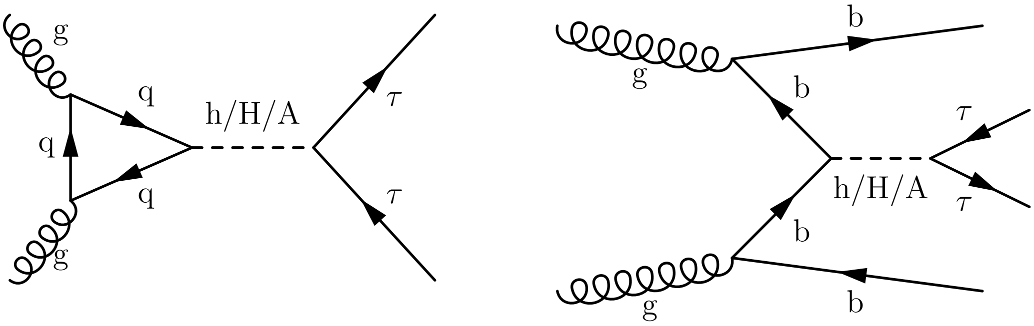
2.3.1 MSSM Neutral Higgs boson decays to $\tau\tau$

In dependence of various SUSY parameters (mostly the mass of the CP-odd[6] Higgs boson m_A and $\tan\beta$), the Higgs couplings in the MSSM can differ significantly from those in the SM. Important production processes, especially for high $\tan\beta$ are gluon fusion and b-quark associated production (see Figure 2.1 for leading order Feynman diagrams). Note that there are other neutral Higgs production processes with involvement of b-quarks. Depending on the analysis cuts, whether and how many b-tagged jets are required, other processes have to be considered

[4]This and the following branching ratios are calculated for the SM Higgs. MSSM Higgs decay cross sections depend on the type of Higgs boson and $\tan\beta$ (ratio of the two vacuum expectation values of the MSSM Higgs fields) and other free MSSM parameters.

[5]b-tagging is the identification of the originating quark flavour of a jet. Bottom-flavoured hadrons have high enough mean life times so that modern particle detectors can often detect the distance traveled between primary vertex (mostly the interaction point) and the secondary vertex (b-hadron decay vertex). The shape and kinematic properties of b-jets can be used for flavour tagging, too. The efficiency of b-tagging at ATLAS is of the order of 60 %.

[6]odd under Charge conjugation Parity symmetry



(a) gluon-gluon fusion production

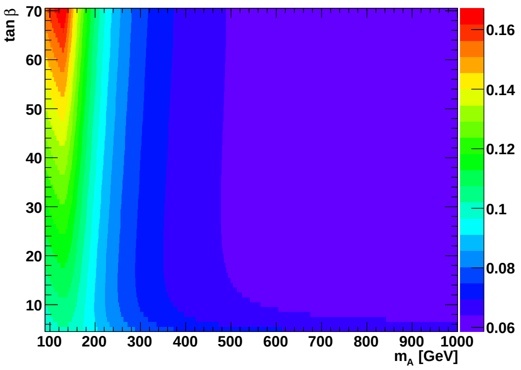
(b) bottom quark associated production

Figure 2.1: Production of neutral MSSM Higgs bosons and decay to a pair of tau leptons

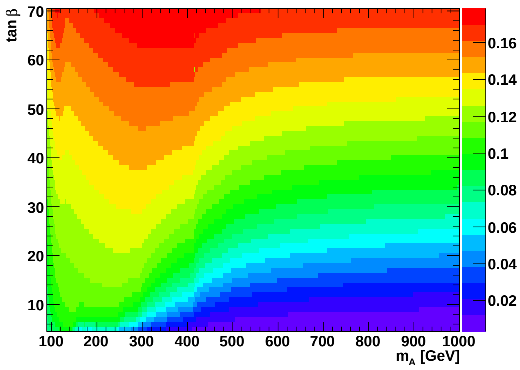
[36]. There are also $b\bar{b} \to \Phi$ and $gb \to \Phi b$[7]. Two final state b-jets can also originate from $q\bar{q} \to g \to b\bar{b}$, where one b emits a Φ. It is not sufficient to only consider top quarks in the fermion loop in Fig. 2.1a, since the coupling of Φ to b-quarks can be greatly enhanced in the MSSM. In Figure 2.2 calculated branching ratios for the decay of neutral MSSM Higgs bosons into pairs of tau leptons are shown. The values were calculated by the LHC Higgs Cross Section Working Group [32] using the FeynHiggs program with the $m_{\mathrm{h}}^{\mathrm{max}}$ scenario[8] [37]. For comparison, also the decay of the CP-odd neutral Higgs A to $b\bar{b}$ is given. Like in the SM, this is the dominant process in most of the parameter space, but it is also suffering from a large QCD background.

[7] Wherever no further specification necessary, neutral MSSM Higgs bosons are going to be called Φ.

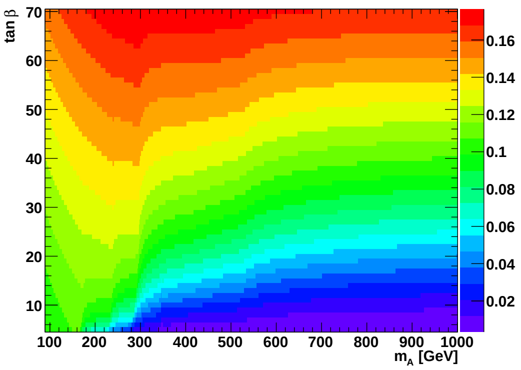
[8] While $\tan\beta$ and m_{A} are the most important parameters of the MSSM, there are lots of other free parameters. For calculation and comparison of values, benchmark scenarios are used.



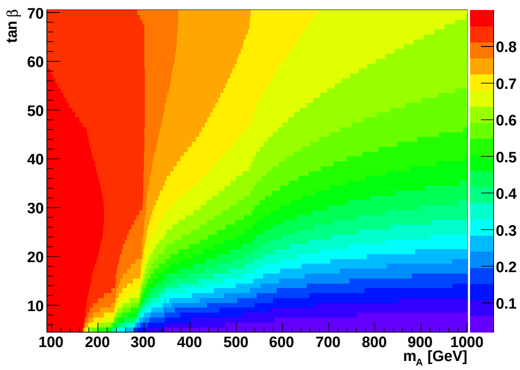
(a) $h \to \tau\tau$



(b) $H \to \tau\tau$



(c) $A \to \tau\tau$



(d) $A \to b\bar{b}$

Figure 2.2: MSSM neutral Higgs boson branching ratios

Chapter 3

Experimental apparatus

3.1 The Large Hadron Collider

Located at the CERN (European Organization for Nuclear Research) laboratory near Geneva (Switzerland), the LHC [13] is the biggest particle accelerator in the world. It is a proton-proton collider, a synchrotron accelerator with the record design center of mass energy of 14 TeV. The nominal beam energy of 7 TeV is planned to be achieved in 2015 after an extended period of maintenance and upgrades (Long Shutdown 1, LS1). After an incident in 2008 and the necessary repairs in 2009 the collider was operated at a beam energy of 3.5-4 TeV from 2010 to 2013. The LHC tunnel, which previously hosted the LEP accelarator [38] (Large Electron Positron Collider, 1989-2000) is situated up to 175 m below the Franco-Swiss border and has a circumference of 27 km (see Figure 3.1).

The counter-rotating proton beams are bent by 1232 superconducting dipole magnets, which have to be kept at 1.9 K (using liquid helium) and are designed to reach magnetic fields of up to 8.3 T. Additionally there are over 500 quadrupole magnets for focusing the beam. The beams consist of up to 2808 bunches of protons, each containing around 10^{11} particles at the beginning of a fill. There is also the possibility to accelerate and collide heavy ions (Pb). The design luminosity is $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, further machine specifications are available in the LHC Design Report [40]. At the end of the 2012 data taking period the LHC had delivered an integrated luminosity of $29\,\mathrm{fb}^{-1}$ to ATLAS, with a peak luminosity of $7.7 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ [41]. The numbers of bunches per beam was increased stepwise from 121 in 2010 to 1380 in 2012.

There are four interaction points, where the beams, which travel through seperate vacuum tubes, are collided. This is were the four main experiments ALICE [42], ATLAS [14], CMS [15] and LHCb [43] are located. Smaller experiments are MoEDAL [44], TOTEM [45] and LHCf [46]. They share interaction points with

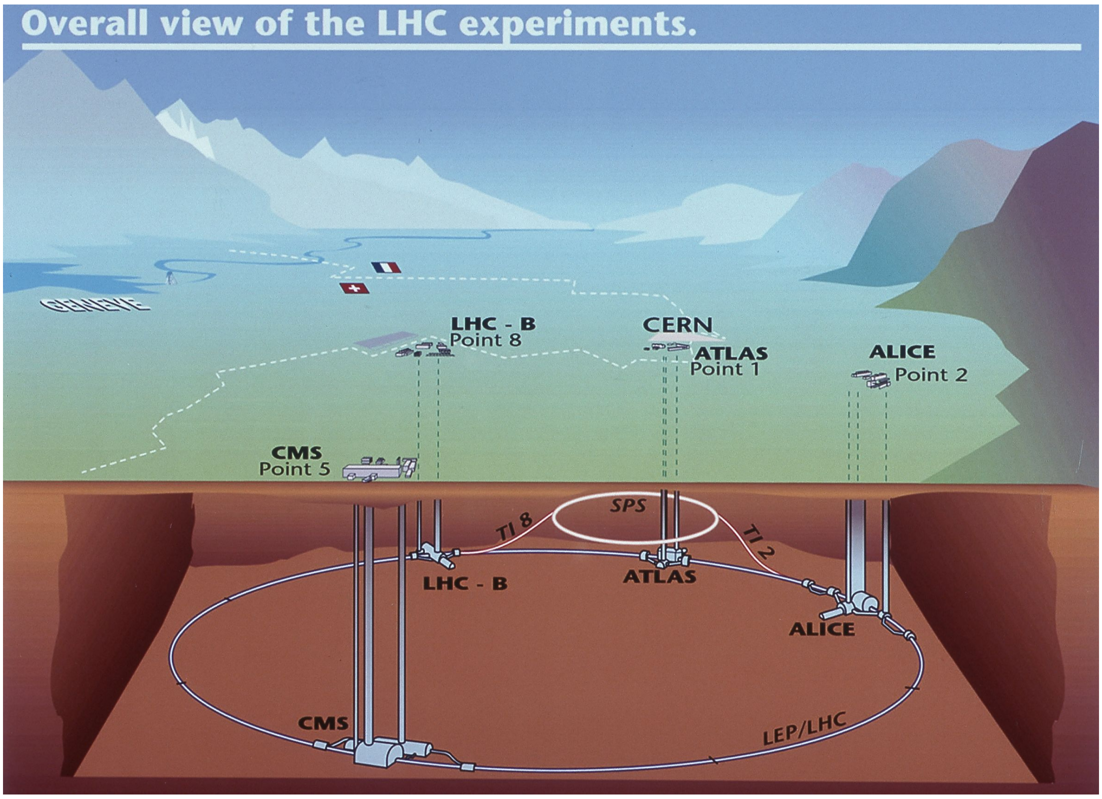


Figure 3.1: Overall view of CERN [39]

the bigger experiments.

The protons are pre-accelerated using previous generations of accelerator infrastructure of CERN. At first the proton energy is boosted to 50 MeV by the linear accelerator LINAC 2. Downstream there is a series of synchrotrons bringing the protons to an energy of 450 MeV before they are injected in the LHC. The oldest part is the Proton Synchrotron PS, which is used since 1959 and is also part of the heavy ion acceleration process. A schematic view of current accelerators at CERN is shown in Fig. 3.2, including systems not directly related to the LHC.

3.2 The ATLAS detector

In this section an overview is given of the various ATLAS detector components, coordinates and trigger system. The ATLAS detector is one of the most complex particle physics experiments in the world and is designed as a multi-purpose detector. To live up to the ambitious requirements of the ATLAS physics program, its multiple layers of detectors cover almost all spatial directions around the interaction point. The goal is to detect as much of the not neutrino-like particles as possible and reliably reconstruct their direction and energy.

One main purpose of the ATLAS detector is the search for the SM Higgs boson, as well as the measurement of its properties. It is also important to repeat

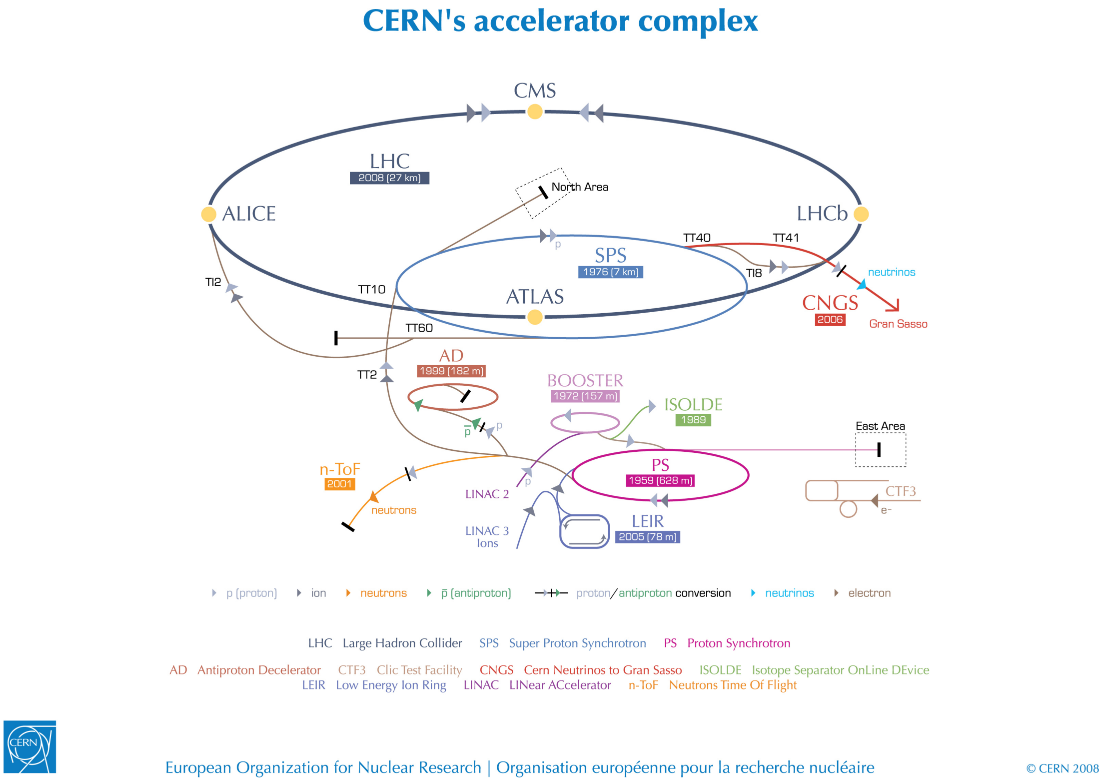


Figure 3.2: Schematic view of accelerators at CERN [47]

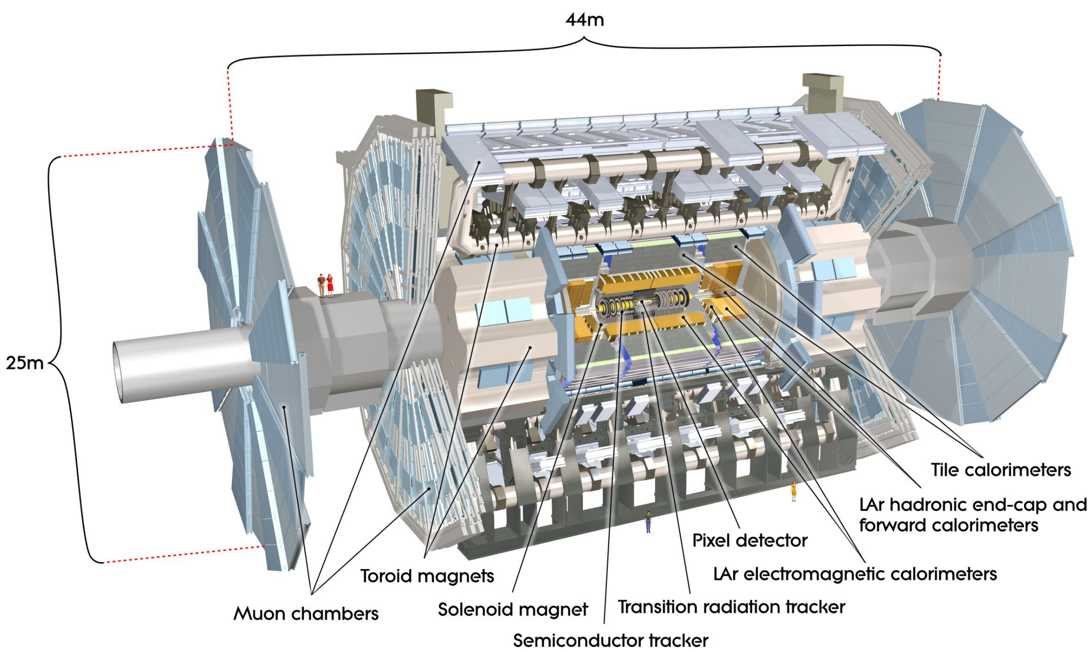


Figure 3.3: Rendered view of the ATLAS detector [14]

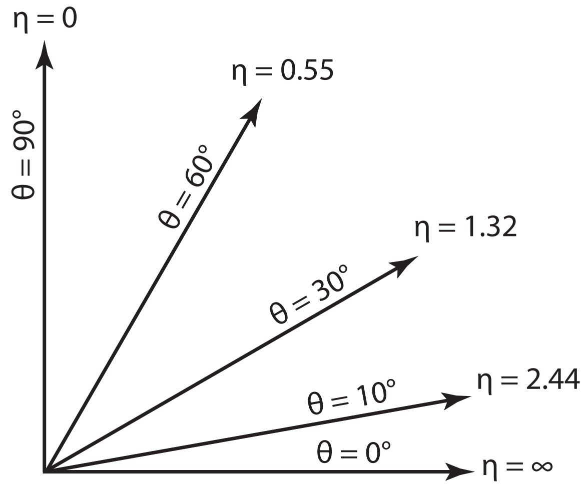


Figure 3.4: Pseudorapidity for several azimuthal angles

measurements of SM parameters and improve their precision. Phenomena like CP-violation are studied, making use of the immense number of recorded events. Another significant field of application is the search for new physics. Various SUSY models are tested, as well as other BSM theories.

A computer rendered view of the detector is shown in Fig. 3.3, the depicted subsystems are going to be described in following sections.

3.2.1 Coordinate System

As a reference, a right-handed cartesian coordinate system (x, y, z) with the origin in the center of the detector, which is also the nominal interaction point, is used. The z axis is parallel to the beams and points towards LHC Point 2, where the ALICE cavern is located. Due to the symmetry of the detector, cylindrical coordinates (r, θ, ϕ) are more widely used. The azimuthal angle ϕ is measured perpendicular to the beam line and is defined by $\tan\phi = {}^{y}/_{x}$. The polar angle θ is defined by: $\cos\theta = {}^{z}/\sqrt{x^2+y^2+z^2}$ (the angle between the position vector and the z axis). Instead of the polar angle, the pseudorapidity η is mostly used:

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

Its functional dependency on the azimuthal angle is depicted in Fig. 3.4.

The particle momenta perpendicular to the beam are of particular interest, because, due to momentum conservation, the sum of the momentum projections in the x,y-plane has to be zero, since the perpendicular momentum of the primary particles is negligible. The longitudinal component of the momentum is not used, because the initial longitudinal momentum of the interacting partons is unknown.

Therefore the transverse momentum p_{T} is defined:

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

Angular separations are expressed by ΔR:

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

3.2.2 Inner Detector

The Inner Detector is built of three detector systems: the Pixel detector, the SemiConductor Tracker (SCT) and the Transition Radiation Tracker (TRT). Each system has components in the barrel region and in the end-caps. The pseudorapidity region $|\eta| < 2.5$ is covered. Its purpose is to very precisely detect charged tracks and vertices. A cutaway view of the Inner Detector is depicted in Fig. 3.5. The whole Inner Detector is surrounded by a solenoid creating a strong magnetic field of 2 T.

While the decay length of tau leptons is too small for them to be detected directly, their decay products leave signals in the Inner Detector. For the hadronically decaying tau lepton the tracks of the charged mesons can be measured and thus, due to the magnetic field, their momentum. Additionally the secondary vertex can often be reconstructed for multi-prong decays of hadronically decaying tau leptons.

The Pixel detector The Pixel consists of three cylindrical layers of silicon pixel modules in the barrel region and three discs of modules at each side of the end-cap region. Charged particles are detected by creation of electron-hole pairs in the silicon pixels due to ionization. It is the detector which is closest to the beamline, installed at radii between 5 and 15 cm. Despite the relatively small volume, there are over 80 million readout channels (1744 modules, about 45,000 pixels each), resulting in the detector system with the highest precision in ATLAS. The intrinsic accuracy is 115 µm in the z-direction and 10 µm in the $r - \phi$-plane. The main purpose of the Pixel is the measurement of the position of the vertices. Its high precision makes it possible to classify certain particles by measuring the distance between primary and secondary vertex.

The SemiConductor Tracker The SCT consists of 4088 silicon strip modules, which are arranged in four cylindrical layers in the barrel region and nine discs of modules on each end-cap side. It is located in the radial range between 30 and 52 cm and has 6.2 million readout channels. Resolutions of 16 µm in $r - \phi$ and

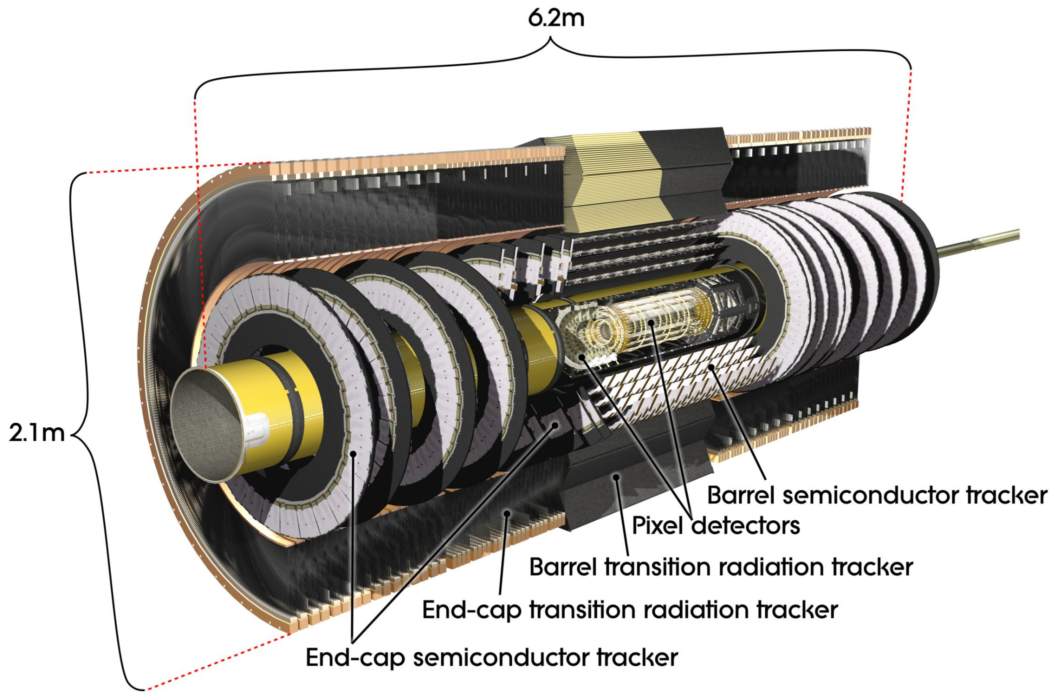


Figure 3.5: Cutaway view of the ATLAS Inner Detector [14]

580 µm in the z-direction are reached. The physical detection principle is the same as for the Pixel, but with strips instead of pixels.

The Transition Radiation Tracker The TRT ist installed at pseudorapidities $|\eta| < 2$. It uses gas filled straw shaped wire chambers with a diameter of 4 mm. The gas mixture consists of xenon (70 %), carbon dioxide (27 %) and oxygen (3 %). Ionizing particles traveling through the straw ionize gas atoms. Resulting ions and electrons are collected by the high voltage potential between the anode wire (gold plated tungsten, diameter 31 µm) and the straw walls. The volume between the straws is filled with radiator material resulting in a higher intensity of transition radiation. The energy deposited in the TRT by transition radiation photons is typically 4-5 times higher than energy deposited by charged particles. This is used for electron identification, since due to being more relativistic in comparison to e.g. pions, electrons cause significantly more transition radiation. Despite the lower number of 420,000 readout channels and an intrinsic resolution of 130 µm the TRT provides a good resolution, because there are 36 track hits on average. The high number of track hits in the TRT result in a better momentum resolution, since the momentum is measured by using the curvature of the tracks in the magnetic field.

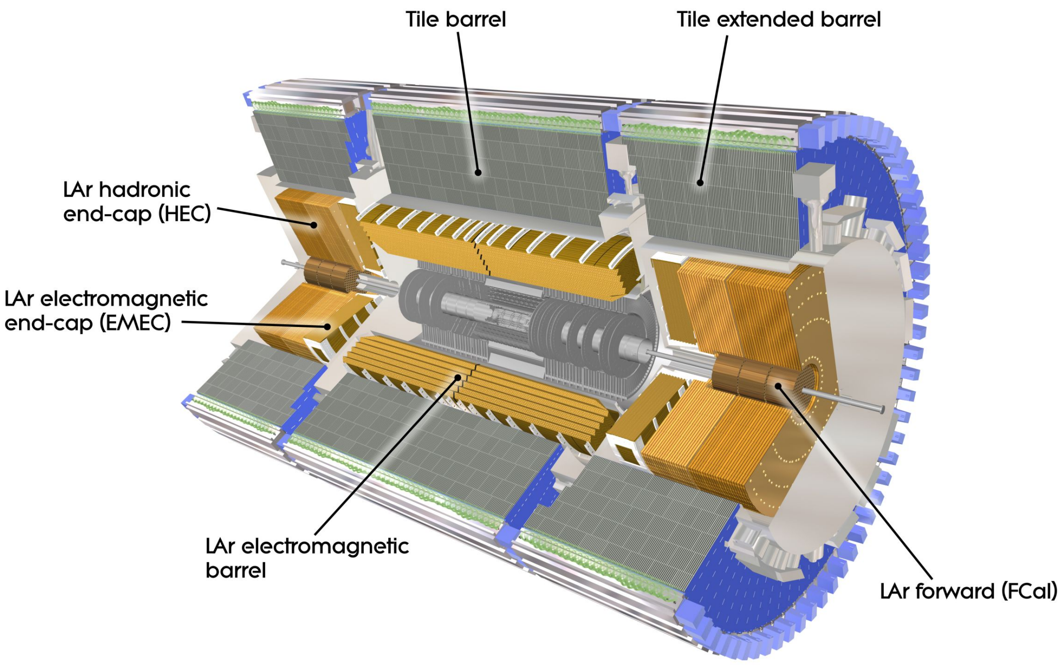


Figure 3.6: View of the ATLAS calorimetry system [14]

3.2.3 Calorimeters

The calorimetry system of ATLAS is built of five different calorimeters (depicted in Fig. 3.6), the LAr electromagnetic barrel and the Tile barrel, as well as LAr end-cap calorimeters and forward LAr calorimeters on both sides. The calorimeters measure the energy of particles by stopping them in their volume and measure the deposited energy. ATLAS calorimeters are divided into the electromagnetic part, which predominantly measures the energy of electromagnetically interacting particles, and the hadronic part, which absorbs remaining particles, except of muons and neutrinos.

The decay products of hadronically decaying tau leptons are stopped in the electromagnetic or the hadronic calorimeter. Contrary to quark or gluon jets, the decay products of a hadronically decaying tau lepton tend to be very collimated with most of the energy within a cone with a radius of $R = 0.2$.

LAr calorimeters There are four liquid argon based calorimeters in ATLAS. The electromagnetic barrel (EMB) and the EM end-cap (EMEC) amount to the electromagnetic calorimeter (ECAL). LAr forward calorimeter (FCAL) and hadronic end-cap (HEC) are part of the hadronic calorimeter system. The LAr calorimeters are sampling calorimeters using different absorber materials and geometries. Energy is measured by collecting the ionization current originating in the active material (LAr). Because the energy deposited in the absorber material can not be measured directly, the energy resolution is inferior compared to homogenous calorimeters. The ECAL uses lead as absorber material, arranged in an accordion geometry. Its LAr gaps are 4 mm thick with copper-kapton electrodes in the middle. To correct for lost energy due to material upstream of the ECAL, a thin presampler strip (LAr) is installed in the pseudorapidity region $|\eta| < 1.8$. This increases the energy resolution of the ECAL. The FCAL is installed at high pseudorapidities ($3.1 < |\eta| < 4.9$), which is important to obtain the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ with high precision. Tungsten and copper are used as absorbers confining the LAr, which is filled in small gaps (0.25 - 0.5 mm) in form of hollow cylinders. The HEC is situated behind the EMEC and uses copper as absorber material, arranged in a sandwich pattern.

Tile calorimeter There are three parts of the Tile calorimeter, a barrel in the center and two barrel extensions, measuring the energy of hadronic showers in the pseudorapidity range of $|\eta| < 1.7$. It is a sampling calorimeter with plastic scintillators as active material and iron as absorber. The scintillation light is collected with wavelength shifting fibres and measured with photomultiplier tubes.

3.2.4 The muon spectrometer

Besides of neutrinos, which are not directly detected at all, muons are the only particles which pass through the calorimeters. The muon spectrometer is the outermost part of the detector. Its purpose is to measure direction and momentum of the muons in the pseudorapidity region $|\eta| < 2.7$. Muons with an energy lower than 5 GeV are most likely stopped before reaching the muon spectrometer. It is the biggest part of the ATLAS detector. There are three cylindrical detection layers in the barrel region, as well as three layers in each end-cap. Coordinates of the muon tracks are measured using monitored drift tubes (MDT). MDTs are arrangements of cylindrical drift chambers with a diameter of 3 cm and are filled with a gas mixture of argon and carbon dioxide (93/7 %). The anode wires are made of a tungsten-rhenium alloy. Drift tubes use the high precision timing of electric signals, caused by electrons created by ionization, to reconstruct the location of the primary particle. At large pseudorapidities ($2.0 < |\eta| < 2.7$) the position measurement is done by cathode strip chambers (CSC) instead. CSCs are multiwire proportional chambers with a gas mixture of argon, carbon dioxide and carbon tetrafluoride (30/50/20 %). The trigger system consists of resistive plate chambers (RPC) in the barrel region and thin gap chambers (TGC) in the end-caps up to $|\eta| = 2.4$. RPCs are installed in front, inbetween and in the back of the MDTs to provide a coincidence signal to veto fake muons. They are ionization chambers with 2 mm distance between the parallel resistive plates. Signal times are below 10 ns, which is one of the reasons for using them as a trigger. TGCs are multiwire proportional chambers with a highly quenching gas. Both RPCs and TGCs are used for measuring additional coordinates to complement the precision measurements of MDTs and CSCs. A toroidal magnetic field is generated by eight large superconducting coils in the barrel region, as well as two in the end-caps. By associating the muon track measured in the spectrometer with the calorimeter energy deposited along the extrapolated track in the calorimeters, the energy resolution is improved.

3.2.5 Trigger system

The nominal collision rate of 40 MHz is significantly too high for the data storage system to handle and the most interesting processes are very rare. Therefore the ATLAS trigger system makes sure to only select and store a fraction of collisions to reach a resulting event rate of approximately 200 Hz. Three levels of the trigger are applied consecutively: L1, L2 (Level 1 and 2) and EF (Event Filter). The Level-1 trigger is implemented in hardware and uses information of many subdetectors, but with decreased granularity. It needs less than 2.5 µs to make a decision and decreases the event rate by a factor over 500. The Level-2 trigger uses more precise detector information on basis of regions of interest recognized by the L1 trigger. The L2 trigger target rate is 3.5 kHz, its run time is around 40 ms. Finally, the EF performs a full offline event reconstruction in about 4 seconds, bringing the event rate down to approximately 200 Hz.

Chapter 4

Theoretical uncertainties in the search for neutral MSSM Higgs bosons in the di-tau decay channel

To interpret the recorded detector signals, it is essential to simulate the model expectations to be compared to data. The procedure of Monte Carlo event generation in general and for ATLAS in particular is outlined in Section 4.1. The goal of the study presented in this chapter is to find the uncertainty on the number of accepted events arising from the dependence of simulated events on generator parameters. To obtain that uncertainty, samples of Monte Carlo generated events are created. This process is repeated, while varying values of generator parameters according to their uncertainty. The generators and parameter variations used are discussed in Section 4.2. The subsequent sections lay out the event selection and the results of the study.

The acceptance is the fraction of events which pass the event selection. Usually the event selection is applied on reconstructed objects after the simulation of the detector response. To obtain small statistical uncertainties on this estimate, it is essential to generate a large number of events per sample. This is difficult for the following reasons. An important free parameter of the event generation is the mass of the proposed neutral MSSM Higgs boson. Studies in search for the bosons sample the parameter space and apply the analysis tools for each chosen set of parameters. Therefore it is necessary to consider different mass points in this study, too. Together with the large number of parameter variations and acceptances of the order of 10^{-2}, this results in a number of generated events required of the order of 10^8. It is not feasible to perform the detector simulation and object reconstruction for that many events for an uncertainty study. Furthermore it is desireable to avoid interference with variation effects due to detector simulation and event reconstruction. Therefore the event selection of the analyses of the considered decay channels[1] $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{lep}}\tau_{\text{lep}}$ [48] is implemented as a truth-level study. The analysis framework Rivet [49] is used to implement the selection, process the events and generate plots of distributions of kinematic variables.

4.1 Monte Carlo event generation at ATLAS

There are two different steps of Monte Carlo event generation. At first the event is generated at particle level, which is also called truth level. Based on numerical and analytical methods the quantum field equations are solved and the output is a sample of events of the given process, where essentially the type, energy and direction of every resulting particle is known. Depending on the type of generator, there may be additional information like unstable intermediate states and the information which particle decayed into which decay products. To achieve this, there is a number of generation steps, which are described in Section 4.1.1. Often this step is followed by simulation of the response of the respective detector. Depending on the complexity of the detector and requirements on the precision, this step can be very computing intensive. The ATLAS detector simulation is discussed in Section 4.1.2.

4.1.1 Particle level Monte Carlo event generation

While the cross section for simple QED processes can be calculated to quite high precision perturbatively given the Lagrangian, this is not feasible for hadron collider experiments, since the primary particles are composed of quarks, anti-quarks and gluons. This leads to a multitude of initial and final states. The high energy involved in LHC interactions results in a large phase space for final state particles. Therefore the matrix element[2] calculation is automated using computer programs. The situation is complicated further, as QCD perturbation theory can only be applied at high scales (momentum transfers), where the constituents of the hadron can be seen as free (asymptotic freedom). This is true in the case of the LHC for the hard interaction in proton-proton collisions, but the softer QCD processes and hadronization can not be treated perturbatively anymore. A mechanism has to be used to connect the scale of the hard scattering with the

[1] Throughout this thesis $\tau_{\text{lep/had}}$ is going to be used to denote leptonically and hadronically decaying tau leptons respectively.

[2] The S-matrix connects the initial states with the final states of a scattering interaction. The elements are called scattering amplitudes and are related to the cross section.

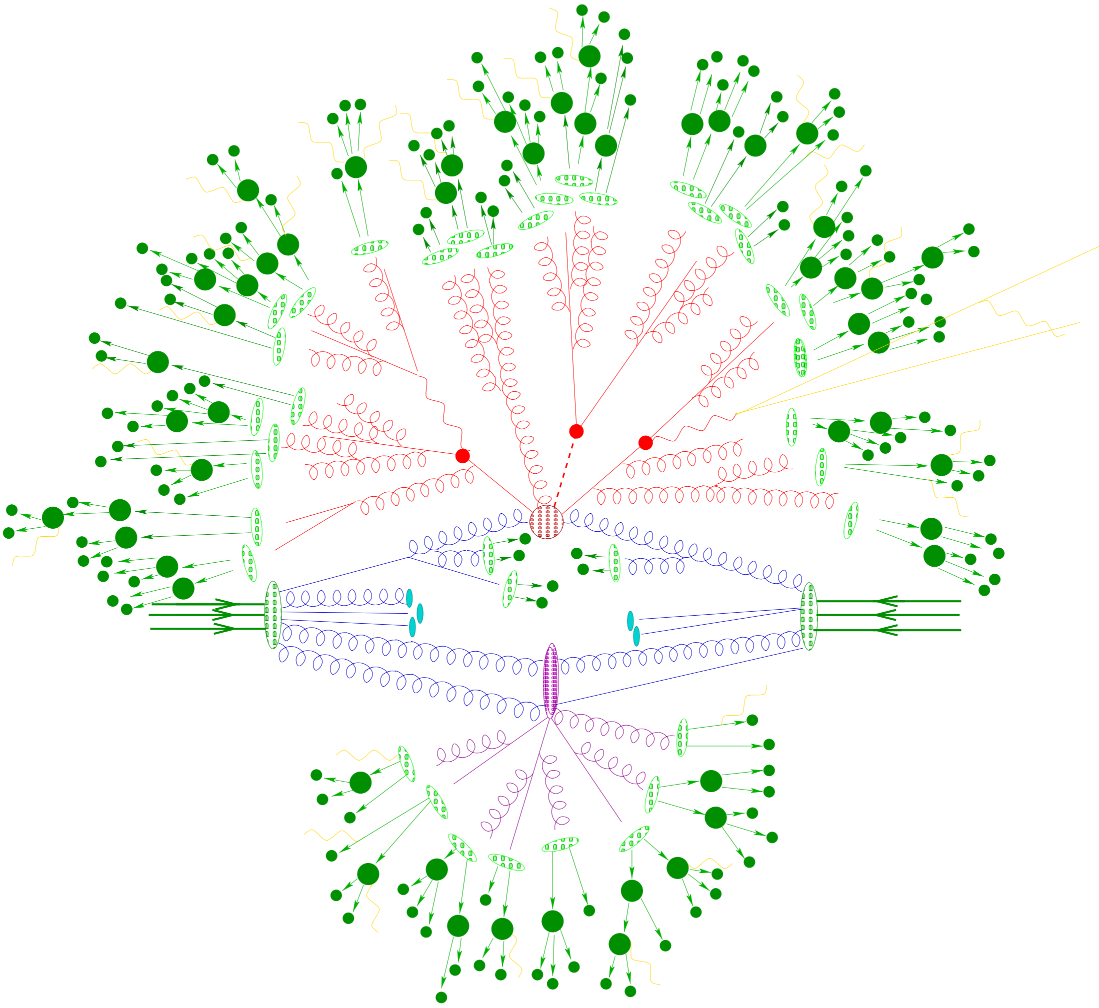


Figure 4.1: Structure of event generation [50]

hadronization scale, where the partons form colorless hadrons. Hence, the event generation is divided in successive steps, which are visualized in Fig. 4.1.

At first the matrix elements for the hard scattering are calculated perturbatively using Feynman rules in combination with a matrix element generator (represented by the dark red circle in the middle of Fig. 4.1). Since this calculation can have many terms (especially for various BSM theories), often approximation procedures are applied or matrix elements are precalculated and are hard-coded in the generator. The Feynman rules for the given process are often generated using a separate program, for instance FeynRules [51]. The cross section for the hard process is usually calculated using factorization. This means that for two interacting hadrons h_1 and h_2 the cross section for the scattering can be written as [52]:

$$\sigma = \sum_{a,b} \int_0^1 \mathrm{d}x_a \mathrm{d}x_b \int \mathrm{d}\Phi_n f_a^{h_1}(x_a, \mu_F) f_b^{h_2}(x_b, \mu_F) \frac{|\mathcal{M}_{ab\to n}|^2(\Phi_n; \mu_F \mu_R)}{2x_a x_b s}. \tag{4.1}$$

This expression depends on the parton density functions [53] of the interacting partons $f_i^{h_j}(x_i, \mu_F)$, the factorization scale μ_F and the renormalization scale μ_R. The uncertainty on these quantities is of major interest in this study. It also depends on the matrix element $\mathcal{M}_{ab\to n}$ for the partonic process $ab \to n$ and the squared center of mass energy s. $\mathrm{d}\Phi_n$ is the differential phase space element, including all configurations of the n final state particles and enforcing momentum conservation.

The initial and final states of the hard subprocess are developing parton showers, that means gluons are emitted and quark-antiquark pairs emerge. This branching leads to a successive decrease of the interaction scale down to the hadronization scale. The parton showers are depicted in light red in Fig. 4.1, initial state radiation is shown directly below the hard process. However, it is a non-trivial task to combine the matrix element calculation of the hard process with the parton shower calculation. The inclusive final state of the hard process has to be connected to the exclusive parton shower state in a sensible way. Double counting of certain phase space regions has to be avoided. There is a multitude of algorithms with different merits and downsides to achieve the matching of the matrix element to the parton shower. The most important matching algorithms, CKKW [54] and MLM [55], are described in Section 4.2.2. After the parton shower process is stopped at a freely selectable interaction scale (usually a few GeV) the hadronization has to be performed. There is no exact theoretical description of such a process, therefore phenomenological models are applied and free parameters are obtained by comparison to data [52]. Another process considered is QED Bremsstrahlung, shown in yellow in Fig. 4.1. The last step is the simulation of the decay of unstable particles. For that, experimentally measured properties of the particles are used, but in some cases are not sufficient for a definite description of the decay. Because of that, additional generator parameters and assumptions arise, which have to be tuned.

4.1.2 ATLAS detector simulation

The stable final state particles, generated as described in Section 4.1.1, are used as input to the simulation of the ATLAS detector response [56]. For the detector simulation Geant4 [57], a commonly used platform implemented in C++ for simulation of particle interaction with matter, is used. It is embedded in the ATLAS offline software framework Athena and interfaced via Python. After the energy deposition in all detector subsystems is simulated, the digitization simulation is performed. Digital readout signals are set to the expected values given the simulated energy content. The resulting event structure has the same form as a recorded data event and therefore the following data processing steps are the same for data and simulation. Sophisticated physics objects reconstruction algorithms aim to restore the particle level information of the event as good as possible. This is discussed in more detail in Chapter 5. To decrease disk space requirements, different event file formats exist with different granularity and variables. The granularity is decreased by merging detector objects. For instance individual cell energies are formed to clusters and hits in the inner detector are used to reconstruct tracks. The so-called D3PD format based on ROOT [58] contains highly specialized variables for certain applications and is designed for use with most high level physics analyses. Low level physics objects like individual cell energies are usually not included.

4.2 Generators and parameter variations

There are two signal processes to consider in neutral MSSM Higgs boson production: bottom quark associated production and gluon fusion. Additionally the background process $Z \to \tau\tau$ is determined, because the study in search for the neutral MSSM Higgs bosons in the semileptonic decay channel uses the Monte Carlo prediction for the $Z \to \tau\tau$ background estimation. A variety of generators with different parameters is used as described in Section 4.2.1. The parameter variations are discussed in Section 4.2.2.

To reduce statistical fluctuations in the event generation process, the random seeds for the generation for all samples which are compared to each other are set to the same values.

4.2.1 Monte Carlo Generators

In this study, the Monte Carlo generators Sherpa, POWHEG BOX, Pythia, Alpgen and Herwig are used.

Sherpa (acronym: Simulation of High-Energy Reactions of PArticles) [50] is a general-purpose generator framework capable of simulating a variety of particle interaction processes. It is designed modularly following the factorization approach outlined in Section 4.1.1. The main module, SHERPA, controls the other simulation modules and takes care of the main effort of the combination of matrix element and parton shower calculations. There are several interfaces for different input and output formats. Sherpa's interface to LHAPDF [59] enables it to use a multitude of recent sets of parton density functions and their uncertainties. The user has the possibility to choose alternative tools for the different steps.

For the matrix element generation AMEGIC++ [60] and Comix [61], Feynman-diagram based generators capable of matrix element calculations, including SM and MSSM, up to NLO (next to leading order in perturbation theory), were developed. APACIC++ [62] is responsible for parton shower generation and part of the merging procedure with the matrix element. The AMISIC++ [63] module was created for simulation of multiple parton interactions. The resulting hadrons are processed by the same simulation tools as the hard scatter hadrons, forming the underlying event. The hadronization step is handled by the AHADIC++ module. Hadron and tau lepton decays are simulated using the HADRONS++ module. PHOTONS++ [64] is Sherpa's package for the simulation of QED radiation. For this study Sherpa was used in version 1.4.1 to generate bottom quark associated Higgs boson production events with the respective tau decay channels. The runtime was greatly reduced by precomputing the matrix elements beforehand and using them as input for the big number of subsamples produced simultaneously on different computing nodes.

POWHEG BOX [65] is an NLO event generator framework implementing the POWHEG (POsitive Weight Hardest Emission Generator) method. By calculating the hard process at NLO the hardest emission is already included. POWHEG BOX is a parton level generator and outputs the resulting particles in the LHE format (Les Houches Events [66]). It makes use of the LHAPDF interface. The result is interfaced with a SMC (Shower Monte Carlo) generator for successive simulation steps. The respective SMC program is supposed to veto emssions which are harder than the emission generated by POWHEG BOX. Pythia 8.1 and POWHEG BOX 1.0 were deployed in this study to generate gluon fusion Higgs boson production events with at least two tau leptons decaying to the respective final states.

Pythia [67] is a general-purpose generator widely used for LHC physics. While having a set of hard processes available itself, usually external matrix element generators are used and interfaced with Pythia using LHE files. Then Pythia takes care of parton showers, hadronization, decay and QED radiation as well as the matching of the different calculation steps. Parton density functions are usually supplied by the LHAPDF interface. Event generation with Pythia follows the same basic structure outlined in Section 4.1.1 and as described for Sherpa, but different approximations and models are used for the individual steps.

ALPGEN [68] is a leading order parton level generator for hadron collision final states with a fixed number of partons. It was used in version 2.14 for this study to generate $Z \to \tau\tau$ LHE files. The generator comes with its own library of parton density functions.

Herwig [69] is the multi-purpose generator used in this study to further process the LHE files produced by ALPGEN to generate $Z \to \tau\tau$ events (version 6.520). Like Pythia, it is often used in combination with an external parton level generator.

4.2.2 Parameter variations

CKKW matching scale As described in Section 4.1.1 it is important to apply a matching algorithm to correctly merge the matrix element and the parton shower calculations. One of the most important matching schemes is CKKW (Catani-Krauss-Kuhn-Webber) matching [54], which is implemented in an extended form [70] by Sherpa. The method splits the phase space into two separate regions - the matrix element domain (jet production) and the parton shower domain (intra-jet evolution). At first the exclusive matrix elements are calculated with the constraint that the scale Q is bigger than the matching scale Q_{cut} and events are generated accordingly. The matrix element shower evolution is tracked backwards to find the most probable evolution kernel. Then the matrix element is reweighted to match the exclusive nature of the parton shower at Q_{cut}. Parton showering is then started at the hard scale, but emissions harder than Q_{cut} are vetoed.

The matching scale Q_{cut} with the default value of $\sqrt{20/E_{\text{CMS}}}$ (for the specific process) was varied up to $\sqrt{30/E_{\text{CMS}}}$ and down to $\sqrt{15/E_{\text{CMS}}}$ to study the effect of its uncertainty. It depends on the total center of mass energy of the collision E_{CMS}. Since the matching scale is an unphysical parameter, there is no sensible theoretically motivated uncertainty. The magnitude of variation follows the usual order of magnitude used in other ATLAS publications and is considered an upper bound.

MLM matching criteria MLM matching [55] is an algorithm for the matching of the matrix element and parton shower used in the $Z \to \tau\tau$ event generation of this study. After the event generation and parton showering a jet cone algortihm is applied. At the event generation, cuts are applied to the transverse momentum and angular separation of the partons. The cut parameter $p_{\text{T,min}}$ was changed to 12 GeV from 15 GeV and indepentently $\Delta R_{\text{min}}(i,j)$ was varied to 0.4 from the default of 0.7. If each jet can be matched to a parton by angular separation ($\Delta R < \Delta R_{\text{max}}$) and each parton matches exactly one cluster, the event is kept. Additionally a minimum transverse momentum of the jets and a minimum angular separation between the jets can be specified. If the process is supposed to be exclusive, events with more jets than hadrons have to be vetoed. In ALPGEN the

MLM matching condition on jet p_T was changed to 15 GeV from its default value of 20 GeV. The minimum angular separation between the jets ΔR was varied to 0.4 from 0.7.

Renormalization scale factor To avoid divergences in Feynman diagram calculations containing loops, a renormalization procedure has to be applied. One possibility is to redefine the theoretical parameters to cancel the UV divergences. They contain the bare quantity and radiative corrections. The renormalized quantities depend on the arbitrarily chosen renormalization scale. The choice of the renormalization scale does not affect the theoretical outcome and therefore has no inherent uncertainty. However, due to the fact that the perturbative calculations are only carried out up to a certain order, a scale dependence is introduced [27]. Usually the renormalization scale is set to the typical energy of the given process and is varied up and down to estimate the uncertainty. The uncertainty decreases with increasing order of the perturbative series calculations. For this study the renormalization scale was changed by 10 % up and down, as performed in [71]. These variations were performed in Sherpa and POWHEG BOX. In ALPGEN the renormalization scale was varied by a factor 2 up and down.

Factorization scale factor Similar to the renormalization scale, a factorization scale is introduced to avoid IR divergences. It separates the domains of soft QCD, which can not be described perturbatively and hard (perturbative) QCD. The factorization scale is varied to study uncertainties due to fixed order calculations. It is changed by a factor of 2 up and down due to convention. These changes were applied in Sherpa, POWHEG BOX and ALPGEN. In POWHEG BOX (for gluon fusion production) it was varied simultaneously with the renormalization scale, according to the usual practice. Additionally in ALPGEN the functional form of the factorization scale definition was changed from the usual $Q = q\sqrt{m_{\tau\tau}^2 + \sum p_{\mathrm{T,jet}}^2}$ to $Q = qm_{\tau\tau}$, where q is the scale factor, $m_{\tau\tau}$ the invariant mass of the two tau leptons and $\sum p_{\mathrm{T,jet}}^2$ the sum of the squared transverse momenta of all jets.

Parton density function Parton density functions describe the probability distribution of the fraction of momentum of partons in a hadron. They most of the time refer to the PDFs of the proton. They are essential for the factorization theorem (see Eq. 4.1.1) and therefore for predictions of QCD processes. PDFs can not be derived from theory and hence are determined experimentally. They depend on the momentum scale of the process examined. Using the DGLAP QCD evolution equation [72, 73, 74] it is possible to calculate the PDFs for other scales. PDFs of different precision and with various parametrizations are provided by multiple collaborations. An important provider of PDFs is the CTEQ collaboration. One of its most recent PDFs, CT10 [75], was used in this study with Sherpa and ALPGEN using the LHAPDF interface. CTEQ also supplied 52 error sets of CT10. They were constructed by varying the 26 free parameters of their PDF parametrization according to their uncertainties. The event generation and analysis were performed for each set. The RMS of the acceptances is taken as total (symmetric) uncertainty arising from the PDF uncertainty. Because ALPGEN does not have access to LHAPDF and error sets are not available with it, two different PDFs were used to estimate the uncertainty. The default PDF for this process, CTEQ6L1 [76], was changed to MRST2001J [77].

Initial/Final state radiation With Pythia it is possible to vary the amount of initial and final state radiation (ISR/FSR). This is achieved by variation of the strong coupling constant α_S for the respective shower, which covers the majority of uncertainty from ISR/FSR. The generator parameters for α_S at the scale of M_Z^2 (M_Z is the mass of the Z boson) are varied up and down by 20 %.

4.3 Rivet event selection

Two decay channels were considered: the semileptonic ($h/H/A \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$) and the purely leptonic ($h/H/A \rightarrow \tau_{\text{lep}}\tau_{\text{lep}}$). The important kinematic cuts of the respective analyses [48] at reconstruction level are implemented at truth level in the Rivet framework as described in the following sections.

4.3.1 Semileptonic decay channel

The analysis in this decay channel is performed separately in the two subchannels $h/H/A \rightarrow e\tau_{\text{had}}$ and $h/H/A \rightarrow \mu\tau_{\text{had}}$. The two (unstable) tau leptons coming from the matrix element are identified by their particle status ID. Their decay trees are analyzed and classified accordingly as leptonic or hadronic. It is made sure that exactly one is hadronic to avoid very rarely occuring events where there are two leptonically or hadronically decaying tau leptons coming from the Higgs boson decay at matrix element level. This supposedly happens due to internal handling of the particles in the generator, when secondary tau leptons have a higher energy than at least one of the primary tau leptons. This effect only appears in Sherpa generated events.

The missing transverse energy $E_{\text{T}}^{\text{miss}}$ is obtained by negatively adding the transverse momentum of all visible final state particles within the pseudorapidity region $|\eta| < 4.9$ (which corresponds to the maximum angular acceptance of the detector). Muons are excluded in this sum, because their energy is not measured by the calorimeters, on which the measurement of $E_{\mathrm{T}}^{\mathrm{miss}}$ is based at ATLAS [78]. To obtain the visible tau decay signatures (all stable[3] particles except neutrinos) of the tau leptons, the sum of the momenta of the stable decay products is calculated. By enforcing a minimum angular separation ΔR between the selected light lepton and visible decay products of the τ_{had}, an overlap is avoided. The following kinematic cuts are applied:

- muon: $p_T > 17\,\mathrm{GeV}, |\eta| < 2.5$
- electron: $p_T > 20\,\mathrm{GeV}, |\eta| < 1.37$ or $1.52 < |\eta| < 2.47$
- tau lepton: $p_T > 20\,\mathrm{GeV}, |\eta| < 2.5$
- angular separation $\Delta R(\tau, e/\mu) > 0.2$

The exluded pseudorapidity range for the electrons is located in the so-called crack region, where the reconstruction of the electron energy is poor due to the geometry of the ECAL.

Then the events are classified according to a variety of categories: b-tag (at least one jet is b-tagged), b-veto (no b-tagged jet), leading b-tag (the leading jet is b-tagged) and leading b-veto (the leading jet is not b-tagged). Additionally there is the inclusive category, as well as each category mentioned before with an additional cut on $E_{\mathrm{T}}^{\mathrm{miss}} > 20\,\mathrm{GeV}$. By introducing b-tagging related categories the sensitivity of the search for neutral MSSM Higgs bosons is improved, because the fraction of the signal and background processes differs between the categories. Most notably the fraction of gluon fusion production events in the b-tagged category is very small, therefore the analysis in this category is customized for b-associated production [48]. The jets are selected using the Anti-kt algorithm [79] with a resolution parameter R of 0.4 and cuts on the jet momentum ($20\,\mathrm{GeV} < p_{\mathrm{T}} < 60\,\mathrm{GeV}$) and its pseudorapidity ($|\eta| < 2.5$). An overlap removal is performed, discarding jets with a jet axis within $\Delta R < 0.2$ around the lepton and the visible τ_{had}. Jets are b-tagged at truth level, if they contain a hadron containing b-flavour. This means there is a tagging efficiency of 100 %, which is clearly different from the actual efficiency at reconstruction level, which depends on the transverse momentum of the jet. In general this study is not supposed to be concerned with experimental effects and mostly it can afford to do so, since most experimental influences have the same effect on the default sample and the sample with parameter variation and therefore have no effect on the acceptance uncertainty. However, it is important that the categorization is as precise as possible since the different resulting uncertainties are

[3]All particles with a decay length significantly bigger than the detector size are considered stable in this context (e.g. muons).

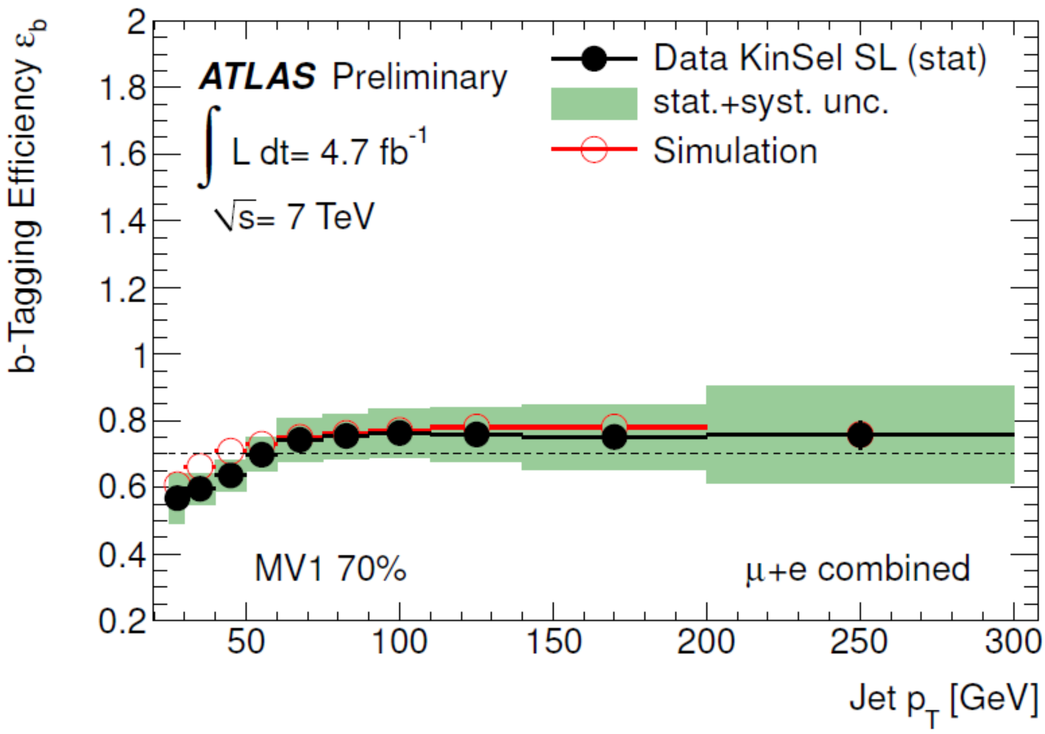


Figure 4.2: Efficiency of b-tagging in dependence of the transverse momentum of the jet in data and simulation samples [80]

applied later to the events in the corresponding category at reconstruction level. The differing tagging efficiency results in different relative numbers of events in the categories, which affects the statistical uncertainty of the acceptance uncertainty. The acceptance uncertainty could change, if a parameter variation results in a different distribution of jet momenta. To avoid these problems, the events in all b-tagging related categories get a weight according to the b-tagging efficiency measured by the ATLAS Top Reconstruction Group [80]. In Fig. 4.2 the b-tagging efficiency is shown with its uncertainties for data and simulation in dependence of transverse momentum of the jet. The tagging was done using `MV1`, a multivariate algorithm combining the output of the `SV0` [81], `I3PD+SV1` and `JetFitterComb` [82] algorithms. The nominal b-tagging efficiency was chosen to be 70 %.

4.3.2 Leptonic decay channel

This channel concentrates on the decay $h/H/A \to e\mu$. It is required to have exactly one muon and one electron passing the following preselection cuts, otherwise the event is vetoed.

- muon: $p_T > 10\,\text{GeV}, |\eta| < 2.5$
- electron: $p_T > 15\,\text{GeV}, |\eta| < 1.37$ or $1.52 < |\eta| < 2.47$

Furthermore the two leptons are required to have opposite charge and an angular separation greater than 0.2. There are three categories, defined similarly to the respective categories in the previous section: b-tag, b-veto and inclusive. The jets are selected using the same algorithm (Anti-kt with $R = 0.4$) and are required to pass the cuts $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $|\eta| < 2.4$. An overlap removal is performed against the selected leptons. Additionally both categories contain the cuts shown in Table 4.1.

	b-tag	b-veto
$E_{\mathrm{T}}^{\mathrm{miss}} + p_T(e) + p_T(\mu) <$	125 GeV	150 GeV
$\Delta\phi(e,\mu) >$	2	1.6
$\sum_{lep=e,\mu} \cos(\Delta\phi(\ell, E_T^{miss})) >$	-0.2	-0.4
$\sum p_T(\mathrm{jet}) <$	100 GeV	-

Table 4.1: Additional cuts of the b-tag and b-veto categories in the leptonic decay channel

4.4 Systematic changes of distributions of kinematic variables

For each considered simulated Higgs boson mass, samples were generated to contain 0.5 million events for each parameter variation using the Monte Carlo generators and parameter variations discussed in Section 4.2. Each of the samples generated with a different PDF error set contained 50,000 events. In the semileptonic decay channel the Higgs masses 90 GeV, 120 GeV, 200 GeV, 450 GeV and 700 GeV were considered, in the leptonic decay channel the masses are 90 GeV, 120 GeV, 200 GeV and 300 GeV. The changes of the di-tau mass shapes due to the considered systematic variations are negligible. As an example in Fig. 4.3, at this statistical level, no significant systematic changes of shape are observed. The same observation is made for the factorization scale variation in Fig. 4.4 and the other variations considered. All plots in this section are normalized to the number of accepted events in the default sample, so that all distributions in each plot have the same integral.

There are, however, distributions which change their shape under parameter variation, especially in the b-veto category and when varying the factorization scale factor. For example, Fig. 4.5 shows the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution for different values of the factorization scale factor. While there are no significant systematic changes of shape in the b-tag category, one can observe a considerable shift of events in the shapes of the b-veto category. Other examples are shown in Fig. 4.6, in which

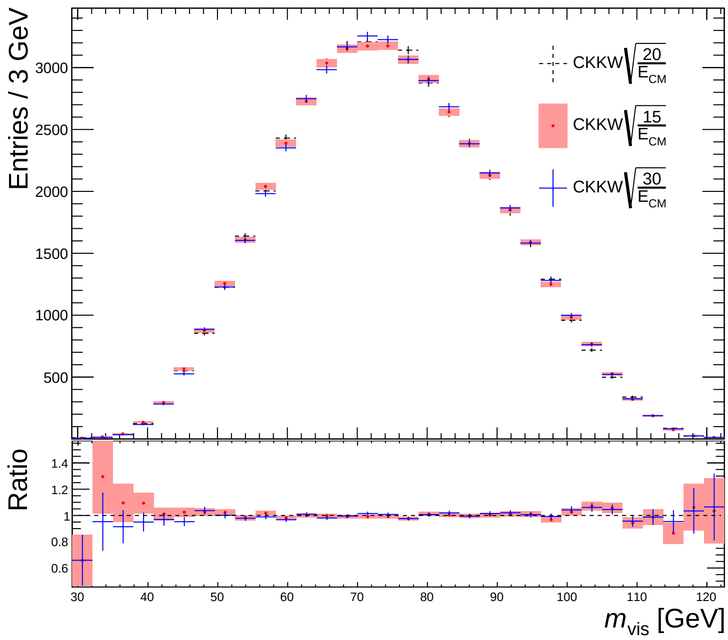


Figure 4.3: Visible mass distributions of the two tau lepton decay products in a b-associated production sample with $m_{\text{Higgs}} = 120\,\text{GeV}$ for different values of the CKKW matching scale in the inclusive category and semileptonic decay channel

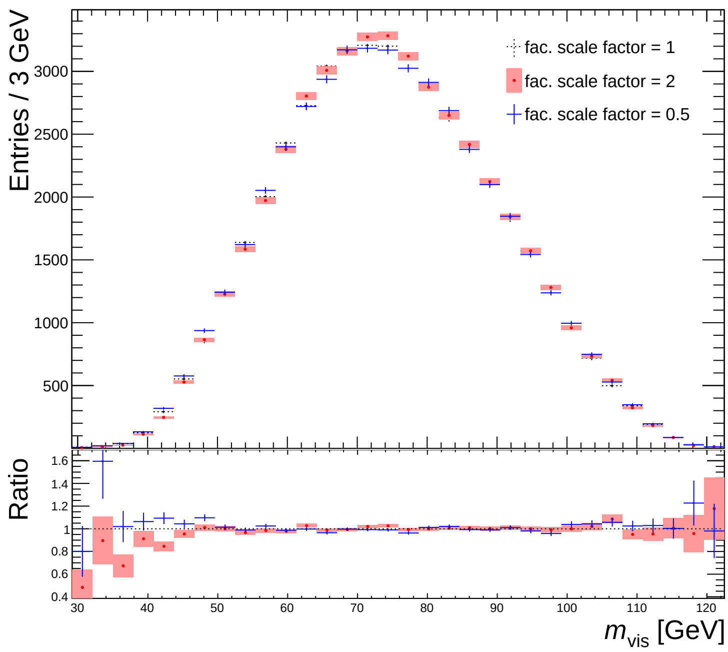
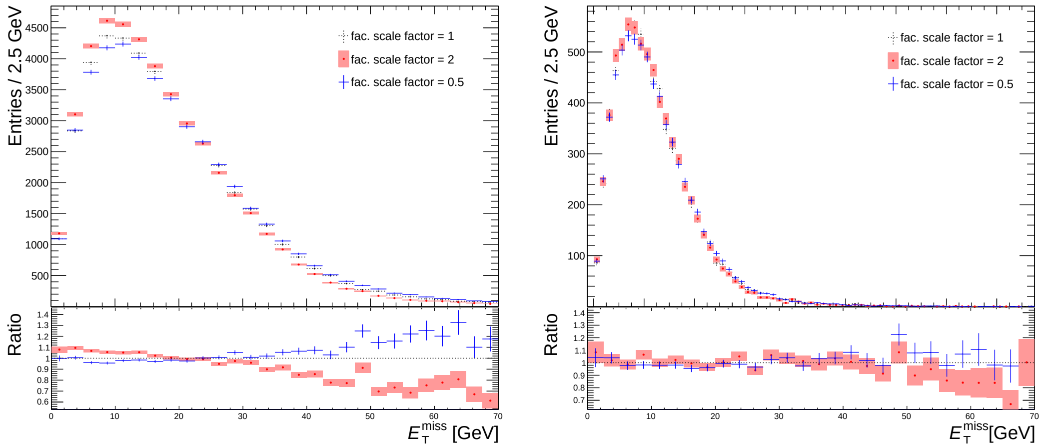


Figure 4.4: Visible mass distributions of the two tau lepton decay products in a b-associated production sample with $m_{\text{Higgs}} = 120\,\text{GeV}$ for different values of the factorization scale factor in the inclusive category and semileptonic decay channel



(a) b-veto category (b) b-tag category

Figure 4.5: Distributions of E_T^{miss} for different values of the factorization scale factor with $m_{\text{Higgs}} = 120\,\text{GeV}$ for b-associated production in the semileptonic decay channel

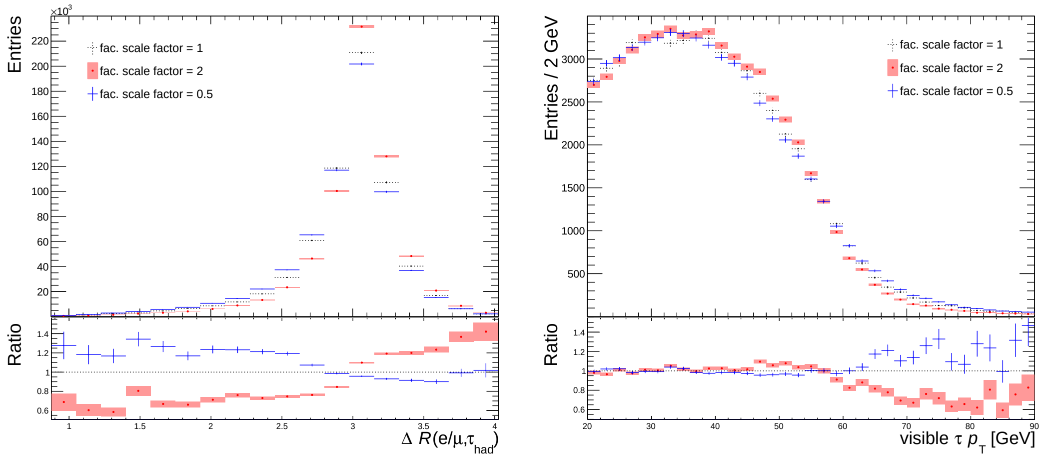


Figure 4.6: Distributions of $\Delta R(\ell, \tau_{\text{had}})$ and $p_{\text{T,had}}^{\text{vis}}$ for different values of the factorization scale factor with $m_{\text{Higgs}} = 120\,\text{GeV}$ for b-associated production in the semileptonic decay channel and b-veto category

the angular separation between lepton and visible τ_{had} as well as the transverse momentum of the visible hadronic tau lepton are shown. Some corresponding plots for the leptonic decay channel are in Appendix A.1.

4.5 Results

The relative uncertainty of the acceptance is defined as in Eq. 4.2, depending on the acceptances A of the sample with modified parameters and of the default sample:

$$\sigma_{\text{acc}}(\text{variation } i) = \frac{A(\text{variation } i)}{A(\text{default})}. \tag{4.2}$$

The absolute statistical uncertainty of the relative systematic uncertainty is derived using Gauss' law for propagation of uncertainties for the two acceptances (see Eq. 4.3):

$$\Delta\sigma_{\text{acc}} = \sqrt{\frac{(\Delta A(i))^2}{A(\text{def})^2} + \frac{A(i)^2}{A(\text{def})^4}(\Delta A(\text{def}))^2}. \tag{4.3}$$

The source of uncertainty of the acceptances ΔA is the statistical uncertainty of the number of events ($A = \frac{N_{\text{accepted}}}{N_{\text{total}}}$). Since the accepted events are a subset of the full set of generated events, the event selection is considered a binomial experiment. The binomial distribution is defined by the total number of events in the given sample and the (true) acceptance A_{true}. It has the variance of $N_{\text{total}}A_{\text{true}}(1-A_{\text{true}})$. The standard deviation is the uncertainty of N_{accepted}, therefore the uncertainty of A is:

$$\Delta A = \frac{\Delta N_{\text{accepted}}}{N_{\text{total}}} = \sqrt{\frac{A(1-A)}{N_{\text{total}}}}.$$

The true acceptance is replaced by its estimator, which is reasonable given the high number of generated events and the precision of the results. This binomial uncertainty approximation is only valid if the number of accepted events is not too close to the total number of events or zero, which is the case for the level of accuracy and the values used in this study. Tables 4.2 - 4.5 show the relative uncertainties of the signal acceptances for a Higgs boson mass of 120 GeV.

Values for the $Z \to \tau\tau$ background are quoted in Table 4.6. Contrary to the other generators used in this study, Alpgen generates samples with a fixed number of hard partons in the matrix element final state. To compare these exclusive samples with data or inclusive simulation samples of other generators, the samples have to be scaled by their respective cross section. This approach often yields more precise results than the parton shower method, because the matrix element evaluation is more exact. Different numbers of jets were used to estimate the dependence of the uncertainty on that value. This is achieved by changing the number of outgoing partons considered by Alpgen in the hard process. There is no statistically significant dependence of the total acceptance uncertainty on the number of jets, which becomes more apparent when symmetrizing the values. It was found that applying b-tagging to all jets or only the leading jet does hardly make a difference for this study. Therefore only the b-tag category is shown in the tables. In the leptonic analysis the acceptance uncertainty of the $Z \to \tau\tau$ background is approximated using a different approach and therefore it is not obtained in this study. The values for all mass points considered are quoted in Appendix A.2.

Total values in the tables are derived by taking the square root of the sum of squared uncertainties for each direction of the deviation. This follows from the law of propagation of uncertainties assuming the different sources of uncertainty are not correlated and the different uncertainties do not have weights different from 1. To correctly calculate the combined uncertainty based on statistical theorems one would have to know the functional dependence of the acceptance on the sources of uncertainty, as well as all elements of the covariance matrix.

The covariance matrix could be determined by repeating the steps of sample generation and applying the event selection for each pair of simultaneous parameter variation, which is strongly limited by computing power at the time of this study.

Event yields	b-tag deviation [%]	b-veto deviation [%]
CKKW down	-4.5 ± 0.8	0.2 ± 0.4
CKKW up	-10.0 ± 0.8	0.9 ± 0.4
Fac. scale up	-22.2 ± 0.7	3.3 ± 0.4
Fac. scale down	14.6 ± 0.9	-3.5 ± 0.4
Ren. scale down	0.2 ± 0.8	-0.6 ± 0.4
Ren. scale up	-0.9 ± 0.8	-0.5 ± 0.4
PDF	± 0.1	± 0.2
Total (positive)	14.6 ± 1.2	3.4 ± 0.7
Total (negative)	-24.7 ± 1.6	-3.6 ± 0.7

Table 4.2: Signal acceptance uncertainties for several systematic deviations of generator parameters for b-associated Higgs production in the semileptonic decay channel for $m_{\text{Higgs}} = 120\,\text{GeV}$

Event yields	b-tag deviation [%]	b-veto deviation [%]
FSR down	-13.4 ± 3.2	0.4 ± 0.3
FSR up	22.6 ± 4.2	-0.1 ± 0.3
ISR down	-15.0 ± 3.2	0.2 ± 0.3
ISR up	21.9 ± 4.1	0.3 ± 0.3
Ren./Fac. scales down	0.9 ± 3.6	-0.2 ± 0.3
Ren./Fac. scales up	0.7 ± 3.6	0.4 ± 0.3
PDF	± 0.0	± 0.1
Total (positive)	31.5 ± 7.8	0.7 ± 0.6
Total (negative)	-20.1 ± 4.5	-0.2 ± 0.4

Table 4.3: Signal acceptance uncertainties for several systematic deviations of generator parameters for gluon fusion production in the semileptonic decay channel for $m_{\text{Higgs}} = 120\,\text{GeV}$

The uncertainties in the b-veto category are significantly smaller than in the b-tag category, because the uncertainties considered are mostly related to

Event yields	b-tag deviation [%]	b-veto deviation [%]
CKKW down	-3.1 ± 0.9	0.4 ± 0.4
CKKW up	-8.3 ± 0.9	2.9 ± 0.4
Fac. scale up	-19.8 ± 0.8	5.6 ± 0.4
Fac. scale down	13.5 ± 1.0	-4.2 ± 0.4
Ren. scale down	0.4 ± 0.9	-0.3 ± 0.4
Ren. scale up	0.8 ± 0.9	0.5 ± 0.4
PDF	± 0.1	± 0.2
Total (positive)	13.5 ± 1.6	6.3 ± 0.8
Total (negative)	-21.7 ± 1.5	-4.2 ± 0.6

Table 4.4: Signal acceptance uncertainties for several systematic deviations of generator parameters for b-associated production in the leptonic decay channel for $m_{\text{Higgs}} = 120\,\text{GeV}$

Event yields	b-tag deviation [%]	b-veto deviation [%]
FSR down	-3.6 ± 6.8	-0.7 ± 0.6
FSR up	16.6 ± 7.8	-0.2 ± 0.6
ISR down	3.6 ± 7.2	0.4 ± 0.6
ISR up	20.3 ± 8.1	-1.2 ± 0.6
Ren./Fac. scales down	2.5 ± 7.1	-0.5 ± 0.6
Ren./Fac. scales up	9.4 ± 7.5	0.0 ± 0.6
PDF	± 0.0	± 0.1
Total (positive)	28.2 ± 16.9	0.4 ± 0.6
Total (negative)	-3.6 ± 6.8	-1.5 ± 1.2

Table 4.5: Signal acceptance uncertainties for several systematic deviations of generator parameters for gluon fusion production in the leptonic decay channel for $m_{\text{Higgs}} = 120\,\text{GeV}$

uncertainties of QCD calculations and events of the b-veto category are less likely to contain QCD jets. Except of very few instances, the results of the semileptonic and the leptonic decay channel are very similar, as the uncertainties are mostly compatible within the given statistical uncertainty intervals. For b-associated MSSM Higgs production the factorization scale uncertainty has the biggest effect, while the influence of the renormalization scale variation is negligible. The CKKW matching scale factor variation leads to significant uncertainties in the b-tag category, but has only a minor impact in the b-veto category. Only very small acceptance uncertainties are observed for the variation of the renormalization scale, which suggests that the variation factor is too small. The uncertainty arising from the PDF variation is negligible. For gluon fusion production the variations of ISR and FSR have the biggest effect in the b-tag category. Increasing ISR or FSR results in bigger acceptance uncertainties than those arising from decreased

Sources	Uncertainty [%]		
$Z \to \tau\tau + \text{jets}$	1 jet	2 jets	4 jets
Matrix element parton 12 GeV	0.5 ± 0.4	-1.8 ± 0.6	-4.4 ± 3.7
Matrix element parton ΔR >0.4	0.0 ± 0.4	0.4 ± 0.6	1.2 ± 2.3
Different scale definition (iqopt2)	0.3 ± 0.4	0.3 ± 0.5	-2.7 ± 1.6
Different PDF set (MRST2001J)	6.4 ± 0.4	3.7 ± 0.5	0.0 ± 1.6
Q^2 Scale (double)	-0.1 ± 0.4	0.2 ± 0.5	1.0 ± 1.6
Q^2 Scale (half)	-0.8 ± 0.4	-1.9 ± 0.5	4.1 ± 1.6
MLM matching condition (p_{T} 15 GeV)	-3.2 ± 0.4	-5.1 ± 0.4	-4.5 ± 1.4
MLM matching condition (ΔR 0.4)	0.7 ± 0.4	1.4 ± 0.5	4.2 ± 2.1
kt factor (double)	1.5 ± 0.4	-1.1 ± 0.5	-0.4 ± 1.5
kt factor (half)	0.2 ± 0.4	0.5 ± 0.5	-3.0 ± 1.7
Total (positive)	6.6 ± 1.1	4.0 ± 1.3	6.1 ± 4.2
Total (negative)	-3.3 ± 0.7	-5.8 ± 1.1	-7.5 ± 4.8

Table 4.6: Systematic uncertainties for the ALPGEN Z+jets Monte Carlo sample in the semileptonic decay channel

ISR/FSR. The variation of the renormalization and factorization scales leads to acceptance uncertainties which are mostly compatible with zero in gluon fusion samples. All uncertainties considered have no significant impact in the b-veto category of gluon fusion production samples. While for gluon fusion production samples the uncertainties for different mass points are, due to their statistical uncertainty, compatible with the hypothesis that there is no mass dependence, there are distinct trends observed for b-associated production events. For instance, the acceptance uncertainty arising from the variation of the factorization scale is decreasing for an increasing mass of the simulated Higgs boson (see Table 4.7 or Appendix A.2).

In this study also the impact of an additional cut on missing transverse energy was analyzed. The respective acceptance uncertainties are quoted in Appendix A.3. Generally the differences between these uncertainties and those without the additional cut are negligible at high Higgs masses. The biggest difference is observed at 90 GeV with the additional cut resulting in slightly bigger acceptance uncertainties, with the exception of the b-veto category in b-associated events, where a descrease of acceptance uncertainty is observed at low Higgs boson masses.

The uncertainties presented in this chapter are used among other uncertainties in the studies in search for neutral MSSM Higgs bosons in the corresponding decay channels [83]. Because the acceptance uncertainties in the b-veto category of gluon fusion production samples are below 1 % and the percentage of b-tagged events in gluon fusion production samples is very small, typically below 0.5 %, the

variation	90 GeV	120 GeV	200 GeV	300 GeV
CKKW down	-2.0 ± 1.1	-3.1 ± 0.9	-6.3 ± 0.9	-6.5 ± 1.3
CKKW up	-19.0 ± 1.0	-8.3 ± 0.9	2.5 ± 0.9	6.7 ± 1.5
Fac. scale down	14.9 ± 1.3	13.5 ± 1.0	8.2 ± 1.1	7.7 ± 1.5
Fac. scale up	-23.0 ± 1.0	-19.8 ± 0.8	-14.1 ± 0.8	-10.1 ± 1.3
Ren. scale down	-0.1 ± 1.1	0.4 ± 0.9	-0.4 ± 0.9	0.3 ± 1.4
Ren. scale up	-1.4 ± 1.1	0.8 ± 0.9	0.8 ± 1.0	1.3 ± 1.4
PDF	±0.1	±0.1	± 0.1	±0.0
total (positive)	14.9 ± 1.3	13.5 ± 1.6	8.6 ± 1.7	10.3 ± 2.9
total (negative)	-29.9 ± 2.4	-21.7 ± 1.5	-15.4 ± 1.5	-12.0 ± 1.8

Table 4.7: Acceptance uncertainties for b-associated production in the leptonic decay channel and b-tag category for different Higgs boson masses

influence of the determined uncertainties on the analyses in search for neutral MSSM Higgs bosons is negligible. The acceptance uncertainty of the $Z \rightarrow \tau\tau$ background process was found to be around 6 %, which is of the same order of magnitude as the uncertainties of the $Z \rightarrow \tau\tau$ cross section and detector related uncertainties in this background process. Monte Carlo acceptance uncertainties in b-associated production samples are of the order of 10 - 30 % in the b-tag category and 2 - 4 % in the b-veto category. They are of the same order of magnitude as the detector related uncertainties as well as the uncertainty of the signal process cross sections and therefore play an important role for the respective neutral MSSM Higgs boson searches.

Chapter 5

Influence of experimental systematic uncertainties on the tau identification BDT score distribution

For the ATLAS data analysis it is essential to have reliable algorithms for reconstruction of physics objects. This includes single particles and vertices and their properties as well as composite objects like jets and the identification of their type. The efficiency of these algorithms has a major impact on physics studies. Therefore it is important to study the influence of experimental uncertainties on these algorithms. The goal of this study is to analyse these influences on the tau identification algorithm (called tau ID) [84]. It was designed to identify hadronically decaying tau leptons in the ATLAS detector, which primarily involves their distinction from QCD jets.

Contrary to other studies on this behalf [85, 84] the uncertainty arising from each source of experimental uncertainties is studied separately. This makes it possible to make quantitative statements about the influence of each of the uncertainties and their contribution to the total uncertainty. As a result not only the integrated uncertainties are given, but also the changes of the shape of the tau ID classifier are presented. In Section 5.1 and 5.2 the algorithms for reconstruction and identification are described. The experimental uncertainties considered are reviewed in Section 5.3. The method of this study is outlined in Section 5.4. Sections 5.5 and 5.6 contain the results in form of systematic changes of the shape of the identification algorithm output and integrated systematic uncertainties.

5.1 The hadronic tau candidate reconstruction procedure

Jets constructed by the Anti-kt algorithm ($R = 0.4$, $|\eta| < 2.5$ and $p_{\mathrm{T,jet}} > 10\,\mathrm{GeV}$) are the input for the reconstruction algorithm [84]. The jet algorithm is seeded with the tracks and topological clusters [30]. Topological clusters are groups of calorimeter cells. The cells adjoining the seed cells or the border of the cluster are added to the cluster as long as their energy surpasses the expected noise. Using the jets as a starting point, the properties of the hadronic tau candidate are calculated. The Tau Jet Vertex Association (TJVA) algorithm finds the vertex with the highest probability of being the primary vertex of the candidate as described in Ref. [85]. Using clusters around the barycenter of all clusters within $\Delta R < 0.2$, the intermediate axis of the candidate is built as a sum of the four-vectors of the clusters recalculated pointing to the TJVA vertex. The reconstructed energy is calibrated at the tau energy scale (TES), which is derived by a Monte Carlo method [86]. The TES is determined in dependence of pseudorapidity, energy of calorimeter clusters corrected by the Local Hadron Calibration scheme [87] and the number of tracks of the hadronic tau candidate. It is derived by comparing the true energy of hadronically decaying tau leptons with the reconstructed energy in simulated events. Around the intermediate axis two regions are defined: the core cone ($\Delta R < 0.2$) and the isolation annulus ($0.2 \leq \Delta R < 0.4$). Tracks in the core cone are associated to the hadronic tau candidate, if they satisfy the following requirements:

- $p_{\mathrm{T,track}} \geq 1\,\mathrm{GeV}$
- number of Pixel hits + number of SCT hits ≥ 7
- number of Pixel hits ≥ 2
- $|d_0| \leq 1.0\,\mathrm{mm}$ (unsigned closest approach of the track to TJVA vertex in the transverse plane)
- $|z_0 \sin\theta| \leq 1.5\,\mathrm{mm}$ (z_0: longitudinal component of the point of closest approach)

The number of tracks in the core cone determines the classification of the tau candidate into 1-prong, 3-prong and/or multi-prong ($\geq$ 2 tracks). Tracks in the isolation annulus satisfying the same requirements are stored in a separate collection of variables to refine the identification process.

5.2 The tau identification algorithms

Currently there are two algorithms for tau identification under discussion [84], both are multivariate. One uses the maximum-log-likelihood estimation method to obtain a discriminating variable. The other one, which this study is focused on, uses boosted descision trees (BDTs) [88]. Unlike the cut-based approach, multivariate methods take multiple variables into account at once, in case of the BDT also including their correlation. A decision tree is a binary tree where at each node an object (defined by a number of variables) is classified into one of the two subsequent nodes according to the cut criterion of this node. The goal is to divide a sample of objects into two classes, called signal and background. In the case of the tau ID BDT a hadronic tau candidate is classified whether it is a real hadronically decaying tau lepton (signal) or not (background). The BDT is trained using a training sample including the Monte Carlo information whether a candidate is a hadronically decaying tau lepton. At each node the variable with the best discrimination power and the respective cut value is determined. The process is terminated when a node is pure, then it is called a leaf, or at a specified termination condition, e.g. maxmimum number of nodes or leaves. To avoid the effect of overtraining, the BDT classification is applied on a testing sample, which is statistically independend from the training sample. The output of the BDT is the BDT score, a scalar variable which is related to the probability whether the given object is of the signal class. It can be used as a weight in the analysis or as a cut criterion, where the cut value determines the resulting sample purity. For the hadronically decaying tau lepton identification each candidate is tagged whether it falls into one of the quality regions: loose, medium and tight, which are defined by a minimal signal efficiency of 70 %, 60 % and 40 % for 1-prong tau candidates and 65 %, 55 % and 35 % for multi-prong tau candidates. Additionally the method of boosting is introduced to avoid instabilities of the tree. This involves training a large number of decision trees with the total BDT score derived as a weighted average of the individual trees. Each time an object is misclassified, its weight is increased (boosted). Another decision tree is trained using the sample with modified weights. As an example for the discriminating power in Fig. 5.1 the distributions of the BDT score is plotted for a signal and a background sample for both 1- and 3-prong candidates.

For the tau identification two BDTs are trained using different sets of variables: one for 1-prong and one for multi-prong tau candidates. The BDT input variables are listed below [84].

- **track radius**: sum of angular separation of charged tracks to the interme-

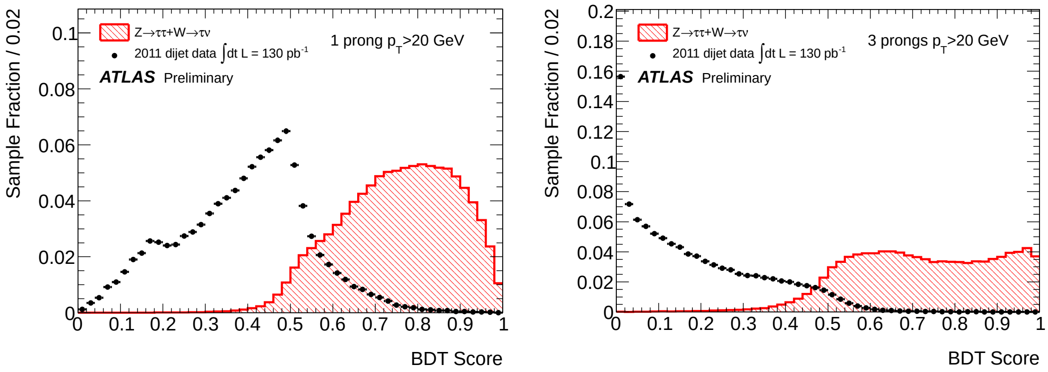


Figure 5.1: BDT score for 1-prong (left) and 3-prong (right) candidates for a signal and a background sample [85]

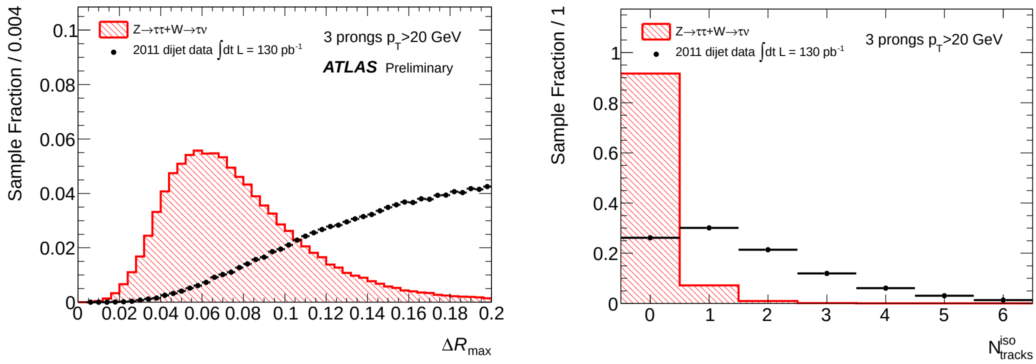


Figure 5.2: Maximum angular separation ΔR_{max} (left) and number of tracks in the isolation annulus $N_{\mathrm{trk}}^{\mathrm{iso}}$ (right) for 3-prong candidates for a signal and a background sample [85]

diate tau axis, weighted by their transverse momentum:

$$R_{\mathrm{track}} = \frac{\sum\limits_{i:\Delta R_i \leq 0.4} p_{\mathrm{T},i} \Delta R_i}{\sum\limits_{i:\Delta R_i \leq 0.4} p_{\mathrm{T},i}} \, .$$

- **leading track momentum fraction**: ratio of the transverse momentum of the leading track in the core cone to the sum of transverse energies in electromagnetic and hadronic calorimeter cells within the core cone, calibrated at the electromagnetic scale.

  Because of the high luminosity most of the time multiple interactions are captured in the same event. This is called pile-up and complicates the event analysis, because the different interactions have to be separated, which in the majority of cases is not completely possible, especially if the detector signals of different interactions overlap. For the leading track momentum fraction there is an additional term for pile-up correction which depends on the number of primary vertices N_{vtx}:

$$f_{\mathrm{track}}^{\mathrm{corr}} = \frac{p_{\mathrm{T,leadtrk,core}}}{\sum\limits_{i:\Delta R_i \leq 0.2} E_{\mathrm{T},i}^{\mathrm{EM}}} + 0.003 \times N_{\mathrm{vtx}} \ .$$

- **core energy fraction**: ratio of energy in the central region ($\Delta R < 0.1$) of the core to the total core energy.

  This variable is corrected for pile-up, if the transverse momentum of the tau candidate is less than 80 GeV.

$$f_{\mathrm{core}} = \frac{\sum\limits_{i:\Delta R_i \leq 0.1} E_{\mathrm{T},i}^{\mathrm{EM}}}{\sum\limits_{i:\Delta R_i \leq 0.2} E_{\mathrm{T},i}^{\mathrm{EM}}}$$

$$f_{\mathrm{core}}^{\mathrm{corr}} = \begin{cases} f_{\mathrm{core}} + 0.003 \times N_{\mathrm{vtx}} & \text{if } p_{\mathrm{T},\tau} < 80\,\mathrm{GeV} \\ f_{\mathrm{core}} & \text{otherwise} \end{cases}$$

- **transverse flight path significance** (multiprong only): secondary vertex decay length divided by its estimated uncertainty:

$$S_{\mathrm{T}}^{\mathrm{flight}} = \frac{L_{\mathrm{T}}^{\mathrm{flight}}}{\delta L_{\mathrm{T}}^{\mathrm{flight}}} \ .$$

- **leading track impact parameter significance** (1-prong only): shortest distance between leading core track and secondary vertex in the transverse plane divided by its estimated uncertainty:

$$S_{\mathrm{leadtrk}} = \frac{d_0}{\delta d_0} \ .$$

- **maximum angular separation $\Delta R_{\mathbf{max}}$** (multiprong only): maximum angular separation between core tracks and intermediate tau axis.

  As shown in Fig. 5.2 the visible decay products of hadronically decaying tau leptons tend to be more collimated than QCD jets.

- **number of tracks in the isolation annulus $N_{\mathbf{trk}}^{\mathbf{iso}}$** (1-prong only)

  Because hadronically tau leptons tend to decay in a narrower cone, their number of tracks in the isolation annulus is significantly smaller on average (see Fig. 5.2).

- **invariant mass of track system $m_{\mathbf{tracks}}$** (multiprong only): mass of the tracks in the core and isolation region.

- **number of π^0 mesons**:

  Because neutral pions do not leave information in the inner detector and their calorimeter energy deposits overlap with those of the charged pions, it is very difficult to reliably reconstruct them. To get an estimate of the number of neutral pions, two additional BDTs are used: the primary and secondary π^0 BDT. They have the same input variables, but different categories. One is for the classification of zero neutral pions, the other distinguishes between one or two of them. The input variables are listed below:

  – number of strip layer cells with more than 200 MeV energy content, N_{strip}

  – fraction of calorimeter energy in the presampler and strip layers

  – inverse leading momentum fraction (not corrected for pile-up)

  – ratio of energy in the electromagnetic calorimeter to the total momentum of tau tracks

  – difference of energy of tau tracks (with charged pion mass) and energy content in the hadronic calorimeter divided by the total energy in the electromagnetic calorimeter

- **visible mass of the hadronic tau candidate**: This variable is calculated using the software module `Pi0Finder`, which was designed to find the most probable π^0 energy clusters.

- **ratio of the transverse momentum of the hadronic tau candidate including the reconstructed neutral pions to the total transverse momentum of the charged pions**

5.3 Sources of experimental uncertainty of tau identification

In this study primarily two sources of experimental systematic uncertainties were considered: the uncertainty on the calorimeter cell energy measurement and on the track reconstruction efficiency. Additionally different shower shape models were tested.

5.3.1 Calorimeter cell energy uncertainty

The energy measurement in the individual calorimeter cells has to be calibrated to give meaningful results. Their energy response is studied using test beams of known energy and intensity as well as live calibration systems like radioactive sources, charge injection systems and lasers [14]. Important parts of the calibration are done by comparisons of recorded data and Monte Carlo simulations. For this study the results of an ATLAS study using minimum bias data[1] were applied [89]. An import source of uncertainty considered is the inaccuracy of the knowledge of the material upstream of the calorimeter, mainly from the tracking detectors and their services. Most of the material between calorimeter and beamline was not in place at the time of the test beam studies and is modelled only approximately. The energy loss in the dead material is simulated. Another source of uncertainty analyzed is the accuracy of the Monte Carlo simulated calorimeter energy response. Additionally the dependence of the Monte Carlo model used for correction of detector effects was studied. The resulting uncertainties of the energy density are given as a function of the pseudorapidity and are listed in Table 5.1. The uncertainties for hadronic and electromagnetic decays were determined separately and averages were obtained, which are listed in Table 5.1. All ATLAS calorimeters were considered.

$\lvert\eta\rvert$	relative uncertainty on cell energy measurement [%]
0.0 - 0.8	$^{+5.1}_{-4.9}$
0.8 - 1.6	$^{+7.9}_{-7.5}$
1.6 - 2.4	$^{+7.4}_{-7.2}$
2.4 - 3.2	$^{+4.6}_{-6.4}$
3.2 - 4.0	$^{+6.0}_{-8.7}$
4.0 - 4.8	$^{+4.6}_{-8.2}$

Table 5.1: Total cell energy density uncertainty [89]

5.3.2 Track reconstruction efficiency uncertainty

Charged tracks in the ATLAS detector are usually reconstructed by applying advanced fitting algorithms to the inner detector hits [14]. Due to high track densities, especially at high pile-up scenarios, the association of detector hits to specific tracks is not always unambiguous. Furthermore a track is discarded, if it fails minimal quality requirements. The track reconstruction efficiency is available as a function of pseudorapidity and transverse momentum. It is measured using simulated event samples and is defined by the number of reconstructed

[1] datasets with minimal trigger requirements

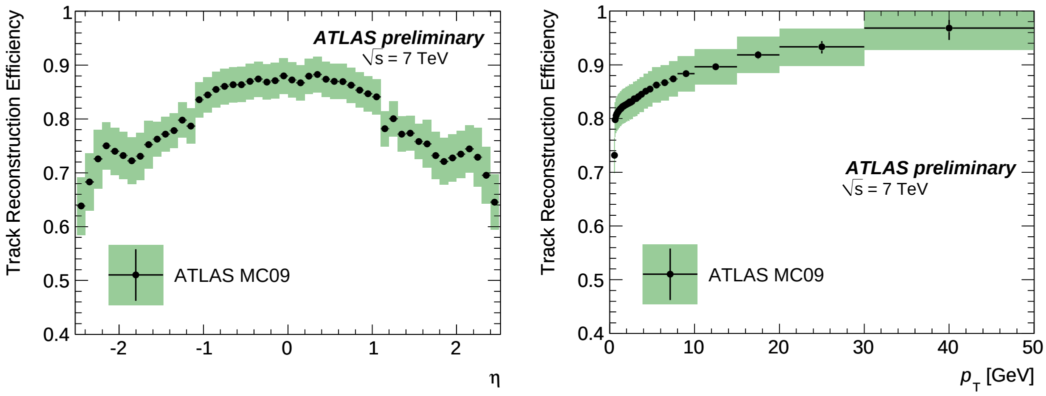


Figure 5.3: Track reconstruction efficiency and its total uncertainty (statistical uncertainty only: black vertical error bars) in dependence of the pseudorapidity and transverse momentum [90]

tracks divided by the number of generated charged particles. The most important uncertainty is the limited accuracy of the material description. It is determined in an ATLAS study [90] by using a variety of configurations with additional detector material and comparing the resulting reconstruction efficiency to the nominal one. The simulated amount of material in the support structure was increased so that the total amount of material in the Inner Detector is 10 %-20 % larger than in the nominal configuration. Additional uncertainties arise due to the track quality cuts. Because these are strongly dependend on the values of the quality cuts, the process studied and the version of the ATLAS reconstruction software, they are not considered in this study, as these influences differ between analyses. Available values for track reconstruction efficiency uncertainties arising from track quality cuts from studies before 2012 can not be directly applied to event samples processed with recent versions of the reconstruction software. The values for the track reconstruction efficiency and its material uncertainty are supplied by the `InDetTrackCorr` software. Fig. 5.3 shows the track reconstruction efficiency and its total uncertainty in dependence of pseudorapidity and transverse momentum.

5.3.3 Shower shape model dependence

In Geant4 [57] there is an option to set the so-called physics list, which defines a string model for the simulation of hadron interactions with matter at high energies as well as models for low energy interactions. It primarily affects the resulting shapes of hadronic showers. By comparison to test beam data and data recorded with the full detector, these models are improved to match the observed shower shapes. However, there is no obvious best model, since some of the models are quite close to each other. Studies analysing the agreement of the different models with data are ongoing and the default ATLAS physics list might change in the future. In this study uncertainties of the shower shape are not covered by the cell energy variation, as it does not reflect changes in the shower profile. To study different shower shapes, other than the default list QGSP_BERT, the lists QGSP and FTFP_BERT were used [91]. QGS stands for quark gluon string model [92]. All models considered use the Geant4 Precompound model (P) for hadronic interactions at lower energies. FTF models use the Fritiof string model [93]. The BERT suffix stands for models implementing the Bertini cascade model [94], which is applied at medium energies.

5.4 Method

The main goal of this particular study is to determine systematic changes of the tau identification BDT score distribution caused by each uncertainty described in the section above. This kind of information is not covered by conventional uncertainty studies, which compare data to simulated events. In this study the uncertainties were applied to simulated events before the identification process. A modified version of the identification module `TauDiscriminant` was developed to allow for changing the relevant event variables on the fly. By this approach the requirement to create samples of modified events beforehand, causing a significant overhead, is avoided.

Cell energy uncertainties are applied by varying cell energy sums and cluster energies coherently in the same direction and the size of the variation depends on the pseudorapidity according to Table 5.1. To cover the track reconstruction efficiency uncertainty, tracks are removed randomly according to a track deletion probablity, defined as

$$\text{track deletion probability} = \frac{\text{absolute material uncertainty of reco. eff.}}{\text{reconstruction efficiency}},$$

as a function of transverse momentum and pseudorapidity. In order to supress the statistical effect of the introduced randomness, the whole classification is repeated 50 times with different random seeds. To prevent an overlap of the uncertainty with the hadronic tau reconstruction efficiency itself, the described process is applied in two different scenarios. In scenario A only tracks in the isolation annulus are considered. Therefore the number of reconstructed hadronic tau candidates is not changed. In scenario B only tracks in the core cone are removed according to the given probability, but it is taken care not to change the number of 1- and 3-prong tau candidates. This is achieved by only looking at candidates with an odd number of core tracks before removal and afterwards the result is normalized to match the total number of 1- and 3-prong candidates of the nominal sample.

Hence, in scenario B only the effect of 3-prong candidates loosing two tracks at once is considered, keeping the total number of reconstructed candidates constant.

It is not feasible to apply the uncertainty arising from the hadronic shower shape directly in `TauDiscriminant`, since it is impracticable to model the change of the event variables. Indeed it is possible to model the changes of lateral and longitudinal shower dimensions using data from test beam studies and by producing special D3PDs containing cell information. However, then the topological clusters and thus the whole tau candidate reconstruction would have to be rebuilt, which happens at a lower reconstruction level and requires a significant amount of additional disk space and computing time. Furthermore it is to be expected that due to the approximative nature of the modelling additional uncertainties would be introduced. Therefore the uncertainty due to the hadronic shower shape is studied by using the official ATLAS event samples with different Geant4 physics lists and process them using the nominal classification.

For this method to work, it is essential that all BDT input variables can be recalculated from the available event variables. This was tested by comparing the recalculated variables to those included in the event sample file in the scenario of no modification of the event variables. In the majority of cases they were in perfect agreement, difficulties are discussed in Section 5.4.1. In addition, `TauDiscriminant` was modified to write out a copy of the input file with a greatly reduced number of variables and with updated values for some of them. The whole process is repeated once on the output file of the first run, since only then the updated scores of the π^0 BDTs are available. Afterwards the BDT shapes are determined under different cut conditions and compared to the nominal shape. Integrated systematic uncertainties are calculated as the relative deviation of the number of tau candidates classified in the individual quality regions loose, medium and tight.

5.4.1 Problems recalculating BDT input variables

While in most of the cases all of the dependencies are included in the standard tau D3PD files and therefore the respective BDT input variable can be recalculated precisely, there are a few cases where information from lower reconstruction levels is not available anymore. In the following the problematic cases are reviewed and it is explained why they do not compromise the method.

Core energy fraction (for the definition see page 44)
Despite the individual cell energies not being available in the standard D3PDs, it is possible to precisely recalculate the variable using event variables containing the sum of cell energies in certain ΔR ranges around the intermediate tau axis.

However, the cell energy uncertainty can not be handled correctly, since the pseudorapidity dependence can only be approximated using the pseudorapidity of the tau axis. It is expected that only a small fraction of events is affected, since the bin width of the given uncertainty ($\Delta\eta = 0.8$) is big compared to the size of the tau cone ($\Delta R = 0.2$). In case all cells in the numerator and denominator are located in the same pseudorapidity bin, the cell energy uncertainty has no effect on the core energy fraction, since the modifying factors cancel out. To study the impact of this problem, the cell energies were modified before reconstruction in the module `tauRec`. Because this requires a relatively large amount of computing time, only small samples were reconstructed. The outcome was compared with the approximative treatment and the BDT score differences were found to be negligible within statistical uncertainties.

$N_{\mathbf{strip}}$ (defined on page 45)
Because of the unavailability of the individual cell information, this variable can not be recalculated at D3PD level. When changing the cell energy it has to be expected that the number of strip layer cells above the energy threshold of 200 MeV changes as well. To estimate the change of N_{strip}, the correlation with the total energy in the prescaler and strip layer shown in Figure 5.4 is used. Since the fraction of energy in prescaler and strip layer does not change due to the modified energy, this variable can be recalculated. The profile of the correlation is approximated analytically. A 3rd degree polynomial fit was found to match very well, with a linear continuation at high energies to match the saturation effect. The slope is relatively small, so that only rarely the energy changes enough to change N_{strip} by 1. Therefore the exact form of the estimation is not of much importance. This approximation was tested using the variation samples produced with `tauRec` and found to match perfectly at the given level of statistical accuracy.

Invariant mass of the associated tracks (defined on page 44)
This variable is originally calculated using floating point values of double precision. However, for recalculation only single precision floating point values are available. This leads to a small fluctuation in BDT score distribution ($< 0.2\,\%$ at 0.05 bin width), which shows no dependence on the BDT score itself.

Leading track impact parameter significance (definition on page 44)
In case the leading track is removed, the leading track impact parameter significance is supposed to change. However, the estimated uncertainty of the impact parameter of the subleading track is not stored, which makes it impossible to model the effect of the uncertainty on this variable. No correlation with available

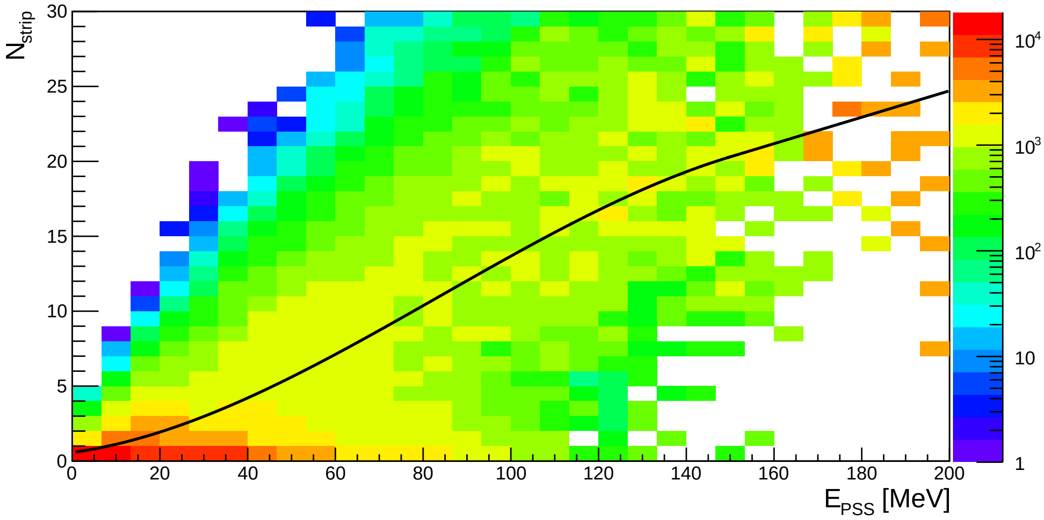


Figure 5.4: N_{strip} vs total energy content in prescaler and strip layers E_{PSS}, reweighted according to the relative frequency of E_{PSS}. A 3rd degree polynomial fit with a linear continuation above 150 MeV is superimposed in black.

variables were found, which prevents an approximation approach similar to the handling of N_{strip}. The variable is recalculated changing only the impact parameter, but not its estimated uncertainty. The effect was studied removing the leading track of all candidates at `tauRec` level and compare this with the outcome of the inaccurate calculation. As shown in Fig. 5.5 in most of the cases the difference of BDT scores is very close to zero. There are a number of events with a score difference bigger than 0.05, especially in the negative direction (smaller scores for the inaccurate way). The whole effect is greatly reduced, because the leading track can only be removed in scenario B, where two tracks are removed. This happens only very rarely and it is yet rarer that the leading track is affected.

5.5 Systematic uncertainties on the shape of the BDT output of the tau identification

The $Z \to \tau\tau$ samples used for this study were generated using the ALPGEN and HERWIG Monte Carlo generators and contained 25,000 events and about 130,000 hadronic tau candidates. Pythia 8 $Z \to \tau\tau$ samples with 25,000 events were used to test different Geant4 physics lists. The following kinematic cuts were applied: $p_{\text{T},\tau} > 20\,\text{GeV}$ and $|\eta| < 2.5$. There is a significant change in the shape of the BDT score distribution due to the variation of cell energies, as shown in Fig. 5.6 and 5.7. It is more easily to spot in the integrated views on the right. For both 1-prong and 3-prong candidates, less energy leads to some candidates getting

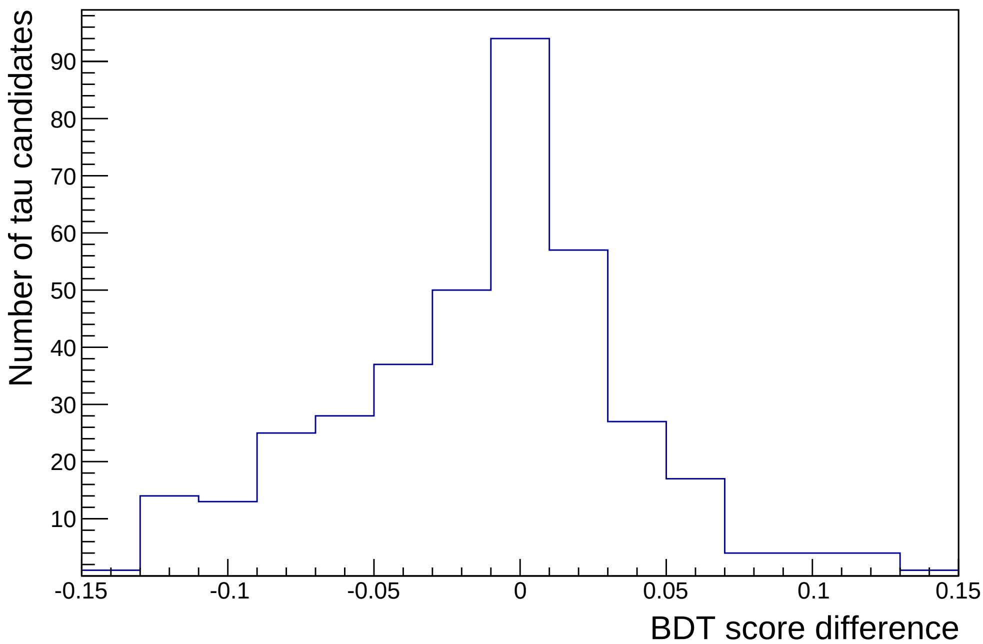


Figure 5.5: Difference of BDT scores of the cases with accurate and inaccurate leading track impact parameter significance calculation. Only tau candidates with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$ are considered.

higher scores and vice versa. Neither one of the track reconstruction efficiency uncertainty scenarios has a visible impact at this statistical level. This can be explained by the relatively small deletion probability for the tracks of the order of 2 %. In scenario B the probability of a candidate loosing two tracks is therefore of the order of 0.04 %. The BDT seems to be very robust against lost tracks in the isolation annulus. Significant systematic changes of shape can be observed when using different Geant4 physics lists, as pictured in Fig. 5.9 and 5.10.

Shown in Fig. 5.8 are the distributions of the π^0 primary and secondary BDT scores for different calorimeter cell energy variations. From these BDT scores an estimation of the number of neutral pions of the tau candidate is calculated. It is zero if the primary BDT score is equal to or greater than 0.465, otherwise it is one for secondary BDT scores above 0.565 or two below that score. Fig. 5.8 shows that an increased cell energy leads to a higher estimated number of neutral pions on average.

The individual systematic uncertainties from cell variation and hadronic shower model are combined quadratically in each direction and then symmetrized. Together with the statistical uncertainties the combined results are shown in Fig. 5.11. Histogram bins outside the pictured range have uncertainties above 100 % due to the very small number of events in that regions and are therefore omitted. As observed for the individual variations, the largest systematic uncertainty is

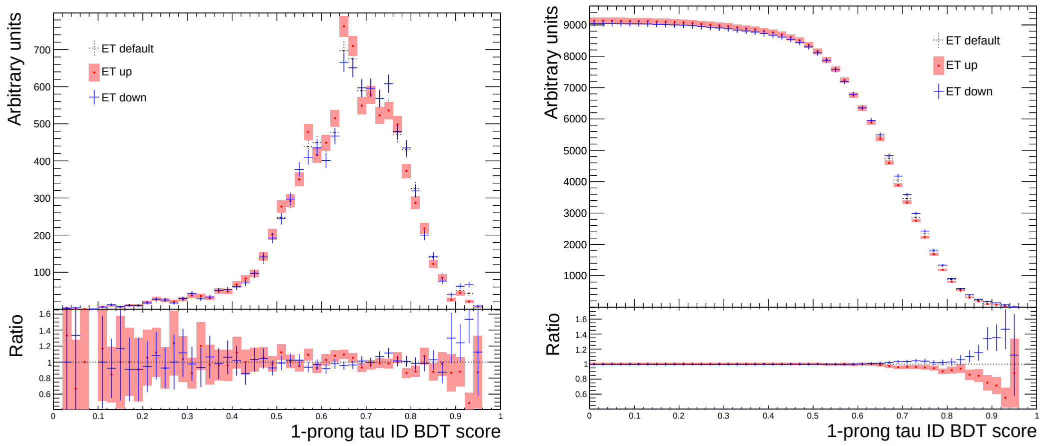


Figure 5.6: Tau ID BDT score for 1-prong candidates under cell energy variation and integrated distribution

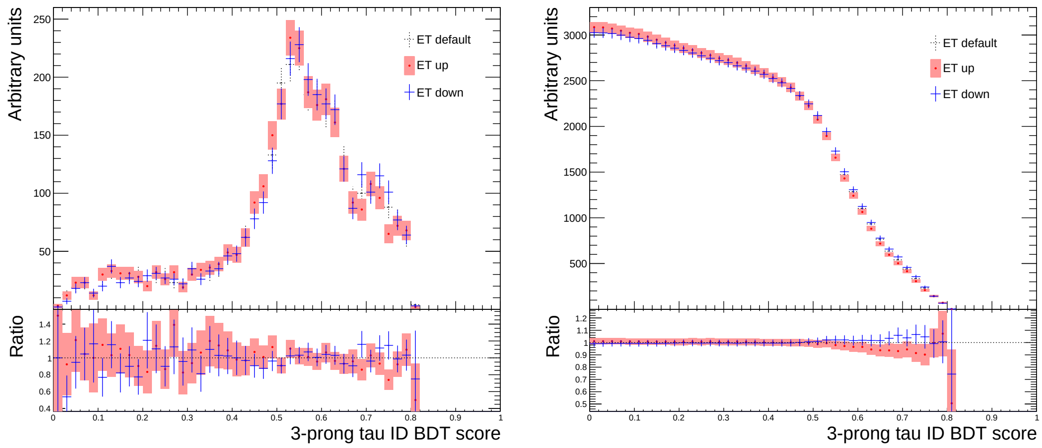


Figure 5.7: Tau ID BDT score for 3-prong candidates under cell energy variation and integrated distribution

observed at high 1-prong BDT scores. Uncertainties of 1-prong hadronic tau candidates with BDT scores between 0.5 and 0.8 are comparably small. Bigger uncertainties for small 1-prong BDT scores are dominated by statistical uncertainties. Overall the dependence of the uncertainty on the 3-prong BDT score is less intense, but the effect of large uncertainties at the highest scores and a drop of uncertainty directly below is observed, too.

5.6 Integrated systematic uncertainties of the tau identification BDT output

In Table 5.2 the relative uncertainties on the total number of 1- and 3-prong candidates $N_{\text{1-prong}} + N_{\text{3-prong}}$ are given for the sources of systematic uncertainties considered. Their absolute statistical uncertainty is of the order of 1.5 %. Conven-

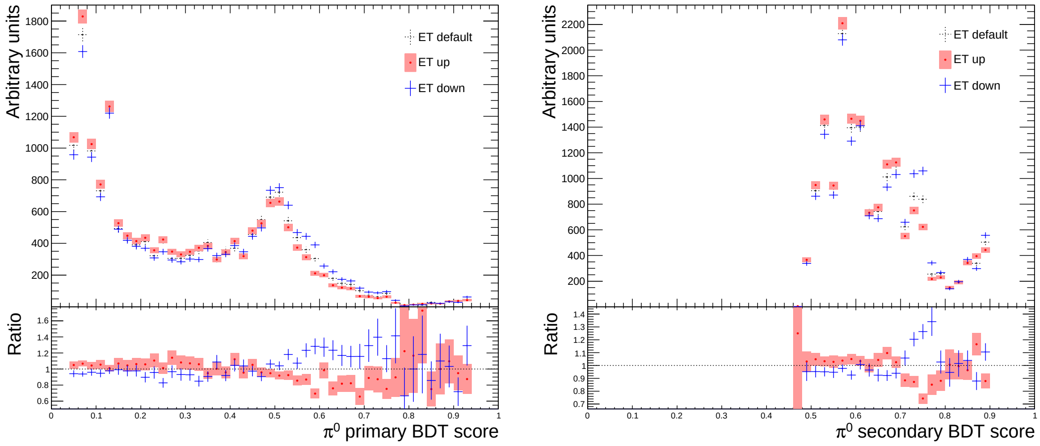


Figure 5.8: π^0 BDT scores for 1- and 3-prong candidates under cell energy variation

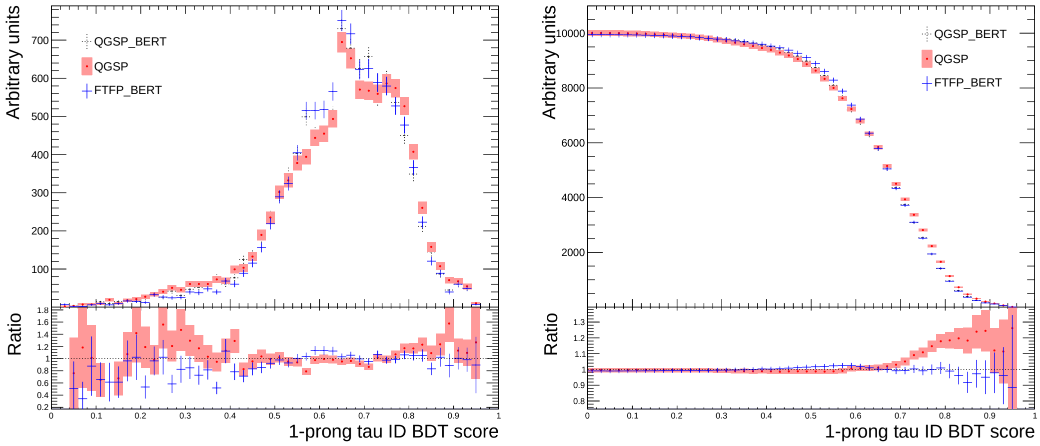


Figure 5.9: Hadronic tau BDT score for different Geant4 physics lists, integrated version on the right, 1-prong

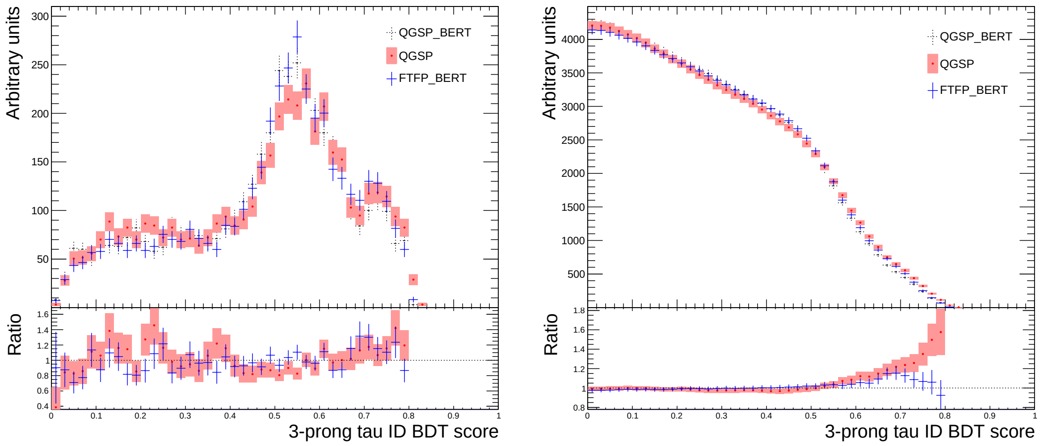


Figure 5.10: Hadronic tau BDT score for different Geant4 physics lists, integrated version on the right, 3-prong

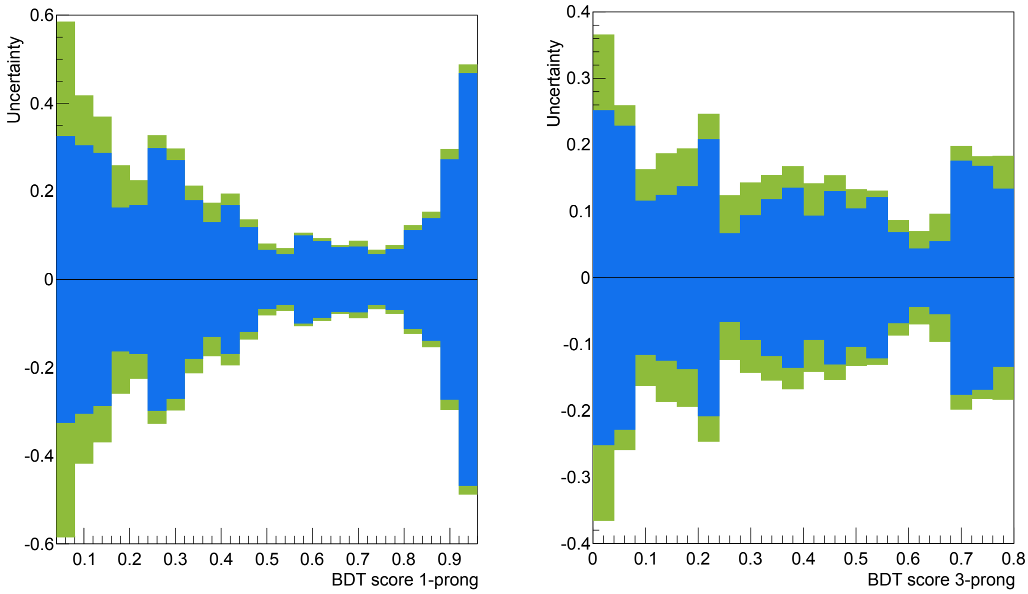


Figure 5.11: Combined symmetrical uncertainties (blue: systematic, green: systematic and statistical)

tional Methods [84] determine the hadronic tau identification efficiency in data and Monte Carlo simulated samples. By comparison of the performance in data and simulation events scale factors are obtained, which are applied as correcting factors in physics analyses. The uncertainties on these scale factors are determined by repeating the process with systematically changed samples and can be compared to the values shown in Table 5.2. These uncertainties are about half as large as the uncertainties obtained in this thesis in the tight quality region and are slightly larger for loose and medium quality tau candidates. Considering the statistical uncertainties on these values, the results compare reasonably well. A merit of the method presented in this thesis is that with the obtained uncertainties the whole range of BDT scores and their distribution can be used in future analyses, rather than only the three predefined quality regions.

In Table 5.3 the integrated uncertainties are listed for 1-prong and 3-prong candidates separately. The relative uncertainties are significantly larger for 3-prong candidates with uncertainties as high as 12.5 %. Because the 1-prong branching ratio is about 6 times higher than the branching ratio of 3-prong decays, the overall uncertainty is dominated by the comparably small uncertainty of 1-prong candidates.

	Loose	Medium	Tight
cell energy up	−0.8%	−1.2%	−4.8%
cell energy down	+0.8%	+1.3%	+3.3%
track reco. eff. A	< 0.1%	< 0.1%	< 0.1%
track reco. eff. B	< 0.1%	< 0.1%	< 0.1%
QGSP	−1.3%	+0.1%	+5.9%
FTFP_BERT	+0.8%	+0.1%	−0.2%
total (positive)	+1.1%	+1.3%	+6.8%
total (negative)	−1.5%	−1.2%	−4.8%

Table 5.2: Relative uncertainties on $N_{\text{1-prong}} + N_{\text{3-prong}}$ for the sources of systematic uncertainty considered

	1-prong			3-prong		
	Loose	Medium	Tight	Loose	Medium	Tight
cell energy up	−0.4%	−0.6%	−4.6%	−2.5%	−3.7%	−5.8%
cell energy down	+0.2%	+0.9%	+3.3%	+3.4%	+3.3%	+3.2%
track reco. eff. A	< 0.1%	< 0.1%	< 0.1%	< 0.1%	< 0.1%	< 0.1%
track reco. eff. B	< 0.1%	< 0.1%	< 0.1%	< 0.1%	< 0.1%	< 0.1%
QGSP	−1.5%	−0.7%	+4.3%	−0.3%	+3.4%	+12.5%
FTFP_BERT	+0.8%	−0.1%	−1.7%	+1.0%	+1.2%	+6.3%
total (positive)	+0.8%	+0.9%	+5.4%	+3.5%	+4.9%	+14.4%
total (negative)	−1.6%	−0.9%	−4.9%	−2.5%	−3.7%	−5.8%

Table 5.3: Relative uncertainties on $N_{\text{1-prong}}$ and $N_{\text{3-prong}}$ for the sources of systematic uncertainty considered

Chapter 6

Summary

After the discovery of a new boson, which has SM-like properties, the focus of particle physics shifts towards physics beyond the Standard Model. A very important candidate for a theory of physics at the TeV scale is supersymmetry. From a phenomenological point of view the minimal implementation in terms of additional Higgs bosons, the MSSM, is a favored model. Four additional Higgs bosons are proposed, two of them electrically neutral. Tau leptons play an important role in the search for these neutral Higgs bosons.

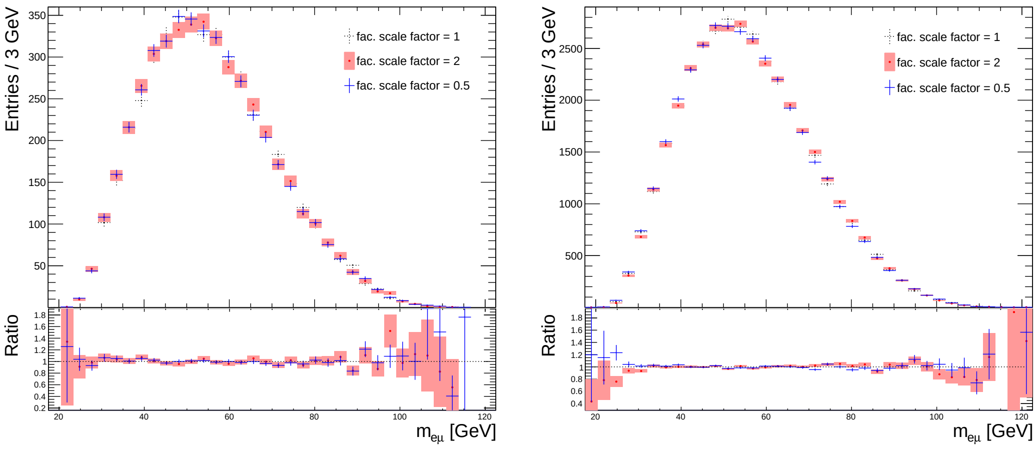
In this study the influence of systematic uncertainties of generator parameters on the signal acceptances of two ATLAS searches for neutral MSSM Higgs bosons in decays to two tau leptons was studied. For the semileptonic analysis also the acceptance uncertainty for the important background process $Z \to \tau\tau + \text{jets}$ was determined. A method to emulate b-quark flavour tagging efficiency at particle level was developed. The acceptance uncertainties for a wide range of generator parameters were presented. For the first time for these analyses the PDF error sets were used to accurately estimate the acceptance uncertainty due to the PDF uncertainty. By increasing the amount of generated events, the statistical uncertainties were greatly reduced compared to those of earlier studies. The systematic uncertainties of the acceptance are found to be in the range of 0 % to 30 %. It was verified that the shapes of the invariant mass of the visible tau lepton decay products are not affected by the uncertainties. Furthermore the systematic changes of the shapes of other observables were studied.

In the second part of this thesis a study on systematic uncertainties of the tau identification algorithm is presented. The respective ATLAS software tools were modified to change the affected event variables according to experimental uncertainties and the input variables of the identification BDTs were recalculated. This method has the merit of making it possible to study the change of the shape of BDT scores for each source of uncertainty individually. Difficulties of recalculating the BDT input variables are described and it was shown how to solve them or why they do not have a notable effect on the study. It was found that while the considered uncertainties of the ATLAS inner detector tracking system have no observable impact on the identification scores at the level of statistical accuracy of the study, there are significant changes in shape due to the uncertainty of the calorimeter cell energy measurement and hadronic shower shapes. These uncertainties were also observed in their integrated form in the three predefined quality regions. The combined uncertainties are in the range of 1 % to 7 % decreasing with tau signal efficiency. The results compare reasonably well with the default method to assess the systematic uncertainties on tau identification efficiency.

Appendix A

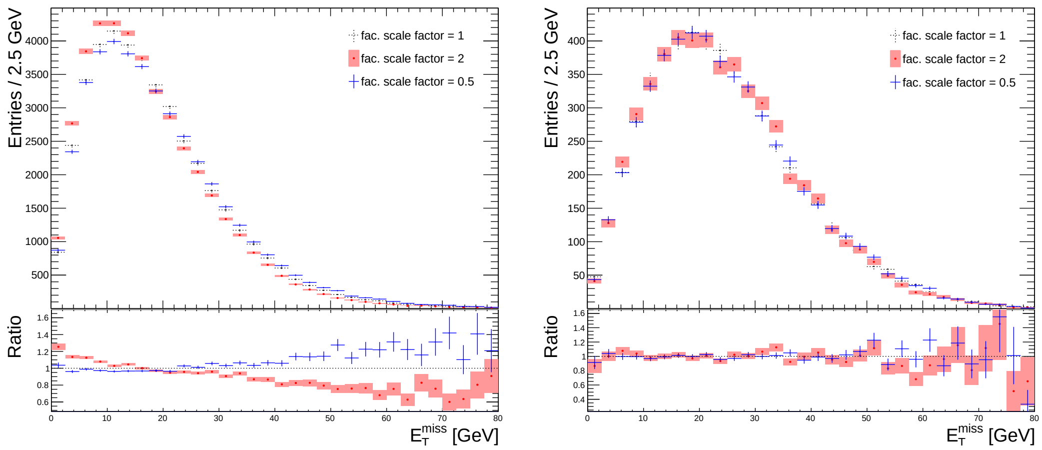
Theoretical uncertainties study

A.1 Systematic changes of shape for the leptonic decay channel



(a) b-tag category (b) b-veto category

Figure A.1: Visible mass distributions of the two selected leptons in a b-associated production sample with $m_{\text{Higgs}} = 120$ GeV for different values of the factorization scale factor in the leptonic decay channel



(a) b-veto category (b) b-tag category

Figure A.2: Distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$ for different values of the factorization scale factor with $m_{\mathrm{Higgs}} = 120$ GeV for b-associated production in the leptonic decay channel

A.2 Acceptance uncertainties for all mass points

variation	90 GeV	120 GeV	200 GeV	300 GeV
CKKW down	-2.0 $\pm$ 1.1	-3.1 $\pm$ 0.9	-6.3 $\pm$ 0.9	-6.5 $\pm$ 1.3
CKKW up	-19.0 $\pm$ 1.0	-8.3 $\pm$ 0.9	2.5 $\pm$ 0.9	6.7 $\pm$ 1.5
fac. sc. down	14.9 $\pm$ 1.3	13.5 $\pm$ 1.0	8.2 $\pm$ 1.1	7.7 $\pm$ 1.5
fac. sc. up	-23.0 $\pm$ 1.0	-19.8 $\pm$ 0.8	-14.1 $\pm$ 0.8	-10.1 $\pm$ 1.3
ren. sc. down	-0.1 $\pm$ 1.1	0.4 $\pm$ 0.9	-0.4 $\pm$ 0.9	0.3 $\pm$ 1.4
ren. sc. up	-1.4 $\pm$ 1.1	0.8 $\pm$ 0.9	0.8 $\pm$ 1.0	1.3 $\pm$ 1.4
pdf	$\pm$0.1	$\pm$0.1	$\pm$ 0.1	$\pm$0.0
total (pos.)	14.9 $\pm$ 1.3	13.5 $\pm$ 1.6	8.6 $\pm$ 1.7	10.3 $\pm$ 2.9
total (neg.)	-29.9 $\pm$ 2.4	-21.7 $\pm$ 1.5	-15.4 $\pm$ 1.5	-12.0 $\pm$ 1.8

Table A.1: Acceptance uncertainties for b-associated production in the leptonic decay channel and b-tag category

variation	90 GeV	120 GeV	200 GeV	300 GeV
CKKW down	-0.3 ± 0.5	0.4 ± 0.4	1.8 ± 0.4	3.5 ± 0.6
CKKW up	3.4 ± 0.5	2.9 ± 0.4	0.1 ± 0.4	-0.9 ± 0.6
fac. sc. down	-5.4 ± 0.5	-4.2 ± 0.4	-4.2 ± 0.4	-3.3 ± 0.6
fac. sc. up	5.1± 0.5	5.6 ± 0.4	6.4 ± 0.4	8.2 ± 0.7
ren. sc. down	-0.3 ± 0.5	-0.3 ± 0.4	-0.8 ± 0.4	-0.3 ± 0.6
ren. sc. up	0.7 ± 0.5	0.5 ± 0.4	0.5 ± 0.4	1.2 ± 0.6
pdf	±0.2	±0.2	±0.2	±0.2
total (pos.)	6.2 ± 0.9	6.3 ± 0.8	6.7 ± 0.8	9.0 ± 1.1
total (neg.)	-5.4 ± 0.9	-4.2 ± 0.6	-4.3 ± 0.6	-3.4 ± 1.0

Table A.2: Acceptance uncertainties for b-associated production in the leptonic decay channel and b-veto category

variation	90 GeV	120 GeV	200 GeV	300 GeV
ISR up	23.7 ± 10.9	20.3 ± 8.1	19.8 ± 7.3	6.1 ± 8.0
ISR down	-13.2 ± 8.3	3.6 ± 7.2	-0.3 ± 6.3	-2.9 ± 7.4
FSR up	5.7 ± 9.6	16.6 ± 7.8	10.3 ± 6.8	20.6 ± 8.7
FSR down	-3.4 ± 9.0	-3.6 ± 6.8	-15.7 ± 5.5	-18.9 ± 6.5
scales up	-2.1 ± 9.1	9.4 ± 7.5	11.1 ± 6.8	4.7 ± 7.9
scales down	16.2 ± 10.4	2.5 ± 7.1	1.9 ± 6.4	-1.7 ± 7.5
pdf	±0.0	±0.0	± 0.0	±0.0
total (pos.)	29.3 ± 17.9	28.2 ± 16.9	25.0 ± 13.7	22.0 ± 14.2
total (neg.)	-13.8 ± 15.3	-3.6 ± 6.8	-15.7 ± 8.4	-19.2 ± 12.4

Table A.3: Acceptance uncertainties for gluon fusion production in the leptonic decay channel and b-tag category

variation	90 GeV	120 GeV	200 GeV	300 GeV
ISR up	-1.1 ± 0.8	-1.2 ± 0.6	-0.5 ± 0.5	-1.5 ± 0.5
ISR down	-0.6 ± 0.8	0.4 ± 0.6	0.8 ± 0.5	0.7 ± 0.6
FSR up	-1.5 ± 0.8	-0.2 ± 0.6	-0.1 ± 0.5	-0.3 ± 0.6
FSR down	-1.4 ± 0.8	-0.7 ± 0.6	0.8 ± 0.5	-0.4 ± 0.5
scales up	-1.1 ± 0.8	0.0 ± 0.6	0.4 ± 0.5	-0.2 ± 0.5
scales down	-1.0 ± 0.8	-0.5 ± 0.6	0.8 ± 0.5	-0.7 ± 0.5
pdf	±0.1	±0.1	± 0.1	±0.1
total (pos.)	0.1 ± 0.0	0.4 ± 0.6	1.4 ± 1.0	0.7 ± 0.6
total (neg.)	-2.8 ± 2.0	-1.5 ± 1.2	-0.5 ± 0.7	-1.7 ± 1.2

Table A.4: Acceptance uncertainties for gluon fusion production in the leptonic decay channel and b-veto category

variation	90 GeV	120 GeV	200 GeV	450 GeV	700 GeV
CKKW down	-0.3 ± 1.2	-4.5 ± 0.8	-6.8 ± 0.5	-6.9 ± 0.4	-7.0 ± 0.4
CKKW up	-18.2 ± 1.2	-10.0 ± 0.8	0.8 ± 0.5	6.9 ± 0.4	9.0 ± 0.4
fac. sc. down	21.7 ± 1.4	14.6 ± 0.7	8.0 ± 0.6	5.1 ± 0.4	4.8 ± 0.4
fac. sc. up	-25.2 ± 1.0	-22.2 ± 0.9	-17.9 ± 0.5	-13.3 ± 0.4	-12.6 ± 0.4
ren. sc. down	0.4 ± 1.2	0.2 ± 0.8	-0.1 ± 0.5	0.4 ± 0.4	-0.1 ± 0.4
ren. sc. up	-2.1±1.2	-0.9 ± 0.8	-0.9 ± 0.5	0.1 ± 0.4	0.0 ± 0.4
pdf	±0.0	±0.1	± 0.1	±0.1	±0.1
total (pos.)	21.7±1.8	14.6 ± 1.2	8.0 ± 0.8	8.6 ± 0.8	10.2 ± 0.7
total (neg.)	-31.2 ±2.3	-24.7 ± 1.6	-19.2 ± 1.0	-15.0 ± 0.6	-14.4 ± 0.7

Table A.5: Acceptance uncertainties for b-associated production in the semileptonic decay channel and b-tag category

variation	90 GeV	120 GeV	200 GeV	450 GeV	700 GeV
CKKW down	-0.7 ± 0.5	0.2 ± 0.4	0.8 ± 0.3	1.7 ± 0.2	1.8 ± 0.2
CKKW up	1.4 ± 0.5	0.9 ± 0.4	-0.3 ± 0.3	-1.7 ± 0.2	-2.4 ± 0.2
fac. sc. down	-3.9 ± 0.5	3.3 ± 0.4	-2.1 ± 0.3	-1.1 ± 0.2	-0.6 ± 0.2
fac. sc. up	3.3 ± 0.6	-3.5 ± 0.4	3.3 ± 0.3	2.6 ± 0.2	2.3 ± 0.2
ren. sc. down	-0.6 ± 0.5	-0.6 ± 0.4	0.1 ± 0.3	-0.1 ± 0.2	0.3 ± 0.2
ren. sc. up	-0.3±0.6	-0.5 ± 0.4	0.0 ± 0.3	0.0 ± 0.2	0.2 ± 0.2
pdf	±0.1	±0.2	± 0.2	±0.2	±0.2
total (pos.)	3.6±0.8	3.4 ± 0.7	3.4 ± 0.6	3.1 ± 0.3	2.9 ± 0.4
total (neg.)	-4.0±1.1	-3.6 ± 0.7	-2.1 ± 0.4	-2.0 ± 0.3	-2.5 ± 0.3

Table A.6: Acceptance uncertainties for b-associated production in the semileptonic decay channel and b-veto category

variation	90 GeV	120 GeV	200 GeV	450 GeV	700 GeV
ISR up	12.5 ± 5.0	21.9 ± 4.1	17.5 ± 2.8	20.6 ± 2.1	18.3 ± 1.8
ISR down	-11.4 ± 4.2	-15.0 ± 3.2	-12.7 ± 2.2	-16.4 ± 1.6	-17.2 ± 1.4
FSR up	22.2 ± 5.4	22.6 ± 4.2	22.6 ± 2.9	27.6 ± 2.1	28.6 ± 1.9
FSR down	-23.7 ± 3.8	-13.4 ± 3.2	-20.8 ± 2.1	-20.0 ± 1.5	-22.9 ± 1.3
scales up	-2.9 ± 4.5	0.7 ± 3.6	1.3 ± 2.5	0.9 ± 1.8	0.1 ± 1.6
scales down	-3.4 ± 4.5	0.9 ± 3.6	-2.3 ± 2.4	-0.2 ± 1.8	-0.5± 1.6
pdf	±0.0	±0.0	± 0.0	±0.0	±0.0
total (pos.)	25.5 ± 7.4	31.5 ± 7.8	28.6 ± 4.7	34.5 ± 3.5	34.0 ± 3.1
total (neg.)	-26.7 ± 8.5	-20.1 ± 4.5	-24.5 ± 3.9	-25.9 ± 2.8	-28.6 ± 2.5

Table A.7: Acceptance uncertainties for gluon fusion production in the semileptonic decay channel and b-tag category

variation	90 GeV	120 GeV	200 GeV	450 GeV	700 GeV
ISR up	0.2 ± 0.5	0.3 ± 0.3	0.0 ± 0.2	-0.2 ± 0.1	-0.2 ± 0.1
ISR down	-0.2 ± 0.5	0.2 ± 0.3	0.2 ± 0.2	0.3 ± 0.1	0.3 ± 0.1
FSR up	-0.1 ± 0.5	-0.1 ± 0.3	-0.4 ± 0.2	-0.2 ± 0.1	-0.4 ± 0.1
FSR down	0.0 ± 0.5	0.4 ± 0.3	0.2 ± 0.2	0.2 ± 0.1	0.5 ± 0.1
scales up	0.0 ± 0.5	0.4 ± 0.3	0.1 ± 0.2	0.3 ± 0.1	0.1 ± 0.1
scales down	-0.3 ± 0.5	-0.2 ± 0.3	0.0 ± 0.2	-0.1 ± 0.1	0.0 ± 0.1
pdf	±0.1	±0.1	± 0.2	±0.2	±0.2
total (pos.)	0.2 ± 0.9	0.7 ± 0.6	0.5 ± 0.4	0.5 ± 0.2	0.6 ± 0.2
total (neg.)	-0.4 ± 0.9	-0.2 ± 0.4	-0.4 ± 0.2	-0.4 ± 0.2	-0.5 ± 0.1

Table A.8: Acceptance uncertainties for gluon fusion production in the semileptonic decay channel and *b*-veto category

A.3 Acceptance uncertainties with additional cut on missing transverse energy in the semileptonic channel

variation	90 GeV	120 GeV	200 GeV	450 GeV	700 GeV
CKKW down	0.3 ± 1.4	-3.8 ± 0.9	-6.9 ± 0.6	-6.8 ± 0.4	-7.0 ± 0.4
CKKW up	-17.6 ± 1.2	-10.1 ± 0.9	0.8 ± 0.6	6.9 ± 0.5	9.1 ± 0.4
fac. sc. down	23.0 ± 1.7	15.1 ± 1.1	8.0 ± 0.6	5.3 ± 0.4	5.0 ± 0.4
fac. sc. up	-26.3 ± 1.1	-22.5 ± 0.8	-18.6 ± 0.5	-13.4 ± 0.4	-12.6 ± 0.4
ren. sc. down	0.2 ± 1.4	-0.2 ± 0.9	-0.3 ± 0.6	0.5 ± 0.4	-0.1 ± 0.4
ren. sc. up	-3.2±1.4	-1.2 ± 0.9	-1.1 ± 0.6	0.1 ± 0.4	0.0 ± 0.4
pdf	±0.1	±0.1	± 0.1	±0.1	±0.1
total (pos.)	23.0±2.6	15.1 ± 1.1	8.0 ± 0.8	8.7 ± 0.9	10.4 ± 0.7
total (neg.)	-31.8 ±2.1	-25.0 ± 2.0	-19.9 ± 1.2	-15.0 ± 0.6	-14.4 ± 0.7

Table A.9: Acceptance uncertainties for *b*-associated production in the semileptonic decay channel and *b*-tag category with a cut on the missing transverse energy

variation	90 GeV	120 GeV	200 GeV	450 GeV	700 GeV
CKKW down	-0.5 ± 0.7	0.0 ± 0.5	0.6 ± 0.3	1.6 ± 0.2	1.7 ± 0.2
CKKW up	0.0 ± 0.7	-0.1 ± 0.5	-0.4 ± 0.3	-1.6 ± 0.2	-2.3 ± 0.2
fac. sc. down	-2.4 ± 0.7	-2.8 ± 0.5	-1.7 ± 0.3	-0.9 ± 0.2	-0.6 ± 0.2
fac. sc. up	0.4 ± 0.7	1.5 ± 0.5	2.5 ± 0.3	2.2 ± 0.2	2.0 ± 0.2
ren. sc. down	0.0 ± 0.7	-0.4 ± 0.5	0.3 ± 0.3	-0.1 ± 0.2	0.2 ± 0.2
ren. sc. up	-1.0±0.7	-0.5 ± 0.5	0.0 ± 0.3	0.0 ± 0.2	0.2 ± 0.2
pdf	±0.1	±0.1	± 0.2	±0.2	±0.2
total (pos.)	0.4±1.2	1.5 ± 0.7	2.6 ± 0.6	2.7 ± 0.3	2.6 ± 0.4
total (neg.)	-2.6±1.2	-2.9 ± 1.0	-1.8 ± 0.4	-1.8 ± 0.3	-2.4 ± 0.3

Table A.10: Acceptance uncertainties for b-associated production in the semileptonic decay channel and b-veto category with a cut on the missing transverse energy

variation	90 GeV	120 GeV	200 GeV	450 GeV	700 GeV
ISR up	10.2 ± 5.4	25.1 ± 4.6	16.6 ± 2.9	20.5 ± 2.1	18.0 ± 1.8
ISR down	-16.6 ± 4.4	-15.6 ± 3.4	-13.4 ± 2.3	-16.5 ± 1.6	-17.8 ± 1.4
FSR up	22.0 ± 5.9	24.4 ± 4.6	21.4 ± 3.0	27.3 ± 2.2	28.5 ± 1.9
FSR down	-27.1 ± 4.0	-15.8 ± 3.4	-22.3 ± 2.2	-20.2 ± 1.5	-23.1 ± 1.3
scales up	-4.8 ± 4.9	1.6 ± 3.9	1.2 ± 2.6	0.8 ± 1.8	-0.2 ± 1.6
scales down	-4.0 ± 4.9	0.9 ± 3.9	-3.5 ± 2.5	-0.6 ± 1.8	-0.7± 1.6
pdf	±0.0	±0.0	± 0.0	±0.0	±0.0
total (pos.)	24.2 ± 8.0	35.1 ± 8.5	27.1 ± 4.9	34.1 ± 3.5	33.7 ± 2.6
total (neg.)	-32.4 ± 9.1	-22.2 ± 4.8	-26.3 ± 4.0	-26.1 ± 2.8	-29.2 ± 3.0

Table A.11: Acceptance uncertainties for gluon fusion production in the semileptonic decay channel and b-tag category with a cut on the missing transverse energy

variation	90 GeV	120 GeV	200 GeV	450 GeV	700 GeV
ISR up	1.1 ± 0.6	0.3 ± 0.4	0.0 ± 0.2	-0.1 ± 0.2	-0.2 ± 0.1
ISR down	-0.6 ± 0.6	-0.8 ± 0.4	0.0 ± 0.3	0.2 ± 0.2	0.2 ± 0.1
FSR up	0.2 ± 0.6	-0.5 ± 0.4	-0.4 ± 0.2	-0.3 ± 0.2	-0.5 ± 0.1
FSR down	0.8 ± 0.6	-0.1 ± 0.4	0.2 ± 0.2	0.2 ± 0.2	0.5 ± 0.1
scales up	0.7 ± 0.6	0.1 ± 0.4	0.1 ± 0.2	0.2 ± 0.2	0.1 ± 0.1
scales down	-0.1 ± 0.6	-0.7 ± 0.4	0.0 ± 0.2	-0.2 ± 0.2	0.0 ± 0.1
pdf	±0.1	±0.1	± 0.2	±0.2	±0.2
total (pos.)	1.5 ± 1.2	0.3 ± 0.6	0.3 ± 0.5	0.4 ± 0.3	0.6 ± 0.2
total (neg.)	-0.6 ± 0.8	-1.2 ± 0.8	-0.4 ± 0.2	-0.4 ± 0.3	-0.6 ± 0.1

Table A.12: Acceptance uncertainties for gluon fusion production in the semileptonic decay channel and *b*-veto category with a cut on the missing transverse energy

Bibliography

[1] J. J. Thomson, *Cathode Rays*, Philosophical Magazine 44.293 (1897).

[2] R. Feynman, *Space-Time Approach to Non-Relativistic Quantum Mechanics*, Reviews of Modern Physics 20.2 (Apr. 1948), pp. 367–387.

[3] Julian Schwinger, *On Gauge Invariance and Vacuum Polarization*, Physical Review 82.5 (June 1951), pp. 664–679.

[4] Sin-Itiro Tomonaga and J. Oppenheimer, *On Infinite Field Reactions in Quantum Field Theory*, Physical Review 74.2 (July 1948), pp. 224–225.

[5] Sheldon L. Glashow, *Partial-symmetries of weak interactions*, Nuclear Physics 22.4 (Feb. 1961), pp. 579–588.

[6] Peter Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Physical Review Letters 13.16 (Oct. 1964), pp. 508–509.

[7] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Physical Review Letters 13.9 (Aug. 1964), pp. 321–323.

[8] Abdus Salam, *Renormalizability of Gauge Theories*, Physical Review 127.1 (July 1962), pp. 331–334.

[9] Steven Weinberg, *A Model of Leptons*, Physical Review Letters 19.21 (Nov. 1967), pp. 1264–1266.

[10] G. Zweig, "An SU(3) Model for Strong Interaction Symmetry and its Breaking. 2." In: *Developments in the Quark Theory of Hadrons*. Ed. by Don Bernett Lichtenberg and S. P. Rosen. Nonantum, MA: Hadronic Press, 1964.

[11] M. Gell-Mann, *A schematic model of baryons and mesons*, Physics Letters 8.3 (Feb. 1964), pp. 214–215.

[12] ALEPH Collaboration et al., *Precision Electroweak Measurements and Constraints on the Standard Model*, (Nov. 2009). arXiv: `0911.2604 [hep-ex]`.

[13] Lyndon Evans and Philip Bryant, *LHC Machine*, Journal of Instrumentation 3.08 (Aug. 2008). Ed. by Lyndon Evans, S08001–S08001.

[14] The ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation 3.08 (Aug. 2008), S08003–S08003.

[15] CMS Collaboration, *The CMS experiment at the CERN LHC*, Journal of Instrumentation 3.08 (Aug. 2008), S08004–S08004.

[16] Stephen P. Martin, “A Supersymmetry Primer”. In: *Perspectives on Supersymmetry*. Ed. by G. L. Kane. 1998. arXiv: `hep-ph/9709356`.

[17] CMS Collaboration, *Search for MSSM Neutral Higgs Bosons Decaying to Tau Pairs in pp Collisions*. CMS Physics Analysis Summaries CMS-PAS-HIG-12-050. Geneva: CERN, 2012.

[18] David Griffiths, *Introduction to Elementary Particles*. New York: Wiley, 2008.

[19] M. Czakon, J. Gluza, and M. Zralek, *Are Neutrinos Dirac or Majorana Particles?*, Acta Physica Polonica B 30 (Nov. 1999), p. 3121. arXiv: `hep-ph/9910357`.

[20] The LEP Working Group for Higgs Boson Searches, ALEPH, DELPHI, L3, and OPAL, *Search for the Standard Model Higgs boson at LEP*, Physics Letters B 565 (July 2003), pp. 61–75. arXiv: `hep-ex/0306033`.

[21] CDF Collaboration and D0 Collaboration, *Higgs Boson Studies at the Tevatron*, (Mar. 2013). arXiv: `1303.6346 [hep-ex]`.

[22] The ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Physics Letters B 716.1 (Sept. 2012), pp. 1–29. arXiv: `1207.7214 [hep-ex]`.

[23] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Physics Letters B 716.1 (Sept. 2012), pp. 30–61. arXiv: `1207.7235 [hep-ex]`.

[24] The ATLAS Collaboration, *Evidence for the spin-0 nature of the Higgs boson using ATLAS data*, (July 2013). arXiv: `1307.1432 [hep-ex]`.

[25] CMS Collaboration, *Properties of the observed Higgs-like resonance using the diphoton channel*. CMS Physics Analysis Summaries CMS-PAS-HIG-13-016. Geneva: CERN, 2013.

[26] Roland de Putter et al., *New Neutrino Mass Bounds from Sloan Digital Sky Survey III Data Release 8 Photometric Luminous Galaxies*, The Astrophysical Journal 761 (2012). arXiv: `1201.1909 [astro-ph]`.

[27] J. Beringer et al., *Review of Particle Physics*, Physical Review D 86.1 (July 2012).

[28] Paul Langacker, *The Standard Model and Beyond (Series in High Energy Physics, Cosmology and Gravitation)*. CRC Press, 2011.

[29] Sally Dawson, "Introduction to the Physics of Higgs Bosons". In: *CP Violation and the Limits of the Standard Model*. Ed. by J. F. Donoghue. 1995. arXiv: `hep-ph/9411325`.

[30] Gianfranco Bertone, Dan Hooper, and Joseph Silk, *Particle dark matter: evidence, candidates and constraints*, Physics Reports 405.5-6 (Jan. 2005), pp. 279–390. arXiv: `hep-ph/0404175`.

[31] Patrick Draper, Patrick Meade, Matthew Reece, and David Shih, *Implications of a 125 GeV Higgs boson for the MSSM and low-scale supersymmetry breaking*, Physical Review D 85.9 (May 2012), p. 095007. arXiv: `1112.3068 [hep-ph]`.

[32] LHC Higgs Cross Section Working Group et al., *Handbook of LHC Higgs Cross Sections: 2. Differential Distributions*, (Jan. 2012). arXiv: `1201.3084 [hep-ph]`.

[33] Abdelhak Djouadi, *The anatomy of electroweak symmetry breaking I: The Higgs boson in the Standard Model*, Physics Reports 457.1-4 (Feb. 2008), pp. 1–216. arXiv: `hep-ph/0503172`.

[34] Abdelhak Djouadi, *The anatomy of electroweak symmetry breaking II: The Higgs bosons in the Minimal Supersymmetric Model*, Physics Reports 459.1-6 (Apr. 2008), pp. 1–241. arXiv: `hep-ph/0503173`.

[35] The ATLAS Collaboration, *Discovery Potential of $A/H \to \tau^+\tau^- \to \ell h$ in ATLAS*. ATLAS Note . Geneva: CERN, July 2010.

[36] Julien Baglio and Abdelhak Djouadi, *Higgs production at the lHC*, Journal of High Energy Physics 2011.3 (Mar. 2011), p. 55. arXiv: `1012.0530 [hep-ph]`.

[37] Sven Heinemeyer, Wolfgang Hollik, and Georg Weiglein, *Constraints on $\tan\beta$ in the MSSM from the upper bound on the mass of the lightest Higgs boson*, Journal of High Energy Physics 2000.06 (June 2000), pp. 009–009. arXiv: `hep-ph/9909540`.

[38] CERN, *LEP design report*. Report CERN-LEP-84-01. Geneva, 1984.

[39] Jean-Luc Caron, *Overall view of LHC experiments*. LHC Project Illustration LHC-PHO-1998-349. May 1998.

[40] Oliver Sim Brüning, Paul Collier, P Lebrun, Stephen Myers, Ranko Ostojic, John Poole, and Paul Proudlock, *LHC Design Report*. CERN Report CERN-2004-003-V-1. Geneva: CERN, 2004.

[41] The ATLAS Collaboration, *Luminosity Public Results*. Last modified July 26, 2013. URL: `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults`.

[42] The ALICE Collaboration, *The ALICE experiment at the CERN LHC*, Journal of Instrumentation 3.08 (Aug. 2008), S08002–S08002.

[43] The LHCb Collaboration, *The LHCb Detector at the LHC*, Journal of Instrumentation 3.08 (Aug. 2008), S08005–S08005.

[44] James Pinfold et al., *Technical Design Report of the MoEDAL Experiment*. Scientific Committee Paper CERN-LHCC-2009-006. MoEDAL-TDR-001. Geneva: CERN, June 2009.

[45] The TOTEM Collaboration, *The TOTEM Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation 3.08 (Aug. 2008), S08007–S08007.

[46] The LHCf Collaboration, *The LHCf detector at the CERN Large Hadron Collider*, Journal of Instrumentation 3.08 (Aug. 2008), S08006–S08006.

[47] Christiane Lefèvre, *The CERN accelerator complex. Complexe des accèlèrateurs du CERN*. CERN PhotoLab CERN-DI-0812015. Dec. 2008.

[48] The ATLAS Collaboration, *Search for the neutral Higgs bosons of the minimal supersymmetric standard model in pp collisions at* $\sqrt{s}$*=7 TeV with the ATLAS detector*, Journal of High Energy Physics 2013.2 (Feb. 2013), p. 95. arXiv: `1211.6956 [hep-ex]`.

[49] Andy Buckley, Jonathan Butterworth, David Grellscheid, Hendrik Hoeth, Leif Lonnblad, James Monk, Holger Schulz, and Frank Siegert, *Rivet user manual*, (Mar. 2010). arXiv: `1003.0694 [hep-ph]`.

[50] T Gleisberg, S Höche, F Krauss, M Schönherr, S Schumann, F Siegert, and J Winter, *Event generation with SHERPA 1.1*, Journal of High Energy Physics 2009.02 (Feb. 2009), pp. 007–007. arXiv: `0811.4622 [hep-ph]`.

[51] Adam Alloul, Jorgen D'Hondt, Karen Causmaecker, Benjamin Fuks, and Michel Rausch de Traubenberg, *Automated mass spectrum generation for new physics*, The European Physical Journal C 73.2 (Feb. 2013), p. 2325. arXiv: `1301.5932 [hep-ph]`.

[52] Andy Buckley et al., *General-purpose event generators for LHC physics*, Physics Reports 504 (July 2011), pp. 145–233. arXiv: `1101.2599 [hep-ph]`.

[53] John C. Collins and Davison E. Soper, *Parton distribution and decay functions*, Nuclear Physics B 194.3 (Jan. 1982), pp. 445–492.

[54] Stefano Catani, Frank Krauss, Bryan R Webber, and Ralf Kuhn, *QCD Matrix Elements + Parton Showers*, Journal of High Energy Physics 2001.11 (Nov. 2001), pp. 063–063. arXiv: `hep-ph/0109231`.

[55] Michelangelo L Mangano, Mauro Moretti, Fulvio Piccinini, and Michele Treccani, *Matching matrix elements and shower evolution for top-pair production in hadronic collisions*, Journal of High Energy Physics 2007.01 (Jan. 2007), pp. 013–013. arXiv: `hep-ph/0611129`.

[56] The ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, European Physical Journal C 70 (May 2010), pp. 823–874. arXiv: `1005.4568 [physics.ins-det]`.

[57] S. Agostinelli et al., *Geant4—a simulation toolkit*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 506.3 (July 2003), pp. 250–303.

[58] R. Brun and F. Rademakers, *ROOT - An object oriented data analysis framework*, Nuclear Instruments and Methods in Physics Research A 389 (Feb. 1997), pp. 81–86.

[59] M. R. Whalley, D. Bourilkov, and R. C. Group, *The Les Houches Accord PDFs (LHAPDF) and Lhaglue*, (Aug. 2005). arXiv: `hep-ph/0508110`.

[60] Frank Krauss, Ralf Kuhn, and Gerhard Soff, *AMEGIC++ 1.0, A Matrix Element Generator In C++*, Journal of High Energy Physics 2002.02 (Feb. 2002), pp. 044–044. arXiv: `hep-ph/0109036`.

[61] Tanju Gleisberg and Stefan Höche, *Comix, a new matrix element generator*, Journal of High Energy Physics 2008.12 (Dec. 2008), pp. 039–039. arXiv: `0808.3674 [hep-ph]`.

[62] F. Krauss, A. Schälicke, and G. Soff, *APACIC++ 2.0. A PArton Cascade In C++*, Computer Physics Communications 174.11 (June 2006), pp. 876–902. arXiv: `hep-ph/0503087`.

[63] S. Alekhin et al., *HERA and the LHC - A workshop on the implications of HERA for LHC physics: Proceedings - Part A*, (Jan. 2006). arXiv: `hep-ph/0601012`.

[64] M. Schönherr and F. Krauss, *Soft photon radiation in particle decays in SHERPA*, Journal of High Energy Physics 2008.12 (Dec. 2008), pp. 018–018. arXiv: `0810.5071 [hep-ph]`.

[65] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, Journal of High Energy Physics 6 (June 2010). arXiv: `1002.2581 [hep-ph]`.

[66] J. Alwall et al., *A standard format for Les Houches Event Files*, Computer Physics Communications 176.4 (Feb. 2007), pp. 300–304. arXiv: `hep-ph/0609017`.

[67] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands, *A brief introduction to PYTHIA 8.1*, Computer Physics Communications 178.11 (June 2008), pp. 852–867. arXiv: `0710.3820 [hep-ph]`.

[68] Michelangelo L Mangano, Fulvio Piccinini, Antonio D Polosa, Mauro Moretti, and Roberto Pittau, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, Journal of High Energy Physics 2003.07 (July 2003), pp. 001–001. arXiv: `hep-ph/0206293`.

[69] Gennaro Corcella, Ian G Knowles, Giuseppe Marchesini, Stefano Moretti, Kosuke Odagiri, Peter Richardson, Michael H Seymour, and Bryan R Webber, *HERWIG 6: an event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, Journal of High Energy Physics 2001.01 (Jan. 2001), pp. 010–010.

[70] Stefan Höche, Frank Krauss, Steffen Schumann, and Frank Siegert, *QCD matrix elements and truncated showers*, Journal of High Energy Physics 2009.05 (May 2009), pp. 053–053. arXiv: `0903.1219 [hep-ph]`.

[71] T Bristow et al., *Search for Neutral MSSM Higgs Bosons $H \to \tau^+\tau^- \to \ell\tau_h$ with the ATLAS Detector in 7 TeV Collisions.* ATLAS Internal Note ATL-COM-PHYS-2012-094. Geneva: CERN, Jan. 2012.

[72] V.N. Gribov and L.N. Lipatov, *Deep inelastic e p scattering in perturbation theory*, Sov.J.Nucl.Phys. 15 (1972), pp. 438–450.

[73] Yuri L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e^+ e^- Annihilation by Perturbation Theory in Quantum Chromodynamics.* Sov.Phys.JETP 46 (1977), pp. 641–653.

[74] G. Altarelli and G. Parisi, *Asymptotic freedom in parton language*, Nuclear Physics B 126.2 (Aug. 1977), pp. 298–318.

[75] Hung-Liang Lai, Marco Guzzi, Joey Huston, Zhao Li, Pavel M. Nadolsky, Jon Pumplin, and C.-P. Yuan, *New parton distributions for collider physics*, Physical Review D 82.7 (Oct. 2010), p. 074024. arXiv: `1007.2241 [hep-ph]`.

[76] Jonathan Pumplin, Daniel Robert Stump, Joey Huston, Hung-Liang Lai, Pavel Nadolsky, and Wu-Ki Tung, *New Generation of Parton Distributions with Uncertainties from Global QCD Analysis*, Journal of High Energy Physics 2002.07 (July 2002), pp. 012–012. arXiv: `hep-ph/0201195`.

[77] A.D. Martin, R.G. Roberts, W.J. Stirling, and R.S. Thorne, *MRST2001: partons and α_S from precise deep inelastic scattering and Tevatron jet data*, The European Physical Journal C - Particles and Fields 23.1 (Mar. 2002), pp. 73–87. arXiv: `hep-ph/0110215`.

[78] The ATLAS Collaboration, *Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 7$ TeV with ATLAS*, The European Physical Journal C 72.1 (Jan. 2012), p. 1844.

[79] Matteo Cacciari, Gavin P Salam, and Gregory Soyez, *The anti- k t jet clustering algorithm*, Journal of High Energy Physics 2008.04 (Apr. 2008), pp. 063–063. arXiv: `0802.1189 [hep-ph]`.

[80] The ATLAS Collaboration, *Measuring the b-tag efficiency in a top-pair sample with 4.7 fb^{-1} of data from the ATLAS detector*. ATLAS Conference Note ATLAS-CONF-2012-097. Geneva: CERN, July 2012.

[81] The ATLAS Collaboration, *Calibrating the b-Tag Efficiency and Mistag Rate in 35 pb^{-1} of Data with the ATLAS Detector* . ATLAS-CONF-2011-089. Geneva: CERN, June 2011.

[82] The ATLAS Collaboration, *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*. ATLAS Conference Note ATLAS-CONF-2011-102. Geneva: CERN, July 2011.

[83] F. Friedrich, A. Goussiou, L. Hauswald, J. Keller, J. Kroseberg, N. Rompotis, T. Vickey, and T. Wang, *Search for Neutral MSSM Higgs Bosons $H \to \tau^+\tau^- \to \ell\tau_h$ produced in 8 TeV Collisions with the ATLAS Detector*. ATLAS Internal Note ATL-COM-PHYS-2013-209. Geneva: CERN, Feb. 2013.

[84] The ATLAS Collaboration, *Identification of the Hadronic Decays of Tau Leptons in 2012 Data with the ATLAS Detector*. ATLAS Conference Note ATLAS-CONF-2013-064. Geneva: CERN, July 2013.

[85] The ATLAS Collaboration, *Performance of the Reconstruction and Identification of Hadronic Tau Decays with ATLAS*. ATLAS Conference Note ATLAS-CONF-2011-152. Geneva: CERN, Nov. 2011.

[86] The ATLAS Collaboration, *Determination of the tau energy scale and the associated systematic uncertainty in proton-proton collisions at $\sqrt{s}$= 7 TeV with the ATLAS detector at the LHC in 2011*. ATLAS Conference Note ATLAS-CONF-2012-054. Geneva: CERN, June 2012.

[87] Paola Giovannini, *Local hadron calibration with ATLAS*, Journal of Physics: Conference Series 293.1 (Apr. 2011), p. 012057.

[88] Byron P. Roe, Hai-Jun Yang, Ji Zhu, Yong Liu, Ion Stancu, and Gordon McGregor, *Boosted decision trees as an alternative to artificial neural networks for particle identification*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 543.2-3 (May 2005), pp. 577–584. arXiv: physics/0408124.

[89] The ATLAS Collaboration, *Measurements of the pseudorapidity dependence of the total transverse energy in proton-proton collisions at $\sqrt{s} = 7$ TeV with ATLAS*, Journal of High Energy Physics 11 (Nov. 2012), p. 33. arXiv: 1208.6256 [hep-ex].

[90] The ATLAS Collaboration, *Charged particle multiplicities in pp interactions at $\sqrt{s} = 7$ TeV measured with the ATLAS detector at the LHC*. ATLAS Conference Note ATLAS-CONF-2010-024. Geneva: CERN, July 2010.

[91] *Geant4 Physics Reference Manual*. Retrieved on August 3, 2013. URL: http://geant4.web.cern.ch/geant4/UserDocumentation/UsersGuides/PhysicsReferenceManual/fo/PhysicsReferenceManual.pdf.

[92] A.B. Kaidalov and K.A. Ter-Martirosyan, *Pomeron as quark-gluon strings and multiple hadron production at SPS-Collider energies*, Physics Letters B 117.3-4 (Nov. 1982), pp. 247–251.

[93] B. Andersson, G. Gustafson, and B. Nilsson-Almqvist, *A model for low-p_T hadronic reactions with generalizations to hadron-nucleus and nucleus-nucleus collisions*, Nuclear Physics B 281.1-2 (Jan. 1987), pp. 289–309.

[94] Hugo Bertini, *Intranuclear-Cascade Calculation of the Secondary Nucleon Spectra from Nucleon-Nucleus Interactions in the Energy Range 340 to 2900 MeV and Comparisons with Experiment*, Physical Review 188.4 (Dec. 1969), pp. 1711–1730.

Acknowledgments

I want to thank Jun.-Prof. Dr. Arno Straessner and Dr. Wolfgang Mader for the excellent supervision. I really appreciate their commitment and support. Also, I am grateful to the whole working group, it was a pleasure to work with you. A special thanks goes to Felix Friedrich for his advice and support. Even in times of stress and full schedules he took his time to discuss my newest findings and further options or to help me finding errors in my source code. Furthermore, I would like to thank Sebastian Wahrmund, whose advice was valuable to me. I also want to thank Marcus Morgenstern and Christian Rudolph for the help and discussions in our group meetings. Another big thanks goes to Fabian Multrus for countless discussions and learning sessions, which contributed significantly to the success of my studies.

Selbstständigkeitserklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Lorenz Hauswald
Dresden, 16.09.2013