Universität Bonn

Physikalisches Institut

A Sensitivity Study for Higgs Boson Production in Vector Boson Fusion in the $H \to \tau\tau \to \ell h + 3\nu$ Final State with ATLAS

Nicolas Möser

For a hypothetical Higgs boson mass between 114.4 GeV and about 135 GeV the production by Vector Boson Fusion and the decay $H \to \tau\tau \to \ell h + 3\nu$ is one of the most promising discovery channels at the LHC. In this thesis, a study of the expected sensitivity of the ATLAS detector for this channel at a centre-of-mass energy of 14 TeV is presented. For the first time, this study includes a full treatment of additional proton-proton interactions, so-called pile-up. The presence of pile-up significantly affects the signal selection efficiency and leads to a deterioration of the reconstructed Higgs boson mass, which is used as a discriminating observable. Two methods have been developed to estimate the dominant background processes from data. By replacing the muons in $Z \to \mu\mu$ events with simulated τ lepton decays, $Z \to \tau\tau$ events can be modelled with high precision. The non-resonant background, $t\bar{t}$ production and W + jets, is estimated by selecting events where lepton and hadronic τ decay have the same electric charge. Assuming a dataset corresponding to an integrated luminosity of 30 fb^{-1}, an expected signal significance between $3.0\,\sigma$ and $4.4\,\sigma$ is obtained for a Higgs boson mass between 115 GeV and 135 GeV. The expected significance decreases to 1.6–2.0 σ in the presence of pile-up.

Physikalisches Institut der
Universität Bonn
Nußallee 12
D-53115 Bonn

BONN-IR-2011-14
November 2011
ISSN-0172-8741

Universität Bonn

Physikalisches Institut

A Sensitivity Study for Higgs Boson Production in Vector Boson Fusion in the $H \to \tau\tau \to \ell h + 3\nu$ Final State with ATLAS

Nicolas Möser

aus Wissen

Dieser Forschungsbericht wurde als Dissertation von der Mathematisch-Naturwissenschaftlichen Fakultät der Universität Bonn angenommen und ist 2012 auf dem Hochschulschriftenserver der ULB Bonn `http://hss.ulb.uni-bonn.de/diss_online` elektronisch publiziert.

1. Gutachter:	Prof. Dr. Markus Schumacher
2. Gutachter:	Prof. Dr. Norbert Wermes
Angenommen am:	15.09.2011
Tag der Promotion:	18.10.2011

Contents

1 Introduction

The *Standard Model of Particle Physics* (*SM*) is among the most successful theories in modern physics. Its predictive power has been demonstrated e.g. by the discovery of the W and Z-bosons at the *Super Proton Synchrotron* at CERN in 1983 and, later on, with the discovery of the *top quark* at the *Tevatron* in 1995. Up to now, it withstood every experimental test, yielding remarkable agreement with precision measurements. The question why particles have mass arises from the fact that within the framework of the SM elementary particles themselves have to be massless. This is in contradiction to the experimentally measured masses of particles like the W and Z bosons or the top quark – the latter mass being of the order of that of a gold atom. Moreover, although the constituents of ordinary matter, electrons, up and down quarks, are nearly massless, a massless electron, e.g., would lead to an infinitely large proton radius.

This apparent contradiction is resolved within the SM by the *Higgs mechanism*. It states that elementary particles are indeed massless but acquire a mass by interacting with an omnipresent field, the *Higgs field*. The mechanism predicts the existence of at least one *Higgs boson*, an excitation of the Higgs field. Thus, this mechanism to generate particle masses can be experimentally tested by searching for the Higgs boson and, once it is found, studying its properties.

A number of experiments already searched for the Higgs boson – so far without success. They established a lower limit on the Higgs boson mass of 114.4 GeV. One of the main reasons to build the *Large Hadron Collider (LHC)* was to provide the means to finally answer the question if Higgs bosons exist. Its centre-of-mass energy is sufficient to produce heavy particles with masses at the TeV scale. The mass of the Higgs boson itself is the last parameter of the SM that has not been measured experimentally, yet. Nevertheless, if the Higgs mechanism is to solve the problem of particle masses within the SM, theoretical bounds can be calculated, which restrict the Higgs boson mass to lie below the TeV scale. Thus, the LHC experiments will be able to answer conclusively whether the Higgs mechanism is indeed the origin of mass.

The ATLAS collaboration has made considerable efforts to determine the discovery potential of the ATLAS detector over the full range of hypothetical Higgs boson masses. The amount of data and therefore time necessary for a discovery varies significantly with the mass of the Higgs boson. In the intermediate mass range, where the Higgs boson decays predominantly into a pair of W or Z bosons, a discovery might be possible in the near future. A mass close to the lower limit of 114.4 GeV, however, as is indeed favoured by the other SM observables, makes a larger data sample necessary. As the Higgs boson decays predominantly into the heaviest particles available, one of the most promising decay modes in this mass regime is $H \to \tau\tau \to \ell h$. The last major study considering the nominal centre-of-mass energy of 14 TeV and a large data sample was performed in [1]. This thesis updates the study in the light of more recent developments and significantly enhances it in several crucial aspects.

The LHC is a discovery machine not only by virtue of high centre-of-mass energy but also by high luminosity. While the former increases the probability of interesting processes (production of heavy particles) to happen *if* a collision occurs, the latter determines the *rate* of collisions. Given an effective area over which the accelerated particles are spread out (determined by beam optics), there are two additional parameters with which luminosity can be adjusted. Instead of accelerating and colliding single protons, they are bunched together. The denser such a bunch is, the more likely it is that a collision takes place. The second parameter is the frequency with which bunches intersect. Increasing luminosity by these two parameters comes at a price known as *pile-up*. At some point the average number of collisions per bunch crossing becomes greater than one. Although the probability of two interesting processes happening at the same time is minimal, this effect, called *in-time pile-up*, increases the activity in the detector. A high bunch crossing frequency becomes a problem if the signal processing time of a detector component is much larger than the time between two bunch crossings. In that case, the detector integrates over several collisions, adding false signatures to the real ones (*out-of-time pile-up*). This thesis presents the first study of VBF $H \to \tau\tau \to \ell h$ that includes pile-up for all relevant processes and investigates the effects in detail.

The second new aspect is the estimation of background processes to this channel from data instead of using simulation. Given a selected data sample, it is a-priori not known what kind of processes it consists of. One strategy is to simulate all processes which are expected to be found in the data sample and apply the same requirements. The accuracy of this depends on the amount of simulated events and how well the simulation models real physics. The second strategy is to obtain such samples from data itself. However, the selected data sample is enriched with signal events and also contains other background processes. The challenge is to find a region of phase space in which events have the same relevant properties as in the signal region but optimally only contains the background process in question. In this thesis, two novel techniques are presented which make it possible to extract all dominant background processes from data.

With the rapid increase of data collected by the LHC experiments many of the assumptions being made in this thesis can now be experimentally tested. The increase in luminosity during data taking in 2011 creates pile-up conditions which are close to the assumed pile-up scenario investigated in this thesis. With this in mind, preparing for future Higgs boson searches in VBF $H \to \tau\tau \to \ell h$ becomes imperative.

2

Theoretical Foundations

Particle physics aims to describe the fundamental building blocks of nature. During the last century, our understanding of these has made remarkable progress. At the end of the 19th century, the *atom* (from Greek átomos for uncuttable) was thought to be elementary. Today we know that atoms in fact consist of a nucleus and an *electron* shell. Electrons are elementary particles, but the nucleus is made up of *protons* and *neutrons*, which in turn are composed of *quarks*.

To our present knowledge, there are 12 fundamental matter constituents. 6 *leptons*: Electrons, *muons*, *taus* and their corresponding *neutrinos*. 6 quarks: *up*, *down*, *charm*, *strange*, *top* and *bottom*. These *fermions*, particles with spin ½, are listed in Table 2.1. They can be grouped into three generations, which have similar properties except for the increasing particle masses. Ordinary, stable matter is made up entirely of electrons, up- and down-quarks, i.e. first generation fermions.

All forces encountered in nature can be traced back to four fundamental *interactions*: *Gravity*, *electromagnetism*, *weak*- and *strong* interaction. Fundamental interactions are mediated by particles, *bosons*, with integer spin. Interactions and associated mediators are shown in Table 2.2. Electromagnetism, with the photon as force carrier, acts on electrically charged particles. The strong interaction, with gluons as mediators, acts only on particles with *colour charge*, – quarks and gluons themselves – whereas the weak interaction acts on all fermions. The $W^{\pm}$ bosons and the Z boson are the mediators of the weak interaction. Photons, gluons W and Z bosons are spin 1 particles. Gravity is believed to be mediated by *gravitons*, a Spin 2 particle. However, attempts to consistently formulate gravity as a quantum field theory have been unsuccessful so far, and the graviton has not yet been discovered.

generation	particle		el. charge [e]	mass
1^{st}	e	electron	-1	511 keV
	ν_e	electron neutrino	0	< 2 eV
	u	up quark	2/3	2.49 MeV
	d	down quark	-1/3	5.05 MeV
2^{nd}	μ	muon	-1	105.7 MeV
	ν_μ	muon neutrino	0	< 0.19 MeV
	c	charm quark	2/3	1.27 GeV
	s	strange quark	-1/3	101 MeV
3^{rd}	τ	tau	-1	1.7768 GeV
	ν_τ	tau neutrino	0	< 18.2 MeV
	t	top quark	2/3	172 GeV
	b	bottom quark	-1/3	4.19 GeV

Table 2.1.: Fundamental fermions and their properties [2]. Throughout this thesis a system of natural units is used, where $\hbar = c = 1$.

interaction	force carrier		mass
electromagnetism	γ	photon	0
weak	W^+, W^-		80.4 GeV
	Z^0		91.2 GeV
strong	g	gluon	0
gravity		graviton	0

Table 2.2.: Carriers of fundamental interactions [2].

2.1. The Standard Model

The line of argumentation presented here closely follows [3] and [4]. The particles and interactions mentioned above are described by the *Standard Model of Particle Physics (SM)*. In a quantum field theory the Lagrangian density, $\mathcal{L}$, describes the dynamics of a system. Global symmetries in a system are connected to conserved quantities. Global gauge symmetries e.g. lead to conserved charges. The SM follows from requiring the Lagrangian density of the free fermion fields to be invariant under *local* gauge transformations. As a consequence of this requirement *gauge fields* have to be introduced, which represent the fundamental interactions described above. In this formalism gauge transformations are represented as groups, with the generator(s) of the group corresponding to the force carriers of the corresponding interaction. As a whole, the SM is invariant under

$$SU(3)_C \times SU(2)_T \times U(1)_Y$$

transformation. The first term is related to the strong interaction, with the superscript C denoting colour charge. The underlying gauge theory, *Quantum Chromo Dynamics (QCD)* plays a major role at hadron colliders. See e.g. [3] for a detailed a detailed description. The remaining two terms form the *electroweak* symmetry group, which unites electromagnetic and weak interactions. Subscripts T and Y denote *weak isospin* and *weak hypercharge* respectively.

2.1.1. Electroweak interactions

The electroweak sector of the SM is a chiral gauge theory: The particles described here are left and right-handed fermions. Charged-current weak interactions act only on left-handed fermions. This part of the electroweak interaction is represented by the $SU(2)_T$ symmetry group. Left-handed fermions are grouped into $SU(2)$-doublets. Q_L and L_L represent left-handed quark and lepton doublets. Right-handed fermions, on the other hand, are singlets under $SU(2)_T$ transformations. The quantum number associated with $SU(2)_T$ is the weak isospin T. An additional $U(1)$ symmetry group is necessary in order to fully describe neutral current interactions in the SM (see below for details). The corresponding quantum number is *weak hypercharge*. The assignment of fermions to doublets and singlets and their corresponding electroweak quantum numbers are summarised in Table 2.3. With these representations the Lagrangian density for free particles reads:

$$\mathcal{L}_{\text{free}} = i\bar{L}_L\gamma^\mu\partial_\mu L_L + i\bar{\ell}_R\gamma^\mu\partial_\mu\ell_R + i\bar{\nu}_R\gamma^\mu\partial_\mu\nu_R + i\bar{Q}_R\gamma^\mu\partial_\mu Q_R + i\bar{u}_R\gamma^\mu\partial_\mu u_R + i\bar{d}_R\gamma^\mu\partial_\mu d_R. \tag{2.1}$$

Here, γ^μ are the γ-matrices (cf. [3]). A local gauge or phase transformation can be formulated the following way:

$$Q_L/L_L \rightarrow e^{i/2(\vec{\alpha}(x)\vec{\tau}+\beta(x)Y)}Q_L/L_L, \tag{2.2}$$

		$SU(2)_T$		$U(1)_Y$
	fermion	T	T_3	Y
$Q_L = \begin{pmatrix} u \\ d' \end{pmatrix}$		1/2	+1/2 -1/2	+1/3
	u_R	0	0	+4/3
	d'_R	0	0	-2/3
$L_L = \begin{pmatrix} \nu_\ell \\ \ell \end{pmatrix}$		1/2	+1/2 -1/2	-1
	ν_R	0	0	0
	ℓ_R	0	0	-2

Table 2.3.: Chiral fermion representations and the corresponding electroweak quantum numbers. Note that d' denotes the electroweak eigenstate of a down-type fermion instead of the mass eigenstate. This designation is dropped in the following for matters of simplicity.

$$u_R/d_R/\ell_R \to e^{i/2\beta(x)Y} u_R/d_R/\ell_R. \tag{2.3}$$

$\vec{\tau}$ contains the three generators of the $SU(2)_T$ group, which can be represented by the Pauli matrices $\vec{\sigma}$. Y is the weak hypercharge operator. The Lagrangian density in Eq. 2.1 is not per se invariant under these transformations. To make it invariant, the derivative ∂_μ has to be replaced by the *covariant derivative*

$$D_\mu = \partial_\mu + ig\frac{\vec{\sigma}}{2}\vec{W}_\mu + ig'\frac{Y}{2}B_\mu. \tag{2.4}$$

g and g' are the coupling constants of the gauge groups, the σ_i are the Pauli matrices. In order to make the Lagrangian density locally gauge invariant, new vector fields have to be introduced. Replacing the derivatives in Eq. 2.1 with the covariant derivative gives a gauge invariant Lagrangian density with new interactions.

$$\begin{aligned}
\mathcal{L} = \quad & \bar{L}_L \left[i\partial_\mu - g\frac{\vec{\sigma}}{2}\vec{W}_\mu - g'\frac{Y}{2}B_\mu \right] L_L + \bar{Q}_L \left[i\partial_\mu - g\frac{\vec{\sigma}}{2}\vec{W}_\mu - g'\frac{Y}{2}B_\mu \right] Q_L \\
+ \quad & \bar{\ell}_R \left[\partial_\mu - g'\frac{Y}{2}B_\mu \right] \ell_R + \bar{u}_R \left[\partial_\mu - g'\frac{Y}{2}B_\mu \right] u_R + \bar{d}_R \left[\partial_\mu - g'\frac{Y}{2}B_\mu \right] d_R \\
+ \quad & i\bar{\nu}_R \partial_\mu \nu_R \\
- \quad & \frac{1}{4} W_{\mu\nu} \cdot W^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}
\end{aligned} \tag{2.5}$$

Terms in the last line denote the field strength tensors $B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$ and $\vec{W}_{\mu\nu} = \partial_\mu \vec{W}_\nu - \partial_\nu \vec{W}_\mu + ig\vec{W}_\mu \times \vec{W}_\nu$. Eq. 2.5 shows the pattern of interactions described above: Only left-handed fermion fields interact with the W_μ. Right-handed neutrinos do not interact at all.

However, the W_μ and B_μ fields are not the physical fields we observe in nature. W_μ^1 and W_μ^2 mix to form the eigenstates of the charge operator:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}\left(W_\mu^1 \mp i W_\mu^2\right) \tag{2.6}$$

The electric charge, Q, is related to the weak isospin and hypercharge by the Gell-Mann–Nishijima formula, $Q = T_W^3 + Y/2$ [5, 6]. The two neutral electroweak gauge bosons, the Z^0 boson and the photon, have to be mixed states of the W_μ^3 and B_μ fields,

$$\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_w & \sin\theta_w \\ -\sin\theta_w & \cos\theta_w \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix} \tag{2.7}$$

with the *electroweak mixing angle* θ_w. The B_μ couples to left-handed neutrinos with hypercharge. Photons on the other hand do not interact with neutrinos, which are electrically neutral. Also, the W_μ^3 cannot be the Z^0 because the latter couples to left-handed and right-handed fermions.

The SM is remarkably successful in describing elementary particles and their interactions. It has been tested experimentally and shown to agree with measurements down to the level of quantum corrections [7]. Furthermore, the SM as a gauge theory is automatically renormalisable [8, 9]. Nonetheless, the Lagrangian density in Eq. 2.5 has what appears to be a serious flaw: both fermion and boson fields are massless, and mass terms cannot be explicitly inserted because they would destroy the gauge invariance. While the photon is indeed massless and first generation fermions are at least comparatively light, W and Z bosons and the top-quark are very massive[1].

2.2. The Higgs Mechanism

To allow for particle masses while leaving the Lagrangian density locally gauge invariant a mechanism is included in the SM to generate masses via interaction with a new scalar field. This scalar field has a non-zero vacuum expectation value, which spontaneously breaks the $SU(2)_T \times U(1)_Y$ symmetry. Such a spontaneous symmetry breaking generates masses for the bosons associated with the broken symmetries. As the photon is massless the underlying $U(1)_{\text{em}}$ symmetry has to remain intact. For the so-called Higgs mechanism [10–12] a scalar field is added to the theory. As the intention is to generate masses by interaction with this scalar field it has to have weak isospin and hypercharge. The simplest way to construct such a scalar field is a single complex isospin doublet:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{2.8}$$

[1] The top quark has nearly the mass of a whole gold atom (183 GeV).

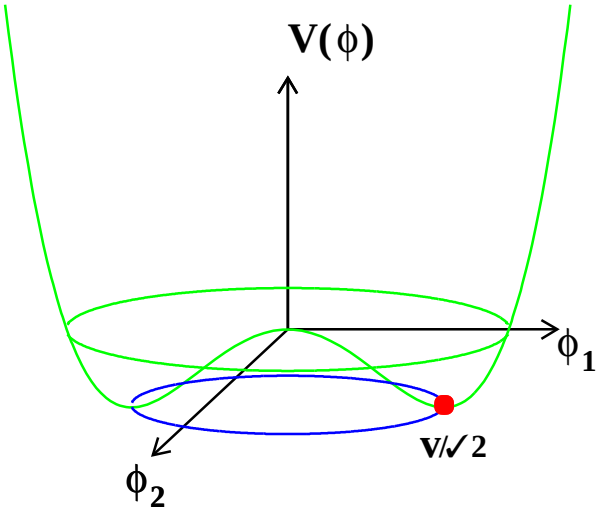


Figure 2.1.: The Higgs potential as a function of two out of its four components.

The Lagrangian density in Eq. 2.5 is thus extended by the kinetic energy terms of the new scalar field, its interaction terms with the $W_\mu^{1,2,3}$ and B_μ fields and a potential term:

$$\mathcal{L}_\phi = \left| \left(i\partial_\mu - g\frac{\vec{\sigma}}{2} \cdot \vec{W}_\mu - g'\frac{Y}{2}B_\mu \right) \phi \right|^2 - V(\phi) \tag{2.9}$$

The potential term has the following form:

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda \left(\phi^\dagger \phi \right)^2 \tag{2.10}$$

The parameter λ has to be larger than zero for the total energy to have a bound from below. Apart from this, the shape of the potential now depends on μ^2. Minima have to fulfil the requirement

$$\left(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2 \right) = -\frac{\mu^2}{\lambda}. \tag{2.11}$$

For $\mu^2 > 0$ the only solution is $\phi_{1,2,3,4} = 0$, leading to a potential with a single ground state. $\mu^2 < 0$ however yields a potential as sketched in Fig. 2.1 (for two parameters of the four components), with an infinite number of ground states with non-vanishing expectation value. They form the surface of a four dimensional sphere - or a circle in the two dimensional example. Thus all these ground states are equally likely.

Any choice will spontaneously break the symmetry of the system, resulting in masses for the heavy gauge bosons. But as the photon is known to be massless, the $U(1)_{\mathrm{EM}}$ symmetry has to remain unbroken. Hence, the ground state, ϕ_{min}, must be invariant under $U(1)_{\mathrm{EM}}$ transformations, or

$$Q\phi_{\mathrm{min}} = 0. \tag{2.12}$$

A ground state with non-vanishing vacuum expectation value therefore has to have an electrically neutral component. The generator of the electric charge fulfils $Q = T_W^3 + Y/2$. This, in principle, leads to two possible configurations of ϕ,

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \text{ or } \phi = \begin{pmatrix} \phi^0 \\ \phi^- \end{pmatrix}, \tag{2.13}$$

with hypercharge $Y = 1$ or $Y = -1$, respectively. The superscripts indicate the electric charge. The ϕ field, however, couples to fermions via terms like $g_d \overline{Q}_L \phi u_R$ (cf. Eq. 2.25 below), which require ϕ to have $Y = 1$. An appropriate choice for a ground state is then

$$\phi_{\text{min}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{2.14}$$

with Y = 1 and $v^2 = -\mu^2/\lambda$. Masses of the gauge bosons can now be extracted by inserting ϕ_{min} into Eq. 2.9.

$$\mathcal{L}_\phi = \left(\frac{1}{2} g v\right)^2 W_\mu^+ W^{-\mu} + \frac{1}{8} v^2 \left(W_\mu^3, B_\mu\right) \begin{pmatrix} g^2 & g g' \\ g g' & g'^2 \end{pmatrix} \begin{pmatrix} W^{3\mu} \\ B^\mu \end{pmatrix} \tag{2.15}$$

The first term can be identified as a mass term for the $W^\pm$ bosons,

$$M_W = \frac{1}{2} v g \tag{2.16}$$

From the non-diagonal mass matrix in the second term of Eq. 2.15 it is clear that the W_μ^3 and B_μ bosons are not mass eigenstates. With

$$\cos\theta_w = g/\sqrt{g^2 + g'^2}, \tag{2.17}$$

replacing W_μ^3 and B_μ by Z_μ and A_μ according to Eq. 2.7 diagonalises the mass matrix.

$$\mathcal{L}_\phi = \left(\frac{1}{2} g v\right)^2 W_\mu^+ W^{-\mu} + \frac{1}{8} v^2 \left(g^2 + g'^2\right) \left(A_\mu, Z_\mu\right) \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} A^\mu \\ Z^\mu \end{pmatrix} \tag{2.18}$$

The photon field A remains massless, while the Z^0 boson is now massive:

$$m_Z = \frac{1}{2} v \sqrt{g^2 + g'^2}. \tag{2.19}$$

Excited states of the scalar field ϕ can be expressed as an expansion around the ground state ϕ_{min}.

$$\phi = e^{i\vec{\sigma}\cdot\vec{\theta}(x)} \left[\phi_{\text{min}} + \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ H(x) \end{pmatrix} \right] \tag{2.20}$$

This parametrisation still has four degrees of freedom as Eq. 2.8. Gauge invariance implies, that an arbitrary gauge may be chosen without changing the physics of the system. The exponential term in Eq. 2.20 has the form of an $SU(2)$ gauge (cf. Eq. 2.2). Hence, a gauge transformation – the so-called unitary gauge – can be applied such that the exponential factor is cancelled. Only the variation of $H(x)$, i.e. in radial direction with respect to the blue circle in Fig. 2.1, represents a change in energy. The $\vec{\theta}$-fields are not physical fields but Goldstone bosons [13]. ϕ now has only one degree of freedom. The additional three are realised in the form of longitudinal polarisations of the now massive gauge bosons. $H(x)$ is a boson with spin 0, the *Higgs boson*. This means, if the Higgs mechanism is indeed realised in nature, an additional particle has to exist, which can be searched for at e.g. collider experiments.

Substituting the rephrased ϕ in Eq. 2.9 leads to a mass term for the Higgs boson and couplings to the gauge bosons.

$$m_H = v\sqrt{2\lambda} \tag{2.21}$$

$$g_{HWW} = \frac{2m_W^2}{v} = g\, m_W \tag{2.22}$$

$$g_{HZZ} = \frac{2m_Z^2}{v} = \frac{g}{\cos\theta_w} m_Z \tag{2.23}$$

The masses of the W and Z bosons can be expressed by known quantities within the SM. The vacuum expectation value v is related to the Fermi coupling constant G_F, which can be measured in nuclear β decays:

$$v = \left(\sqrt{2}G_F\right)^{-1/2} \approx 246\,\text{GeV} \tag{2.24}$$

Fermion masses are not directly generated by the Higgs mechanism. However, gauge invariant terms can be included in the Lagrangian density by hand:

$$\mathcal{L}_{m_u,m_d} = -g_d \bar{Q}_L \phi u_R - g_u \bar{Q}_L i\sigma_2 \phi^* d_R \tag{2.25}$$

An equivalent term gives masses to the leptonic fields. This results in the following fermion masses and couplings to the Higgs boson:

$$m_f = \frac{g_f v}{\sqrt{2}} \tag{2.26}$$

$$g_{Hff} = \frac{m_f}{v} = \frac{g m_f}{2m_W} \tag{2.27}$$

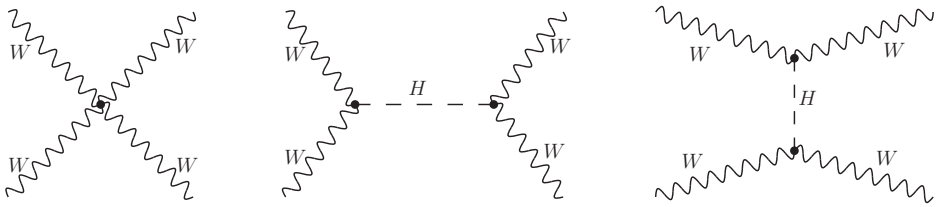


Figure 2.2.: Some tree-level Feynman diagrams for scattering of W bosons.

It is worth noting that the Higgs boson coupling to fermions is proportional to the fermion mass. Therefore the Higgs boson will decay predominantly into the pair of fermions with the highest mass. Fermion masses themselves are not predicted by the Higgs mechanism. They have to be measured.

2.3. Bounds on the Higgs Boson Mass

Although the Standard Model predicts the vacuum expectation value, it makes no prediction of the Higgs boson mass. m_H is the only free parameter of the model. Nevertheless limits can be established from theoretical input, precisions measurements of other SM parameters and direct searches.

2.3.1. Theoretical Bounds

The theoretical bounds presented in the following are not actual limits on m_H as e.g. the relation between m_W and m_Z. Rather, they mark boundaries below or above which the Higgs mechanism would not be compatible with the SM. A more detailed discussion can be found in [14].

Tree-level Unitarity

Interactions of longitudinal components of W and Z bosons lead to cross sections which increase with energy. At some point this would violate unitarity. This behaviour is cancelled by including the Higgs boson into the theory. Fig. 2.2 shows some diagrams contributing to W^+W^- scattering as an example. By decomposing the scattering amplitudes into partial waves of orbital angular momentum an upper bound on m_H can be derived at which the Higgs boson still restores unitarity:

$$m_H \lesssim 710\,\text{GeV} \tag{2.28}$$

Equivalently, a too heavy Higgs boson will effectively not contribute to the theory and thus not restore unitarity. It should be noted, however, that the Higgs boson self-coupling λ be-

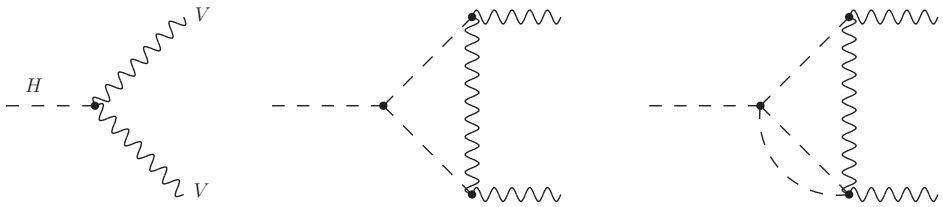


Figure 2.3.: Generic diagrams of Higgs boson decays into vector bosons with one- and two-loop corrections.

comes strong at high masses. Therefore radiative corrections can become large, rendering the theory non-perturbative. The argument is valid only assuming the Standard Model remains perturbative.

Perturbativity in Higgs Boson Decays

Requiring the Standard Model to remain a perturbative theory also leads to constraints from decays of the Higgs boson itself. Fig. 2.3 shows generic diagrams for one- and two-loop corrections of decays into heavy vector bosons. They involve the quartic coupling λ, which grows as a function of m_H. For $m_H \sim \mathcal{O}(10\,\text{TeV})$ the one-loop term grows close to the leading order term, which means the perturbative series is not convergent. Two-loop contributions become as large as one-loop contributions already at $m_H \sim \mathcal{O}(1\,\text{TeV})$.

Triviality and Stability Bounds

More stringent bounds can be established from the energy scale dependence of the already mentioned quartic coupling λ. Taking into account only contributions of the Higgs boson itself to λ – Fig. 2.4 shows typical diagrams up to the one-loop level – the solution to the renormalisation group equation, again at one-loop level, reads:

$$\lambda(Q^2) = \frac{\frac{m_H^2}{2v^2}}{1 - \frac{3}{4\pi^2}\frac{m_H^2}{2v^2}\log\frac{Q^2}{v^2}}. \tag{2.29}$$

Here, $m_H = v\sqrt{2\lambda(v^2)}$. $\lambda(Q^2)$ varies logarithmically with Q^2 and becomes infinite at the so-called Landau pole.

$$\Lambda_{\text{pole}} = v \exp\left(\frac{4\pi^2 v^2}{3m_H^2}\right) \tag{2.30}$$

The position of this divergence depends on m_H. Thus, the upper bound on the Higgs mass depends on the energy scale up to which the Standard Model is supposed to valid.

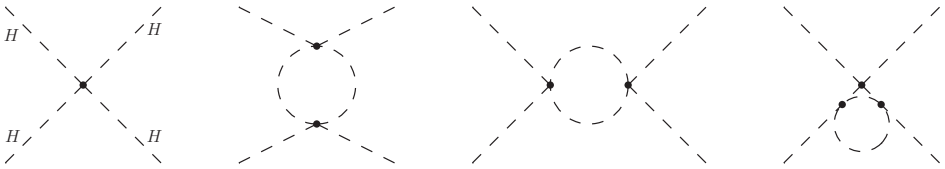


Figure 2.4.: Generic diagrams of Higgs boson decays into vector bosons with one- and two-loop corrections

To ignore all contributions except those from the Higgs boson itself is only valid for large values of λ. In general, also contributions from fermions and gauge bosons have to be considered. An approximate solution to the renormalisation group equation, including only the top quark and heavy gauge bosons and assuming $\lambda \ll \lambda_t, g_1, g_2$ (λ_t: top quark Yukawa coupling) is given by

$$\lambda(Q^2) = \frac{m_H^2}{2v^2} + \frac{1}{16\pi^2}\left[-12\frac{m_t^2}{v^4} + \frac{3}{16}\left(2g_2^4 + (g_2^2+g_1^2)^2\right)\right]\log\frac{Q^2}{v^2}. \tag{2.31}$$

Again, λ varies logarithmically with Q^2 and becomes negative for

$$m_H^2 < \frac{v^2}{8\pi^2}\left[-12\frac{m_t^2}{v^4} + \frac{3}{16}\left(2g_2^4 + (g_2^2+g_1^2)^2\right)\right]\log\frac{Q^2}{v^2}. \tag{2.32}$$

But from Eq. 2.10 it is clear that $\lambda < 0$ leads to an unstable vacuum as it is not bounded from below. Therefore requiring vacuum stability constrains the Higgs mass from below.

Fine-Tuning

Radiative corrections to the Higgs boson mass include at one-loop level diagrams as shown in Fig. 2.5. Cutting off the loop integral momenta at a scale Λ the result for m_H is approximately

$$m_H^2 = (m_H^0)^2 + \frac{3\Lambda^2}{8\pi^2 v^2}\left[m_H^2 + 2m_W^2 + m_Z^2 - 4m_t^2\right], \tag{2.33}$$

where m_H^0 is the bare mass contained in the unrenormalised Lagrangian. While other divergences encountered in the Standard Model are of logarithmic nature the Higgs mass is quadratically divergent. As pointed out in the previous sections, consistency of the SM demands a physical mass of $100\,\text{GeV} \lesssim m_H \lesssim 1\,\text{TeV}$. If the cut-off scale Λ is large, e.g. at the Grand Unification Scale $\sim 10^{16}\,\text{GeV}$ [2], a fine-tuning of 16 digits between the bare mass and the radiative corrections is needed in order to obtain a physical mass in the desired range. If m_H fulfils the *Veltman condition*,

$$m_H^2 = 4m_t^2 - 2m_W^2 - m_Z^2. \tag{2.34}$$

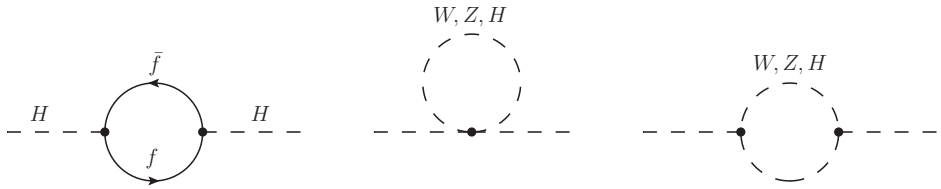
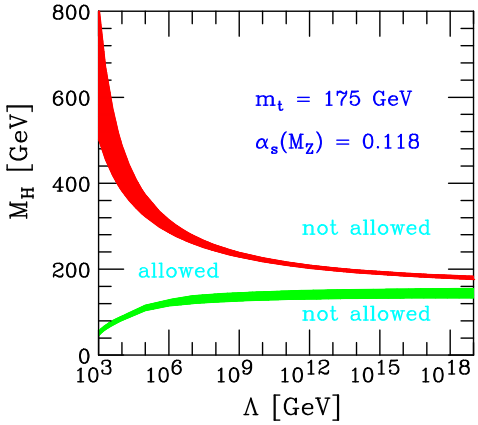
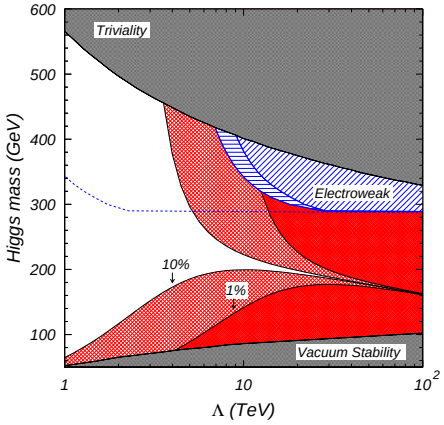


Figure 2.5.: One-loop corrections to the Standard Model Higgs boson mass



(a) Upper and lower bounds on m_H from triviality of the Landau pole and vacuum stability respectively.



(b) Contours for the fine-tuning parameter Δ_{FT} as function of m_H and Λ. The region labelled 1% (10%) represents a fine-tuning of greater than 1 part in 100 (10).

Figure 2.6.: Theoretical bounds on the Higgs boson mass as a function of the scale Λ up to which the SM is to be valid, from [14].

the quadratic divergences in Eq. 2.33 cancel. Moreover, since m_t, m_W and m_Z have already been measured this would predict $m_H \approx 310\,\mathrm{GeV}$. Unfortunately Eq. 2.33 is obtained at the one-loop level and the cancellation no longer occurs at higher-order corrections.

Although fine-tuning is unsatisfying from a scientific point of view, it cannot be ruled out entirely. Furthermore, the amount of fine-tuning which is considered acceptable is rather arbitrary. Defining $\Delta_{\mathrm{FT}} = \Delta m_H^2(\Lambda^2)/m_H^2$ as a measure of fine-tuning, the weak scale is fine-tuned to one part in Δ_{FT}. For larger values of the cut-off scale Λ the range of allowed Higgs boson masses for which the fine-tuning is smaller than a given level becomes increasingly smaller.

Combined Theoretical Bounds

Figure 2.6 shows the theoretical bounds on m_H as a function of Λ with and without considering fine-tuning. If Λ is of the order of a few tens of TeV, $m_H \sim 200\,\mathrm{GeV}$ is a preferred value. Not

considering fine-tuning m_H has to lie in the range

$$130\,\text{GeV} < m_H < 180\,\text{GeV} \tag{2.35}$$

if the SM is to be valid up to the Grand Unification scale.

2.3.2. Experimental Bounds

Direct Limits

Direct limits on the Higgs boson mass come from searches for the Higgs boson at various experiments. The LEP experiments set an upper limit on the mass of a SM Higgs boson of

$$m_H > 114.4\,\text{GeV}, \tag{2.36}$$

at 95% confidence level (Fig. 2.7) [7], while Tevatron experiments excluded $156 < m_H < 177\,\text{GeV}$ (Figures 2.7 and 2.8) [15]. Recently published results from the LHC experiments establish an even larger range excluded values of m_H. ATLAS excludes a SM Higgs in most of the range between $m_H = 146\,\text{GeV}$ and $466\,\text{GeV}$ at 95% confidence level (Fig. 2.9) [16]. CMS excludes most values of m_H between $145\,\text{GeV}$ and $400\,\text{GeV}$ (Fig. 2.10) [17]. Given the dramatic increase of LHC luminosity new limits or a discovery are expected in the near future.

Indirect Limits

The Higgs boson mass is linked to the other observables of the SM via loop corrections. Assuming the Standard Model to be the correct theory of nature, it is possible to deduce the most probable Higgs boson mass by fitting all Standard Model parameters to the very precisely measured electroweak observables. The $\Delta\chi^2$ of this fit to the precision observables is shown in Fig. 2.7. Under the hypothesis that the SM is a valid theory in this energy regime, the lowest $\Delta\chi^2$ is obtained for $m_H = 89\,\text{GeV}$. An upper limit on the Higgs boson mass can be derived as $161\,\text{GeV}$ at the 95% confidence level.

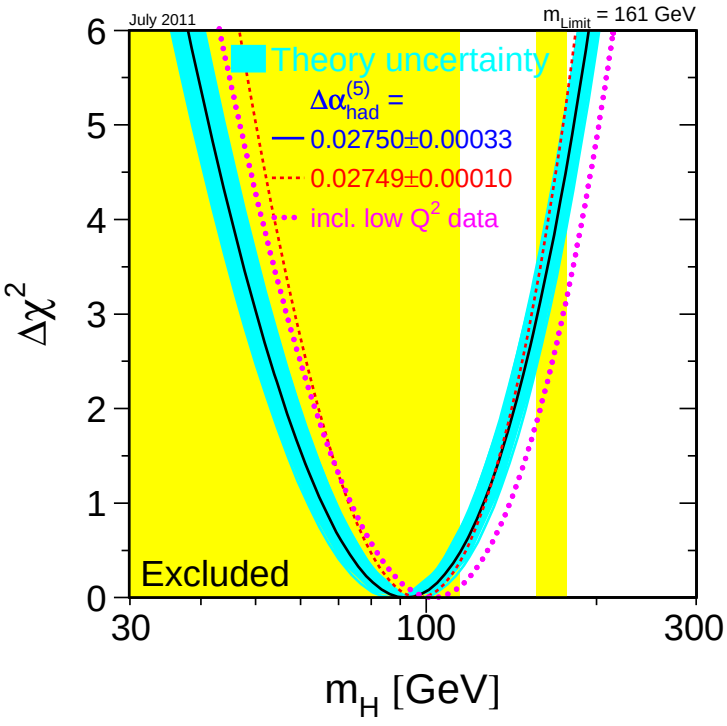


Figure 2.7.: *Direct limit:* The yellow area shows the range of m_H excluded by direct searches. *Indirect limit:* The lines give the $\Delta\chi^2$ of a global fit of parameters of the Standard Model to data as a function of the assumed Higgs boson mass (from [7]).

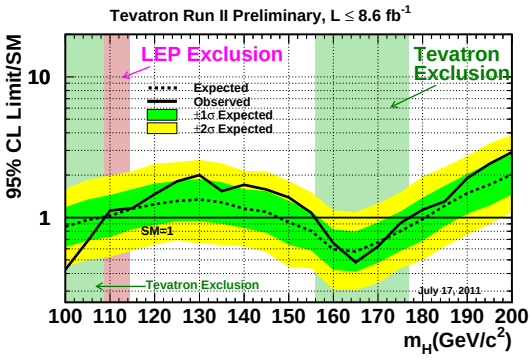


Figure 2.8.: Combined Tevatron Higgs boson exclusion limit as a function of the hypothetical Higgs boson mass. [15]

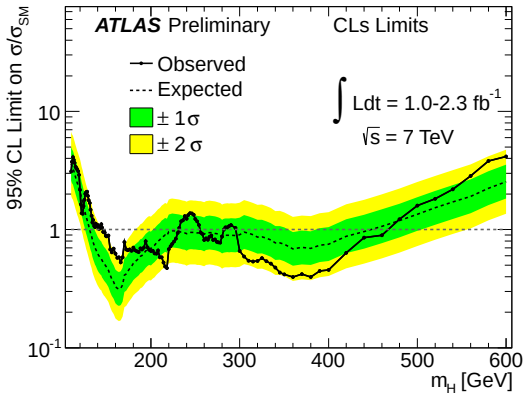


Figure 2.9.: Combined ATLAS exclusion limit for a Standard Model Higgs boson. [16]

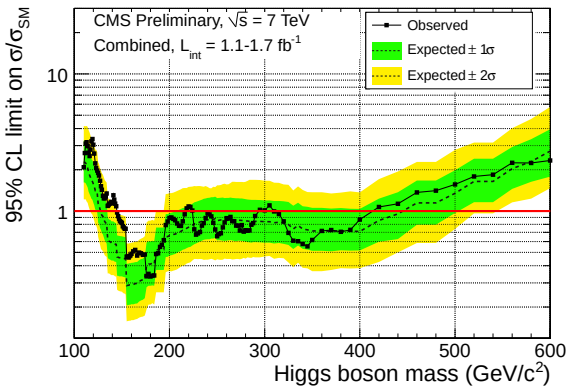


Figure 2.10.: Combined CMS exclusion limit for a Standard Model Higgs boson. [17]

3
The Experiment

Theories in high energy physics often predict the existence of new particles. These new particles usually have either high masses or small couplings to already known particles, which is why they have not been seen yet. One way to study heavy particles is to produce them resonantly by colliding two particles with a centre-of-mass energy at least as large as the mass of the new particle.

At a hadron collider machine like the LHC the most basic process is the collision of two colour-charged elementary particles, i.e. quarks or gluons. Most of the time, both particles undergo an elastic scattering at a small angle with respect to the beam. On rare occasions, for instance when a heavy particle decays, particles are produced at larger angle and, correspondingly, high momentum perpendicular to the beam axis[1]. Such a process is usually referred to as *hard process*. However, the picture is much more complex: A scattering of particles means that (colour-) charged objects are accelerated, which possibly leads to bremsstrahlung. Emissions can come from the incoming particles (*initial state radiation or ISR*) or from the outgoing ones (*final state radiation, FSR*). Quarks and gluons are confined in hadrons. This affects the incoming and the outgoing part of the collision. The outgoing particles form new hadrons if they carry colour charge. At the LHC, the incoming particles are protons. At relativistic energies, quarks and gluons can be interpreted as quasi-free *partons* within the proton which carry a statistically distributed fraction of the proton momentum [18]. To make a collision an even more complex event, more than one parton within each proton can interact. Interactions in addition to the hard process are often referred to as *underlying event*. Because accelerating and colliding individual protons is not economic, *bunches* of protons are accelerated and brought to collision at once. Depending of the density of the bunch and the number protons it consists of, several proton-proton collision can happen during the same *bunch crossing*. This effect is called *pile-up*.

At the LHC, protons are brought to collision at four different points. To study the properties of the produced particles, large detectors have been constructed at each of the four points of

[1] The momentum component perpendicular to the beam axis is referred to as transverse momentum (p_{T}).

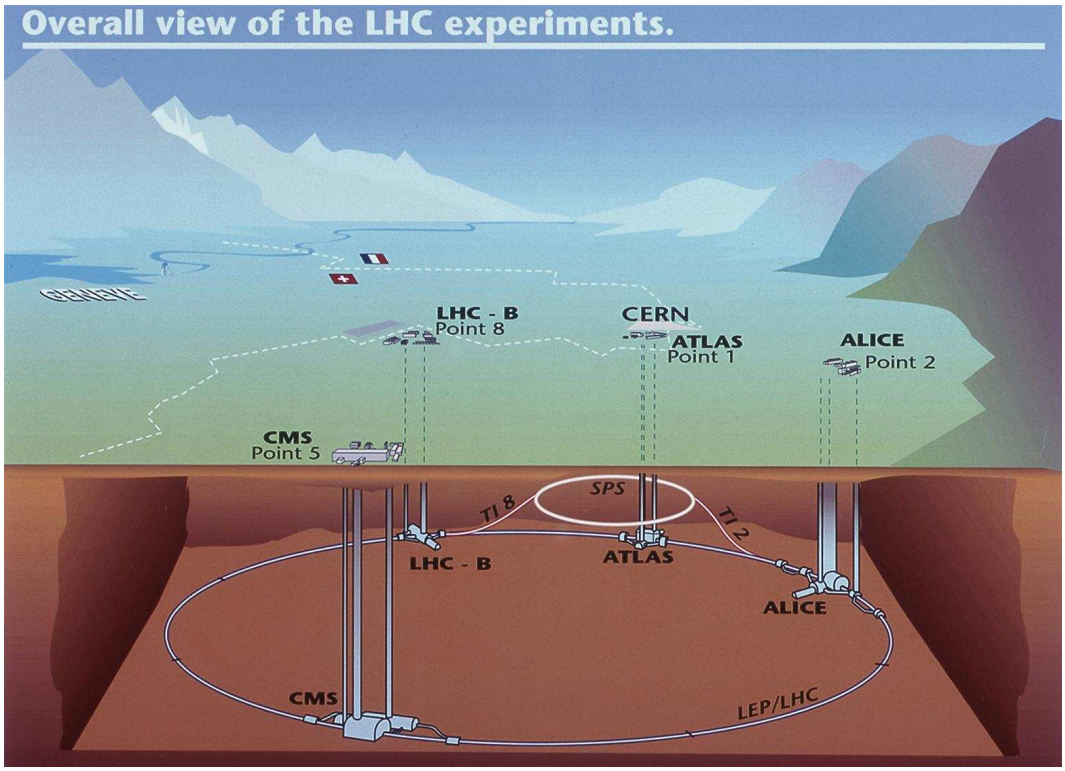


Figure 3.1.: The Large Hadron Collider and the four (largest) experiments [25].

collision. All four detectors are actually a combination of several sub-detectors. ATLAS [19] and CMS [20] are multi-purpose detectors. ALICE [21] was built to investigate heavy-ion collisions resulting in a quark-gluon plasma. The purpose of LHCb [22] is to study B-physics at the LHC.

This thesis describes a specific search for the Higgs boson in an assumed data sample of $30\,\mathrm{fb}^{-1}$ (corresponding to three years of data taking) gathered at a centre-of-mass energy of 14 TeV with the ATLAS detector. Neither the centre-of-mass energy nor the size of the data sample have been delivered by the LHC, yet. Hence, instead of using real LHC collision data, the study has to be carried out using a simulated sample generated by several *Monte-Carlo* event generators. The ATLAS detector is replaced by a very detailed GEANT4 [23] [24] simulation. The algorithms used to reconstruct particles from detector signals are identical in both cases.

3.1. The Large Hadron Collider

The *Large Hadron Collider (LHC)* [26] is situated at the *European Centre of Particle Physics (CERN)* near Geneva, Switzerland (Fig. 3.1). It has a circumference of about 27 km and is designed to deliver a centre-of-mass energy of 14 TeV. It consists of a system of superconducting dipole magnets with a field strength of over 8 T to bend the beams for their circular orbit. To

achieve this, the whole 27 km are cooled down to a temperature of 1.9 K using supra-fluid helium. In addition to the dipole magnets, quadrupole magnets with gradients of about 230 T/m are used to focus the beams. Cooling is a complex and expensive task, thus it makes sense to keep the accelerator in one cryostat vessel. On the other hand, particles of the same charge are accelerated in both directions. Thus, two independent magnetic channels have to be enclosed into the same vessel.

The LHC itself is the last stage in a series of accelerators that accelerates protons to an energy of up to 7 TeV. It starts with a mundane bottle of hydrogen gas. Once stripped down to protons, these are accelerated to 50 MeV by the Linac2. From there they are further accelerated by the Booster (PSB), the Proton Synchrotron PS and the Super Proton Synchrotron SPS to the injection energy of the LHC of 450 GeV.

The LHC first started operation in fall 2008. After a severe malfunction only days after commissioning it had to be shut down. Since late 2009 it is operational again. Due to safety considerations, the LHC is currently running at a reduced centre-of-mass energy of 7 TeV. The design energy is planned to be reached around 2014.

3.1.1. Luminosity

One of the key parameters of any collider experiment is the luminosity. It is related to the event rate, $\dot{N}$, of a given process via

$$\mathscr{L} = \frac{\sigma}{\dot{N}}, \tag{3.1}$$

where σ denotes the cross section and $\mathscr{L}$ the luminosity. In general, discovering new particles requires them to be produced with a sufficiently high rate. While the production cross section is fixed for a given centre-of-mass energy, luminosity is determined by several machine parameters and thus can be tuned. For a particle collider employing bunched beams with an equal number of particles per bunch in both beams luminosity is given by

$$\mathscr{L} = f\frac{n^2}{A_{int}}. \tag{3.2}$$

f is the collision frequency, n is the number of particles per bunch. Two direct measures to increase luminosity are therefore increasing the number of particles per bunch and decreasing the space between two bunches in a beam, increasing f. A_{int} in Eq. 3.3 is the effective cross section area of the colliding beams in which interactions take place. It is determined by the beam optics of the collider:

$$A_{int} = 4\sqrt{\varepsilon_x \beta_x^* \varepsilon_y \beta_y^*} \tag{3.3}$$

Assuming particles deviate from the nominal beam position in the transverse plane according to a Gaussian distribution, β^*, the so-called beam waist, gives the variance of that distribu-

tion at the interaction point. The emittance ε is proportional to the area of the ellipse which envelopes the phase space distribution of beam particles.

To increase luminosity by tuning either of the parameters is a challenge in itself. Furthermore, such an increase comes at a price. These effects as they are directly reflected in the simulated data samples used for this study.

3.1.2. Pile-Up

The term *pile-up* is not unambiguously defined within the high energy physics community. Within the ATLAS collaboration proton-proton interactions are classified in the following way:

- *hard process*: a hard scattering of to partons, usually leading to high transverse momenta of the decay products
- *underlying event*: other reactions of partons from the same proton-proton collision as the hard process
- *pile-up (in-time)*: additional proton-proton collisions occurring in the same bunch crossing as the hard process
- *pile-up (out-of-time)*: reactions happening in other bunch crossings; recorded due to the finite integration time of detector components.

Collisions contributing to in-time pile-up are usually governed by *minimum-bias* events. The term refers to events with just enough transverse momentum to be seen by the detector. The number of minimum-bias events per bunch crossing follows a Poisson distribution with mean

$$\langle N_{\mathrm{MB}} \rangle = \frac{\mathscr{L} \cdot \sigma_{\mathrm{inelastic}}}{f}, \tag{3.4}$$

where $\sigma_{\mathrm{inelastic}}$ is the cross section for inelastic, non-single-diffractive proton-proton scattering.

3.2. Event Generation

In order to simulate the outcome of particle collisions *Monte Carlo* event generators are commonly used in high energy physics. Given a Lagrangian density of the interaction, the Feynman rules can be used as a prescription to calculate processes via perturbation theory. Feynman rules define how to draw *Feynman diagrams* (many of which are shown in this thesis). For a given initial and final state, the simplest such graph(s) represents the *leading order (LO)* of the perturbation series. Higher orders are represented by additional lines, with internal lines being *virtual corrections* and external lines *real corrections*. The first order beyond LO

is usually referred to as *next-to-leading order (NLO)*. With this tool a *matrix element* can be calculated for a hard process. It gives probability densities for the kinematics of the process, which allows the generation of random events using Monte Carlo techniques.

3.2.1. Parton Shower

Matrix element calculations have some limitations. As described above, a scattering processes involves accelerating colour charges which leads to QCD bremsstrahlung. Fixed order matrix element calculations, however, diverge for collinear as well as soft emissions. These divergences have to be cancelled by higher order virtual corrections.

Instead of relying on matrix element calculations only, the strategy is to factorise the whole process into the hard processes calculated by a matrix element, ISR and FSR. The latter two are described by an approximate method, the so-called *parton shower (PS)*. The PS describes the evolution of a parton from a scale that is associated with the hard process down to the scale of hadronisation by consecutive emissions.

To implement this, the probability that a parton a splits into partons b and c at given scale Q^2 is needed. It can be obtained by approximating matrix element calculations in the soft/collinear limit. One obtains a set of differential equations, the so-called DGLAP equations [27],

$$\mathrm{d}\mathcal{P}_a = \frac{\alpha_s}{2\pi}\frac{\mathrm{d}Q^2}{Q^2}P_{a\to bc}(z)\mathrm{d}z, \tag{3.5}$$

where z denotes the fraction of energy b receives from a. $P_{a\to bc}(z)$ are so-called splitting kernels and give the probability for the type of splitting, i.e. $q \to qg$, $g \to gg$ or $g \to q\bar{q}$. In these probabilities the soft and collinear divergences are still present. This leads to the unpleasant fact that the total probability of a parton to split can be greater than one.

This is cured by the second ingredient of the PS. The *Sudakov form factor* [30] gives the probability that no splitting has occurred between two scales. With Eq. 3.5 the probability of a splitting in an infinitesimal scale-interval δQ^2 is given by:

$$P_{\text{splitting}} = \sum_{b,c}\int_z \frac{\mathrm{d}\mathcal{P}_a(z')}{\mathrm{d}z'}\mathrm{d}z' \cdot \delta Q^2 \tag{3.6}$$

The probability that no splitting happens follows from probability conservation as $P_{\text{no splitting}} = 1 - P_{\text{splitting}}$. Integrating over scale-intervals then gives the Sudakov form factor:

$$\Delta(Q^2_{\text{max}}, Q^2) = \exp\left(-\sum_{b,c}\int_{Q^2}^{Q^2_{\text{max}}}\frac{\mathrm{d}Q'^2}{Q'^2}\int_z \frac{\alpha_s}{2\pi}P_{a\to bc}(z')\mathrm{d}z'\right) \tag{3.7}$$

More details can be found in [31].

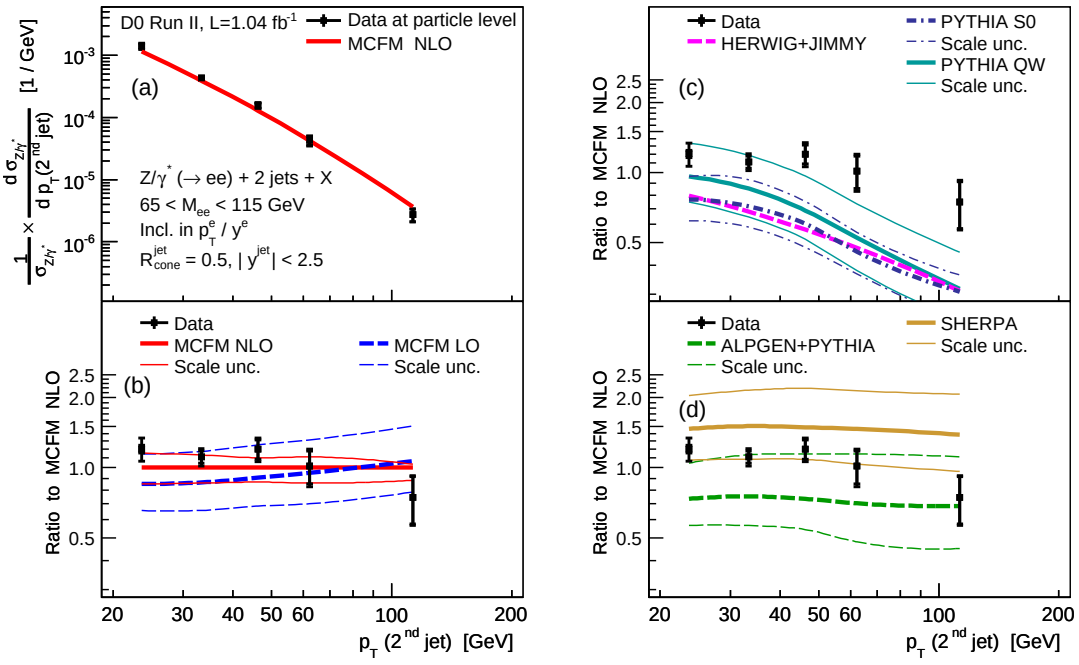


Figure 3.2.: p_T spectrum of the 2nd jet in Z+2jets events measured by D0 [35].

3.2.2. Parton Shower to Matrix Element Matching

Although parton shower MC like PYTHIA [32] or HERWIG [33] works well for a multitude of processes it fails for an important class of processes. Production of Z or W bosons in association with two or more hard jets is the dominant background to the signal process studied in this thesis (see Chapter 4). Here parton shower MC usually generates jets with a too soft p_T spectrum (see Fig. 3.2c). This failure is generally attributed to the nature of the parton shower itself, as it approximates soft and collinear emissions. However, the example of multi-jet production (Fig. 3.3) shows that the p_T spectrum of additional jets generated by a parton shower is not necessarily too soft. Parton showers generate emissions starting from a selected hard process. An event with a Z boson and two jets is produced e.g. by generating a Z boson and a recoiling jet initiated by a parton from a matrix element calculation. The second jet is then produced by the parton shower [32]. However, this procedure misses an entire class of events: Di-jet events in which a Z boson is radiated off a quark [34]. An example is shown in Figure 3.4.

To remedy this shortcoming, MC generators like ALPGEN [36] and SHERPA [37] use automated matrix element generators. These calculate matrix elements for all processes with the desired final state. The resulting events can then be used as input for the established parton showers of PYTHIA and HERWIG. But care has to be taken to avoid double counting with this approach. If e.g. Z boson production with additional jets is to be generated, double counting can occur when two jets are produced via the matrix element calculation or one jet comes from the matrix element calculation and an additional hard gluon is radiated off a parton by the PS. There are several methods to properly merge PS and matrix element. The *MLM matching* as employed by ALPGEN works like this [34]:

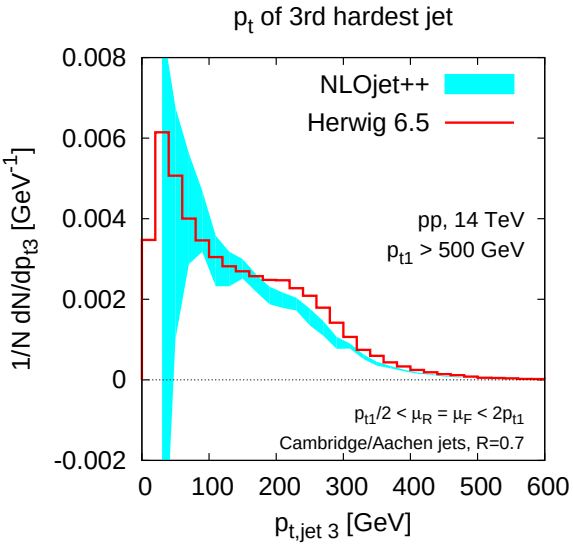


Figure 3.3.: p_{T} spectrum of the third jet in QCD multi-jet production. From [34].

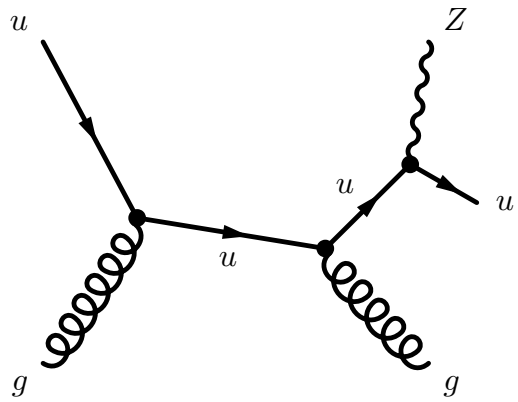


Figure 3.4.: Example Feynman diagram of a Z+2 jets event, in which the Z boson is radiated off a quark.

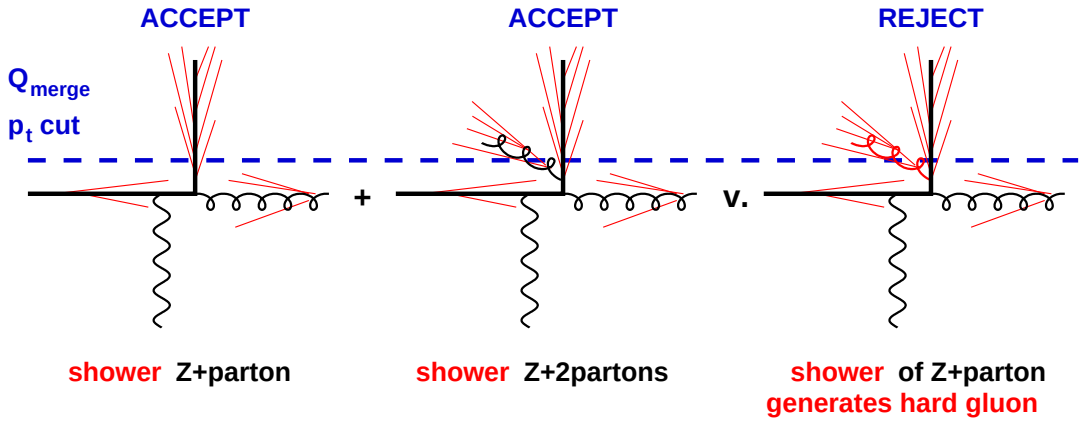


Figure 3.5.: Sketch of MLM matching in a Z+2 jets event. From [34]. Red lines indicate particles generated by a parton shower. Black lines symbolise partons from a matrix element calculation. The middle and right diagrams differ in the origin of the hard gluon radiation. The dashed blue line represents the Q_{merge} p_{T} cut.

Produce events from LO matrix element calculations, for e.g. Z + 1 parton, Z + 2 partons, ..., Z + N partons. The procedure is illustrated in Fig. 3.5. All partons are required to have $p_{\mathrm{T}} > Q_{\mathrm{ME}}$ and be separated by R_{ME}. These requirements restrict the usage of matrix elements to the region of phase-space where they work well. Each of these events is then processed by a PS program (e.g. HERWIG). Particle jets with $p_{\mathrm{T}} > Q_{\mathrm{merge}}$, with $Q_{\mathrm{merge}} \gtrsim Q_{\mathrm{ME}}$ are then identified with a jet algorithm. If each parton is close to a jet (in angle) and no additional jets exist above Q_{merge}, the event is accepted.

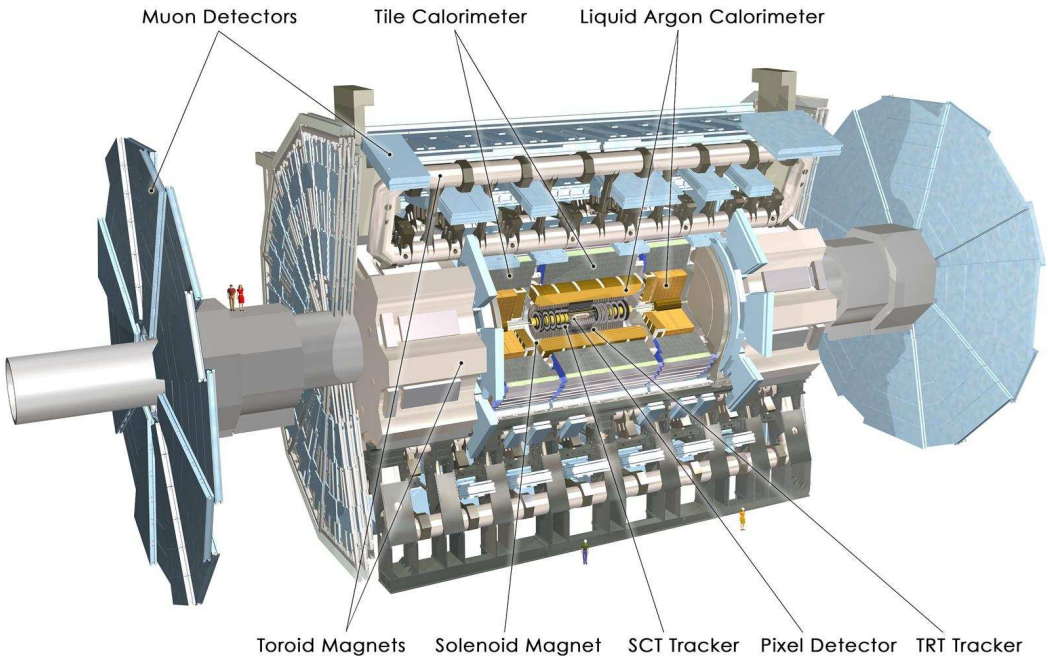


Figure 3.6.: Cross section of the ATLAS detector with all major sub-systems [38].

3.2.3. Simulation of Pile-up

Pile-up is included into the simulation: in-time and out-of-time pile-up are simulated by overlaying the hard processes and the underlying event with additional soft di-jet events (minimum-bias). Radiation background from the ATLAS underground cavern is taken into account. Neutrons and photons from this source might degrade the performance of the muon spectrometer. Two additional effects are considered: The beam pipe is not completely evacuated, thus interactions of protons with residual gas particles can occur (beam gas events). Furthermore, although the beam is well collimated, some protons can end-up being in a trajectory further away from the nominal beam spot. These can hit e.g. collimators and thus initiate signals in the detector (beam halo events).

3.3. The ATLAS Experiment

ATLAS[2] is a multi-purpose detector built to cover a physics program ranging from Higgs boson searches and supersymmetry to heavy-ion physics. It is 22 m high and 42 m long, weighing roughly 7000 metric tons. It has hermetic calorimetry, allowing particle detection as close as one degree away from the beam axis. The main detector volume has the form of a *barrel*, the ends are closed by the *endcaps*. A computer generated cross section of the detector is shown in Fig. 3.6. For further information on the detector see e.g. [39].

[2] A Toroidal LHC ApparatuS

The design of each detector component was driven by the LHC parameters and the physics program: High particle multiplicities require radiation hard components close to the beam pipe. The high frequency of collisions means that the *trigger* system (see Sec. 3.5) has to reach a decision whether or not to record an event very fast. In addition, the integration times of components need to be short. To satisfy the demanding physics program the individual sub-detectors need to have a high resolution in addition to the necessary robustness.

3.3.1. The Coordinate System

The coordinate system used by the ATLAS collaboration is a Cartesian system defined as follows: The *x*-axis points upwards, the *y*-axis towards the centre of the LHC-ring and the *z*-axis points into the direction of the beam pipe, such that a the resulting coordinate system is right-handed. ϕ is the azimuthal angle defined as the angle to the *x*-axis in the $x-y$-plane. The pseudorapidity η is defined via the polar angle θ:

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right) \tag{3.8}$$

A particle going perpendicular to the beam pipe has $\eta = 0$, while particles going into the direction of the beam pipe have $\eta = \pm\infty$. For massless particles, the pseudorapidity is equal to the rapidity. Differences in rapidity are invariant under Lorentz boosts. QCD processes are supposed to produce particle densities flat in η. Three-dimensional distances are measured in

$$\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}. \tag{3.9}$$

Momentum balance is only given in the transverse plane because the momenta of the colliding partons are not equal to the proton momenta and statistically distributed. Thus, transverse momentum, p_{T}, is usually used instead of total momentum.

3.3.2. Inner Detector

The purpose of the *inner detector* is to reconstruct trajectories of charged particles, called *tracks*. Tracks are used to measure particle momenta from their curvature - the inner detector is situated inside a solenoidal magnet with a field of 2 T. Tracks can be used to reconstruct the primary vertex of a process. Secondary vertices appear if long(er) lived particles decay. Thus, *B* and *D* hadrons, as well as τ leptons can be identified. The inner detector consists of three sub-detectors, Fig. 3.7 shows a schematic of the components.

The Pixel Detector

The pixel detector is the innermost sub-detector, its first layer being at a radius of only 5 cm. Pixel cells are silicon sensors, which act as diodes, depleted by the applied voltage. If a

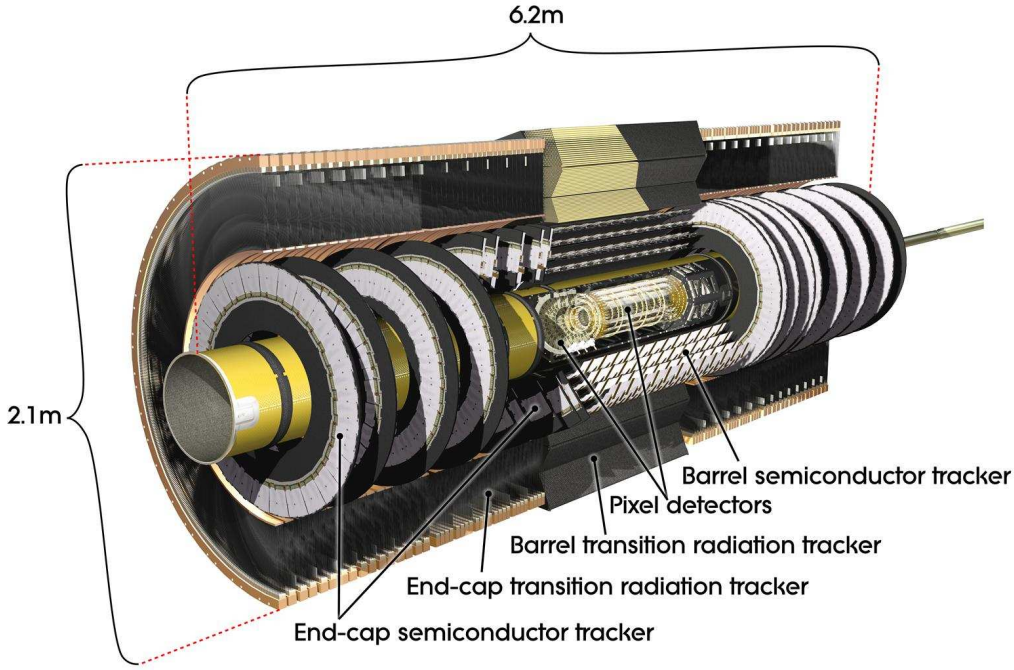


Figure 3.7.: The inner detector with its three sub-detectors [40].

charged particle traverses the sensor it creates electron-hole-pairs in the depleted material. The electrons drift towards the cathode, where they are collected to form a signal. Each cell has a size of $50 \times 400\,\mu\text{m}^2$. In the barrel region, pixel cells are arranged in three layers with a radius of 5.05 cm, 8.85 cm and 12.25 cm. The endcap region is made up of three disks on each side. In total the pixel detector has about 80 million readout channels.

The Semi Conductor Tracker

A second silicon detector encloses the pixel detector. The *Semi Conductor Tracker (SCT)* has four double-layers in the barrel and nine disks on each side of the barrel layers. It is segmented into strips with a width of $80\,\mu$m and a length of 12.8 cm. To provide better resolution in z-direction, two layers are arranged back-to-back with a stereo angle of 40 mrad between the strips. The four layers in the barrel are situated at radii between 30 cm and 52 cm and cover $|\eta| < 1.4$. The disks extend the coverage to $|\eta| < 2.5$.

The Transition Radiation Tracker

The *transition radiation tracker (TRT)* is the third and outermost sub-detector of the inner detector. It is less precise than the two silicon tracking detectors but provides on average additional 36 measurements per track. It consists of straw tubes filled with a mixture of xenon, carbon dioxide and oxygen. A high positive voltage is applied to the wire at the centre of a

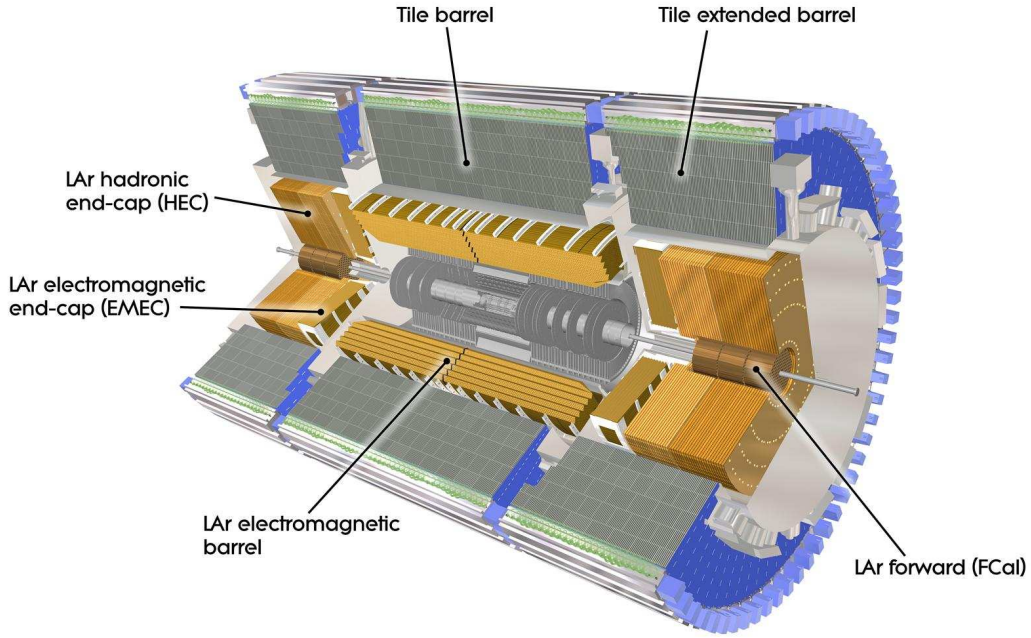


Figure 3.8.: Calorimeter components of the ATLAS detector [41].

tube. If a charged particle traverses a tube, it ionises the gas. The electrons drift to the wire producing an avalanche of secondary electrons close to the wire which amplify the signal. Since the drift velocity is known and constant, a drift circle can be reconstructed. Each tube has a diameter of 4 mm with a maximum length of 144 cm. In the barrel the tubes are arranged along the z-direction at radii between 56 cm and 107 cm and $|\eta| < 0.7$. It is complemented by disks with tubes in radial direction. In total the TRT has about 350,000 readout channels.

In addition to space-point measurements, the TRT provides particle identification. A radiator material surrounds the straw tubes forming boundaries with a different index of refraction on each side. Ultra-relativistic particles, i.e. $\gamma \gtrsim 1000$, emit transition radiation photons at these boundaries at small angles with respect to the flight direction of the particle. These photons are absorbed by the gas inside the tube via the photoelectric effect. This results in an additional signal in the straw tube, providing a so-called high-threshold hit. Due to their low mass, electrons are much more likely to produce such hits than pions. This can be used in the identification of electrons.

3.3.3. Calorimeters

As can be seen in Fig. 3.8, ATLAS calorimeters can be subdivided into three different types: An electromagnetic calorimeter, a hadronic calorimeter and forward calorimeters.

Electromagnetic Calorimeter

The *electromagnetic calorimeter (ECAL)* is a sampling calorimeter. *Liquid argon (LAr)* serves as active material, lead is used as absorber. When an electron traverses the absorber, it emits photons via bremsstrahlung. Photons convert into electron-positron pairs. In that way electromagnetic shower cascades are generated. Electrons and positrons produced in such a shower deposit energy by ionisation in the active material. A high-voltage field draws off the deposited charges to the electrodes. The collected energy is proportional to the the energy of the original particle.

In order to ensure full coverage in ϕ-direction without cracks, the absorber plates and the electrodes are arranged in an accordion-shape. The calorimeter is divided into a barrel calorimeter *(EMB)* and two endcap calorimeters *(EMEC)*. The segmentation varies for different η regions. The ECAL consists of two to three samplings in radial direction. It is designed to fully contain electromagnetic showers. Therefore its thickness corresponds to 24 radiation length[3] in the barrel region and 26 in the endcaps. Material in front of the ECAL already corresponds to 2.3 radiation length. To account for this, a presampler precedes the ECAL.

Hadronic Calorimeters

The hadronic calorimeter (HCAL) is made up of two types of calorimeters. The *scintillator tile calorimeter (TileCal)* covers the barrel region. It is a sampling calorimeter with iron as absorber and plastic scintillator plates as active material. Traversing particles initiate electromagnetic or hadronic showers. The latter occur through inelastic hadronic interactions with the material. Secondary particles excite the scintillator material, which in turn emits light that is transported to photo-multipliers to generate a signal. The endcap region has to withstand higher radiation doses. Therefore the *hadronic endcap calorimeter (HEC)* employs LAr as active material, which is intrinsically radiation harder. Copper is used as absorber. At $\eta = 0$ the HCAL corresponds to 9.7 hadronic interaction length – sufficient to provide a good energy resolution for hadronic jets and shielding for the muon spectrometer.

Forward Calorimeter

Both ECAL and HCAL provide coverage up to $|\eta| < 3.2$. This already corresponds to an angle of only 4.7 degrees with respect to the beam pipe. However, many processes, including the signal process studied in this thesis, deposit significant energy at even smaller angels. The *forward calorimeter (FCal)* covers an additional region in $|\eta|$ between 3.1 and 4.9 (approx. 0.9 degrees from the beam axis). It has to withstand considerable radiation doses. Like in the HEC, LAr is used as active material. The FCal is subdivided into three sections on each side. Each section is made of metal with regularly spaced longitudinal channels. The channels

[3] One radiation length is the distance in a material over which the energy of a high-energetic electron is reduced to 1/e.

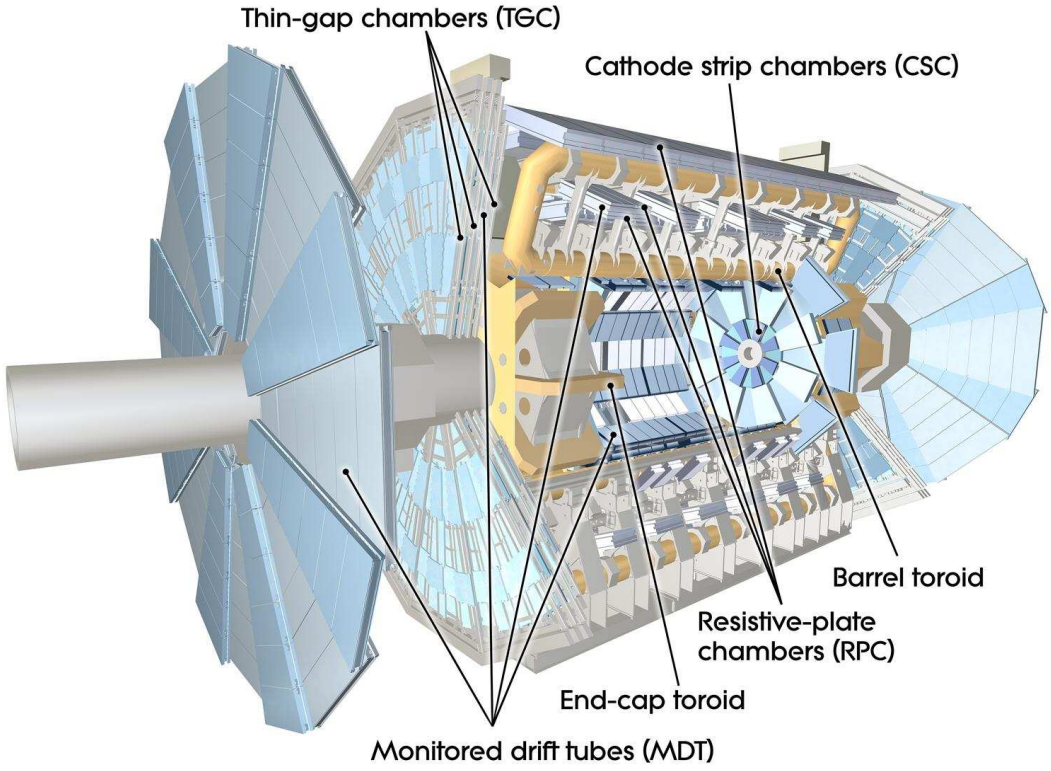


Figure 3.9.: The ATLAS muon spectrometer [42].

contain concentric rods and tubes with LAr in the gaps. The rods are at a positive high voltage with respect to the tubes and the enclosing metal.

3.3.4. The Muon Spectrometer

The ATLAS muon spectrometer (Fig. 3.9) forms the outermost layer of the detector and constitutes the largest part of its volume. It has several functions. First of all it identifies muons, since most other particles are stopped in the calorimeter. The muon is a minimum ionising particle. Therefore its calorimeter deposition cannot be used to measure momenta. High-energy tracks have only a small curvature, so the inner detector alone would provide bad momentum resolution. The muon spectrometer functions as an additional lever arm – an additional measurement far from the interaction point – and thus improves momentum resolution. Finally it is used for triggering.

The muon spectrometer consists of chambers arranged in three concentric rings around the beam axis in the barrel (stations) and four wheels perpendicular to the beam axis in each endcap region. The muon spectrometer has a gap at $\eta = 0$ which is needed for services. Four different detector technologies are employed in the chambers. The first two provide precision measurements, the second two are used for triggering.

Monitored Drift Tube chambers

*MDT*s contain tubes, roughly 3 cm in diameter, filled with pressurised Ar/CO_2 gas. Electrons resulting from a charged particle that traversed the tube and ionised the gas are collected at a central wire at a potential of 3 kV. Each individual tube provides a drift circle (cf. Sec. 3.3.2). A disadvantage of this technology is the maximum drift time from wall to wire of about 700 ns. A particle passing close to the wire can generate a series of signal pulses of up to these 700 ns, while only the electrons generated closest to the wire a relevant to form a drift circle. An adjustable dead-time has been implemented to prevent this phenomenon.

Cathode-Strip Chambers

MDTs have a limit for safe operation at counting rates of about $150\,Hz/cm^2$. This will be exceeded in the first layer of the endcap at $|\eta| > 2$. Therefore MDTs are replaced by *CSC*s in this region, which can safely handle counting rates of up to $1000\,Hz/cm^2$. CSCs are multiwire proportional chambers. The central wire of a chamber is oriented in the radial direction, with the other wires parallel to the central one. Each layer of wires has a layer of cathode strips in front and behind it, one of which has strips parallel to the wires, the other perpendicular. Information is gathered only from the strips, wire signals are not read out.

Resistive Plate Chambers

*RPC*s provide input to the trigger system in the barrel. An RPC consists of two parallel resistive plates kept at a distance of 2 mm by insulating spacers. The gap is filled with a gas mixture. A particle crossing an RPC ionises the gas, creating primary electrons. The electrons are accelerated towards the anode plate by the strong electric field of 4.9 kV/mm in the gap, causing avalanches on their way. At nominal operating voltage a signal with a width of about 5 ns is generated.

Thin Gap Chambers

In the endcap region, *thin gap chambers (TGC)* are used for triggering. TGCs work similar to multi-wire proportional chambers. Information is read out from wires for the radial coordinate and from radial strips for the azimuthal coordinate. The chambers are operated in saturation mode, which allows for the quick responses necessary for trigger decisions.

3.3.5. GEANT4 Simulation

For simulated events, the actual detector is replaced by a detailed detector simulation based on GEANT4. GEANT4 works similar to a ray tracing program for 3D graphics generation. It

relies on a detailed geometrical description of the ATLAS detector. For particles traversing the detector volume, hits, i.e. deposited energy in the detector, and secondary particles produced in interactions are simulated. Secondary particles are added to the list of particles and simulated the same way. The simulation goes on until all particles are either stopped within the detector or have left the volume. Instead of simulating fundamental interactions, GEANT4 uses transport models – macroscopic parametrisations of energy loss mechanisms or analytical calculations – to simulate the effects of particles traversing matter.

The hits are then digitised. The response of the smallest detector components (cells, pixels, etc.) to deposited energy is simulated in this step. Afterwards, the output format is identical to that of the real ATLAS detector and can be processed by the same reconstruction algorithms.

Full simulation of proton-proton collisions is extremely complex: Even on modern CPU cores it takes about 10-20 minutes to simulate a single event.

3.4. Reconstruction Algorithms

The raw detector data, such as energy depositions in calorimeter cells or hits in tracking devices, has to be transformed into higher-level physics objects in order to perform physics analyses. For every object the ATLAS software includes a set of algorithms which represent different approaches to reconstruct such an object. I will concentrate, therefore, on algorithms relevant for this thesis. For further details see Ref. [1, 39]. In addition, the ATLAS software is still constantly under development. Information provided here represents a snapshot of the software that was used to reconstruct the simulated data samples used in this thesis. Usually, algorithms distinguish between reconstruction and identification. Reconstruction means creating an object and collecting all components. As an example, an electron consists energy depositions in the calorimeter and a measured trajectory of a particle. These have to be found, matched and the information merged into one object. The purpose of identification is to make sure that the reconstructed object is actually what it was reconstructed as. For instance, pions and electrons both provide a track in the inner detector and matching energy in the calorimeter.

3.4.1. Cluster finding

A calorimeter cluster is a group of calorimeter cells which is spatially coherent. Its combined energy deposition would ideally reflect the energy that a single particle lost in the calorimeter. The ATLAS software provides several means to find calorimeter clusters:

Sliding-Window Clustering

The sliding-window algorithm forms clusters from rectangles (windows) in $\eta - \varphi$ space of out calorimeter towers, i.e. cells in several layers of the calorimeter forming a unit. It searches for seed clusters by summing up the deposited transverse energy in a window of pre-defined size.

The window consecutively slides over the whole $\eta - \varphi$ grid in question. If the total transverse energy in the window exceeds a threshold and if it is a local maximum, i.e. no other window surrounding it has a higher total energy, a cluster is formed.

Topological Clustering

Sliding-window clusters have a fixed size. Topological clustering aims to reconstruct 3-dimensional shower shapes and thus creates clusters (*TopoClusters*) of variable size and form. Three thresholds define which cells are used for cluster finding: a seed threshold t_s, a neighbour threshold t_n and a general threshold t_g. Thresholds are given in energy over the standard deviation of noise, E/σ_{noise}, assuming that noise is Gaussian distributed. In this study, thresholds are set to $t_s = 4$, $t_n = 2$, $t_g = 0$.

In a first step, calorimeter cells are identified as seeds if they pass t_s. For every seed, a cluster is created consisting of the single seed cell. In the next step, the cluster is grown: for each seed cell, neighbouring cells are added to the cluster if they pass t_g and are not already part of a cluster. Neighbours are defined as all eight cells in a rectangular grid of the same calorimeter sampling. In addition, all cells in adjacent samplings, which have at least a partial overlap with the seed cell in (η, ϕ)-space are considered. If a neighbouring cell passes t_n it is added to a list. If a neighbouring cell already belongs to a cluster but passes t_n, both clusters are merged. In the next iteration, cells in the list of cells that passed t_n are processed the same way. This is repeated until no more new cells above t_n are found.

Ideally, each TopoCluster corresponds to one primary particle that entered the calorimeter. To achieve this level of resolution an additional splitting algorithm is necessary. This algorithm looks for local maxima in parent clusters. A local maximum is defined as a cell with $E > 500\,\text{MeV}$ that has no neighbouring cells with higher energy. It has to be surrounded by at least four cells in the parent cluster and has to be situated in the second or third sampling of the ECAL or the first sector of the FCal. Other parts are considered for secondary local maxima. Secondary maxima are used only if they do not overlap in (η, ϕ) with a primary maximum. Parent clusters are then split according to the neighbouring relations, making sure no two local maxima end up in the same cluster. Cells at the border of two sub-clusters are shared. Their energy is added to both clusters with a weight

$$w_1 = \frac{E_1}{E_1 + rE_2} \tag{3.10}$$

where $w_2 = 1 - w_1$ and $r = \exp(d_1 - d_2)$. The d_i are the distances to the cluster centres in units of a typical shower scale in the ECAL (5 cm in this study).

3.4.2. Electron reconstruction

An electron candidate, in its most general form, is a calorimeter cluster with a track in the inner detector, which points towards it. Electron reconstruction is seeded by a sliding-window

		no. of cells	
	η-region	5×5	3×7
barrel			
	< 1.40	0.125×0.125	0.075×0.175
	$1.40 < \lvert\eta\rvert < 1.475$	0.375×0.375	0.225×0.125
endcap			
	$1.375 < \lvert\eta\rvert < 1.425$	0.25×0.125	0.15×0.125
	$1.425 < \lvert\eta\rvert < 2.5$	0.125×0.125	0.075×0.175
	$2.5 < \lvert\eta\rvert < 3.2$	0.5×0.5	0.3×0.7

Table 3.1.: Window sizes in $\eta\times\phi$ for different detector regions used by the sliding window algorithm in the electron reconstruction.

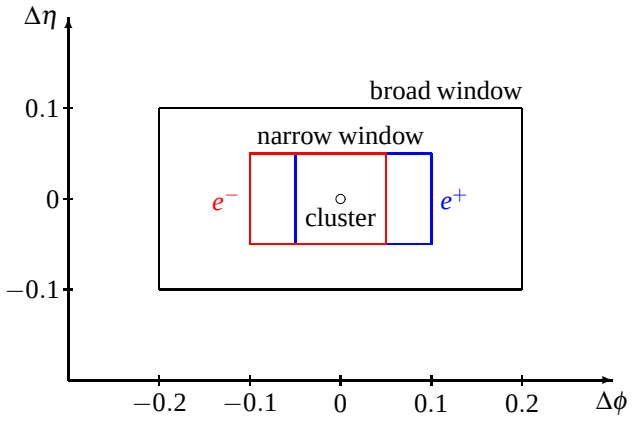


Figure 3.10.: Cluster/Track matching in the sliding window algorithm used by the electron reconstruction.

cluster with a window size of 5×5 cells in the middle layer of the electromagnetic barrel or endcap calorimeters and a transverse energy of at least 3 GeV. A track is searched for in a broad rectangular window of size 0.2×0.4 in $(\Delta\eta,\Delta\phi)$ around the barycentre of the cluster. It is considered a match if it has at least three silicon (Pixel + SCT) hits, has $E/p < 10$ and lies within an asymmetric narrow window of $(-0.05,-0.1),(0.05,0.05)$ for tracks with negative charge and $(-0.05,-0.0.05),(0.05,0.1)$ for positive charge around the centre of the cluster (see Fig. 3.10). If the seed cluster is in the barrel calorimeter it is replaced by a cluster of 3×7 cells around the original centre. In the overlap region between barrel and endcap, the amount of energy in the second layer decides in which category the cluster falls. The actual size of an electron cluster thus depends on the actual detector region it is in (see Table 3.1).

Reconstructed electron candidates can either be rejected in the electron identification steps or fall into one of three categories relevant for this analysis: Every electron candidate qualifies as a *loose* electron if it passes some cuts on shower shape variables in the middle layer of the ECAL and has a ratio of energy in the ECAL over energy in the HCAL (hadronic leakage)

below a certain threshold. In addition to all loose criteria, *medium* identification requires an isolated cluster in the calorimeter and uses additional shower shape information from the η strip layer. Medium identification also requires a tighter match between cluster and track and applies some track quality cuts but does not use information from transition radiation. *Tight* identification requires the candidate to pass medium cuts plus tighter track quality and track-cluster matching cuts. It also uses information on transition radiation from the TRT.

3.4.3. Muon reconstruction

ATLAS employs a variety of strategies to reconstruct and identify muons. There are two families of algorithms, with one algorithm per family for each strategy. Only the STACO family (Muonboy, Staco, MuTag, CaloMuon) is used in this analysis.

Muonboy: Standalone Muons

Standalone muons are muons reconstructed exclusively in the muon spectrometer. Muonboy [43] starts by identifying regions of activity, using information from the trigger chambers. A region of roughly $\Delta\eta \times \Delta\phi = 0.4 \times 0.4$ is centred around least one hit in each coordinate. Next, track segments are reconstructed. Muonboy tries to combine each MDT hit in one multilayer with each MDT hit of the other multilayer of the same station or an adjacent one. Each pair of hits is required to point loosely into the direction of the interaction point. A segment is then a straight line between two hits. Such pairs of hits are close enough in space that a straight line is an appropriate approximation. As MDT hits are actually drift circles and a segment should be a tangent to both circles there exist four solutions for each pair (see Fig. 3.11). To solve this ambiguity all four segment candidates are matched with other hits in the corresponding MDT chambers. A segment is declared valid if its quality factor, which is a combination of standard χ^2 for found hits and a penalty for missing ones, is sufficiently small. In a final step, track segments are combined to form tracks. Track segments are extrapolated to the other stations. A track has to consist of at least two segments.

Staco: Combined Muons

Combined muons consist of two tracks, one from the muon spectrometer and one from the inner detector. The Staco algorithm calculates the weighted sum of both parameter vectors to obtain the combined track:

$$P_{\text{combined}} = \left(C_{\text{ID}}^{-1} + C_{\text{MS}}^{-1}\right)\left(C_{\text{ID}}^{-1} P_{\text{ID}} + C_{\text{MS}}^{-1} P_{\text{MS}}\right) \tag{3.11}$$

Here P denotes a vector of parameters that describe a track, C is the respective covariance matrix. The subscripts denote inner detector and muon spectrometer respectively. The corresponding χ^2 is used to describe the goodness of the combination. Using combined muons

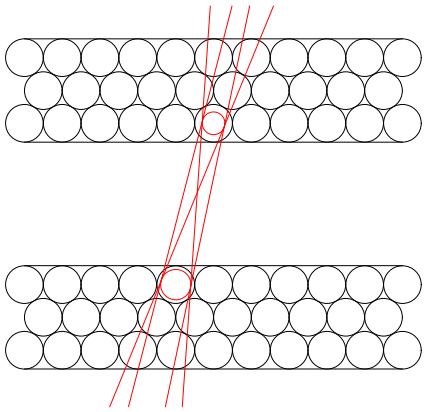


Figure 3.11.: Ambiguity in track segment reconstruction. Black circles symbolise tubes in an MDT chamber. Red circles represent measured drift circles. The red lines mark all possible track segments that can be constructed from the two drift circles.

suppresses muons not coming from the interaction point like muons from pion or kaon decays in the calorimeter or cosmic muons. It also improves momentum and impact parameter resolution.

MuTag: Tagged Muons

The MuTag algorithm specifically addresses the reconstruction of muons with low transverse momentum. These particles often fail to reach the middle station of the spectrometer. Since a track in the muon spectrometer requires at least two track segment - and therefore hits in two muon station - such a low-p_{T} muon will not be reconstructed by either Muonboy or Staco. MuTag extrapolates tracks from the inner detector to the first station of the muon spectrometer. In some regions, where a particle would only traverse one station, it extrapolates to the middle station in order to increase the efficiency there. MuTag defines a χ^2 using the difference between the predicted track and nearby track segments not already used by Staco. Only the inner detector track is used to evaluate track parameters.

CaloMuon: Calorimeter Muons

Although muons are minimum ionising particles, they deposit energy in the calorimeter. Calorimeter muons combine a track in the inner detector with an extrapolated trajectory in the calorimeter. Using this kind of reconstruction is used to cover e.g. the gap region in the muon spectrometer at $\eta = 0$, where information from the spectrometer is not available.

3.4.4. Jet reconstruction

In this study a cone algorithm was used to reconstruct jets. The parameters are adjustable and reflect the choice for this channel in [1]. The ATLAS cone algorithm works in the following way: TopoClusters above a threshold of $p_{\mathrm{T}} > 1\,\mathrm{GeV}$ are sorted according to their transverse momentum in descending order. This list is used as seeds for the algorithm. A seed defines the initial jet axis in η and ϕ. All four-momenta of clusters with the centre inside a cone of $\Delta R < 0.4$ around the jet axis are summed up (E-scheme). The resulting four-vector is the new jet axis. This is done iteratively until either the jet axis is stable, i.e. the difference in η and ϕ between new and previous jet axis are each smaller than 0.05 or the new axis leaves the acceptance region of the detector ($|\eta| > 5.0$). In the latter case the jet is rejected. In the former the jet is accepted if the distance in η and ϕ to already reconstructed jets is, again, smaller than 0.05. In case reconstructed jets overlap, i.e. share clusters, they are either split or merged, depending on their shared energy: If the overlapping jet shares more than 50% of E_{T} the jets get merged. Note that this fraction is always with respect to the jet with more E_{T}. If the fraction is smaller than 50% the clusters in question are removed from the jet which is farther away in ΔR. The jet energy is calibrated according to global weights taken from detailed simulations ("H1-style calibration") [44]. It should be noted, although the cone algorithm itself is *collinear safe*, the seeding makes it unsafe: If the highest energetic constituent gets split, e.g. by calorimeter effects, and thus falls below the threshold the jet might not be reconstructed or split in two jets. Since, however, the p_{T} threshold for jets in this analysis is at 20 GeV (see Sec. 7.1.4) the seed threshold of 1 GeV should have only a minimal effect.

Since jets are built entirely from calorimeter objects, the jet energy resolution depends on the energy resolution of the calorimeter. The jet energy resolution, $\sigma(E)$, is therefore given by:

$$\sigma(E) = a\sqrt{E} \oplus b \oplus cE. \tag{3.12}$$

The first term reflects fluctuations in the number of particles produced in a shower and is proportional to $\sqrt{n}$ and thus to $\sqrt{E}$, where n is the number of particles. The second term is due to effects like noise and pile-up, which can be considered constant on average. The last term results e.g. from cracks in the calorimeter and dead cells, and is therefore proportional to E.

3.4.5. Reconstruction of Hadronic τ Decays

τ leptons have a mean lifetime of $2.9 \times 10^{-13}\,\mathrm{s}$ [2]. This means they decay within a short distance to the interaction point. Therefore τ leptons can only be identified via their decay products. If the τ decays into an electron or muon, it cannot easily be distinguished from other (direct) electrons or muons – although, if the corresponding track does not point directly to the primary vertex, this is an indication of the lepton originating from a τ decay. Hadronic decays account for about two thirds of all τ decays. See Sec. 4.5 for a detailed discussion

of hadronic τ decay modes. Although hadronic τ decays are real hadronic jets, they can be distinguished from quark- or gluon-induced jets. Apart from possible kinematic differences (due to the τ-mass), jets from quarks or gluons have on average a higher track multiplicity and cover a larger area in the detector. Two approaches are used to reconstruct hadronic τ decays ("τ_{had} candidates").

The *track-seeded* algorithm starts from good-quality tracks. The leading track is required to have $p_{\text{T}} > 6\,\text{GeV}$. Around this track, additional tracks with $p_{\text{T}} > 1\,\text{GeV}$ are searched for in a core region of $\Delta R < 0.2$. If only one additional track is found, the closest lower-quality track (if available) in the core region is used as well. The direction of the τ_{had} candidate is calculated as the p_{T}-weighted average of coordinates of the associated tracks. Energy is reconstructed by an energy flow algorithm [45].

The *calorimeter-seeded* algorithm is less complex: It takes reconstructed jets with $p_{\text{T}} > 10\,\text{GeV}$ as τ_{had} candidates. All good-quality tracks within $\Delta R < 0.3$ from the barycentre of the jet are associated with the candidate. Coordinates and energy are those of the seed-jet. Energy is then calibrated by a dedicated τ_{had} calibration.

The τ_{had} identification is based on a number of shower-shape and track related observables. These are combined into several multivariate and cut based discriminants. For this study, a log-likelihood method is used, based on 17 input variables [46]. In this context, EM cells means cells from the presampler and the first and second sampling of the ECAL. Photons tend to be fully contained in the first two layers of the ECAL, while the third layer is large in some areas of the detector and thus already contains significant energy depositions from hadronic showers. HAD cells therefore refer to the HCAL plus the third sampling of the ECAL. If not stated otherwise, energy is collected within a radius of $\Delta R < 0.4$. The following list shows all observables that enter the log-likelihood calculus.

Observables used by the **calorimeter-seeded** algorithm:

- *emRadius*: E_{T}-weighted radius of depositions in EM cells:

$$\text{emRadius} = \frac{\sum \Delta R(\tau_{\text{had}}, \text{cell}) E_{T,\text{cell}}}{\sum E_{T,\text{cell}}}$$

- *isoFrac*: Ratio of E_{T} in the isolation region over E_{T}:

$$\text{isoFrac} = E_{\text{T}}(0.1 < \Delta R < 0.2)/E_{\text{T}}(\Delta R < 0.4)$$

- *stripWidth2*: Width of the energy deposition in the first sampling of the ECAL (strip cells):

$$\text{stripWidth2} = \frac{(\sum \Delta\eta(\tau_{\text{had}}, \text{cell}))^2 \cdot E_{\text{T}}(\text{cell})}{\sum E_{\text{T}}(\text{cell})} - \frac{(\sum \Delta\eta(\tau_{\text{had}}, \text{cell}) \cdot E_{\text{T}}(\text{cell}))^2}{(\sum E_{\text{T}}(\text{cell}))^2}$$

- *numStripCells*: Number of strip cells with $E_{\text{T}} > 200\,\text{MeV}$.

- *etEM2etTracks*: Calibrated E_{T} in EM cells over the scalar p_{T} sum of up to three tracks.
- *etHad2etTracks*: Same as above but using calibrated HAD cells.
- *signD0Trk3P*: Signed impact parameter of leading track.

Observables used by the **track-seeded** algorithm:

- *rWidth2Trk3P*:

$$\mathrm{rWidth2Trk3P} = \frac{\sum \Delta R^2(\tau_{\mathrm{had}}, \mathrm{trk}) p_{T,\mathrm{trk}}}{\sum p_{T,\mathrm{trk}}} - \frac{\left(\Delta R(\tau_{\mathrm{had}}, \mathrm{trk}) p_{\mathrm{T,trk}}\right)^2}{\left(\sum p_{\mathrm{T,trk}}\right)^2}$$

- *massTrk3P*: Invariant mass of associated tracks.
- *nAssocTracksIsol*: No. of associated tracks of τ_{had} candidates in $0.1 < \Delta R(\tau_{\mathrm{had}}, \mathrm{trk}) < 0.2$.
- *mVisEflow*: Visible mass from energy flow.
- *z0SinThetaSig*:

$$\mathrm{z0SinThetaSig} = \frac{z0 \sin\theta}{\sigma(z0 \sin\theta)}$$

- *trFlightPathSig*: Signed transverse distance between primary and secondary vertex over the uncertainty of the secondary vertex.

Observables used by **both** algorithms:

- *etTracks2et*: Scalar p_{T} sum of up to three tracks over τ_{had} E_{T}.
- *dRmin*: minimum $\Delta R(\tau_{\mathrm{had}}, \mathrm{track})$ of tracks within $\Delta R < 0.2$
- *dRmax*: maximum $\Delta R(\tau_{\mathrm{had}}, \mathrm{track})$ of tracks within $\Delta R < 0.2$
- *ratioET*: Ratio of scalar p_{T} sum of tracks excluding tracks associated to the τ_{had} candidate over total scalar p_{T} sum within $\Delta R < 0.2$.

3.4.6. Missing Transverse Energy

Neutrinos traverse the detector without interacting with the material. They can only be detected indirectly by a seeming non-conservation of total momentum. Ideally, the vectorial sum of all neutrino momenta is equal to the negative vectorial sum of all other momenta. Measuring momenta directly is only possible for charged particles and only in the limited range of the tracking sub-detectors of the inner detector. Because of this, energy depositions in the calorimeter are used instead. In proton collisions the actual momenta of the colliding partons cannot be determined. Therefore only the transverse momentum balance can be used.

The resulting quantity is called missing transverse energy ($\not{E}_{\mathrm{T}}$), defined as:

$$\not{E}_x = -\sum_i E_{x,i}, \quad \not{E}_y = -\sum_i E_{y,i}, \tag{3.13}$$

where $\not{E}_{x/y}$ is the x and y-components of $\vec{\not{E}}_{\mathrm{T}}$ and $E_{x/y,i}$ being x and y-projections of energy depositions in calorimeter cells.

The actual calculation of $\not{E}_{\mathrm{T}}$ is usually more involved than summing up energy in calorimeter cells, as additional information can be used to improve the $\not{E}_{\mathrm{T}}$ resolution. One approach is the refined calculation. For this, energy depositions in calorimeter cells associated with different reconstructed objects and muon momenta are summed up. The algorithm keeps track of cells already used in one of the sub-algorithms to avoid double counting. Cell energies are calibrated according to the reconstructed object they belong to. E.g. an electron candidate might overlap with a jet. Cells belonging to the electron cluster would enter the calculus calibrated as electron. The remaining hit cells of the jet enter the sum using jet calibration. Calorimeter cells are processed by the following sub-algorithms (in this order):

- *MET RefEle*: identified (tight) electrons with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$
- *MET RefGamma*: identified (tight) photons with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$
- *MET RefMuon*: cells around tracks of non-isolated (overlap in $\Delta R < 0.3$ with jet) combined or spectrometer-only muons
- *MET RefTau*: τ_{had} candidates identified with a cut-based method
- *MET RefJet*: H1-style calibrated cone-jets ($\Delta R < 0.4$) based on TopoClusters with $p_{\mathrm{T}} > 5\,\mathrm{GeV}$
- *MET CellOut*: cells in TopoClusters not associated with above objects
- *MET Muonboy*: isolated combined muons, MuTag-muons around $|\eta| = 1.3$, calorimeter muons in $|\eta| < 0.1$
- *MET Cryo*: cryostat energy from jets in $|\eta| < 3.2$

At a hadron collider the resolution of the $\not{E}_{\mathrm{T}}$ measurement depends mainly on the uncertainty of the jet energy. Therefore, in the absence of particles which escape the detector, the resolution can be parametrised as [47]:

$$\sigma(\not{E}_{\mathrm{T}}) = a\sqrt{\Sigma E_{\mathrm{T}} - d} \oplus b \oplus c(\Sigma E_{\mathrm{T}} - d) \tag{3.14}$$

Corresponding to the terms in the jet energy uncertainty (cf. 3.4.4), the first term is due to sampling and purely statistical fluctuations. b reflects electronic noise, pile-up and the underlying event. The last term stems from detector inhomogeneities, cracks and dead cells. ΣE_{T} is the scalar sum of energy depositions in the calorimeter. Noise, pile-up, etc. can lead to an offset in ΣE_{T} that is taken into account by d.

3.5. Trigger

At a hadron collider the vast majority of collisions consists of processes which are generally considered ”uninteresting”. Writing out and processing every single event would, thus, be a waste of disk space and processing time. Moreover, such a strategy would be impossible from a technical point of view. A bunch spacing of 25 ns means collisions at a rate of 40 MHz. Given that a single event amounts to roughly 100 MB of raw data, the ATLAS detector would produce several petabytes of data per second. Therefore, a trigger system is necessary, which reduces the data to a manageable rate while preserving as many interesting events as possible.

The ATLAS trigger system is based on the identification of objects which are supposed to be produced in processes of interest. Usually these are leptons or photons, but also τ_{had} candidates, highly energetic jets and large $\not{E}_{\text{T}}$ are searched for. This identification procedure is divided into three trigger stages: *Level 1 (L1)*, *level 2 (L2)* and *event filter (EF)*, the latter two forming the *high-level trigger (HLT)*. The L1 trigger uses reduced-granularity information from the calorimeter and the muon spectrometer (RPCs and TGCs) to quickly find signatures of the desired objects. A decision has to reach the front-end electronics within 2.5 μs. From these signatures *regions of interest (RoI)* are built, which seed the L2 trigger. There are two types of HLT algorithms: Reconstruction algorithms reconstruct objects with a level of sophistication depending on the stage. Hypothesis algorithms basically perform particle identification, testing if the reconstructed object corresponds to the one implied by the trigger chain. The average processing time at L2 is 40 ms. A final decision is taken by the event filter, which uses algorithms close to their off-line counterparts. The average processing time is on the order of several seconds. Trigger rates should combine to no more than 200 Hz after EF.

3.5.1. Electron Trigger

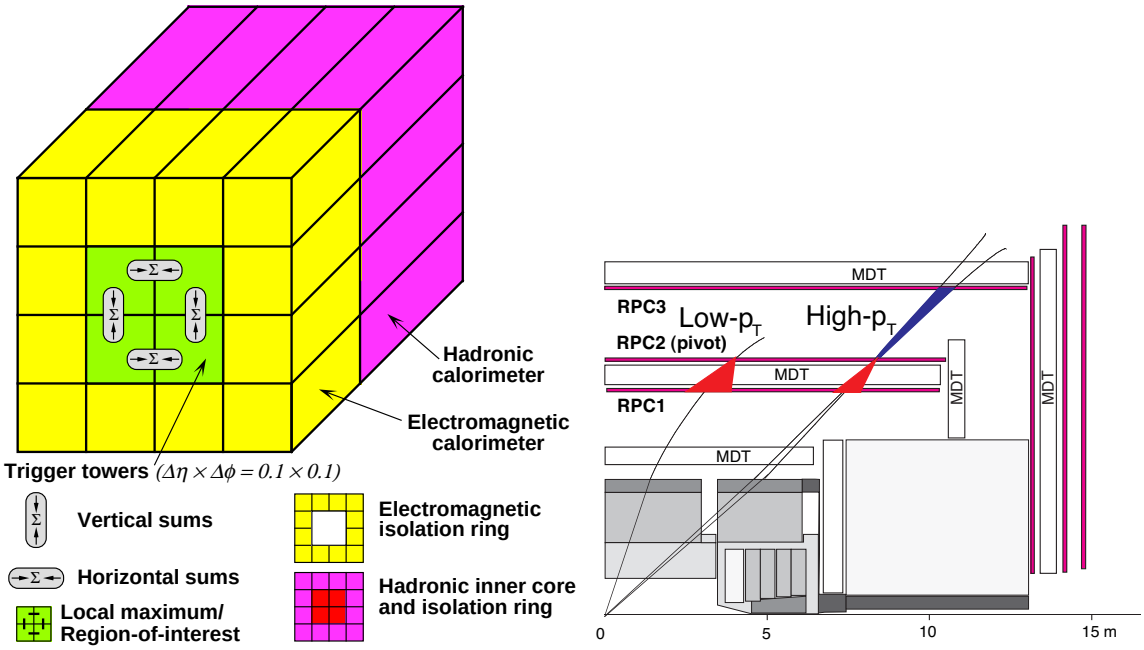
The standard electron trigger chain used in this analysis is

`L1 EM23I` $\rightarrow$ `L2 e25i medium1` $\rightarrow$ `EF e25i medium1`.

The first part of each trigger label gives the level (L1,L2 or EF). The second part gives the trigger object (e.g. e for electron), the threshold in p_{T} and an ”i” if an isolation requirement is imposed. The third part gives additional information on identification criteria, e.g. medium electron identification.

Level 1 and Level 2

On level 1, the electron algorithm works as sketched in Fig. 3.12a. The calorimeter is divided into trigger towers, i.e. a block of cells of all calorimeter layers with a size of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. The algorithm uses a window of 4×4 towers. Energy is summed up for all four possible combinations of 1×2 and 2×1 towers in the central 2×2 trigger towers. One of these four combinations has to pass the E_{T} threshold of the trigger (*here:* 23 GeV). Three



(a) Electron trigger algorithm. **(b)** Muon trigger algorithm.

Figure 3.12.: Schematic view of level 1 electron/photon/τ_{had} and muon trigger algorithms [39].

additional criteria may be applied. Cuts on the maximum transverse energy in the isolation region (12 towers around the central 4 towers) can be applied separately for the ECAL and the HCAL. These cuts are 4 GeV and 3 GeV, respectively, for the `L1 EM23I`. Moreover, a maximum E_{T} in the HCAL can be specified for the central four towers. This is set to 2 GeV for `L1 EM23I`. To avoid ambiguities, the sum of the energy in the inner region must be a local maximum.

In the first step of the L2 electron trigger, the L1 calorimeter cluster is refined. In a window ($\Delta\eta \times \Delta\phi = 0.2 \times 0.2$) around the centre of the L1 RoI the calorimeter cell in layer 2 of the ECAL with the highest energy is identified. This position is further refined by calculating the E_{T} weighted η and ϕ in a window of 3×7 cells in the same layer. Identification is based on a number of shower shape variables. In the second step, tracks are built from Pixel and SCT space points. In the final step, the cluster is matched to a track.

3.5.2. Muon Trigger

The muon trigger chain used in this analysis is:

`L1 MU20` $\rightarrow$ `L2 mu20` $\rightarrow$ `EF mu20`

Level 1 and Level 2

Muon triggering on level 1 makes use of RPC chambers in the barrel and TGC chambers in the endcaps. Although details of the implementation differ, the principle is the same in both regions. Upon a hit in the pivot plane, a straight line is calculated from that hit to the interaction point (see Fig. 3.12b). This line represents the trajectory of a muon with infinite momentum. Deviations from this line can be interpreted as being due to the (finite) p_{T} of the muon. Thus, searching for additional hits within a ”road” defined by a trajectory of a given p_{T} effectively represents a cut on the p_{T} of a track. In the barrel region the pivot plane is RPC2. Depending on the trigger mode (high-p_{T} or low-p_{T}) additional hits are searched for in RPC3 (high) or RPC1 (low). In the endcap region, the pivot plane is the outermost plane in z-direction. From there, low- and high-p_{T} tracks are extrapolated back to the interaction point.

In a first step oon L2, a tracking algorithm is run in the muon spectrometer alone. RPC hits that formed the L1 track candidate are used to define a road through the MDT chambers around the muon trajectory. A track fit is done using MDT hits within this road. p_{T} is estimated by using look-up tables. In the second step, algorithms described in Sec. 3.5.1 identify tracks and vertices in the inner detector within the L1 RoI. Finally, tracks in the inner detector are combined with muon spectrometer tracks. To avoid the time consuming extrapolation of inner detector tracks to the muon stations, parametrised analytical functions are used instead.

4

The Signal Process

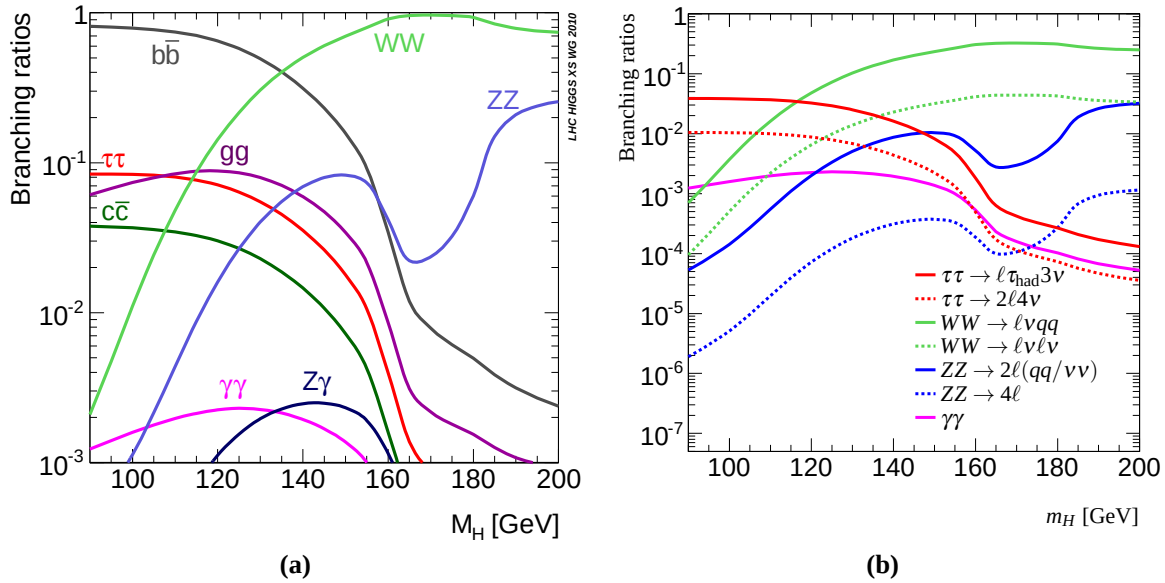
General properties of the Higgs boson and means of reconstructing and distinguishing between final state particles were discussed in Chapters 2 and 3. Based on these two chapters, I will describe in this section the special properties of the signal process and how it is simulated.

4.1. Higgs Boson Decays

Couplings of the Higgs boson to fermions and gauge bosons are proportional to their respective masses (Eqs. 2.22, 2.23 and 2.27). Hence, the Higgs boson predominantly decays into particles with the highest accessible mass. As the highest accessible mass obviously depends on m_H, the various final states are sensitive to different mass regions. Branching ratios of the individual decay modes are shown in Fig. 4.1a. Below about $2m_W$ the heaviest accessible particles are τ leptons and b quarks. Photons and gluons, although massless, contribute significantly since they can couple indirectly to the Higgs boson via loops which contain heavy particles – mainly top quarks in the case of $H \to gg$ and W bosons for $H \to W^+W^-$. Above the corresponding thresholds W and Z bosons have by far the highest masses, thus they dominate the decay modes at large m_H. Although the W boson is lighter than the Z boson, decays into W bosons have a higher branching ratio owing to the fact that W^+ and W^- are distinguishable particles.

These primary branching ratios, however, do not directly translate into experimental sensitivity. Most of the decay products are unstable themselves. Hadronic final states pass lepton triggers very inefficiently at best, while jet triggers have either high thresholds or only a fraction of the triggered events is read out in order to keep the rates low. This usually rules out direct searches for hadronic Higgs decay modes like $H \to b\bar{b}$[1] and gg, as well as entirely

[1] More recent studies in HW/HZ looking for jet substructure in highly boosted events as proposed in [49] have shown promising results.



(a) **(b)**

Figure 4.1.: Higgs boson branching ratios as a function of m_H. a) shows branching ratios into primary decay products [48]. b) is a selection of branching ratios into detectable final states.

hadronic decay modes of $H \to WW$ and ZZ. Figure 4.1b shows branching ratios into final states with at least one electron, muon or photon. At low m_H, apart from WW, especially $H \to \tau^+\tau^-$ and $H \to \gamma\gamma$ are important channels.

4.2. Higgs Production Mechanisms

The Higgs boson coupling to SM particles also determines which production processes are dominant. Direct production by annihilation of two quarks is strongly suppressed. Instead, as the Higgs boson couples preferentially to W and Z bosons and top quarks, the main production mechanisms are *gluon-gluon fusion*, *vector boson fusion*, *associated production with W/Z bosons* and *associated production with heavy quarks*. The corresponding cross sections and Feynman diagrams are shown in Figures 4.2 and 4.3. All cross sections in Fig. 4.2, except for $pp \to ttH$, have been calculated up to NNLO in QCD with electroweak corrections up to NLO. $pp \to ttH$ is an NLO calculation in QCD. The uncertainties include variations of the factorisation and renormalisation scales, uncertainties in α_s and the PDF uncertainties.

4.3. Vector Boson Fusion (VBF)

Although the cross section for Higgs boson production via vector boson fusion is about an order of magnitude smaller than for gluon-gluon fusion, the process has a characteristic signature that can be exploited to distinguish it from many large QCD background processes. The

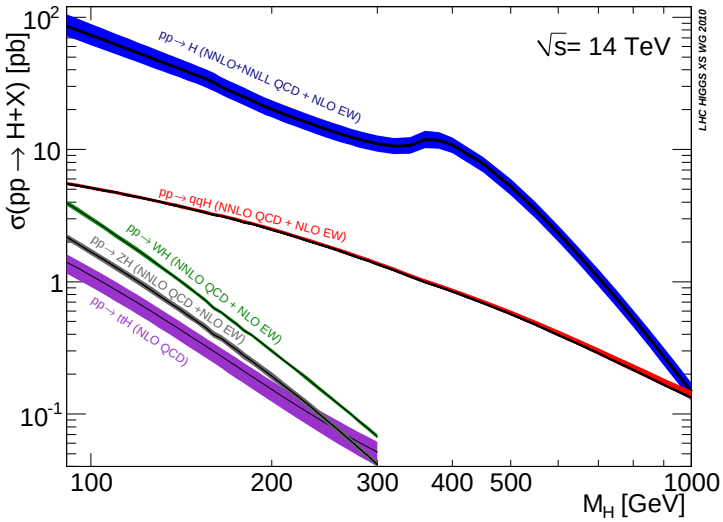
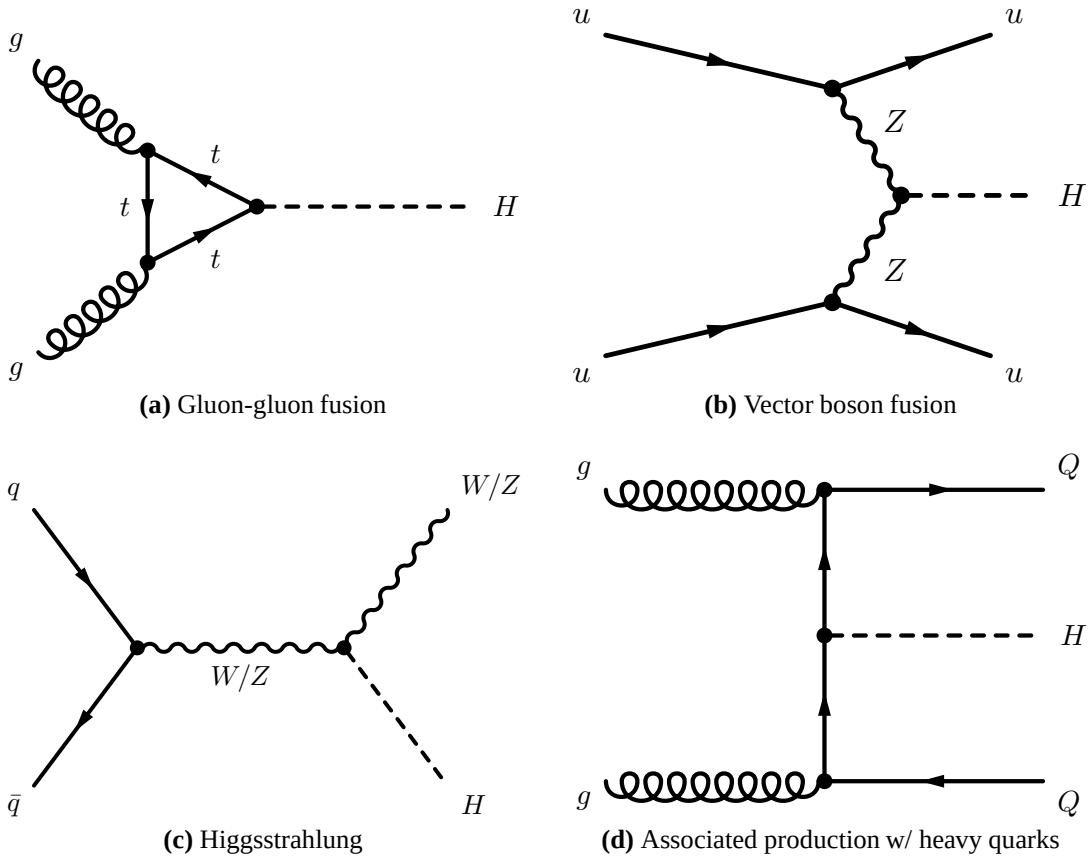


Figure 4.2.: Cross sections of dominant Higgs boson production processes at the LHC ($\sqrt{s} = 14\,\text{TeV}$) as a function of m_H [48].



(a) Gluon-gluon fusion

(b) Vector boson fusion

(c) Higgsstrahlung

(d) Associated production w/ heavy quarks

Figure 4.3.: Dominant Higgs boson production processes at the LHC.

two initial quarks each emit a W or Z boson which fuse to produce the Higgs boson. Therefore, the vector bosons must have an energy of $\mathcal{O}\left(\frac{1}{2}m_H\right)$. On the other hand, the W or Z boson tends to carry off only a small fraction of the energy of the initial quark [14]. Correspondingly, it follows that the outgoing quark must have a very large energy. Transverse momenta of the outgoing quarks, however, are set by the vector boson propagators in the matrix element to $p_{\mathrm{T}} \sim m_{W,Z}$. Although large on an absolute scale, the p_{T} is small compared to the total energy. This translates into small scattering angles with respect to the beam axis and, as a result, large separation in η. To summarise, the two *tagging jets* caused by the outgoing quarks have high p_{T} (with respect to the QCD jet spectrum), a large separation in η and a large di-jet mass.

An additional important feature of VBF Higgs boson production is suppressed hadronic activity in the central region of the detector. The suppression is a consequence of the lack of colour flow between the two initial quarks. Gluons are typically emitted from the quarks at small angles. In contrast, QCD background with colour exchange often emits gluons into the central region. This feature, in conjunction with the separation of tagging jets in η, is usually referred to as *rapidity gap*. It can be exploited by rejecting events with additional high-p_{T} central jets.

In the context of $H \to \tau^+\tau^-$ searches, VBF is of special interest for another reason. As the Higgs boson has to recoil against two jets with significant p_{T}, it receives transverse momentum. This ensures that the Higgs boson decay products are not back-to-back in the transverse plane, which is a prerequisite for mass reconstruction. See Sec. 4.6 for details.

4.4. $H \to \tau^+\tau^- \to \ell h + 3\nu$

As discussed above, the $\tau^+\tau^-$ decay mode is one of the important channels for Higgs boson masses close to the LEP exclusion limit. The total branching ratio ranges between 7.7% at $m_H = 115\,\mathrm{GeV}$ and 4.5% at $m_H = 135\,\mathrm{GeV}$ [48]. In the following, a Higgs boson mass of $m_H = 120\,\mathrm{GeV}$ will be assumed. In WW-related analyses ($H \to WW$, $t\bar{t}$), the semi-leptonic decay mode is usually considered a compromise between the high rates of the fully hadronic decay mode and the "cleanness" of the fully leptonic decay mode. And like in the fully leptonic decay mode, a lepton is available to trigger on. The τ lepton, however, has a higher leptonic branching ratio than the W boson: $\tau^- \to \mu^-\bar{\nu}_\mu\nu_\tau$ or $\tau^- \to e^-\bar{\nu}_e\nu_\tau$ account for 35.21% of the cases. Hence, with about 46% of all $\tau^+\tau^-$ decays the semi-leptonic channel has the highest branching ratio (see Table 4.1).

Due to the small ratio $m_\tau/m_H = 1.78\,\mathrm{GeV}/120\,\mathrm{GeV}$ the τ leptons are highly boosted. This is important for mass reconstruction (Sec. 4.6). In the ℓh channel[2], a high-p_{T} electron or muon is available for the trigger. The other τ decays hadronically and must be reconstructed and identified according to Sec. 3.4.5. The τ decay products tend to be central. In combination with VBF this can be exploited by requiring the decay products to be situated between the tagging jets in η. Three neutrinos in the final state come exclusively from the τ decays, which

[2] In the following, the signal process is denoted as $H \to \tau\tau \to \ell h$, where h represents a hadronic τ lepton decay.

decay mode	branching ratio
$\ell\ell$	12.40%
ℓh	} 45.63%
$h\ell$	
hh	41.97%

Table 4.1.: $\tau\tau$ decay modes and their respective branching ratios [2]. ℓ and h denote leptonic and hadronic τ decays respectively.

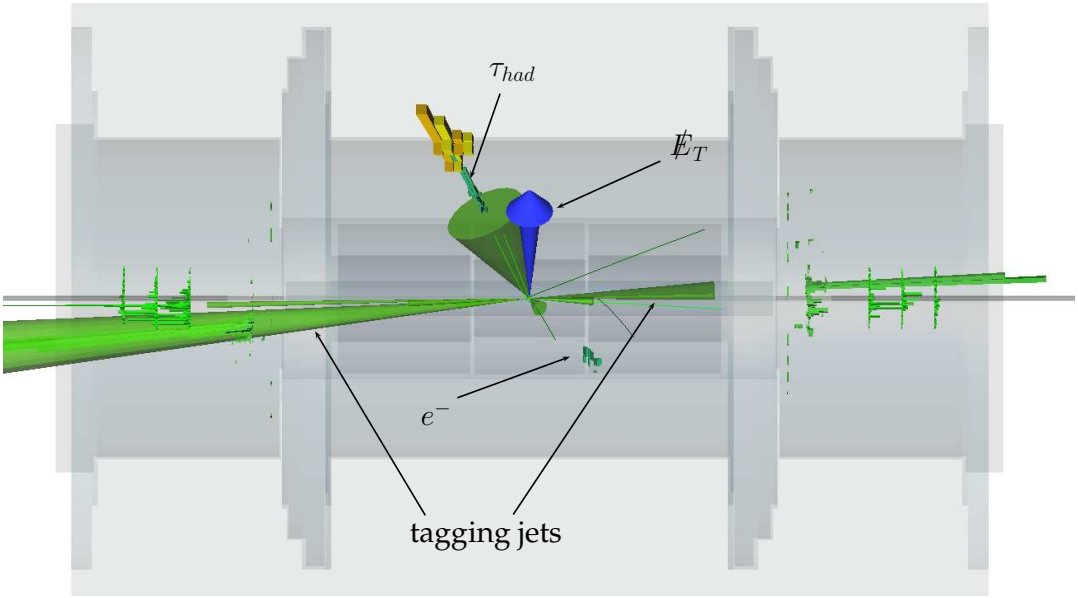


Figure 4.4.: ATLAS event display of a typical VBF $H \to \tau\tau \to \ell h$ event. Green cones represent reconstructed jets with length proportional to energy. The blue arrow stands for $\not{E}_{\mathrm{T}}$. Green and yellow cuboids symbolise energy depositions in ECAL/HEC/FCal and barrel HCAL respectively.

again is important for mass reconstruction. They provide considerable $\not{E}_{\mathrm{T}}$, which can be used to discriminate between signal and QCD multi-jets background or $Z \to \ell\ell$. An event display of a typical $H \to \tau\tau \to \ell h$ event is shown in Figure 4.4.

4.5. Hadronic Tau Decays

τ leptons decay via a vertex of the weak interaction, $\tau \to W^-\nu_\tau$ (Fig. 4.5). The decay mode is determined by the decay of the W boson. In 65% of the cases it decays into a pair of quarks which form hadronic final states. Valid final states are limited by $m_\tau = 1776.82 \pm 0.16\,\mathrm{MeV}$ [2]. Thus, only decays into mesons are allowed. Baryons have to be produced in pairs to

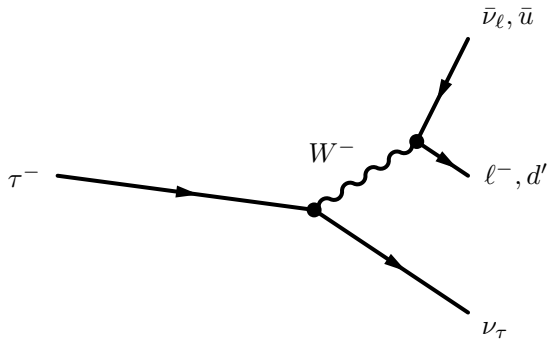


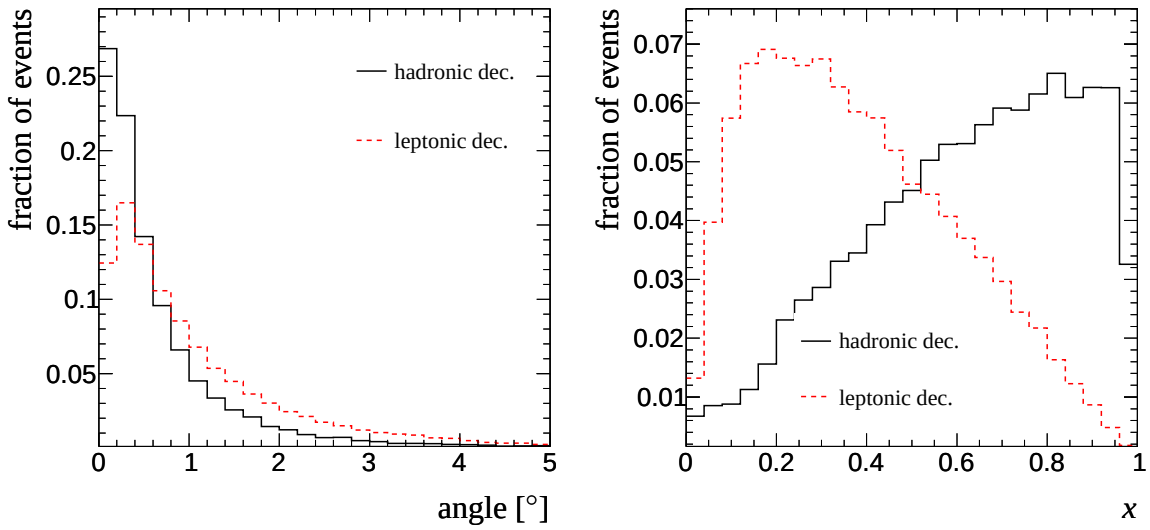
Figure 4.5.: Feynman graph for τ^- decays.

τ^- decay mode	branching ratio
$\pi^- \nu_\tau$	11.06%
$\rho^- \nu_\tau$	25.02%
$a_1^- \nu_\tau$	18.38%
$K^- \nu_\tau$	2.80%
others	7.54%

Table 4.2.: Branching ratios of hadronic τ decays [50].

conserve baryon number, with two times the proton mass being already greater than m_τ. To conserve the electric charge of the τ lepton, an odd number of charged particles has to be produced. Although in principle up to 11 charged pions are allowed kinematically, this process is very rare. 77.38% of all hadronic τ decays contain one charged particle (*1-prong*) [2] and zero or more neutral particles. Decays with three charged particles (*3-prongs*) and possibly neutral particles account for another 22.47%. Up to now, decays into five charged particles were observed, representing 0.15% of all hadronic τ decays.

Table 4.2 shows branching fractions of resonances in hadronic τ decays. Hadronic τ decays are subject to a number of selection rules (see [50] for details). Decays into a single charged π or K meson would be favoured by phase space. But the W boson decays into a left handed d quark and a right handed $\bar{u}$ quark. In the limit of massless quarks, helicity equals chirality. Therefore only spin 1 states are allowed. This would exclude pions and kaons, which are spin 0 particles. As quarks are not massless, decays into a single pion or kaon are possible but suppressed. The amplitude for kaon production is additionally suppressed by $\sin\theta_c$, the Cabbibo angle. The only resonances which are not suppressed are the $\rho(770)$ and the $a_1(1260)$. They subsequently decay into $\pi^-\pi^0$ and $\pi^-\pi^-\pi^+/\pi^-\pi^0\pi^0$ respectively.



(a) Angle between τ lepton and visible decay product. **(b)** Fraction x of the τ lepton energy carried away by the visible decay product.

Figure 4.6.: Distributions of τ lepton decays on generator level obtained from the signal process. No cuts are applied.

4.6. Mass Reconstruction

One of the key features of VBF $H \to \tau\tau$ is the possibility to reconstruct the Higgs boson mass. Although the signal process has three neutrinos in the final state, the neutrino momenta can be reconstructed if all decay products of a τ lepton go into the same direction. Due to the large Higgs boson mass and the comparatively small τ lepton mass this is a fairly good approximation (cf. Figure 4.6a). This approximation is called the *collinear approximation*. The known quantities entering the calculation of the neutrino momenta are the two momenta of the visible τ decay products, $\vec{p}_\ell$ and $\vec{p}_h$ as well as $\not{E}_\mathrm{T}$. As additional approximation it is assumed that $\not{E}_\mathrm{T}$ results only from the neutrinos of the τ decays: $\vec{\not{E}}_\mathrm{T} = \sum \vec{p}_\nu$.

Consider the triangle defined by $\not{E}_\mathrm{T}$, $\vec{p}_{\mathrm{T},\nu_\ell}$ and $\vec{p}_{\mathrm{T},\nu_h}$ (Fig. 4.7), where $\vec{p}_{\mathrm{T},\nu_\ell}$ is the sum of the neutrino transverse momenta in the leptonic τ decay, and $\vec{p}_{\mathrm{T},\nu_h}$ is the transverse momentum of the neutrino in the hadronic τ decay. With the collinear approximation, i.e.

$$\frac{\vec{p}_{\mathrm{T},\nu_\ell}}{p_{\mathrm{T},\nu_\ell}} = \frac{\vec{p}_{\mathrm{T},\ell}}{p_{\mathrm{T},\ell}} \quad \text{and} \quad \frac{\vec{p}_{\mathrm{T},\nu_h}}{p_{\mathrm{T},\nu_h}} = \frac{\vec{p}_{\mathrm{T},h}}{p_{\mathrm{T},h}},$$

also the ϕ-coordinates are equal. Therefore the Law of Sines gives:

$$\frac{p_{\mathrm{T},\nu_\ell}}{\sin\Delta\phi(\vec{\not{E}}_\mathrm{T}, \vec{p}_{\mathrm{T},h})} = \frac{p_{\mathrm{T},\nu_h}}{\sin\Delta\phi(\vec{p}_{\mathrm{T},\ell}, \vec{\not{E}}_\mathrm{T})} = \frac{\not{E}_\mathrm{T}}{\sin\left(\pi - \Delta\phi(\vec{p}_{\mathrm{T},\ell}, \vec{\not{E}}_\mathrm{T}) - \Delta\phi(\vec{\not{E}}_\mathrm{T}, \vec{p}_{\mathrm{T},h})\right)} \tag{4.1}$$

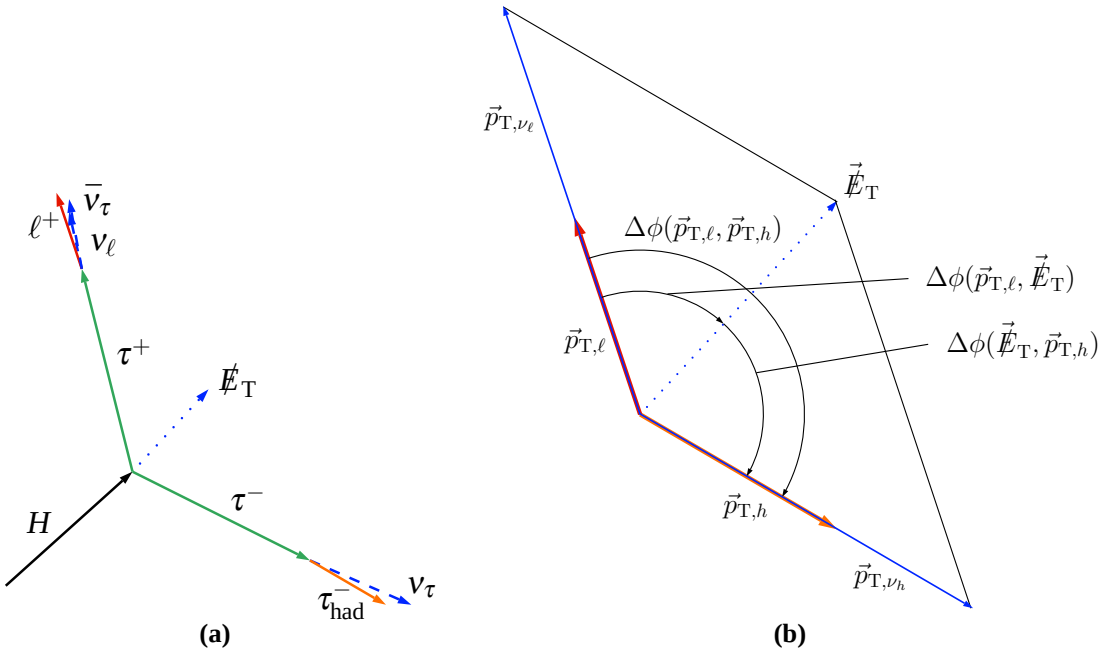


Figure 4.7.: Collinear approximation. a) Sketch of a $H \to \tau\tau \to \ell h$ decay. b) Vector decomposition of $\not{E}_T$.

The third sine can be contracted to $\sin\Delta\phi(\vec{p}_{T,\ell}, \vec{p}_{T,h})$[3]. With this, the transverse momenta of the neutrinos can be expressed by known quantities:

$$p_{T,\nu_\ell} = \not{E}_T \frac{\sin\Delta\phi(\vec{\not{E}}_T, \vec{p}_{T,h})}{\sin\Delta\phi(\vec{p}_{T,\ell}, \vec{p}_{T,h})} \tag{4.2}$$

$$p_{T,\nu_h} = \not{E}_T \frac{\sin\Delta\phi(\vec{p}_{T,\ell}, \vec{\not{E}}_T)}{\sin\Delta\phi(\vec{p}_{T,\ell}, \vec{p}_{T,h})} \tag{4.3}$$

In the collinear approximation, the decay products are parallel to the original τ. Thus, it follows that

$$p_{T,\ell} = x_\ell \, p_{T,\tau_1} = x_\ell \left(p_{T,\ell} + p_{T,\nu_\ell}\right), \tag{4.4}$$

where x_ℓ is the fraction of the τ momentum the leptonic decay product carries. x_h can be defined in the same way for the hadronic branch. With equations 4.2 and 4.3 x_ℓ and x_h can be

[3]Note that here $\Delta\phi$ has a sign determined by the direction of rotation.

expressed by measured quantities.

$$x_\ell = \frac{p_{\mathrm{T},\ell}}{p_{\mathrm{T},\ell} + p_{\mathrm{T},\nu_\ell}} = \frac{1}{1 + \dfrac{\not{E}_\mathrm{T} \sin\Delta\phi(\vec{\not{E}}_\mathrm{T}, \vec{p}_{\mathrm{T},h})}{p_{\mathrm{T},\ell} \sin\Delta\phi(\vec{p}_{\mathrm{T},\ell}, \vec{p}_{\mathrm{T},h})}} \tag{4.5}$$

$$x_h = \frac{p_{\mathrm{T},h}}{p_{\mathrm{T},h} + p_{\mathrm{T},\nu_h}} = \frac{1}{1 + \dfrac{\not{E}_\mathrm{T} \sin\Delta\phi(\vec{p}_{\mathrm{T},\ell}, \vec{\not{E}}_\mathrm{T})}{p_{\mathrm{T},h} \sin\Delta\phi(\vec{p}_{\mathrm{T},\ell}, \vec{p}_{\mathrm{T},h})}} \tag{4.6}$$

According to the theorem on intersecting lines, the ratio $p_{\mathrm{T},\tau_1}/p_{\mathrm{T},\ell}$ is equal to the ratio p_{τ_1}/p_ℓ. Therefore, both total τ momenta are now known:

$$\vec{p}_{\tau_1} = \frac{\vec{p}_\ell}{x_\ell} \tag{4.7}$$

$$\vec{p}_{\tau_2} = \frac{\vec{p}_h}{x_h} \tag{4.8}$$

Neglecting the τ lepton mass and the masses of the visible decay products, the invariant mass of the $\tau\tau$ system is:

$$\begin{aligned} M^2_{\tau\tau} &= 2p_{\tau_1}p_{\tau_2}\left[1 - \cos\angle(\vec{p}_{\tau_1}, \vec{p}_{\tau_2})\right] \\ &= \frac{2p_\ell p_h\left[1 - \cos\angle(\vec{p}_\ell, \vec{p}_h)\right]}{x_\ell x_h} \\ &= \frac{M^2_{\ell h}}{x_\ell x_h} \end{aligned} \tag{4.9}$$

Note that x_ℓ transforms into x_h if ℓ and h are exchanged. Although the formulae were derived from a specific order of the three vectors, the result is universal. From the definition of x_ℓ and x_h in Eq. 4.4 it is clear that physically meaningful values lie in the range $0 < x < 1$. Fig. 4.6b shows the true distributions (i.e. the fraction of energies not reconstructed using above formulae) of x_ℓ and x_h. The position of the peak is clearly different for hadronic and leptonic decays. The latter are three-body decays without intermediate resonances with the visible decay, the lepton, being one out of three particles. The visible hadronic tau decay ideally sums up all particles except for the ν_τ leading to higher x values.

The reconstruction of x_ℓ and x_h with the collinear approximation has two special cases in which no solution exists: a) Both decay products are parallel in the transverse plane and $\not{E}_\mathrm{T}$ points into the same direction. b) $\not{E}_\mathrm{T}$ is equal to the p_T of one of the visible decay products and points into the opposite direction. Two additional cases have a solution but yield zero and thus lead to a division by zero in Eq. 4.9: the visible decay products are either c) parallel or d) back-to-back in the transverse plane, and $\not{E}_\mathrm{T}$ is parallel to neither. All these cases, however, are of mathematical nature and should not occur in a real experiment.

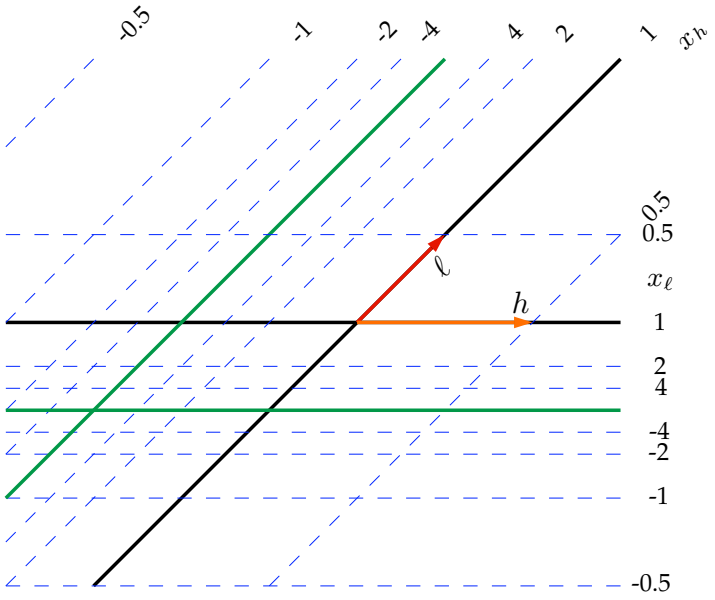


Figure 4.8.: Values of x_ℓ,x_h as a function of the $\not{E}_{\mathrm{T}}$-vector for fixed vectors $\vec{p}_\ell$ and $\vec{p}_h$. The $\not{E}_{\mathrm{T}}$-vector pointing along lines parallel to $\vec{p}_\ell(\vec{p}_h)$ results in constant values of $x_h(x_\ell)$. Solid black lines mark $x_\ell = 1$ and $x_h = 1$. Solid green lines are poles of x_ℓ and x_h respectivley, where the values approach $\pm\infty$.

A geometric example of possible configurations for fixed vectors $\vec{p}_\ell$ and $\vec{p}_h$ is shown in Fig. 4.8. For the special case where $\vec{\not{E}}_{\mathrm{T}}$ is just a multiple of $\vec{p}_\ell$, it is simple to derive values of x_ℓ: Then Eq. 4.5 reduces to $x_\ell = 1/(1+r)$ with $r = \not{E}_{\mathrm{T}}/p_{T,\ell}$. If $\vec{\not{E}}_{\mathrm{T}}$ points into the same direction as $\vec{p}_\ell$, i.e. $r \geq 0$, x_ℓ lies between 0 and 1. It approaches 0 for $\not{E}_{\mathrm{T}} \to \infty$ or $p_{T,\ell} \to 0$. $x_\ell = 1$ means $\not{E}_{\mathrm{T}} = 0$ or $p_{T,\ell} \to \infty$. If $\vec{\not{E}}_{\mathrm{T}}$ points into the opposite direction ($r < 0$) two cases can be distinguished. For $-1 < r < 0$ x_ℓ is positive and greater than 1. At $r = -1$ x_ℓ has a pole. $r < -1$ leads to negative values of x_ℓ. Generalising the special case where $\vec{\not{E}}_{\mathrm{T}} = a \cdot \vec{p}_\ell$ results in lines of constant x_ℓ in the $x-y$ plane parallel to $\vec{p}_h$. Requiring $0 < x_{\ell,h} < 1$ thus restricts $\vec{\not{E}}_{\mathrm{T}}$ to the area enclosed by the two black lines in the upper right part of Fig. 4.8.

Topologies close to points where no solution exists should be avoided in order to improve the resolution of $M_{\tau\tau}$. In Fig. 4.8, solutions for x_ℓ and x_h close to the green lines, case b), change rapidly. Hence, the $\not{E}_{\mathrm{T}}$ resolution has an especially large impact on the mass resolution in this regime. Case b) can be rejected by restricting x_ℓ and x_h to the range $0 < x_{\ell,h} < 1$. The other cases can be excluded by imposing an upper or lower limits on $\Delta\phi(h,\ell)$. The resulting mass distribution is shown in Fig. 4.9. After the full selection, discussed in Chapter 7, a mass resolution of about 10 GeV can be expected, neglecting effects of pile-up (see Chapter 8).

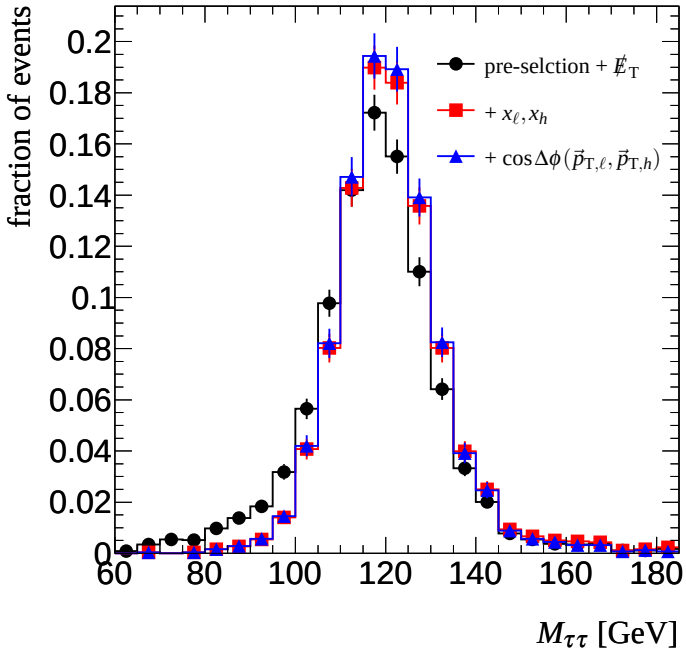


Figure 4.9.: Effect of cuts on the $M_{\tau\tau}$ distribution. Shown are $M_{\tau\tau}$ after pre-selection (cf. Sec. 7.2) and a cut on $\not{E}_T$ (black), after additionally requiring $0 < x_\ell < 0.75$, $0 < x_h < 1$ (red) and after additionally requiring $\cos\Delta\phi(\vec{p}_{T,\ell}, \vec{p}_{T,h}) > -0.9$ (blue). signum(x_ℓ) = signum(x_h) is implied already in the black distribution to avoid negative values under the square root in Eq. 4.9. All distributions are normalised to unity.

sample	σ_{prod} [pb]	BR($H \to \tau\tau$)	$\varepsilon_{\text{filter}}$	N_{MC}	$\int \mathscr{L}$ [fb^{-1}]
$m_H = 115$ GeV	4.436	0.0765	0.468	49,425	322
$m_H = 115$ GeV (pile-up)				48,800	318
$m_H = 120$ GeV	4.259	0.0711	0.469	99,946	726
$m_H = 120$ GeV (pile-up)				98,241	713
$m_H = 125$ GeV	4.100	0.0637	0.476	49,650	410
$m_H = 125$ GeV (pile-up)				47,725	394
$m_H = 130$ GeV	3.948	0.0549	0.479	49,850	491
$m_H = 130$ GeV (pile-up)				49,350	485
$m_H = 135$ GeV	3.801	0.04443	0.483	42,075	516
$m_H = 135$ GeV (pile-up)				40,600	497

Table 4.3.: Signal MC samples

4.7. Simulation

The signal samples are generated with HERWIG/JIMMY [33, 51, 52]. The process $q\bar{q} \to q'\bar{q}'H$ includes, apart from the VBF diagram, an additional s-channel graph. This, however, is not generated by HERWIG. It can be neglected if a jet topology as described in Sec. 4.3 is required. The sample is filtered on generator level for at least one electron or muon ("lepton filter") with $p_{\text{T}} > 5\,\text{GeV}$ and $|\eta| < 2.7$. The sample includes ℓh and $\ell\ell$ final states as well as some hh in cases where a suitable muon or (more likely) an electron came from a source other than the Higgs boson decay. Events are available for Higgs boson masses of $m_H = 115, 120, 125, 130, 135\,\text{GeV}$. τ decays are simulated by the TAUOLA package [53], QED radiative corrections by PHOTOS [54].

Production cross sections, σ_{prod}, the number of available events, N_{MC}, and the corresponding integrated luminosity can be found in Table 4.3. $\varepsilon_{\text{filter}}$ gives the efficiency of the applied filter. The cross sections are taken from [55] and do not take into account the s-channel graph. Branching ratios are taken from [48]. All simulated signal samples are at least ten times larger than expected for an integrated luminosity of $30\,\text{fb}^{-1}$.

5

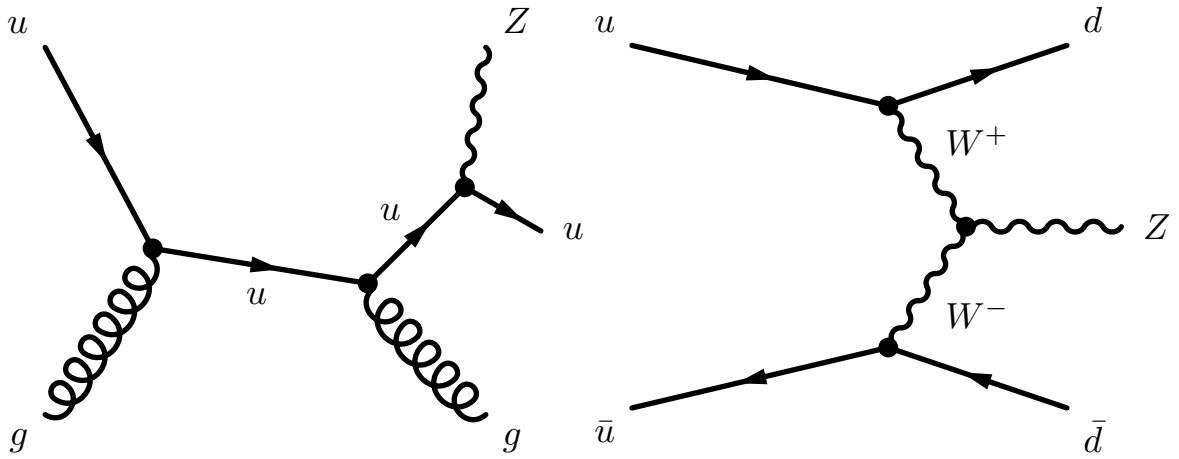
Background Processes

Usually, information on what kind of hard process has occurred during a proton-proton collision is experimentally inaccessible – it can only be identified indirectly by measuring its final state particles. In most cases, the final state of a process is not unique. Even if they are unique, limited resolution of the detector, misidentification of particles, inefficiencies or high-energetic contributions from pile-up or the underlying event can cause processes with different final states to end up in a signal selection.

Processes with the potential to contaminate a signal selection are referred to as *background processes*. Background is generally divided into two categories. *Irreducible* background means processes with a final state identical to the signal process. As such it can only be distinguished from a signal processes by kinematic criteria. *Reducible* background, on the other hand, refers to processes with different final states, that mimic a signal event by one of the effects described above.

It stands to reason that in principle any process is at least a reducible background. Whether it is likely to significantly contribute to a specific signal selection depends on the probability with which it is produced (the cross section) and the probability to mimic the final state in question (selection efficiency). The former can be calculated, the latter is usually determined by simulation. The dilemma here is that in principle one needs to know beforehand, which processes are likely to contribute because simulation is very time-consuming. This is a problem for processes with, for instance, low selection efficiency but a high cross section. A partial solution is to apply filters before the simulation step to accumulate topologies which are more likely to end up in a certain selection.

To summarise, any estimation of background from simulation alone will be more or less incomplete – a final assessment can only be done by comparison to data. In the following, I will give an overview of the background processes considered to be dominant and a description of the ways they most likely enter the signal selection. Descriptions are supplemented with information on which Monte-Carlo generators were used and how many simulated events this



(a) QCD **(b)** EW

Figure 5.1.: Example Feynman graphs for Z boson production.

study is based on. Note that the generators used for the individual processes are the same as in [1]. If not stated otherwise, TAUOLA was used for τ decays, PHOTOS for electromagnetic final state radiation.

5.1. $Z \to \tau\tau$

$Z \to \tau\tau$ is the dominant background to VBF $H \to \tau\tau$. It is an irreducible background process, i.e. if accompanied by two jets, the final state is identical. $Z \to \tau\tau$ can be subdivided into QCD and electroweak (EW) contributions, the labelling being based on whether there is colour exchange between the incoming partons. Two examples of Feynman diagrams are shown in Fig. 5.1. The EW contribution contains diagrams (and thus kinematics) which are similar to the signal process. This can be seen in Figures 5.2-5.4. The separation in η of the two leading jets is on average higher for EW induced processes than for QCD. For the former, the invariant mass of the tagging jets is very signal-like. And, like the signal process, EW $Z \to \tau\tau$ has suppressed hadronic activity in the central region. Although especially QCD $Z \to \tau\tau$ has different kinematical properties, the expected number of events after all cuts is substantial. The only way to distinguish those events from the signal process is by reconstructing the di-tau mass as described in Sec. 4.6.

$Z \to \tau\tau$ events corresponding to QCD diagrams are generated with ALPGEN plus HERWIG/JIMMY. No $\tau\tau$ decay mode is specified. Instead, a one-lepton filter with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ and $|\eta| < 2.7$ is applied. To enrich VBF like events, an additional filter (*VBF filter*) is applied, acting on jets reconstructed from generated final state particles with a cone of $\Delta R < 0.4$. The filter requires at least two jets which fulfil:

- $p_{\mathrm{T}} > 15\,\mathrm{GeV}$, $|\eta| < 5$

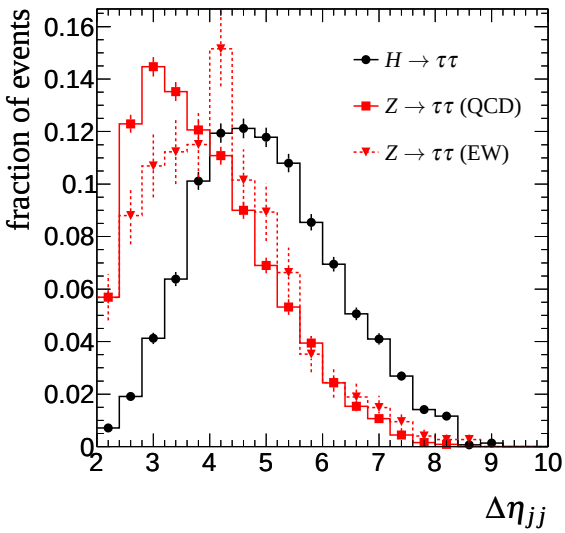


Figure 5.2.: η separation of tagging jets

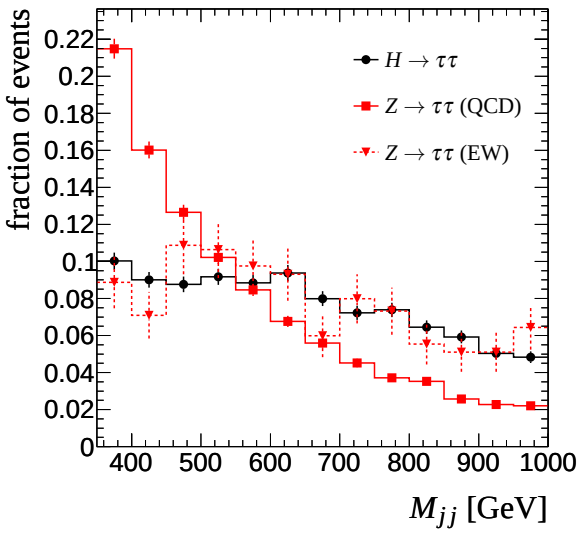


Figure 5.3.: invariant mass of the tagging jets

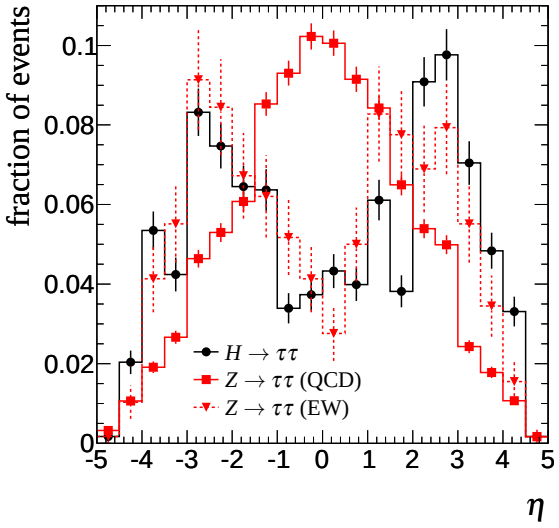


Figure 5.4.: pseudorapidty of additional jets

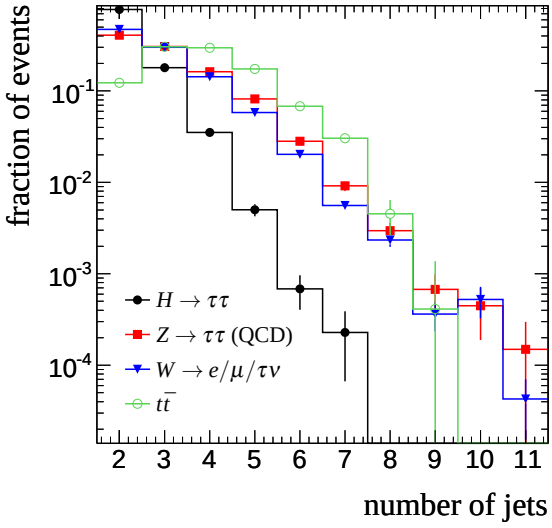


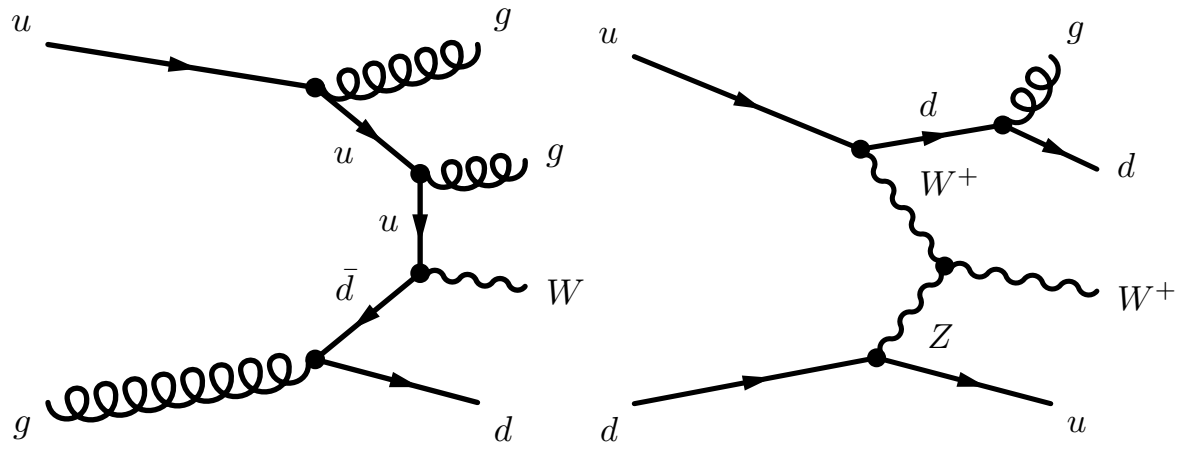
Figure 5.5.: jet multiplicity after pre-selection (cf. Sec. 7.2)

sample	σ_{prod} [pb]	$\varepsilon_{\text{filter}}$	N_{MC}	$\int \mathscr{L}$ [fb^{-1}]
ALPGEN QCD Np0	1606.98	0.00517	92,840	11.2
ALPGEN QCD Np0 (pile-up)			92,691	11.2
ALPGEN QCD Np1	412.32	0.0124	80,154	15.7
ALPGEN QCD Np1 (pile-up)			78,933	15.5
ALPGEN QCD Np2	149.89	0.0404	327,983	54.2
ALPGEN QCD Np2 (pile-up)			310,745	51.3
ALPGEN QCD Np3	50.63	0.0945	282,197	59.0
ALPGEN QCD Np3 (pile-up)			281,447	58.8
ALPGEN QCD Np4	15.41	0.167	132,031	51.4
ALPGEN QCD Np4 (pile-up)			132,729	51.7
ALPGEN QCD Np5	6.014	0.245	44,000	29.9
ALPGEN QCD Np5 (pile-up)			43,572	29.6
Sherpa EW	1.791		159,269	88.9

Table 5.1.: $Z \to \tau\tau$ MC samples

- no electron/photon/τ_{had} candidate with a relative difference in p_{T}, $(p_{T,\text{jet}} - p_{T,e/\gamma/\tau_{\text{had}}})/p_{T,e/\gamma/\tau_{\text{had}}}$, less than 0.3 within $\Delta R < 0.05$.
- $M_{jj} > 300\,\text{GeV}$
- $\Delta\eta_{jj} > 2$

$Z \to \tau\tau$ samples produced by ALPGEN are available with zero to five partons from the matrix element (*Np0-5*). Events from EW diagrams are produced by SHERPA which is also responsible for the decays of τ leptons. No filter is applied. The EW sample is only available without pile-up. The production cross sections, σ_{prod} are taken from ALPGEN and SHERPA for the corresponding samples. In accordance with [56], the ALPGEN cross sections are scaled such that the sum of the individual cross sections is equal to the fully inclusive NNLO cross section calculated with FEWZ [57, 58] for $60 < M_{\ell\ell} < 200\,\text{GeV}$. This corresponds to a *k-factor*[1] of 1.1. No k-Factor is applied to the SHERPA cross section. Apart from the ALPGEN Np0 and Np1 samples – which do not contribute much when applying full VBF cuts – all samples have at least the number of events expected for $30\,\text{fb}^{-1}$ (Table 5.1).



(a) QCD **(b)** EW with additional gluon radiation

Figure 5.6.: W boson production with three additional jets.

5.2. $W \to (\ell/\tau)\nu$

W boson production (Fig. 5.6) is a reducible background because the final state does not contain isolated electrons or muons and real τ_{had} candidates at the same time. In fact, the probability that a QCD jet is misidentified as a hadronic τ decay is much higher than the probability that it produces an isolated lepton[2]. Therefore the hadronic decay mode of $W \to \tau\nu$ practically does not contribute as background since it lacks a good electron or muon in the final state. Hence, the dominant topology is a W boson which decays into an electron or muon, accompanied by at least three jets. In contrast to $Z \to \tau\tau$, the additional jet is necessary to provide the τ_{had} candidate. In order to be reconstructed as τ_{had} candidate this jet has to be within $|\eta| < 2.5$, the region covered by the tracking detectors. Because of that, the average difference in η of the two tagging jets is higher than in (QCD-) $Z \to \tau\tau$. A jet identified as τ_{had} does not trigger the jet veto (Sec. 7.3.2), which makes this requirement less effective in suppressing W production. On the other hand, this means that the contribution of electroweak diagrams similar to the signal process is smaller than in $Z \to \tau\tau$: The lack of colour flow between the initial partons suppresses central jets needed to provide τ_{had} candidates.

The transverse mass of the lepton-$\not{E}_{\text{T}}$ system,

$$M_{\text{T}} = \sqrt{2 p_{\text{T},\ell} \cdot \not{E}_{\text{T}} \left(1 - \cos\Delta\phi(\ell, \not{E}_{\text{T}})\right)}, \tag{5.1}$$

is a poweful tool to distinguish leptonic W decays from other processes. Leptonic W decays typically have large M_{T} as can be seen in Figure 5.8. An upper limit on M_{T} can be used to suppress W background. Leptonic $W \to \tau\nu_\tau$ decays, however, have an additional neutrino

[1] The ratio of a cross section calculated at higher order over the LO cross section is often referred to as k-factor.

[2] Exact numbers depend on the details of the identification (see Sec. 7.1). Typical orders of magnitude are 10^{-2} for reconstructed τ_{had} candidates and 10^{-4} for leptons.

sample		σ_{prod} [pb]	$\varepsilon_{\text{filter}}$	N_{MC}	$\int \mathscr{L}$ [fb^{-1}]
$W \to e\nu$ QCD	Np0	20885	0.0101	19,950	0.118
	Np0 pile-up			19,926	0.118
	Np1	9334	0.0285	29,998	0.271
	Np1 pile-up			29,998	0.271
	Np2	5078	0.0826	99,999	0.905
	Np2 pile-up			99,590	0.901
	Np3	2476	0.209	149,997	1.642
	Np3 pile-up			146,031	1.598
	Np4	1112	0.364	69,956	1.401
	Np4 pile-up			69,956	1.401
	Np5	453	0.550	29,975	1.045
	Np5 pile-up			29,975	1.045
$W \to \mu\nu$ QCD	Np0	20886	0.00757	19,999	0.158
	Np0 pile-up			19,999	0.158
	Np1	9331	0.0253	29,949	0.309
	Np1 pile-up			29,949	0.309
	Np2	5054	0.0759	98,000	0.969
	Np2 pile-up			99,750	0.987
	Np3	2474	0.198	149,748	1.735
	Np3 pile-up			149,248	1.729
	Np4	1111	0.359	69,999	1.425
	Np4 pile-up			69,999	1.425
	Np5	454	0.539	29,749	1.046
	Np5 pile-up			29,999	1.054
$W \to \tau\nu$ QCD	Np0	20884	0.00140	10,000	0.429
	Np0 pile-up			10,000	0.429
	Np1	9331	0.0583	15,000	0.668
	Np1 pile-up			15,000	0.668
	Np2	5065	0.0175	50,000	2.145
	Np2 pile-up			50,000	2.145
	Np3	2470	0.0423	72,700	3.904
	Np3 pile-up			72,104	3.873
	Np4	1112	0.0839	35,000	3.053
	Np4 pile-up			34,750	3.031
	Np5	456	0.134	15,000	2.111
	Np5 pile-up			15,000	2.111

Table 5.2.: $W \to e/\mu/\tau$+jets MC samples

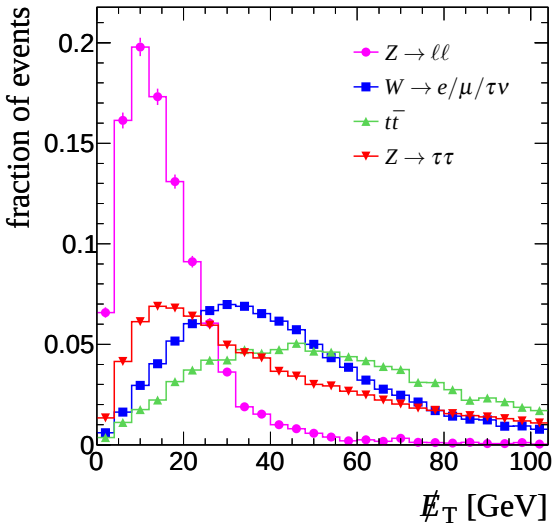


Figure 5.7.: Missing transverse energy after pre-selection

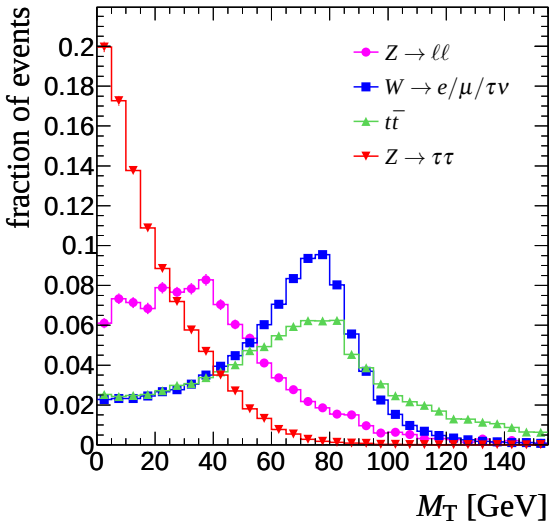


Figure 5.8.: Transverse mass after pre-selection

at a smaller angle to the lepton. An upper limit on M_T thus enriches $W \to \tau\nu_\tau$ compared to $W \to \mu\nu_\mu$ and $W \to e\nu_e$.

Monte-Carlo samples are generated with ALPGEN plus HERWIG/JIMMY. Again, only the QCD diagrams are included. Lepton- and VBF-filters are applied identical to the $Z \to \tau\tau$ samples generated with ALPGEN. A k-factor of 1.15 is obtained using FEWZ as described in Sec. 5.1. As can be seen in Table 5.2, the number of available events is much too small for the targeted integrated luminosity. $W \to e\nu$ and $W \to \mu\nu$ miss at least a factor of 17, while the situation for $W \to \tau\nu$ is slightly better (roughly a factor of 8 too small, at best). Samples including EW diagrams are not available but can be considered negligible due to the above reasons.

5.3. $Z \to \ell\ell$

Except for the Z decay products, $Z \to ee$ and $Z \to \mu\mu$ are identical to $Z \to \tau\tau$ as described in Sec. 5.1. A perfectly reconstructed $Z \to \ell\ell$ event is not a background to the signal process because the final state includes two well-isolated leptons. Apart from suppressing $Z \to \ell\ell$, such events have to be rejected in order to separate the $H \to \tau\tau \to \ell h$ measurement from $H \to \tau\tau \to \ell\ell$.

Accordingly, one of the leptons is either outside of the tracking region or otherwise fails to be reconstructed. This leaves two possible scenarios. One of the leptons is not reconstructed as electron or muon, but gets misidentified as a τ_{had}. This class of events will be similar to $Z \to \tau\tau$. It makes up two thirds of all $Z \to \ell\ell$ events with exactly one identified lepton, one τ_{had} candidate and at least two jets (objects defined as in Sec. 7.1). The second case (one third

sample		σ_{prod} [pb]	$\varepsilon_{\text{filter}}$	N_{MC}	$\int \mathscr{L}$ [fb^{-1}]	as ESD
$Z \to ee$ QCD	Np0	1607.18	0.00927	39,000	2.617	
	Np0 pile-up			39,000	2.617	
	Np1	410.81	0.0243	75,000	7.504	
	Np1 pile-up			74,750	7.479	
	Np2	149.01	0.0728	104,996	9.676	
	Np2 pile-up			104,746	9.653	
	Np3	50.48	0.176	99,998	11.232	
	Np3 pile-up			99,748	11.204	
	Np4	15.82	0.324	50,000	9.768	
	Np4 pile-up			49,750	9.719	
	Np5	5.816	0.459	29,954	11.215	
	Np5 pile-up			29,954	11.215	

Table 5.3.: $Z \to ee$ MC samples.

of events) occurs when instead of one lepton one of the jets is identified as τ_{had}. This class is similar to $W \to (\ell/\tau)\nu$.

The process does not include neutrinos, except from the decay of mesons in jets. However, the ATLAS detector has a finite $\not{E}_{\text{T}}$ resolution. In general, the presence of jets, which are usually not as well measured as other objects, adds to the $\not{E}_{\text{T}}$ measurement. In addition, effects as described in the previous paragraph contribute to $\not{E}_{\text{T}}$. Therefore, the $\not{E}_{\text{T}}$ distribution for this process has a most probable value which is greater than zero (cf. Fig. 5.7). Since $\not{E}_{\text{T}}$ comes from the jets or a lepton escaping detection, the $\not{E}_{\text{T}}$ vector will not point into the direction of the lepton or close to it. Compared to e.g. $Z \to \tau\tau \to \ell h$ where two of the three neutrinos are approximately parallel to the lepton, the larger angle between lepton and $\not{E}_{\text{T}}$ leads to a larger transverse mass (Fig. 5.8).

Both, $Z \to ee$ and $Z \to \mu\mu$, are provided by ALPGEN plus HERWIG/JIMMY. As with the other ALPGEN samples, only QCD diagrams are included. Lepton- and VBF filter are identical to $Z \to \tau\tau$ with the exception that here two leptons (electron/muon) are required. Cross sections provided by ALPGEN are used, multiplied by a k-factor of 1.1. $Z \to \mu\mu$ samples are also needed in ESD format (see Sec. **??**). Tables 5.3 and 5.4 shows the number of available events. Apart from Np0/1 all samples have roughly one third of the expected events.

sample		σ_{prod} [pb]	$\varepsilon_{\text{filter}}$	N_{MC}	$\int \mathscr{L}$ [fb^{-1}]	as ESD
$Z \to \mu\mu$ QCD	Np0	1607.71	0.00574	39,748	4.307	4,249
	Np0 pile-up			39,748	4.307	38,998
	Np1	411.21	0.0205	79,999	9.478	8,250
	Np1 pile-up			79,478	9.416	73,358
	Np2	148.59	0.0634	98,899	10.503	99,649
	Np2 pile-up			98,753	10.487	98,649
	Np3	50.57	0.161	99,750	12.263	97,750
	Np3 pile-up			99,500	12.232	96,091
	Np4	15.70	0.299	48,850	10.422	28,350
	Np4 pile-up			48,861	10.424	49,100
	Np5	5.902	0.443	30,000	11.477	3,250
	Np5 pile-up			29,750	11.381	27,500

Table 5.4.: $Z \to \mu\mu$ MC samples.

5.4. Top Quark Pair Production

If the W bosons in top quark pair production (cf. Fig. 5.9) decay into a lepton and τ_{had} respectively, the final state - including two b-jets to provide to high-p_{T} jets - is very similar to the signal process. Neutrinos from the W decays provide considerable $\not{E}_{\text{T}}$. This configuration of W decays, however, makes up only roughly 50% of all events passing the pre-selection (Sec. 7.2). The probability of a W decay resulting in a τ_{had}, $\text{BR}(W \to \tau\nu) \cdot \text{BR}(\tau \to \tau_{\text{had}}\nu)$, is approximately 7%. By contrast, $BR(W \to q\bar{q})$ is 68%. The efficiency to identify a QCD jet as a τ_{had} is on the percent level. This means that the probability to see an event with one of the four or more jets identified as a τ_{had} is comparable to events with a real τ_{had}.

The isolated lepton comes almost always from a W boson. Accordingly, the distribution of transverse mass (Fig. 5.8) is similar to that of $W \to (\ell/\tau)\nu$ events. As with the other QCD-processes considered here, hadronic activity in the central region is not suppressed. On the contrary, $t\bar{t}$ production has a high jet multiplicity (cf. Fig. 5.5), making a veto on additional central jets a powerful tool against this process.

The $t\bar{t}$ sample is produced by MC@NLO plus HERWIG/JIMMY. A filter is applied which selects events where both top quarks decay into a W boson and a b quark. At least one of the W boson has to decay to an electron, muon or τ lepton with p_{T} of the charged lepton above 1 GeV. The approximative NNLO cross section given in [59] is used. MC@NLO generates events with negative weights, i.e. the effective number of events is lower than the total number ($N_{\text{eff}} = N_{\text{tot}} - 2N_{\text{neg}}$). The sample is a factor of 10 smaller than 30 fb^{-1}, a factor of 20 with pile-up (Table 5.5).

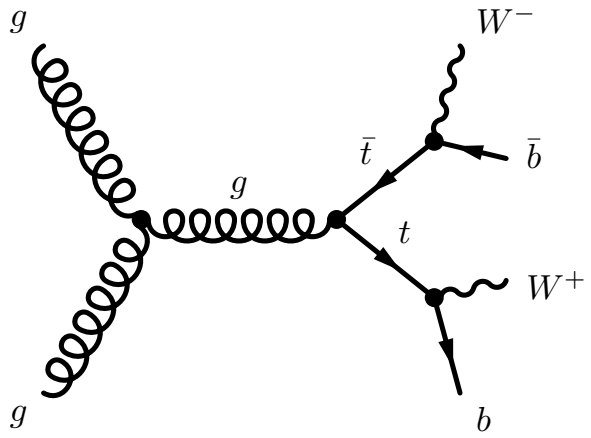


Figure 5.9.: Top quark pair production.

samples		σ_{prod} [pb]	$\varepsilon_{\text{filter}}$	$N_{\text{MC,eff}}$	$\int \mathscr{L}$ [fb^{-1}]
MC@NLO	no pile-up	883	0.561	1,449,607	2.926
	pile-up			696,371	1.406

Table 5.5.: $t\bar{t}$ MC samples

5.5. Di-Boson Production

Di-boson production (Fig. 5.10) refers to processes with two weak gauge bosons: WW, WZ or ZZ. The range of topologies is rather large and depends on the decay modes of the gauge bosons. WW is similar to $t\bar{t}$ production with the exception that it does not come with two

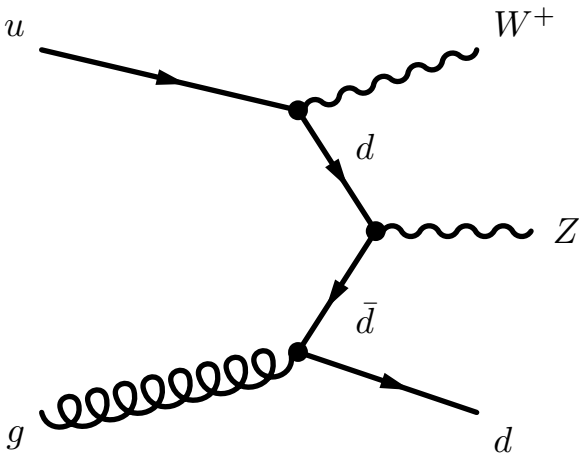


Figure 5.10.: Example of a di-boson graph. *Here: WZ* production with one additional jet.

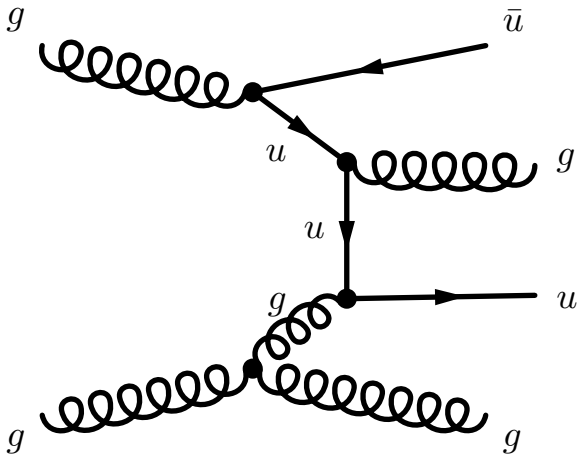


Figure 5.11.: A Feynman graph for multi-jets production with four jets.

sample		σ_{prod} [pb]	$N_{\text{MC,eff}}$	$\int \mathscr{L}$ [fb^{-1}]
WW		111.6	34,015	0.305
	pile-up		34,169	0.306
ZZ		14.75	33,431	2.267
	pile-up		23,373	1.585
W^+Z		29.37	35,894	1.222
	pile-up		35,894	1.222
W^-Z		18.39	36,117	1.964
	pile-up		36,117	1.964

Table 5.6.: Di-boson MC samples

additional jets at LO. Considering only the lepton and the τ_{had} candidate in the final state ZZ is comparable to single Z production. Events where both Z bosons decay leptonically will most likely be rejected due to the number of charged leptons. On the other hand, a Z boson decay cannot provide both tagging jets, as these have an invariant mass close to m_Z. WZ is similar to either W or Z production, depending on which boson produces the charged lepton.

Monte-Carlo samples are generated by MC@NLO plus HERWIG/JIMMY. No additional filters are applied. Cross sections are provided by MC@NLO. As for $t\bar{t}$, there are events with negative weights, so that the effective number of events is smaller than the number of generated events. It should be noted, that MC@NLO may not be the optimal choice for producing background to VBF $H \to \tau\tau$, because only one jet comes from a matrix element calculation. As shown in Table 5.6, the integrated luminosity of all samples is very small. (Table 5.6).

5.6. QCD Multi-Jet Production

Multi-jet production, an example graph is shown in Fig. 5.11, means processes including only quark and gluon jets in the final state. The event must consist of at least four jets, two of which need to fulfil the VBF criteria, one that produces an isolated high-energetic lepton and one that is identified as a τ_{had}. Multi-jets are a background by virtue of large numbers alone: While the probability for an event to mimic the signal is extremely low, the cross section is huge. In addition, like $Z \to \ell\ell$, multi-jet events have no genuine $\not{E}_{\text{T}}$.

Including these samples into the analysis is meant to give an impression of what to expect in data. Samples are produced with ALPGEN without generator-level filters at a centre-of-mass energy of 10 TeV. It should be noted, that ALPGEN produces multi-jet events including W and Z boson production, so there is some overlap with above processes. Although more than 11

million events are available this results in very low integrated luminosity for all sub-samples (cf. Table 5.7).

sample			σ_{prod} [pb]	N_{MC}	$\int \mathscr{L}$ [pb^{-1}]
QCD multi-jets	J2	Np2	30114237	2,920,123	0.097
		Np3	9835390	967,891	0.098
		Np4	1497279	148943	0.099
		Np5	226778	27,495	0.121
QCD multi-jets	J3	Np2	1116549	1,114,589	0.998
		Np3	1486726	1,475,443	0.992
		Np4	511243	551,024	1.078
		Np5	162795	187,775	1.153
QCD multi-jets	J4	Np2	31872	317,920	9.975
		Np3	64535	642,593	9.957
		Np4	50203	439,412	8.753
		Np5	24147	221,655	9.179
		Np6	11973	218,424	18.243
QCD multi-jets	J5+	Np2	751	226,742	302.081
		Np3	1920	587,532	280.063
		Np4	2173	642,301	295.528
		Np5	1432	407,899	284.925
		Np6	968	285,704	295.179
total				11,383,465	

Table 5.7.: QCD Multi-jets MC samples

6
Analysis Overview

Chapters 4 and 5 provided information on the characteristic properties of the signal process and the dominant background processes. Building on that information, the aim of this chapter is to give an overview of the analysis strategy, which is developed in the following chapters, and tools which are needed carry out the analysis.

6.1. Analysis Strategy

The analysis is kept close to [1] but differs distinctly in the treatment of control samples and pile-up.

Discriminating Observable: The reconstructed di-τ mass $M_{\tau\tau}$ is at the centre of the analysis. Discrimination between signal and background is based on good knowledge of the $M_{\tau\tau}$ distribution of the background processes.

Data: A future analysis on real data will be collected using a single electron or muon trigger. If necessary, a combination of electron/muon trigger and a τ_{had} trigger may be used in order to achieve a lower trigger rate or to lower the p_{T} threshold (cf. Sec. 7.4). For the purpose of this analysis, the data sample is assumed to have an integrated luminosity of $30\,\text{fb}^{-1}$. Such a data sample will be collected over an extended period of time and thus contain various configurations of luminosity and pile-up. Here, $\mathscr{L} = 10^{33}\,\text{s}^{-1}\text{cm}^{-2}$ and 6.9 simultaneous minimum-bias events are assumed for the whole sample.

Control Regions: In order to gather adequate knowledge of $M_{\tau\tau}$ distributions as claimed above, a core concept of this study is to extract all major backgrounds from data. Towards this end, two control regions are constructed. $Z \to \mu\mu$ can be selected virtually free of signal. This can be exploited to study $Z \to \tau\tau$, which otherwise would be impossible to collect without a significant contamination of signal close to a hypothetical signal peak

in $M_{\tau\tau}$. By collecting $Z \to \mu\mu$ and replacing the muons with simulated τ decay products, $Z \to \tau\tau$ can be estimated from data with only minimal input from MC (Sec. 9.1).

The other two major background processes are $W+$jets and $t\bar{t}$ production. In contrast to $H/Z \to \tau\tau$, here the lepton-τ_{had} pair is not produced resonant. The $M_{\tau\tau}$ distribution is shaped by the selection criteria. Assuming that for this reason the shape is independent of the charge of the lepton or τ_{had}, $M_{\tau\tau}$ can be extracted from a control region constructed by requiring equal charge instead of opposite charge (Sec. 9.2).

Selection: The cut-based selection is following [1] with only minor modifications. It can be subdivided into three parts: The pre-selection ensures that all required final state objects are present and correspond to the trigger channel. The first set of selection criteria acts on $\not{E}_{\text{T}}$ and the presumed τ decay products and thus makes use of the kinematics of the $H \to \tau\tau$ decay. The second set exploits the VBF topology of the tagging jets. Some modifications have to be implemented when considering pile-up.

Signal Extraction: The extraction of the signal significance is based entirely on $M_{\tau\tau}$ shapes. After the selection, three shapes are available: total data, $Z \to \tau\tau$ and $W+$jets/$t\bar{t}$. The signal significance is calculated using the *profile likelihood method* [60]. Signal and background shapes are parametrised and fit to the data, the two background shapes being fit to the respective control region simultaneously. Normalisations are free parameters of the fit, thus systematic uncertainties are shape uncertainties only. Systematic uncertainties enter the calculation as nuisance parameters in the profile likelihood.

6.2. Definition of Efficiency, Rejection and Purity

In the following chapters, the terms *efficiency*, *rejection* and *purity* will be used a number of times. Efficiency, when used in the context of a cut selection, is a synonym for acceptance, meaning number of events passing the selection over the total number of events. The reconstruction efficiency is defined with respect to generator-level objects:

$$\varepsilon_{\text{reco}} = \frac{\#\text{reconstructed objects, matched to generator-level object}}{\#\text{generator-level objects}} \tag{6.1}$$

In the case of τ_{had} candidates, the generator-level object is the visible τ decay, i.e. the four-vector of the ν_τ is subtracted from the four-vector of the τ. Generator-level jets are constructed by a cone algorithm running on generated particles. Identification efficiency is calculated relative to reconstructed objects:

$$\varepsilon_{\text{ID}} = \frac{\#\text{reconstructed and identified objects, matched to generator-level object}}{\#\text{reconstructed objects, matched to generator-level object}} \tag{6.2}$$

Rejection is the inverse efficiency, $R = 1/\varepsilon$ and is commonly used for background instead efficiency. The purity, in the context of reconstruction and identification, is defined as:

$$\text{purity} = \frac{\#\text{reconstructed and identified objects, matched to generator-level object}}{\#\text{reconstructed objects}} \tag{6.3}$$

6.3. Information from Low-Statistics Samples

In Chapter 5, the list of available MC events manifests a discrepancy between the size of the $t\bar{t}$ and $W+$jets samples and the target integrated luminosity. Resources available for long-term projections are naturally limited, especially if the machine parameters deviate from the current configuration. As mentioned in Chapter 3, the full GEANT4 simulation of an event is very time-consuming. Given the huge cross sections of both processes, generating samples corresponding to the assumed integrated luminosity of $30\,\text{fb}^{-1}$ was not feasible. On the other hand, because both processes are considered major backgrounds, this severely limits the ability to determine the sensitivity of the ATLAS detector for the signal process and the validity of the background extraction methods. In order to mitigate this effect, two methods are used in this study to extract additional information from low-statistics samples.

6.3.1. τ_{had} ID Factorisation

In samples without real hadronic τ decays, one of the steps with the least acceptance in the analysis is to require an identified τ_{had}. In the context of this study, practically all τ_{had} candidates in $W+$jets are misidentified QCD jets. Assuming that this misidentification is a purely statistical process, dropping τ_{had} identification and instead weighting the event with the probability to pass the identification should yield the same result. A very similar method was already successfully used in [61]. The probability, i.e. identification efficiency, clearly depends on the transverse momentum of the candidate. Identification probabilities are determined using p_{T}, the number of tracks, the type of particle which initiated the jet (up-type quark, down-type quark or gluon) and the electric charge of the candidate. The identification efficiency is processes dependent. Therefore, it is taken from the sample to which the method is to be applied. Events are then processed as follows:

- Treat every permutation of identified/not identified candidates as a separate event.
- Weight every such event with $\prod_{i}^{\text{ID}} \varepsilon_i \cdot \prod_{j}^{\not\text{ID}} \left(1-\varepsilon_j\right)$.

Here, the first product runs over all τ_{had} candidates assumed to be identified in the given permutation. ε_i is the identification probability of a specific candidate as a function of the parameters presented above. The second product runs over all τ_{had} candidates assumed to have failed the identification in the given permutation. An example is given in Table 6.1.

event	identification		weight
	passed	failed	
1	$\tau_{\text{had},1}$, $\tau_{\text{had},2}$		0.02
2	$\tau_{\text{had},1}$	$\tau_{\text{had},2}$	0.18
3	$\tau_{\text{had},2}$	$\tau_{\text{had},1}$	0.08
4		$\tau_{\text{had},1}$, $\tau_{\text{had},2}$	0.72
			1.0

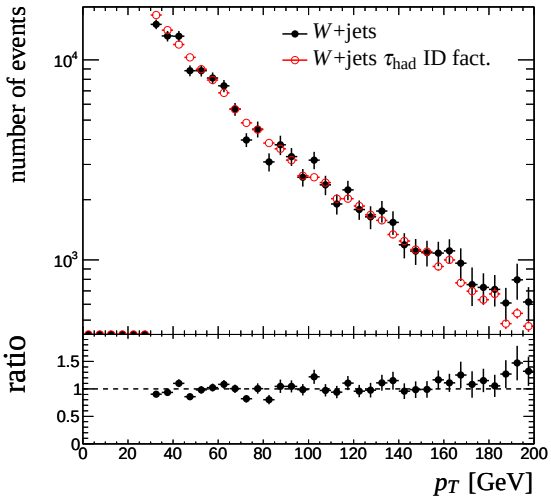
Table 6.1.: An example of τ_{had} ID factorisation of a single event containing two reconstructed τ_{had} candidates with identification probabilities $\varepsilon_1 = 0.2$ and $\varepsilon_2 = 0.1$. The event is split into four separate events with weights according to the assumed result of the τ_{had} identification. The sum of weights of all four events is equal to 1.

Figures 6.1a and 6.1b show τ_{had} p_{T} and $M_{\tau\tau}$ for standard τ_{had} identification compared to factorised identification. See appendix B for additional distributions. Both distributions agree in shape and normalisation. This is true for most other relevant distributions. Therefore this methods provides reliable results.

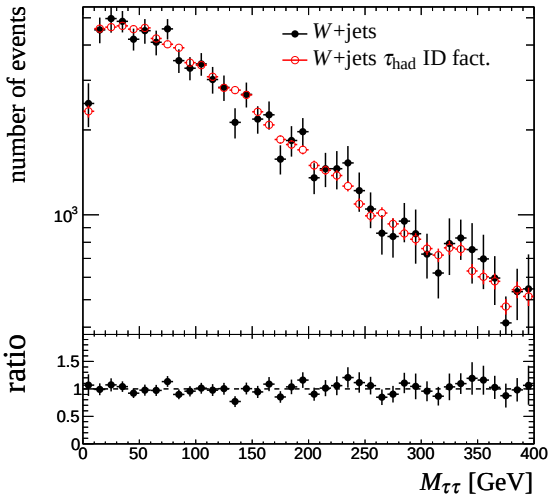
Two exceptions can be seen in Figs. 6.1c and 6.1d. The discrepancy in the charge combinations of lepton and τ_{had} has to be corrected for. This observable directly affects the efficiency of requiring opposite sign events as well as the size of the $W+$jets control sample (cf. Sec. 9.2). The reason for the deviation is probably an inaccurate parametrisation of the type of jet, due to the ambiguities encountered in matching generated partons to reconstructed objects. To take this effect into account, the number of events obtained with this method is scaled to reproduce the number of events in the non-factorised sample after the lepton-τ_{had} charge requirement. A second deviation is present in the η distribution of τ_{had} candidates. The τ_{had} identification efficiency is not flat in η (see Sec. 7.1.3). Because the identification efficiency is not parametrised as a function of η, the distributions cannot be expected to agree. A full parametrisation of the efficiency in p_{T} and η is not feasible due to the limited number of available events. This directly affects the lepton-τ_{had} centrality requirement (cf. Sec 7.3.2). It requires τ_{had} and lepton to be between the tagging jets in η. The factorisation appears to produce τ decay products less central than the real distribution. One can argue that this leads to a lower acceptance of the requirement and thus the factorisation yields a conservative estimate.

6.3.2. Cut Factorisation

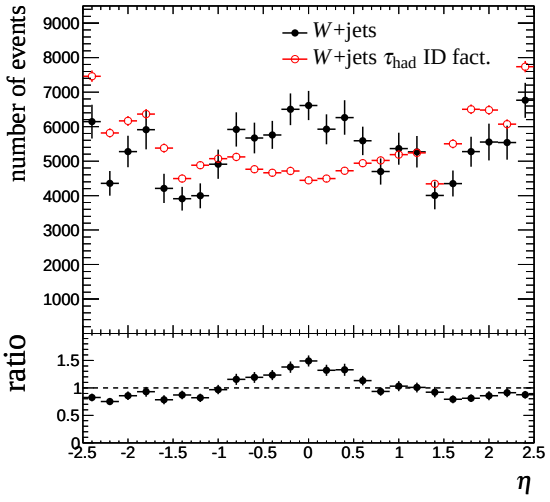
Factorisation of τ_{had} identification works less well for samples with a significant admixture of real hadronic τ decays. The identification process can no longer be assumed random and the gain in statistics is much lower since the acceptance of the corresponding requirement is



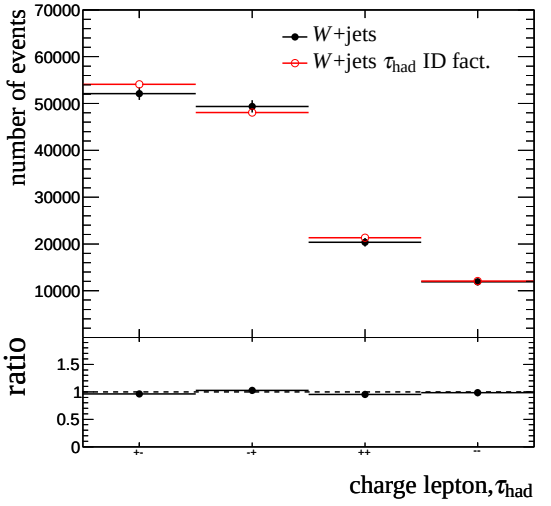
(a) τ_{had} p_T [GeV]



(b) $M_{\tau\tau}$ [GeV]



(c) τ_{had} η



(d) lepton/τ_{had} charge combination

Figure 6.1.: Comparison of distributions with and without τ_{had} ID factorisation after requiring exactly one identified τ_{had}.

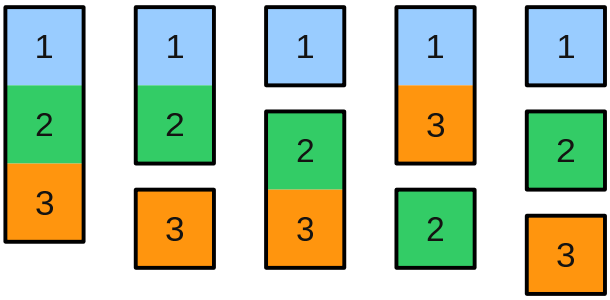


Figure 6.2.: There are five ways in which to factorise three cuts. Adjacent rectangles represent consecutively applied cuts, separated rectangles are factored out. The leftmost partition is equivalent to not employing cut factorisation.

higher. A more general approach to increase the effective size of a sample is cut factorisation.

Assuming that in a cut-based selection two cuts with efficiencies ε_1 and ε_2 are uncorrelated, the total efficiency ε_{tot} is given by $\varepsilon_{\text{tot}} = \varepsilon_1 \cdot \varepsilon_2$. In that case, ε_{tot} can be determined by applying cuts 1 and 2 separately to a sample. The obvious advantage is the gain in statistics since both cuts are applied to the full sample. In general, of course, a typical selection consists of cuts with various degrees of correlation to other cuts. The task then is to arrange correlated cuts into sets such that these sets have minimal correlations with the other sets. This poses two challenges: on the one hand, the number of ways a set of n elements (cuts) can be partitioned into nonempty subsets grows rapidly with n. In fact, the number corresponds to the Bell numbers B_n [62]. For 10 cuts there are already 115975 possible arrangements. See Figure 6.2 for an example. On the other hand, in a scenario in which cut factorisation is necessary, it is impossible to directly estimate the goodness of the approximation.

Several approaches are commonly used in order to find out which factorisation works best, ranging from fast simulation to employing common sense. Here I demonstrate a systematic approach to this problem. The general idea is to test all possible factorisations and sort them according to a figure of merit extracted from a set of the same cuts with softer thresholds.

The method acts on samples after a proper pre-selection. This is necessary for two reasons: Firstly, the limiting factor for this approach is the number of cuts to be factorised. Secondly, there are obviously cuts which make sense only after a pre-selection. Requiring τ decay products to be between the tagging jets in η means to require two jets, an identified τ_{had} and an identified lepton, in the first place. The pre-selection corresponds to the one described in Sec. 7.2. In order to reduce the set of remaining cuts further, the cuts related to the collinear approximation are combined into one cut. The purpose of this method is to find a cut factorisation which comes as close as possible to the result obtained applying all cuts consecutively. Hence, a suitable figure of merit is:

$$\text{f.o.m.} = \frac{\prod_i \varepsilon_{\text{cut i}}}{\varepsilon_{\text{tot}}} \tag{6.4}$$

Good values are then close to one. The analysis is performed on two samples:

1. **Loose cuts test sample:** This is the sample for which the best cut factorisation needs to be found. In order to be able to compare factorisation and standard cut analysis, a loose cut selection (cf. Appendix D) is applied.

2. **Tight cuts reference sample:** In order to see if the cut factorisation works as well with tighter cuts, a sample is needed, which is not too different from the sample in question but for which enough events are available to perform the normal analysis. On this sample the cut factorisation is done with the usual cuts (Sec. 7.3).

A good cut factorisation is defined as one with the figure of merit in both samples deviating no more than a pre-defined fraction from one. By selecting several good factorisations, which have the additional feature that the figure of merit in the loose cut sample is close to the one in the reference sample, it is possible to derive a rough measure of uncertainty introduced by this method. An application of this method to $t\bar{t}$ is shown in Sec. 7.5.2.

7
Signal Selection

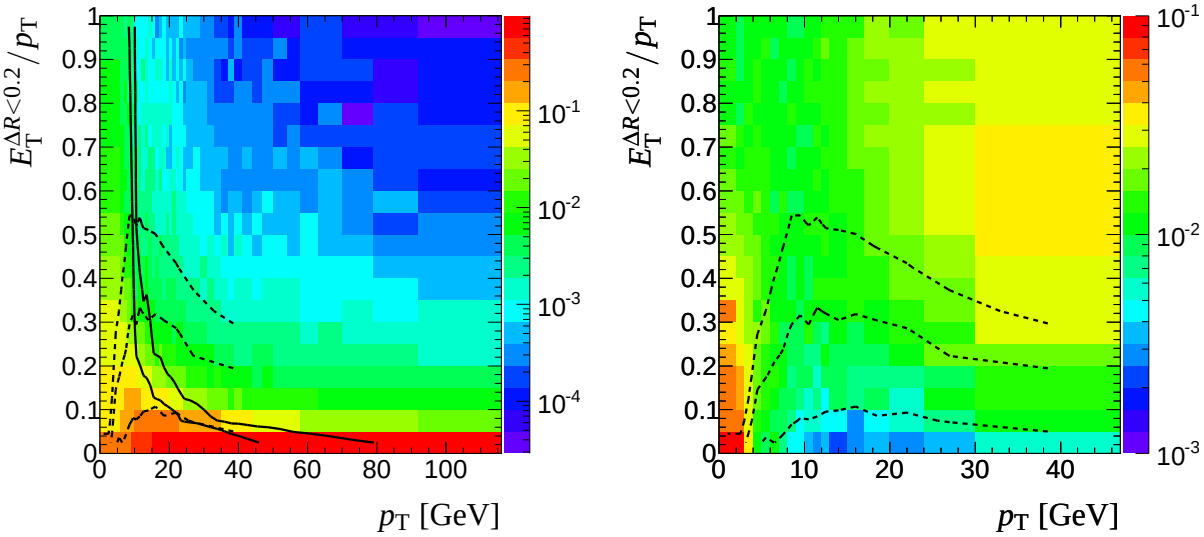
This chapter covers the requirements (cuts) applied to data in order to select signal-like events. The cuts are mostly identical to a selection proposed in [63] on a theoretical basis and optimised for use in ATLAS in [1]. As the latter study was performed on a considerably larger statistical basis than this thesis, I decided against a full re-optimisation of the selection. This chapter is an update on previous results using samples without pile-up. It also serves as a reference for Chapter 8, in which the effects of pile-up are presented.

7.1. Object Definition

Most of the simulated samples used for this analysis where filtered requiring a high-energetic electron or muon on generator level. This approach is valid only if the selection ensures a high enough purity. Object identification has to reflect this condition instead of maximising the signal significance. Electrons, muons, τ_{had} candidates and jets are considered as objects in the following.

7.1.1. Muons

A muon candidate is defined as a combined muon reconstructed by the STACO algorithm (cf. Sec. 3.4.3). No additional identification criteria are applied, except for requiring the combined track to be the best match of a track in the inner detector to the track in the muon spectrometer. Here, "signal" muons are defined as originating from Z/W bosons or τ leptons. They are produced outside of jets and are therefore usually well *isolated*: only a small amount of energy is deposited in the calorimeter around a muon. Background comes mainly from real muons in jets or long-lived hadronic objects like kaons. In such cases, muon candidates will be less well isolated. Fig. 7.1b shows E_{T} deposited in the calorimeter around signal muons.



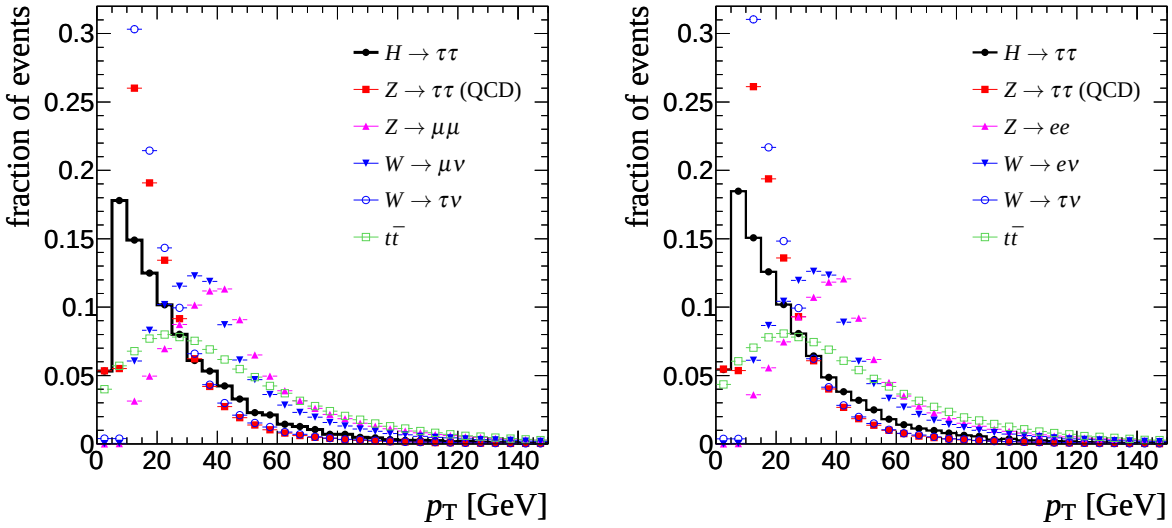
(a) Muons from leptonic Z, W and τ decays.

(b) Muons produced in jets.

Figure 7.1.: Relative isolation of muons. Bins are normalised such that the integral over each column of p_T is 1. Solid lines reflect the 90% and 95% contours of muons from Z/W bosons and τ decays. Dashed lines show the 1%,5% and 10% contour lines of muon candidates from jets.

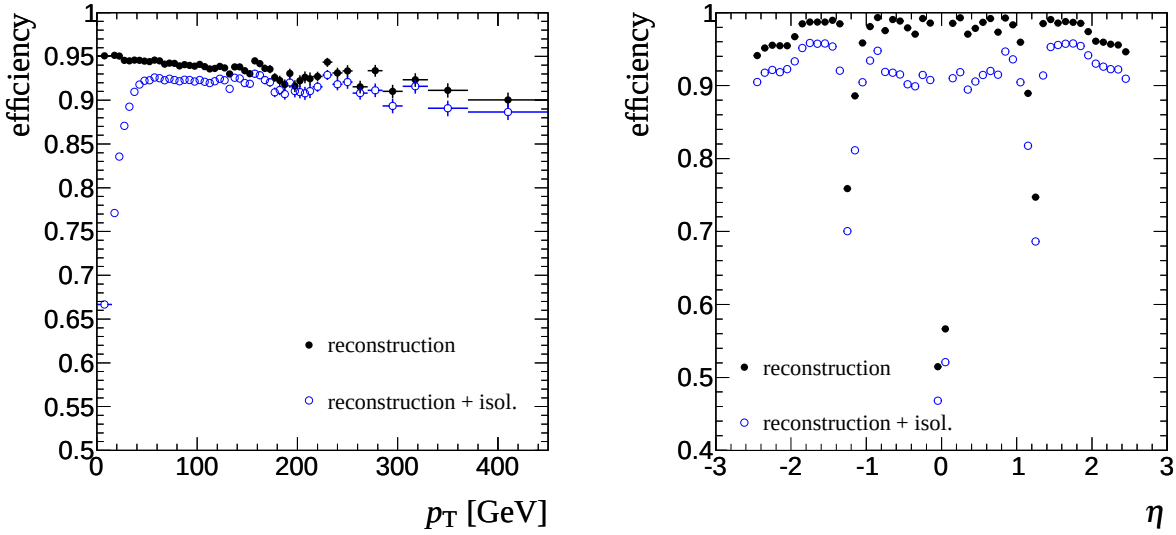
The corresponding p_T spectra can be seen in Fig. 7.2a. With a p_T between 10 and 100 GeV, 90% of these muons have an E_T below 2.5 GeV in a cone of $\Delta R < 0.2$. In this analysis, muons are required to have E_T in $\Delta R < 0.2$ over p_T of the muon to be smaller than 0.1 ("relative isolation"). While preserving 95% of signal muons with $p_T > 40$ GeV, the purity is expected to be well above 95%. Below a p_T of roughly 10-15 GeV the 90% contour level rises to high values of relative isolation, so candidates are required to have $p_T > 10$ GeV.

The selection efficiency is shown in Figs. 7.3a and 7.3b. It approaches a maximum of approx. 92% around $p_T = 50$ GeV and decreases towards smaller values. This decrease is due to relative isolation: for muons with low p_T the allowed energy gets smaller, which makes it more probable that the muon is rejected due to noise or activity from the underlying event. For high muon momenta, the probability to produce electromagnetic showers even in the muon chambers increases. These showers mask hits caused by real muons and explain the slowly decreasing efficiency for large p_T [19]. The efficiency as a function of η shows three pronounced dips, which reflect the gap in the muon spectrometer around $\eta = 0$ and the transition between barrel and endcap chambers at $|\eta| \approx 1.3$.



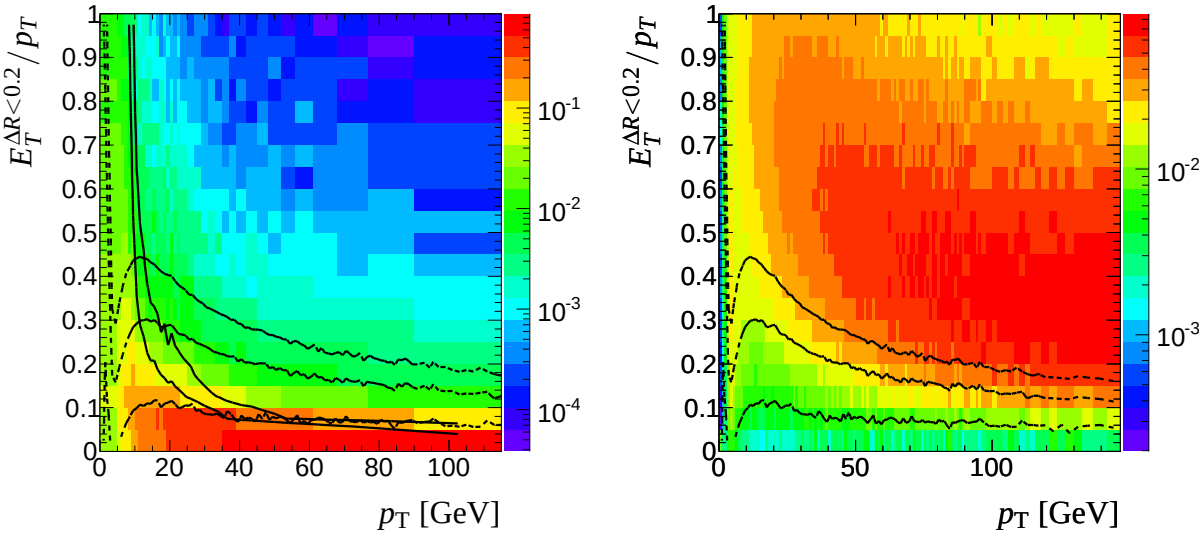
(a) Generated muons. **(b)** Generated electrons.

Figure 7.2.: p_T spectra of generated electrons and muons from W, Z or τ decays.



(a) efficiency vs p_T **(b)** efficiency vs η

Figure 7.3.: Efficiency to reconstruct muons from $Z \to \mu\mu$ decays with and without additional isolation requirement.



(a) Electrons from leptonic Z, W and τ decays.

(b) Electrons produced in jets.

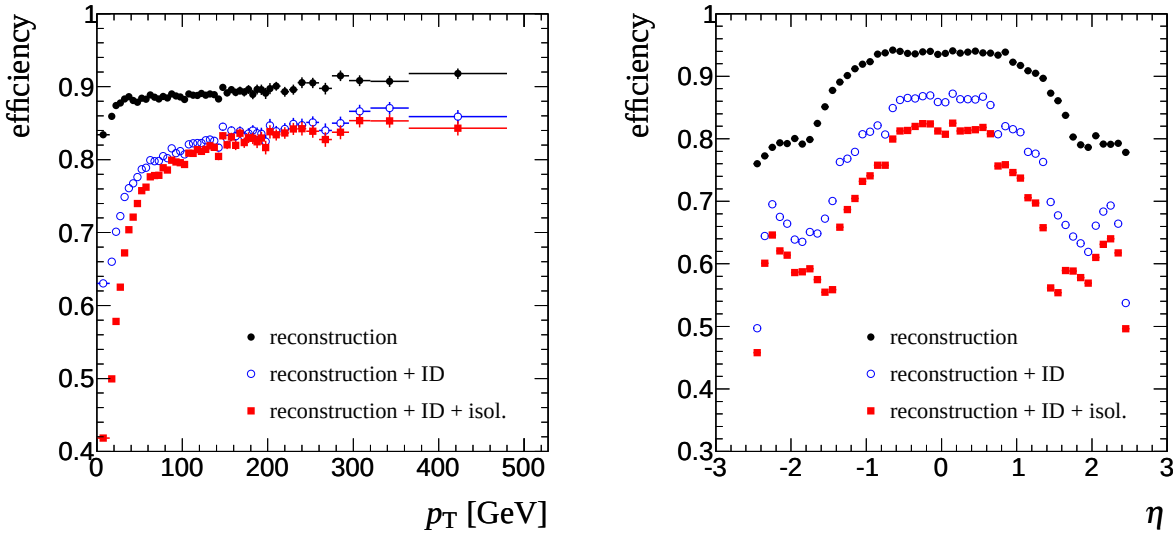
Figure 7.4.: Relative isolation of electrons. Bins are normalised such that the integral over each column of p_T is 1. Solid lines reflect the 90% and 95% contours of electrons from Z/W bosons and τ decays. Dashed lines show the 1%,5% and 10% contour lines of electron candidates from jets.

7.1.2. Electrons

Electron candidates are objects reconstructed by the calorimeter-based electron algorithm with $p_T > 15\,\text{GeV}$ and passing the medium identification criteria (cf. Sec. 3.4.2). There are several sources of background. Photons convert to electron-positron pairs. Hence, a dedicated conversion finder is included in the electron reconstruction. Similar to muons, additional background comes from hadronic objects or real electrons in jets, both of which are not isolated (cf. Fig 7.4b). Therefore, electrons have to pass the same isolation requirement, $E_T^{\Delta R<0.2}/p_T < 0.1$. p_T spectra of generator-level electrons from W, Z or τ decays can be seen in Fig. 7.2b. The electron identification efficiency varies with p_T (and thus η) as can be seen in Figs. 7.5a and 7.5b. It ranges from approximately 60% at $p_T = 10\,\text{GeV}$ to greater than 80% for electrons above 100 GeV. Relative isolation additionally reduces the efficiency by up to 20 percentage points for low-p_T electrons.

7.1.3. Hadronic τ decays

τ_{had} candidates are reconstructed by either the calorimeter-seeded or track-seeded algorithm (or both) (cf. Sec. 3.4.5). The reconstruction step of the calorimeter-seeded algorithm just selects reconstructed jets with p_T above a certain threshold. Hence, in contrast to electrons and muons, the reconstruction step does not significantly reduce the misidentification probability. Neglecting the very rare 5-prong τ decays, a hadronic τ decay contains either one or three

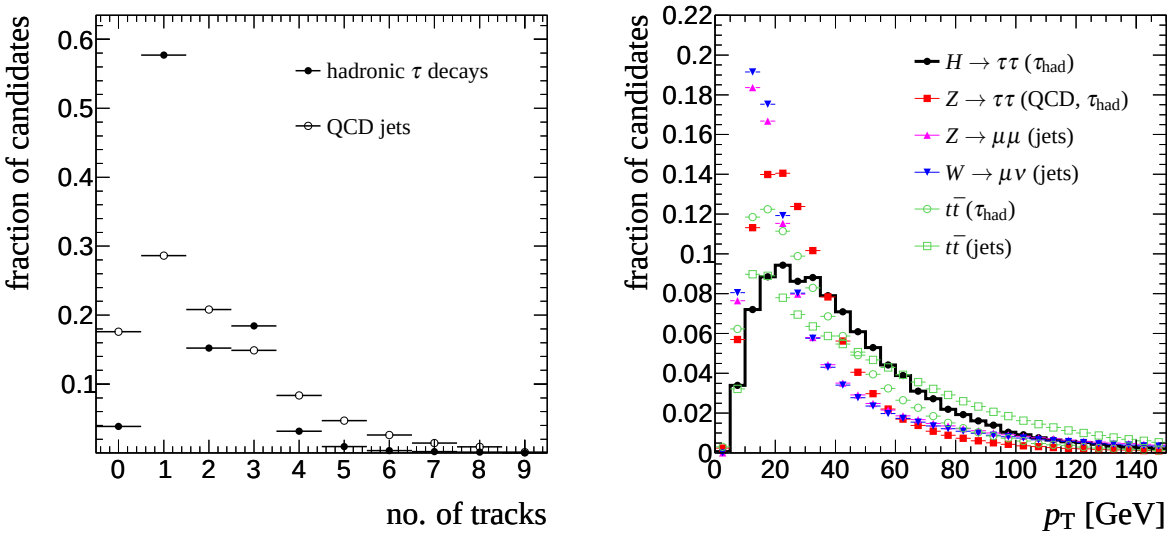


(a) efficiency vs p_T **(b)** efficiency vs η

Figure 7.5.: Electron reconstruction, identification and isolation efficiency for electrons from $Z \to ee$ decays as a function of p_T and η.

charged particles. The track multiplicity is shown in Fig. 7.6a. τ_{had} candidates are required to have one or three associated tracks. This suppresses QCD jets, which often have a higher track multiplicity. Photon conversions are suppressed by vetoing candidates with two tracks. The p_T threshold of 30 GeV helps to further suppress background from QCD jets. The p_T spectra of τ_{had} candidates from various sources is shown in Fig. 7.6b.

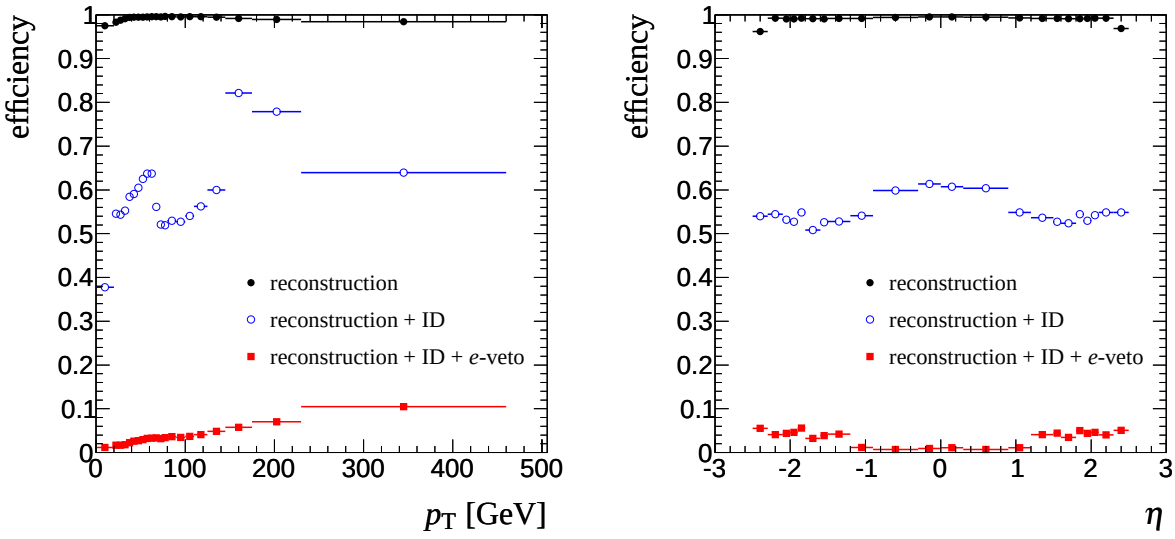
As Figures 7.7 and 7.8 demonstrate, electrons and muons have a high probability to be reconstructed and even identified as a hadronic τ decay: electrons have one track and a well confined shower in the calorimeter. Muons also have a track in the inner detector and deposit energy in the ECAL and the HCAL. Misidentification of electrons or muons is suppressed by several means: If a τ_{had} candidate overlaps with an identified electron or muon within $\Delta R < 0.2$ it is discarded ("overlap removal"). Muons are minimum ionising particles, i.e. they only deposit small amounts of energy in the calorimeter. Usually this means that muons are not reconstructed by a jet algorithm or are rejected by the p_T threshold for τ_{had} candidates. The track-seeded algorithm, however, reconstructs candidates based on the p_T of the leading track. If such a muon is not also reconstructed by the calorimeter-seeded algorithm, the energy calculus is taken from the track-seeded candidate. Thus, muons with p_T above the threshold of the track-seeded algorithm but still sufficiently low to deposit energy in the calorimeter below the threshold of the calorimeter-seeded algorithm have a high probability to be reconstructed as hadronic τ decay (cf. Fig. 7.8a). Because of the non-negligible identification efficiency of electrons and muons, overlap removal alone is insufficient. The τ_{had} identification provides dedicated vetoes: The muon veto is implemented by requiring an energy of at least 5 GeV in the ECAL. The electron veto rejects electrons based on a dedicated electron identification.



(a) Number of tracks associated with τ_{had} candidates.

(b) p_T of τ_{had} candidates on generator level.

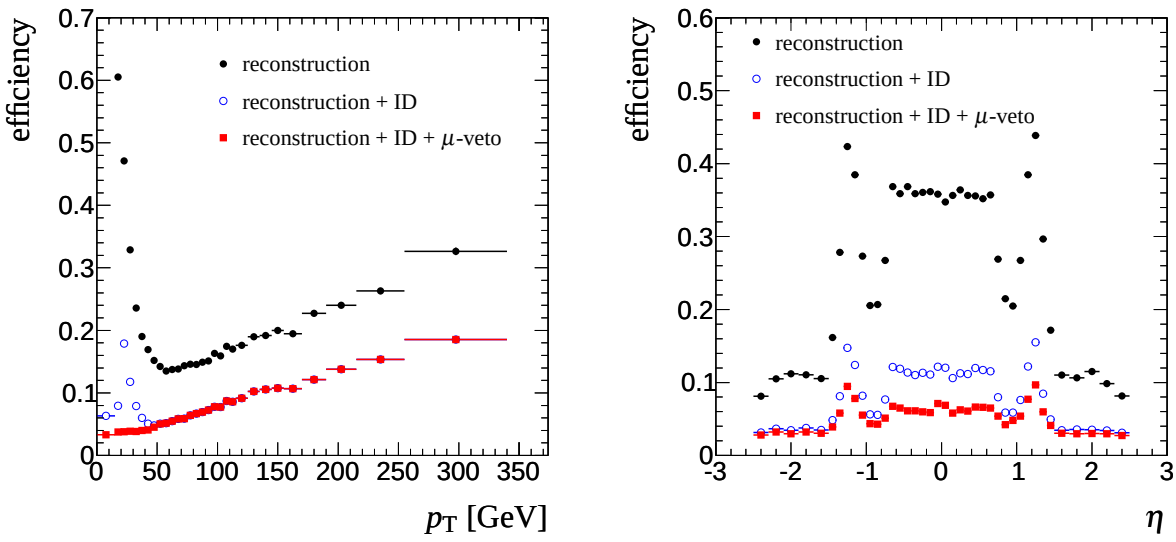
Figure 7.6.: Properties of τ_{had} candidates. Left: p_T distribution. Right: Number of reconstructed tracks associated with τ_{had} candidates.



(a) efficiency vs p_T

(b) efficiency vs η

Figure 7.7.: Efficiency to reconstruct and identify electrons from $Z \to ee$ decays as τ_{had} candidate.



(a) efficiency vs p_T

(b) efficiency vs η

Figure 7.8.: Efficiency to reconstruct and identify muons from $Z \to \mu\mu$ decays as τ_{had} candidate.

The τ_{had} identification itself is implemented by a cut on the log-likelihood discriminant (LLH). By varying the cut, the efficiency can be selected over a wide range. Accessible working points are shown in Fig. 7.9. An estimate of the signal efficiency and the background rejection based on data from 2010 can be found in [64][1]. I developed a procedure to optimise the cut on the samples with low statistics, which is documented in [65,66]. The LLH threshold is optimised such that it maximises $S/\sqrt{B}$ in a sample extracted with a subset of the selection criteria.[2] Contrary to the documentation, for this thesis the threshold is optimised for 1-prong and 3-prong candidates separately. LLH $> 3(9)$ for 1(3)-prong candidates is used in this analysis. Fig. 7.10 shows $S/\sqrt{B}$ as a function of the LLH threshold. Since the peaks are relatively broad, choosing a threshold not too far from the optimal value does not have a large impact on the final signal significance. The resulting τ_{had} identification efficiency is shown in Fig. 7.11.

[1] A direct comparison is difficult, because in [64] a definition of signal efficiency is used that mixes reconstruction and identification

[2] S and B are the expected number of signal and background events after cuts.

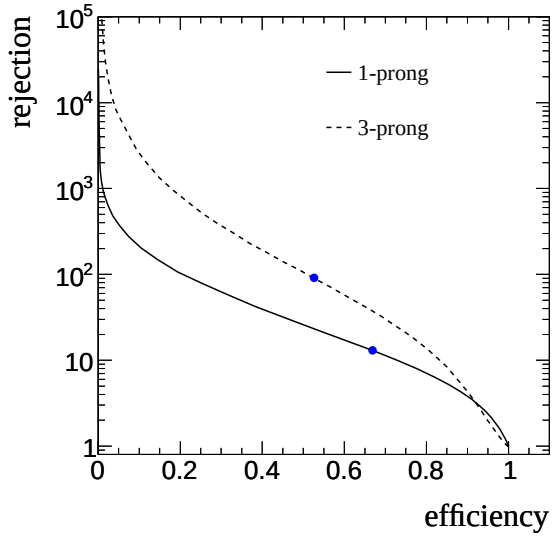


Figure 7.9.: QCD jet rejection vs hadronic τ decay identification efficiency for τ_{had} candidates with $p_{\text{T}} > 30\,\text{GeV}$. Blue dots mark the working point used in this analysis.

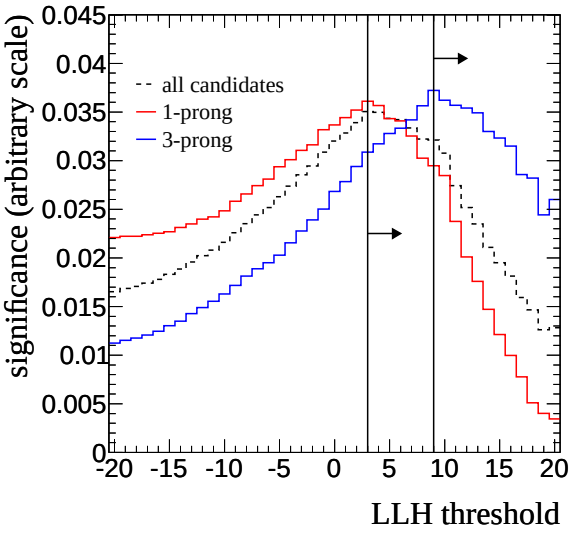
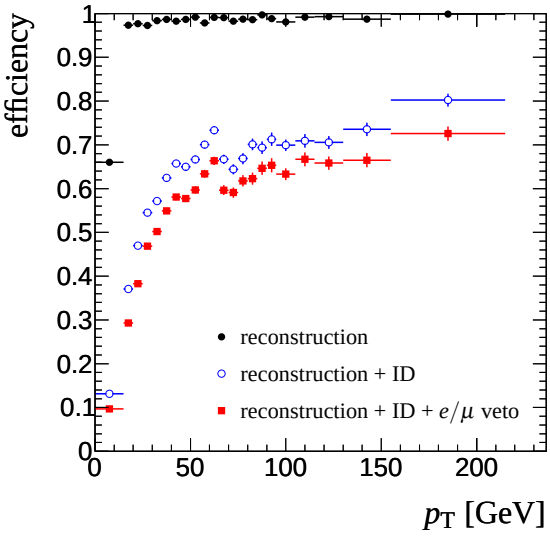
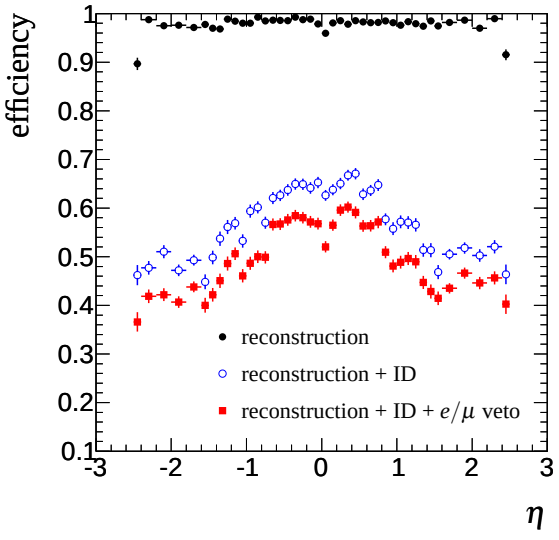


Figure 7.10.: Relative signal significance ($S/\sqrt{B}$) LLH discriminant threshold. The LLH requirements for 1-prong and 3-prong candidates are indicated by arrows.



(a) efficiency vs p_{T}



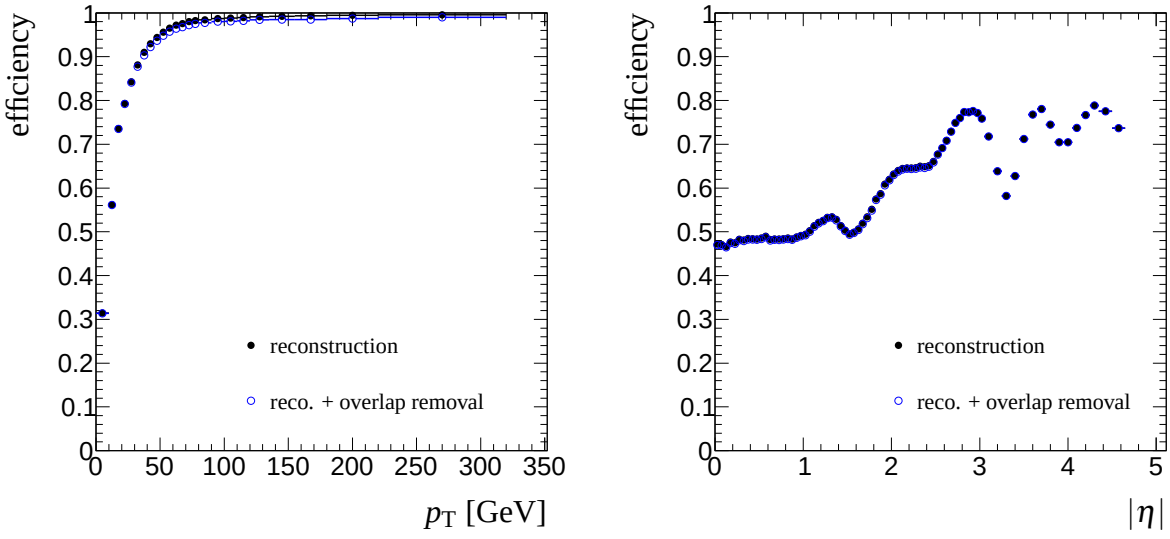
(b) efficiency vs η

Figure 7.11.: τ_{had} reconstruction, identification and electron/muon veto efficiency for candidates from $H \to \tau\tau$ decays.

7.1.4. Jets

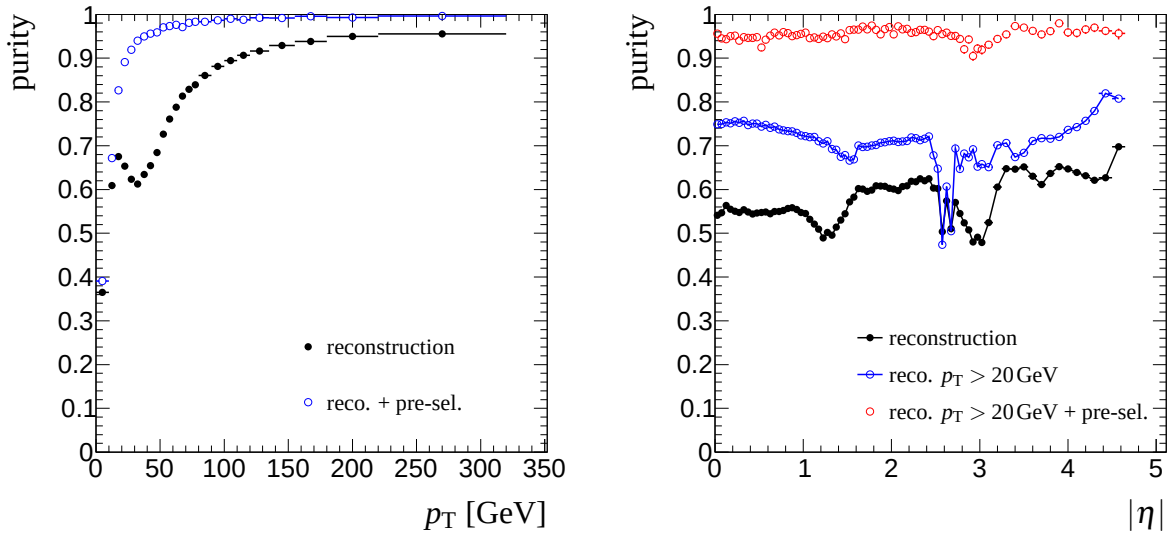
Jets are defined as objects reconstructed by a seeded cone algorithm (Sec. 3.4.4) with a cone size of 0.4. In the context of this analysis the intended meaning of ”jet” is quark- or gluon-induced jet. The cone algorithm reconstructs jets from TopoClusters and thus does not discriminate between the possible sources of a jet. Hence, there is substantial overlap with other reconstructed objects. Jets overlapping with identified electrons, muons or τ_{had} candidates within $\Delta R < 0.4$ are rejected.

This overlap removal has only a minimal impact on the jet reconstruction efficiency, as can be seen in Fig. 7.12. Above about 100 GeV, the jet reconstruction efficiency is approximately one. Dips in the η dependence occur at transitions between the different regions of the ATLAS calorimeter. At $|\eta| \approx 1.5$ the LAr barrel calorimeter ends. $|\eta| \approx 3.3$ marks the limit of the EMEC and HEC. Purity above about 100 GeV is close to one after pre-selection (see Sec. 7.2), as shown in Fig. 7.13. The p_{T} dependence of the purity is due to the inefficiency of τ_{had} selection. Firstly, τ_{had} candidates are rejected below 30 GeV. Secondly, the identification efficiency is only about 60% at medium p_{T}. Thus, in a sample with real τ_{had} candidates a significant amount of jets are actually caused by hadronic τ decays which failed the identification. After the pre-selection at least one τ_{had} candidate is identified and therefore removed from the list of jets. Events with more than one identified τ_{had} candidate are rejected. However, additional real hadronic τ decays which are reconstructed but fail the identification still end up in the list of jets. The purity as a function of η before pre-selection mirrors the reconstruction efficiency. Regions of the detector that generate more noise add energy to real jets, possibly pushing the jet over the threshold. On the other hand noise can generate new jets, especially in the low-p_{T} regime.



(a) efficiency vs p_T **(b)** efficiency vs η; integrated over full p_T range

Figure 7.12.: Jet reconstruction efficiency with and without overlap removal. Jets are extracted from the $Z \to \tau\tau$ sample.



(a) purity vs p_T **(b)** purity vs η

Figure 7.13.: Purity of jet reconstruction after overlap removal. The pre-selection requires one identified τ_{had} and one identified electron or muon. Jets are extracted from the $Z \to \tau\tau$ sample. Lines are meant to guide the eye.

7.2. Pre-Selection

Having defined the objects, the purpose of the pre-selection is to select the final state objects expected from the signal process. Moreover, leptons and τ_{had} candidates are selected such that the ℓh channel is separated from the $\ell\ell$ and hh final states. The pre-selection comprises the following requirements:

- *exactly one* charged **lepton**

  electron with $p_{\text{T}} > 25\,\text{GeV}$ if an electron trigger fired

  muon with $p_{\text{T}} > 20\,\text{GeV}$ if a muon trigger fired

- *exactly one* τ_{had} with $p_{\text{T}} > 30\,\text{GeV}$
- **lepton** and τ_{had} have opposite charge
- *at least two* **jets** with $p_{\text{T}} > 20\,\text{GeV}$

7.3. Selection Requirements

The selection consists of two parts: Jet related cuts exploit the jet topology of VBF. Lepton, τ_{had} and $\not{E}_{\text{T}}$ related cuts make use of the kinematics of the $H \to \tau\tau$ decay.

7.3.1. Lepton, τ_{had} and $\not{E}_{\text{T}}$ Requirements

The final state of the signal process contains three neutrinos. The resulting $\not{E}_{\text{T}}$ distribution can be seen in Fig. 7.14a. Apart from suppressing $Z \to ll$ and part of the di-boson background, the main purpose of this requirement is to be safe from background processes such as QCD multi-jet production, which are not considered in this analysis.

The observables x_ℓ and x_h represent the fraction of the τ momentum carried by its visible decay product (see Sec. 4.6). The physically meaningful range for both is $0 < x < 1$. In fact, inserting a negative value into Eq. 4.9 leads to an unphysical mass. From a mathematical point of view, values greater than one are allowed. And since the x-values only approximately represent the fraction of the τ momentum and taking into account $\not{E}_{\text{T}}$ resolution, values greater than one are to be expected. On the other hand, restricting the range of allowed x-values makes for a narrower $M_{\tau\tau}$ distribution and has better separation power, especially against $Z \to ll$ (cf. Figs. 7.14b and 7.14c).

Restricting $\Delta\phi$ between the lepton and the τ_{had}, again, has two functions: Firstly, if lepton and τ_{had} are exactly back-to-back, the collinear approximation does not work. But even for values close to 180 degrees the result strongly depends on the $\not{E}_{\text{T}}$ resolution. Small variations in $\not{E}_{\text{T}}$ can then cause large variations in $x_{\ell/h}$. Secondly, as can be seen in Fig. 7.14d, a cut on

$\Delta\phi$ can be used to suppress background since the processes with misidentified τ_{had} candidates tend to produce lepton-τ_{had} pairs which are back-to-back.

As described in Chapter 5, a high transverse mass is characteristic of processes where lepton and $\not{E}_{\text{T}}$ come from a W boson decay (or similar topologies). An upper limit on M_{T} is a powerful cut to suppress these backgrounds (Fig. 7.14e).

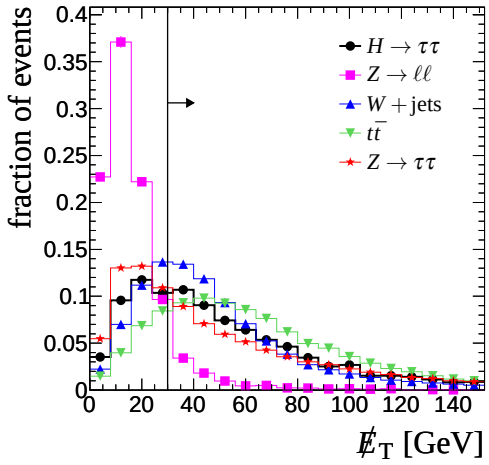
To summarise, the following requirements for τ decay products and $\not{E}_{\text{T}}$ are applied:

- $\not{E}_{\text{T}} > 30\,\text{GeV}$,
- $0 < x_\ell < 0.75$,
- $0 < x_h < 1$,
- $\cos\Delta\phi(\ell, \tau_{\text{had}}) > -0.9$,
- $M_{\text{T}} < 30\,\text{GeV}$.

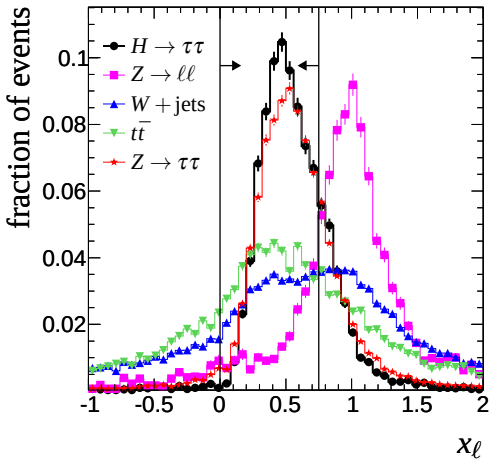
7.3.2. Jet Requirements

The tagging jets topology in VBF has been discussed in Chapter 4. The initial partons each emit a heavy gauge boson and are therefore expected to have considerable p_{T} (Fig. 7.15a). Hence, the two jets with the highest p_{T} are defined as tagging jets. In the following, indices 1 and 2 for jets refer to the tagging jets according to their p_{T}. The angle between tagging jet and beam axis is supposed to be small. The tagging jets should be found in different hemispheres of the detector and have a large separation in η (Figures 7.15b and 7.15c). The high energy of the partons and the large opening angle translate into a large di-jet mass. This is shown in Figure 7.15d. Decay products of the Higgs boson tend to be central and between the tagging jets in η. Fig. 7.15e shows this lepton-τ_{had} centrality. Due to the lack of colour flow between the initial partons, hadronic activity in the central region of the detector is suppressed. A veto on additional jets in the central region of the detector (Fig. 7.15f) is used in order to reject events from many QCD processes. It should be noted that the η distribution of additional jets, and therefore the efficiency of the jet veto, varies significantly between Monte-Carlo generators (cf. [67]). The following jet related requirements are applied:

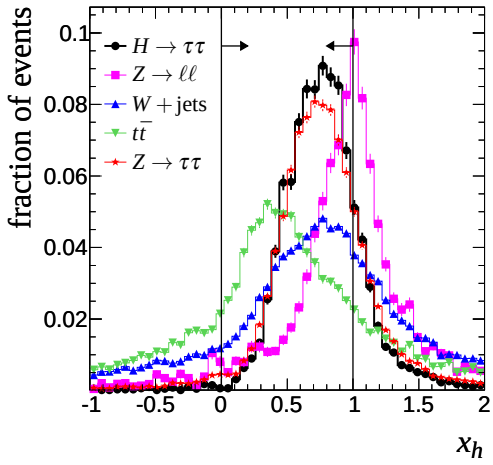
- $\eta_{j_1} \cdot \eta_{j_2} < 0$,
- $p_{T,j_1} > 40\,\text{GeV}$,
- $\Delta\eta_{j_1,j_2} > 4.4$,
- $\min(\eta_{j_1}, \eta_{j_2}) < \min(\eta_\ell, \eta_{\tau_{\text{had}}}) < \max(\eta_\ell, \eta_{\tau_{\text{had}}}) < \max(\eta_{j_1}, \eta_{j_2})$,
- $M_{jj} > 700\,\text{GeV}$,
- reject events with additional jets with $p_{\text{T}} > 20\,\text{GeV}$ in $|\eta| < 3.2$.



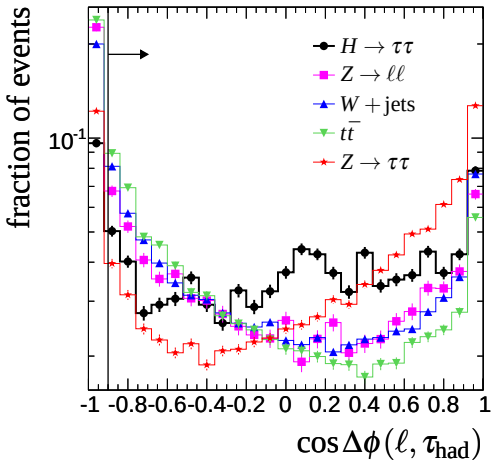
(a) $\not{E}_T$ after pre-selection.



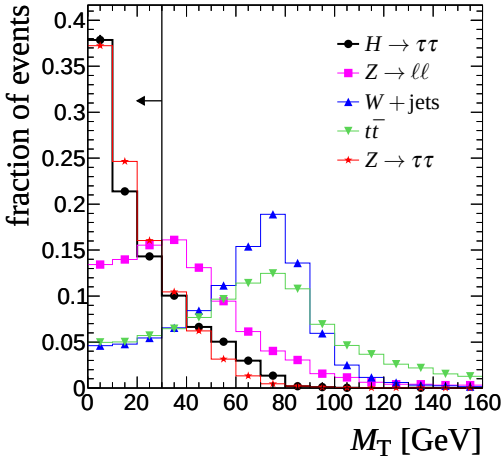
(b) x_ℓ after pre-selection.



(c) x_h after pre-selection.

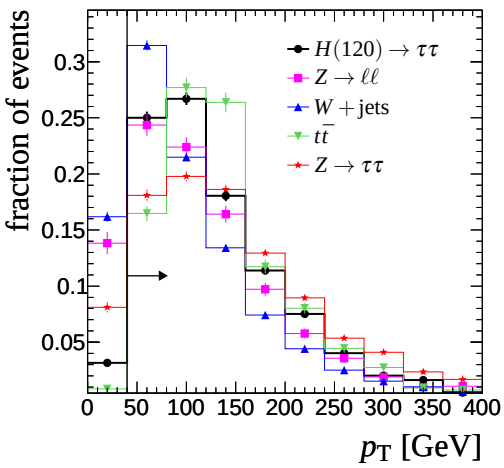


(d) $\cos\Delta\phi(\ell, \tau_{\text{had}})$ after pre-selection.

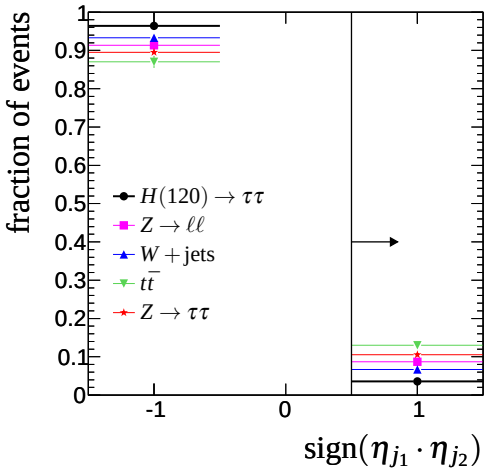


(e) M_T after pre-selection.

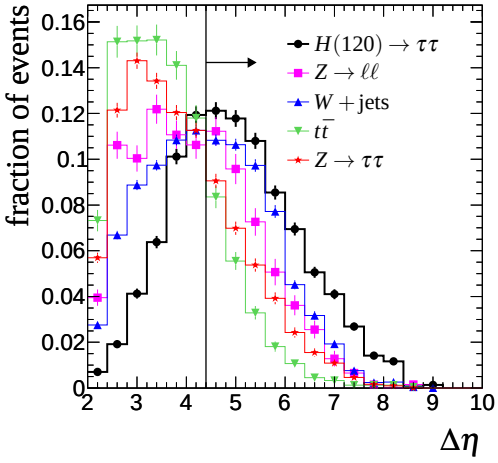
Figure 7.14.: $H \to \tau\tau$ decay related observables used in the cut selection. In 7.14a and 7.14e pre-selection does not include a negative charge product.



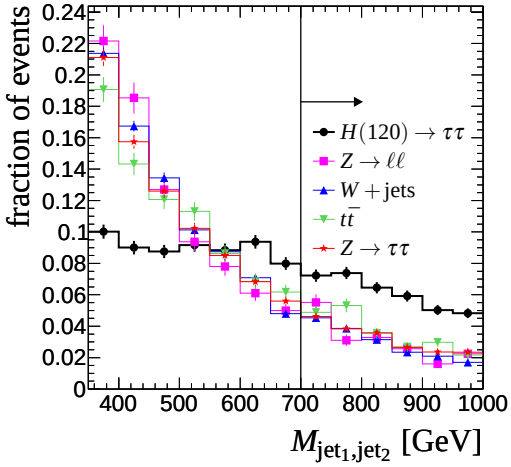
(a) p_T of the leading tagging jet



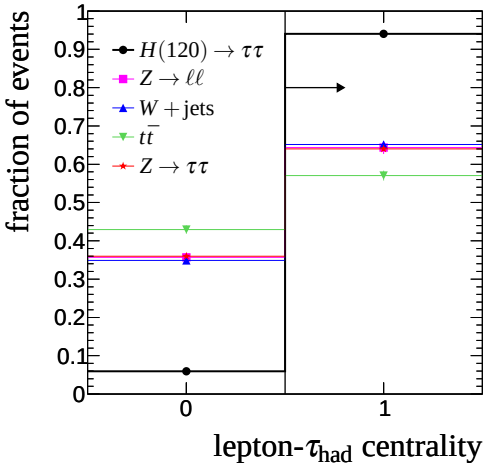
(b) sign$(\eta_{j_1} \cdot \eta_{j_2})$



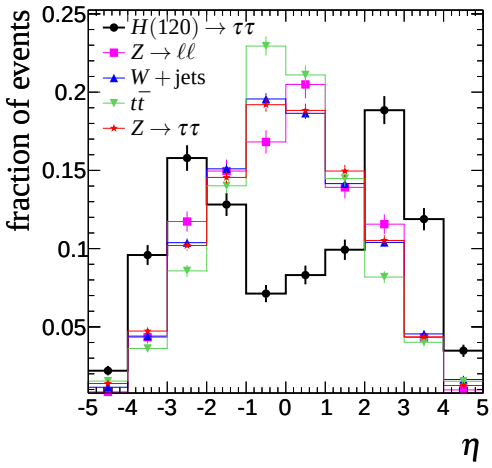
(c) $\Delta\eta_{jj}$ after jet veto.



(d) Di-jet mass after jet veto.



(e) Lepton-τ_{had} centrality



(f) η distribution of additional jets after jet veto.

Figure 7.15.: Jet related observables used in the cut selection.

cut	$H \to \tau\tau$		$Z \to \tau\tau$	
	# events	ε_{rel}	# events	ε_{rel}
total	4260(10)	-	8.06(1) $\times 10^5$	-
filter cut	1147(7)	0.27	9.48(3) $\times 10^4$	0.12
pre-selection	185(3)	0.16	5700(60)	0.06
$\not{E}_{\text{T}}$	131(2)	0.71	3470(50)	0.61
x_ℓ,x_h	95(2)	0.73	2200(40)	0.63
$\cos\Delta\phi(\ell, \tau_{\text{had}})$	93(2)	0.97	2160(40)	0.98
M_{T}	71(2)	0.76	1830(30)	0.85
sign$(\eta_{j1}\eta_{j2})$	69(2)	0.97	1640(30)	0.9
p_{T} (lead. jet)	68(2)	0.99	1630(30)	0.99
$\Delta\eta_{jj}$	43(1)	0.63	360(20)	0.22
centrality	42(1)	0.98	280(10)	0.78
M_{jj}	37(1)	0.89	190(10)	0.68
jet veto	32(1)	0.85	84(7)	0.44

Table 7.1.: Expected number of events for an integrated luminosity of $30\,\text{fb}^{-1}$ after cuts and relative efficiency for the signal and $Z \to \tau\tau$.

Several background samples are filtered on generator level. In order to present comparable numbers in the following tables, an additional filter requirement ("filter cut") that resembles the VBF and lepton filter (see Chapters 4 and 5) is applied to all samples before any other selection criteria:

- no. of jets ≥ 2
- the two hardest (p_{T}) jets: $M_{jj} > 350$ GeV, $\Delta\eta_{jj} > 2.5$
- at least one identified muon or electron

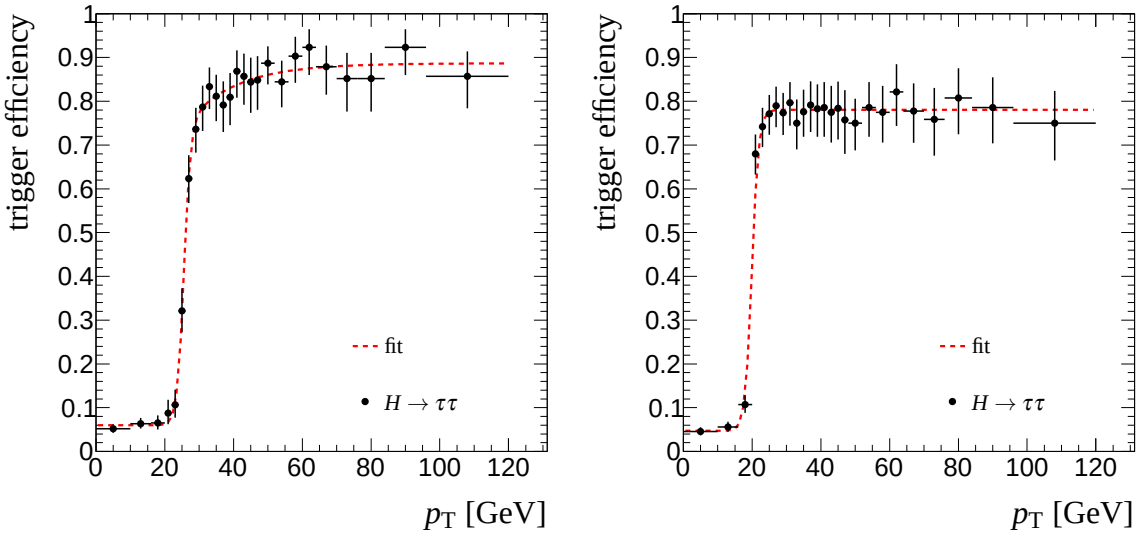
In order to discuss the effects of different trigger configurations, Tables 7.1-7.3 show the number of expected events after the selection omitting the trigger requirement.

cut	$W+$jets		$t\bar{t}$	
	# events	ε_{rel}	# events	ε_{rel}
total	3.464(6) $\times 10^7$	-	1.486(1) $\times 10^7$	-
filter cut	4.87(2) $\times 10^6$	0.14	8.04(3) $\times 10^5$	0.05
pre-selection	1.52(6) $\times 10^4$	0.0	2.23(6) $\times 10^4$	0.03
$\not{E}_{\text{T}}$	1.06(5) $\times 10^4$	0.7	1.92(5) $\times 10^4$	0.86
x_ℓ,x_h	1700(200)	0.16	5700(300)	0.3
$\cos\Delta\phi(\ell,\tau_{\text{had}})$	800(100)	0.49	3600(200)	0.63
M_{T}	190(50)	0.23	900(100)	0.26
sign$(\eta_{j1}\eta_{j2})$	160(50)	0.82	900(100)	0.93
p_{T} (lead. jet)	140(40)	0.89	900(100)	1.0
$\Delta\eta_{jj}$	60(30)	0.43	180(50)	0.21
centrality	30(20)	0.54	120(40)	0.67
M_{jj}	30(20)	1.0	80(30)	0.67
jet veto	30(20)	1.0	0(0)	0.0

Table 7.2.: Expected number of events for an integrated luminosity of $30\,\text{fb}^{-1}$ after cuts and relative efficiency for the major non-resonant background processes.

cut	$Z \to \ell\ell$		di-boson	
	# events	ε_{rel}	# events	ε_{rel}
total	2.503(3) $\times 10^6$	-	5.22(2) $\times 10^6$	-
filter cut	4.68(1) $\times 10^5$	0.19	5.1(2) $\times 10^4$	0.01
pre-selection	3500(100)	0.01	600(300)	0.01
$\not{E}_{\text{T}}$	360(30)	0.1	500(200)	0.8
x_ℓ,x_h	40(10)	0.12	300(200)	0.71
$\cos\Delta\phi(\ell,\tau_{\text{had}})$	26(8)	0.59	200(100)	0.67
M_{T}	11(5)	0.41	20(20)	0.07
$\text{sign}(\eta_{j1}\eta_{j2})$	11(5)	1.0	20(20)	1.0
p_{T} (lead. jet)	11(5)	1.0	20(20)	1.0
$\Delta\eta_{jj}$	4(4)	0.33	0(0)	0.0
centrality	4(4)	1.0	0(0)	-
M_{jj}	0(0)	0.0	0(0)	-
jet veto	0(0)	-	0(0)	-

Table 7.3.: Expected number of events for an integrated luminosity of $30\,\text{fb}^{-1}$ after cuts and relative efficiency for the minor non-resonant background processes.



(a) Electron trigger. **(b)** Muon trigger.

Figure 7.16.: Efficiency of the default triggers for charged leptons from the Higgs boson. Electrons and muons within $|\eta| < 2.5$ are considered.

7.4. Trigger

In the ℓh channel the simplest choice is a single-lepton trigger. Default triggers should therefore be the electron/muon trigger with the lowest threshold that is not pre-scaled[3]. An artificial constraint for this study is the fact that only a very limited trigger menu is available in the simulated samples. `EF e25i medium1` and `EF mu20` are used as default (see Sec. 3.5 for details). The trigger efficiency is defined as number of generator electrons (muons) from a $H \to \tau\tau$ decay, matched to an EF electron (muon) in an event which passes the electron (muon) trigger, over the total number of generator electrons (muons). The *turn-on curve* of the trigger is not a step function with step at the nominal threshold. The efficiency as function of p_T is shown in Fig. 8.2. To fit the data points, a sigmoid function

$$f(p_T) = p_1 + \frac{p_2}{1+\exp(-x+p_3)} \cdot \frac{1}{1-\exp(-p_4 x)} \tag{7.1}$$

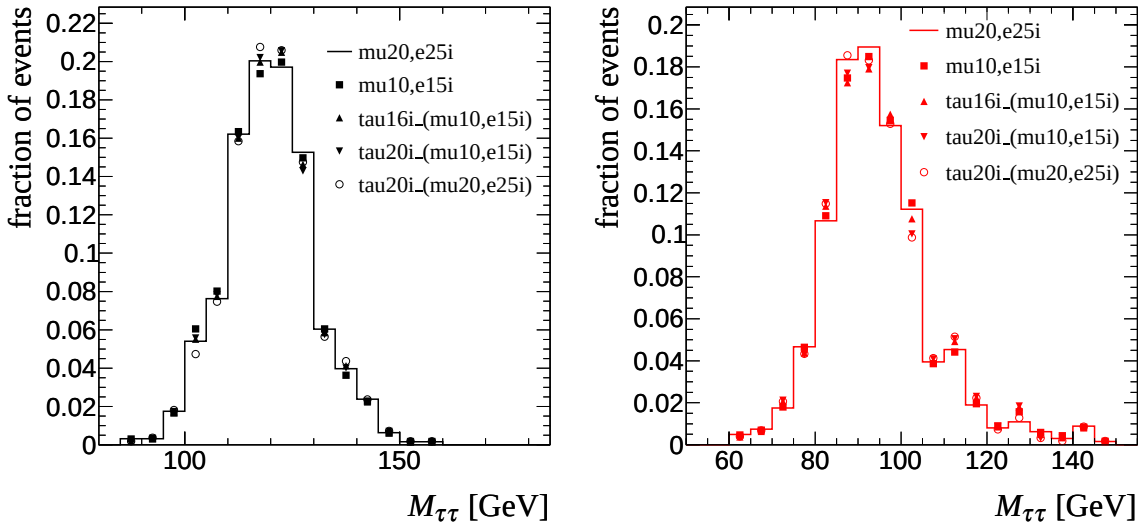
is used. The additional term reflects the asymmetric behaviour of the trigger turn-on which is not correctly described by an error function. For electrons and muons within $|\eta| < 2.5$ the maximum efficiency is 89% and 78% respectively. Both triggers reach 95% of their maximum at higher values than 25(20) GeV: 42 GeV for electrons, 23 GeV for muons.

While single-lepton triggers are the simplest choice, also a combination of lepton- and τ_{had} trigger is of interest for this channel. Such an `AND` trigger obviously does not increase the trig-

[3] Pre-scaling refers to (randomly) selecting only a fraction of events that would pass the trigger in order to reduce the trigger rate.

triggers	$H \to \tau\tau$	$Z \to \tau\tau$
`mu20 | e25i`	27(1)	72(7)
`mu10 | e15i`	28(1)	75(7)
`tau16i & (mu10 | e15i)`	25(1)	60(6)
`tau20i & (mu10 | e15i)`	24(1)	58(6)
`tau20i & (mu20 | e25i)`	23(1)	57(5)

Table 7.4.: Expected number of for an integrated luminosity of $30\,\text{fb}^{-1}$ after the cut selection and different trigger configurations.



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

Figure 7.17.: $M_{\tau\tau}$ shape after cuts for different trigger configurations. $Z \to \tau\tau$ distributions are extracted skipping the last two cuts to prevent large effects due to statistical fluctuations.

ger efficiency if it has the same threshold for the lepton. However, it might provide an option to either lower the threshold or prevent increasing it, in case the default trigger needs to be pre-scaled. Results after cuts given different trigger requirements are shown in Table 7.4. Using a τ_{had} trigger with a threshold of 20 GeV in combination with the default triggers reduces the signal acceptance by 15% ($Z \to \tau\tau$: 21%). Note that object definition and pre-selection remain unchanged, i.e. electrons (muons) are still required to have $p_{\text{T}} > 25(20)\,\text{GeV}$, which explains why the gain in reducing the trigger thresholds turns out to be very small. With a threshold below the offline cut of $p_{\text{T}} > 30\,\text{GeV}$, τ_{had} triggers do not significantly influence the $M_{\tau\tau}$ shape (cf. Fig. 7.17).

7.5. Background Estimation from Simulation

Except for $Z \to \tau\tau$, for which an adequate amount of MC events is available, tables 7.1-7.3 show zero events or a number with very high statistical uncertainty for all background processes. A conservative approach is to calculate the upper limit of a poisson confidence interval for zero observed events and scale it with the highest event weight of the sample. The number obtained in this way is, however, not useful in order to give an estimate of the future sensitivity for a channel.

7.5.1. W+jets Production

In the case of $W+$jets, τ_{had} ID factorisation as described in Sec. 6.3.1 can be used to improve the estimate. The result can be seen in Table 7.5. The expected number of events obtained this way is 14. As explained in Sec. 6.3.1, the distribution of partons from which jets originate is difficult to model. To correct for the remaining discrepancy, the efficiency of the charge product requirement is scaled by 0.96.

7.5.2. Top Quark Pair Production

$t\bar{t}$ production is separated into two samples in order to improve the estimate. τ_{had} ID factorisation can be used with events in which the τ_{had} candidate is a misidentified QCD jet. For true τ_{had} candidates, cut factorisation as proposed in Sec. 6.3.2 is used.

Table 7.6 shows the number of events after selection for true and misidentified τ_{had} candidates as well as the result obtained from τ_{had} ID factorisation. To correct for the discrepancy in the charge product, the the τ_{had} ID factorised sample is scaled by 1.21. The procedure yields 5 expected events after all cuts.

A cut factorisation is considered valid, if $0.9 <$ f.o.m. < 1.1. I.e. for both the loose cut selection for $t\bar{t}$ and the reference selection for $Z \to \tau\tau$, the product of factorised efficiencies have to be within 10% of the result obtained by sequentially applying all cuts. Out of 4139 possible factorisations, five fulfil this criterion. The final expectation ranges between 7 and 12 events, with the factorisation which gives the best figure of merit at 9. This is a plausible result as the ratio between the true τ_{had} sample and the (factorised) misidentified τ_{had} sample is roughly 2:1 for the last one or two cuts before the jet veto. The total estimated number of expected $t\bar{t}$ events after all cuts is thus 14.

cut	$W+$jets		$W+$jets, factorised	
	# events	ε_{rel}	# events	ε_{rel}
total	$3.464(6)\times10^7$	-	$3.464(6)\times10^7$	-
trigger	$1.803(4)\times10^7$	0.52	$1.803(4)\times10^7$	0.52
trigger lepton	$1.622(4)\times10^7$	0.9	$1.622(4)\times10^7$	0.9
no. of leptons	$1.620(4)\times10^7$	1.0	$1.620(4)\times10^7$	1.0
no. of τ_{had}	$1.34(2)\times10^5$	0.01	$1.356(6)\times10^5$	0.01
no. of jets	$8.7(2)\times10^4$	0.65	$8.91(4)\times10^4$	0.66
charge product	$6.5(1)\times10^4$	0.74	$6.33(4)\times10^4$	0.71
$\not{E}_{\text{T}}$	$4.3(1)\times10^4$	0.67	$4.35(3)\times10^4$	0.69
x_ℓ	$1.39(5)\times10^4$	0.32	$1.42(2)\times10^4$	0.33
x_h	7300(400)	0.52	7300(100)	0.52
$\cos\Delta\phi(\ell,\tau_{\text{had}})$	4100(300)	0.57	3920(80)	0.54
M_{T}	1500(200)	0.36	1330(40)	0.34
$\text{sign}(\eta_{j1}\eta_{j2})$	600(100)	0.42	650(30)	0.49
p_{T} (lead. jet)	600(100)	0.97	630(30)	0.97
$\Delta\eta_{jj}$	70(30)	0.11	70(10)	0.11
ℓ-τ_{had} centrality	20(20)	0.37	49(8)	0.69
M_{jj}	20(20)	1.0	28(7)	0.58
jet veto	20(20)	1.0	14(5)	0.51

Table 7.5.: Expected number of events after cuts for an integrated luminosity of $30\,\text{fb}^{-1}$ and relative efficiency for $W\to e/\mu/\tau\nu$ + jets, with and without τ_{had} ID factorisation. Errors are purely statistical and do not reflect additional uncertainties due to factorisation.

cut	$t\bar{t}$ (true τ_{had})		$t\bar{t}$ (fake τ_{had})		fact. $t\bar{t}$ (fake τ_{had})	
	# events	ε_{rel}	# events	ε_{rel}	# events	ε_{rel}
total	$7.51(3)\times10^5$	-	$1.411(1)\times10^7$	-	$1.411(1)\times10^7$	-
trigger	$1.37(1)\times10^5$	0.18	$6.916(1)\times10^6$	0.49	$6.916(1)\times10^6$	0.49
trigger lepton	$1.01(1)\times10^5$	0.74	$6.177(9)\times10^6$	0.89	$6.177(9)\times10^6$	0.89
no. of leptons	$1.01(1)\times10^5$	0.99	$5.434(9)\times10^6$	0.88	$5.434(9)\times10^6$	0.88
no. of τ_{had}	$1.00(1)\times10^5$	1.0	$9.8(1)\times10^4$	0.02	$8.92(3)\times10^4$	0.02
no. of jets	$9.3(1)\times10^4$	0.92	$9.5(1)\times10^4$	0.97	$8.65(3)\times10^4$	0.97
charge product	$9.2(1)\times10^4$	0.99	$7.5(1)\times10^4$	0.79	$7.47(3)\times10^4$	0.86
$\not{E}_{\text{T}}$	$7.8(1)\times10^4$	0.85	$5.80(9)\times10^4$	0.77	$5.79(3)\times10^4$	0.77
x_ℓ	$3.04(7)\times10^4$	0.39	$2.45(6)\times10^4$	0.42	$2.48(2)\times10^4$	0.43
x_h	$2.17(5)\times10^4$	0.71	$1.31(4)\times10^4$	0.54	$1.30(1)\times10^4$	0.52
$\cos\Delta\phi(\ell,\tau_{\text{had}})$	$1.14(4)\times10^4$	0.53	7300(300)	0.56	7380(90)	0.57
M_{T}	3000(200)	0.26	2000(200)	0.27	2370(50)	0.32
$\text{sign}(\eta_{j1}\eta_{j2})$	1500(100)	0.51	900(100)	0.43	950(30)	0.4
p_{T} (lead. jet)	1400(100)	0.96	800(100)	0.99	940(30)	0.99
$\Delta\eta_{jj}$	130(40)	0.09	30(20)	0.04	59(8)	0.06
ℓ-τ_{had} centrality	100(30)	0.77	10(10)	0.33	45(7)	0.76
M_{jj}	80(30)	0.8	10(10)	1.0	43(7)	0.95
jet veto	0(0)	0.0	0(0)	0.0	5(2)	0.11

Table 7.6.: Expected number of events after cuts for an integrated luminosity of $30\,\text{fb}^{-1}$ and relative efficiency for $t\bar{t}$ production. The first sample, $t\bar{t}$ (true τ_{had}) contains $t\bar{t}$ events with true τ_{had} candidates. The second sample contains $t\bar{t}$ events with misidentified QCD jets as τ_{had} candidates. The third sample is the same as the second but with τ_{had} ID factorisation applied. Errors are purely statistical and do not reflect additional uncertainties due to factorisation.

7.5.3. $Z \to \ell\ell$ and Di-boson

In the $Z \to \ell\ell$ sample, 12 events are expected before applying the jet requirements. The jet kinematics are probably similar to $Z \to \tau\tau$ or $W +$ jets. They cannot be entirely equal to $Z \to \tau\tau$ since a significant number of τ_{had} candidates does not come from leptons from the Z boson but are QCD jets. The jet requirements have an acceptance of 0.045 for $Z \to \tau\tau$ events (note that this sample includes EW diagrams). The acceptance is even lower for $W +$ jets events. Taking the larger acceptance yields an expectation of below one event. Hence, $Z \to \ell\ell$ can be safely neglected.

There is no good estimation for di-boson production. Only a handful of actual MC events survives the pre-selection; i.e. there is no room for cut factorisation. The expected 600 events after pre-selection are comprised of two thirds WW events and one third WZ events. The former is most similar to $t\bar{t}$. This is estimated to have an acceptance after pre-selection of about 10^{-4}. Assuming the same acceptance for WW, the result is a comfortable factor of 25 below one event. The actual expectation value might be higher since the jet multiplicity in the WW sample is lower than in the $t\bar{t}$ sample. This will lead to less events being rejected by the jet veto. In the WZ sample, most of the τ_{had} candidates come from the Z boson. The corresponding lepton candidate comes from either the W or Z boson. A worst case scenario would be that WZ is entirely $Z \to \tau\tau$-like. In this case, the expected number of events after cuts would be roughly 5. However, an estimate in [1] using a cut factorisation results in a suppression of about 2×10^{-8} for the combined sample. With this factor the total expectation is 0.1 events. To summarise, di-boson production can be neglected as background.

7.5.4. QCD multi-jets

Given the limitations mentioned in Sec. 5.6, an estimate of the multi-jet background can only be carried out on an approximate level. The acceptance of the cut selection is shown in Table 7.7. The second set of numbers was obtained using τ_{had} ID factorisation. Although this method works well for $W +$ jets and $t\bar{t}$, the deviations in the multi-jets sample are large. This sample consists of many subsamples with low statistical power (Sec. 5.6). Many of the subsamples do not survive the pre-selection. Hence, the table probably underestimates the true cut acceptance. τ_{had} ID factorisation applied to a di-jet sample of only 25,000 events but filtered to contain a lepton with $p_{\text{T}} > 15\,\text{GeV}$ on generator level agrees with the factorised multi-jets sample for the last three non-vanishing steps (up to $p_{\text{T},j_1} > 40\,\text{GeV}$). Taking into account that no trigger is included and that the jet veto has an acceptance of not more than 50% for all other background samples, the result is about two orders of magnitude below one. A final answer has to come directly from data. A method to measure multi-jets as background to $H \to \tau\tau$ searches from data independent of the $W +$ jets estimate was recently proposed and demonstrated in [68].

cut		QCD		QCD, factorised
	# events	ε_{rel}	# events	ε_{rel}
total	5.260(4) $\times 10^{11}$	-	5.260(4) $\times 10^{11}$	-
trigger	5.260(4) $\times 10^{11}$	1.0	5.260(4) $\times 10^{11}$	1.0
trigger lepton	8.6(5) $\times 10^{7}$	0.0	8.6(5) $\times 10^{7}$	0.0
no. of leptons	8.6(5) $\times 10^{7}$	1.0	8.6(5) $\times 10^{7}$	1.0
no. of τ_{had}	6.0(4) $\times 10^{4}$	0.0	2.1(5) $\times 10^{5}$	0.0
no. of jets	2600(500)	0.04	1.8(6) $\times 10^{4}$	0.09
charge product	1700(400)	0.68	5000(2000)	0.3
$\not{E}_{\text{T}}$	500(200)	0.29	1000(600)	0.18
x_ℓ	200(100)	0.4	130(30)	0.13
x_h	0(0)	0.0	50(20)	0.37
$\cos\Delta\phi(\ell,\tau_{\text{had}})$	0(0)	-	28(9)	0.59
M_{T}	0(0)	-	11(5)	0.38
$\text{sign}(\eta_{j1}\eta_{j2})$	0(0)	-	5(3)	0.46
p_{T} (lead. jet)	0(0)	-	5(3)	1.0
$\Delta\eta_{jj}$	0(0)	-	0.8(5)	0.16
ℓ-τ_{had} centrality	0(0)	-	0.4(3)	0.54
M_{jj}	0(0)	-	0.2(2)	0.5
jet veto	0(0)	-	0(0)	0.0

Table 7.7.: Expected number of events after cuts for an integrated luminosity of $30\,\text{fb}^{-1}$ and relative efficiency for QCD mulit-jet production. The trigger is not included. Errors are purely statistical and do not reflect additional uncertainties due to factorisation.

	$H \to \tau\tau$	$Z \to \tau\tau$	$W+\text{jets}$	$t\bar{t}$	$Z \to \ell\ell$	di-boson
after cuts	27(1)	72(7)	14(5)	14(4)	0	0

Table 7.8.: Expected number of events after the selection for an integrated luminosity of $30\,\text{fb}^{-1}$.

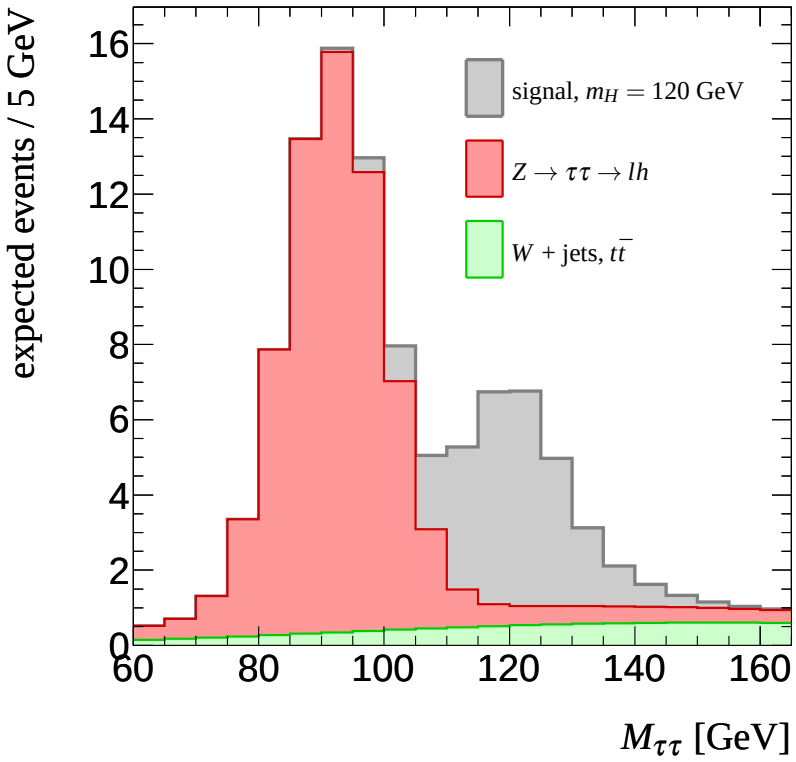


Figure 7.18.: $M_{\tau\tau}$ distributions for an integrated luminosity of $30\,\text{fb}^{-1}$. The histograms are taken from parametrisations discussed in Sec. 10.2

7.6. Results

The expected event yield after the full selection is summarised in Table 7.8. The final signal to background ratio is $S/B = 0.24$. The $M_{\tau\tau}$ distribution is shown in Fig. 7.18. Histograms are smoothed using a parametrisation as described in Chapter 10. Assuming a Higgs boson mass of 120 GeV, the signal peak is well separated from the Z peak. $W+\text{jets}$ and $t\bar{t}$ production together sum up to 35% of the expected background. The reconstructed collinear mass, however, is spread over a larger mass range. Thus, both processes contribute less to the signal region than the pure number of events suggests.

8

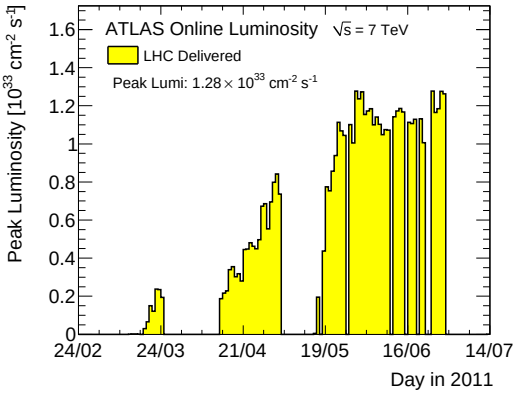
The Influence of Pile-Up

An event, as it is measured by a detector at a hadron collider, contains more reaction products than just from the collision of of two elementary particles. The previous chapter is based on the simulation of single proton-proton interactions. This approximation is valid only in the limit of small luminosity and large times between collisions. For higher luminosities and smaller bunch spacings, pile-up, as described in Chapter 3, cannot be neglected. In-time pile-up refers to additional proton-proton interactions taking place during the same bunch crossing. The average number of interactions per bunch crossing scales linearly with luminosity and bunch spacing. Assuming 23 simultaneous interactions at $\mathscr{L} = 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and a bunch spacing of 25 ns (cf. Chapter 3), $\mathscr{L} = 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and 75 ns leads to 6.9 interactions on average. The latter set of parameters is simulated in this study. A snapshot of the interaction parameters at ATLAS during data taking in 2011, Figure 8.1, shows an already similar pile-up configuration. Out-of-time pile-up is simulated for ± 11 bunch crossings or ± 825 ns.

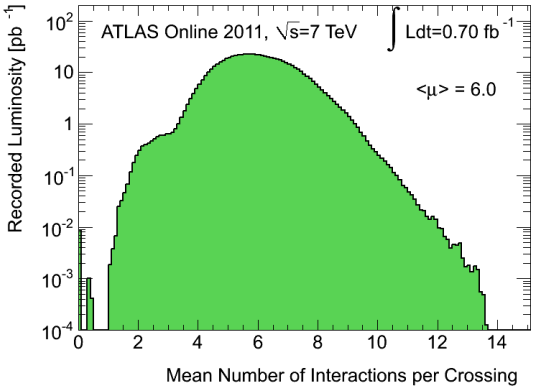
Table 8.1 shows the acceptance of the selection applied to the signal sample with and without pile-up. The first difference to note is a drop in the total acceptance of about 50%. A closer look reveals three items of the selection with a major deviation of the relative acceptance: The Trigger, the τ_{had} identification and the jet veto. These are discussed in the following.

8.1. Trigger

Already at the trigger level about 10% less events, with respect to the no-pile-up scenario, pass the selection (see Table 8.1). Fig. 8.2 shows the turn-on behaviour of electron and muon trigger. The threshold behaviour of both turn-on curves is almost unchanged. However, the maximum efficiency reached by both triggers is decreased: `EF e25i medium1` drops from 88% to 82%, `EF mu20` from 78% to 68%. As discussed in Sec. 3.5, ATLAS has a three-



(a) Maximum instantaneous luminosity per day

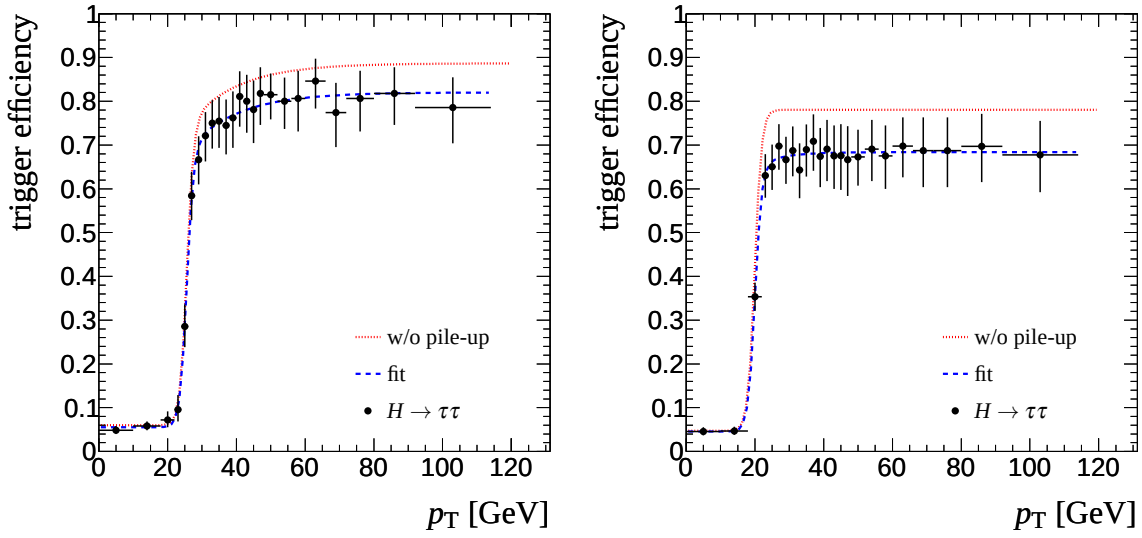


(b) Maximum number of events per beam crossing per day

Figure 8.1.: Interaction parameters at ATLAS [69] from during data taking in 2011.

cut	$H \to \tau\tau \to \ell h$ (pile-up)		$H \to \tau\tau \to \ell h$	
total	3100(10)	-	3100(10)	-
trigger	1007(7)	0.33	1133(7)	0.37
trigger lepton	874(6)	0.87	1016(7)	0.9
no. of leptons	868(6)	0.99	1009(7)	0.99
no. of τ_{had}	189(3)	0.22	299(4)	0.3
no. of jets	156(3)	0.82	231(3)	0.77
charge product	153(3)	0.98	226(3)	0.98
$\not{E}_{\text{T}}$	104(2)	0.68	152(3)	0.67
x_ℓ	86(2)	0.82	131(2)	0.86
x_h	66(2)	0.77	108(2)	0.83
$\cos\Delta\phi(\ell, \tau_{\text{had}})$	63(2)	0.95	105(2)	0.97
M_{T}	45(1)	0.72	81(2)	0.77
$\text{sign}(\eta_{j1}\eta_{j2})$	38(1)	0.83	68(2)	0.84
p_{T} (lead. jet)	37(1)	0.99	67(2)	0.99
$\Delta\eta_{jj}$	21(1)	0.57	37(1)	0.55
ℓ-τ_{had} centrality	20.7(9)	0.98	37(1)	0.98
M_{jj}	18.6(9)	0.9	32(1)	0.87
jet veto	13.2(8)	0.71	27(1)	0.84

Table 8.1.: Comparison of cut acceptances in the signal sample with and without. Samples are filtered on generator level to contain only the ℓh channel.



(a) Electron trigger.

(b) Muon trigger.

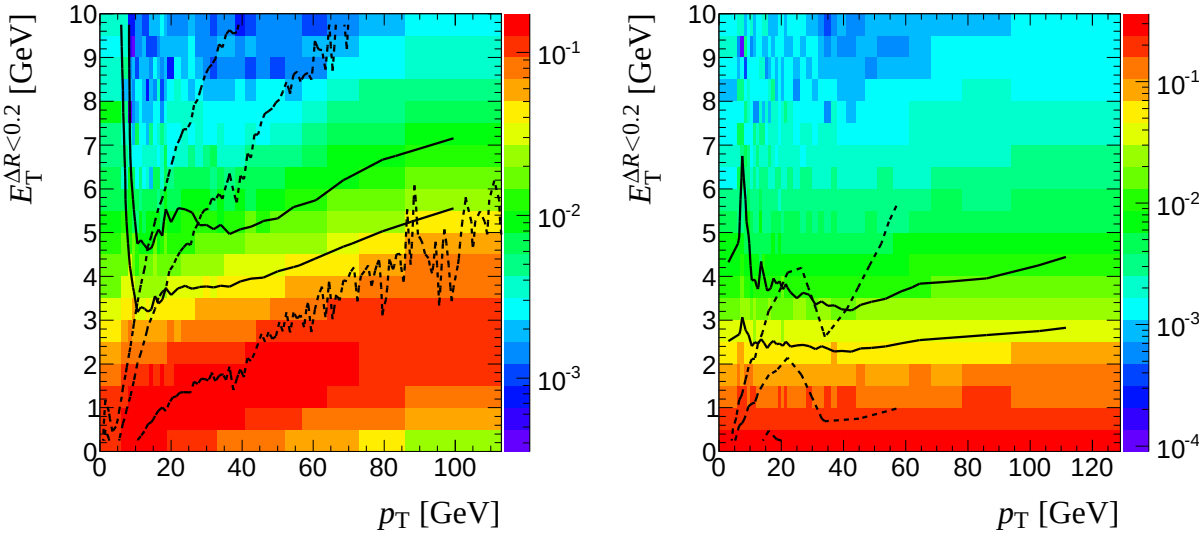
Figure 8.2.: Efficiency of the default triggers, estimated as described in Sec 7.4. Electrons and muons within $|\eta| < 2.5$ are considered.

stage trigger system. Table 8.2 shows the efficiency of the different trigger stages. Both L1 triggers are unaffected by pile-up. L1 trigger algorithms are designed to be relatively robust. Given that the `L1_EM23I` trigger imposes an isolation requirement on electron candidates, one might expect the efficiency to deteriorate in the presence of pile-up. The requirement of $E_T < 4\,\text{GeV}$ in the isolation region, however, affects only a small fraction of electrons in the relevant kinematic region (Fig. 8.3).

A large difference in efficiency occurs at L2. At this stage, pattern recognition and hypothesis testing algorithms are used. Also, the L1 muon trigger uses only information from the RPC and TGC chambers on L1. L2 uses additional information from MDT chambers which suffer from out-of-time pile-up due to the long integration time. EF algorithms are, again, unaffected by pile-up. However, candidates that did not pass the L2 trigger due to isolation, additional tracks, or out-of-time pile-up were already discarded, leaving candidates similar to the non-pile-up case.

8.2. τ_{had} identification

The most dramatic change in acceptance due to pile-up is introduced by the τ_{had} identification. With the settings described in Sec. 7.1.3, the identification efficiency decreases by roughly one third. The effect is shown in more detail in Fig. 8.4a. The reconstruction is unaffected over most of the p_T range. Pile-up increases the acceptance of the electron and muon veto: The electron veto is a dedicated electron identification, which will suffer from additional particles

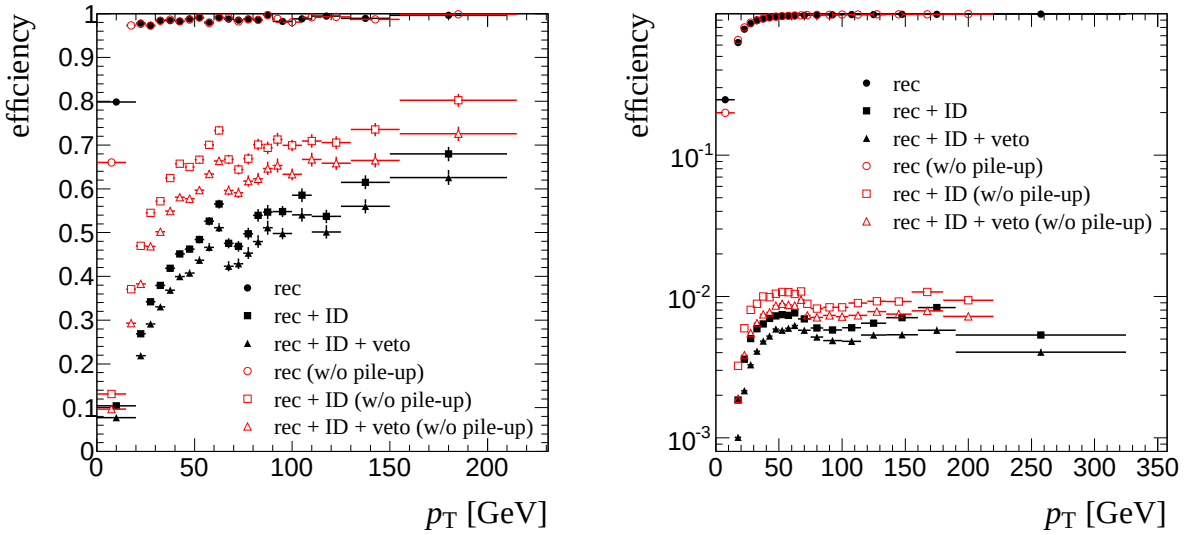


(a) Electrons. **(b)** Muons.

Figure 8.3.: Absolute isolation of electrons and muons from Z/W bosons and τ decays. Bins are normalised such that the integral over each column of p_T is 1. Solid lines reflect the 90% and 95% contours, dashed lines show the 1%,5% and 10% contour lines of candidates from jets.

trigger level	$H \to \tau\tau \to \ell h$				$H \to \tau\tau \to \ell h$ (pile-up)			
	eh		μh		eh		μh	
total	1544(8)	-	1554(8)	-	1542(8)	-	1555(8)	-
level 1	800(6)	0.52	879(6)	0.57	784(6)	0.51	884(6)	0.57
level 2	588(5)	0.74	649(5)	0.74	540(5)	0.69	569(5)	0.64
event filter	507(5)	0.86	611(5)	0.94	464(4)	0.86	529(5)	0.93

Table 8.2.: Number of events expected to pass the electron (muon) trigger chain in the eh (μh) channel and relative efficiencies. The difference in total events with and without pile-up is due to statistical fluctiations in the simulated samples.



(a) Hadronic τ decays from $H \to \tau\tau$.

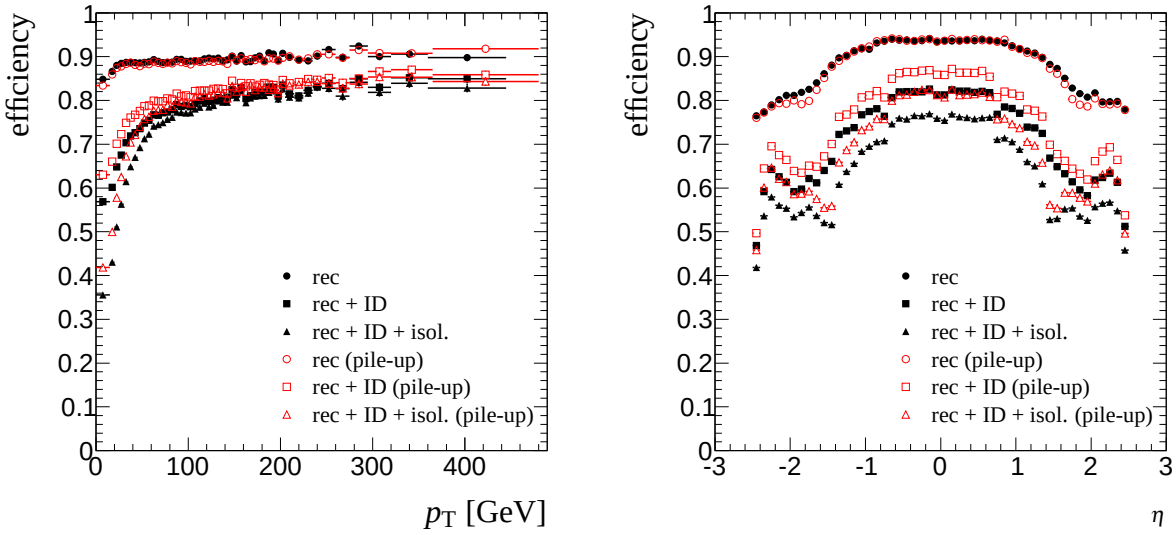
(b) QCD jets from $W +$ jets.

Figure 8.4.: τ_{had} reconstruction and identification efficiency. Candidates are matched to generated hadronic τ decays or jets, respectively.

and energy that are not correlated to an electron. This effect can be seen in Figure 8.5, showing the efficiency of the standard electron reconstruction in the presence of pile-up. The muon veto is a lower limit on energy in the ECAL. Pile-up makes it more likely to pass the threshold. The identification efficiency on the other hand deteriorates, especially in the medium-p_T regime. While the misidentification probability also decreases (Fig. 8.4b), the number of candidates from jets nearly doubles in the 1-prong bin and increases by about 40% in the 3-prong bin (Fig. 8.9).

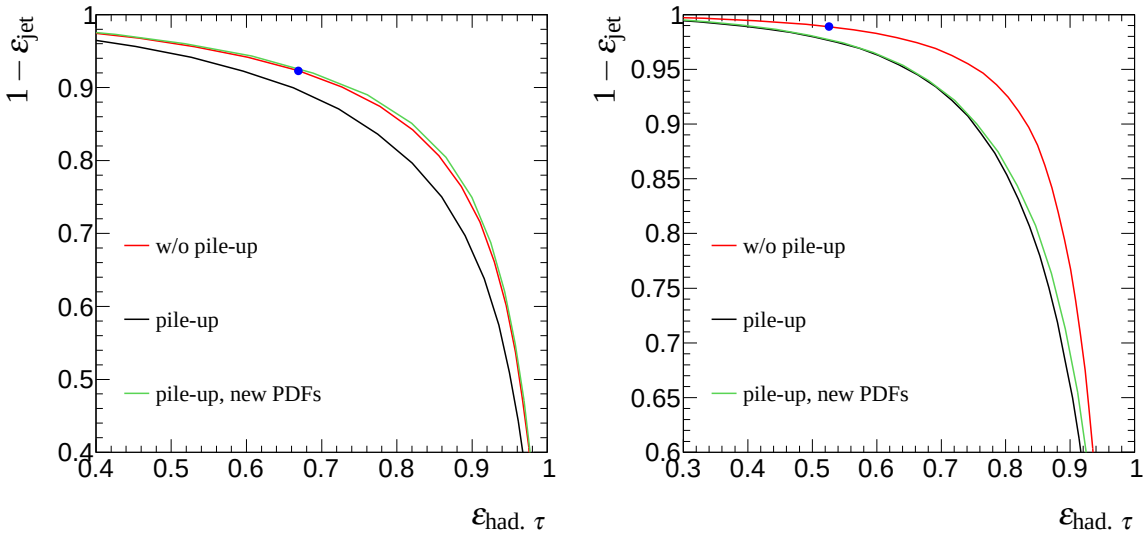
Figures 8.7a and 8.7b depict the distributions of the log-likelihood discriminant for signal and background. In both cases the distributions are shifted towards smaller values. The τ_{had} reconstruction acts on TopoClusters, which have built-in noise suppression. Therefore out-of-time pile-up can be expected to have small effects. In-time pile-up causes the production of additional particles which in turn can add tracks and energy to a τ_{had} candidate.

Tracks of additional charged particles in the vicinity of a candidate reconstructed as n-prong failed the track quality criteria, or else n would have been incremented. Quality criteria in particular include limits on the point of closest approach in z-direction to the primary vertex. Figure 8.8a shows that the mean number of additional low-quality tracks around candidates reconstructed by both algorithms scales approximately linearly with the number of primary vertices. The position of a candidate reconstructed by the calorimeter-seeded algorithm is determined by the barycentre of the cluster. For these candidates, the minimal ΔR to the closest track is usually different from 0. The maximum ΔR depends on whether there are additional tracks close to the candidate. The maximum ΔR of 1-prong calorimeter-seeded candidates is shown in Fig. 8.8b. Pile-up and non-pile-up distributions are a superposition of



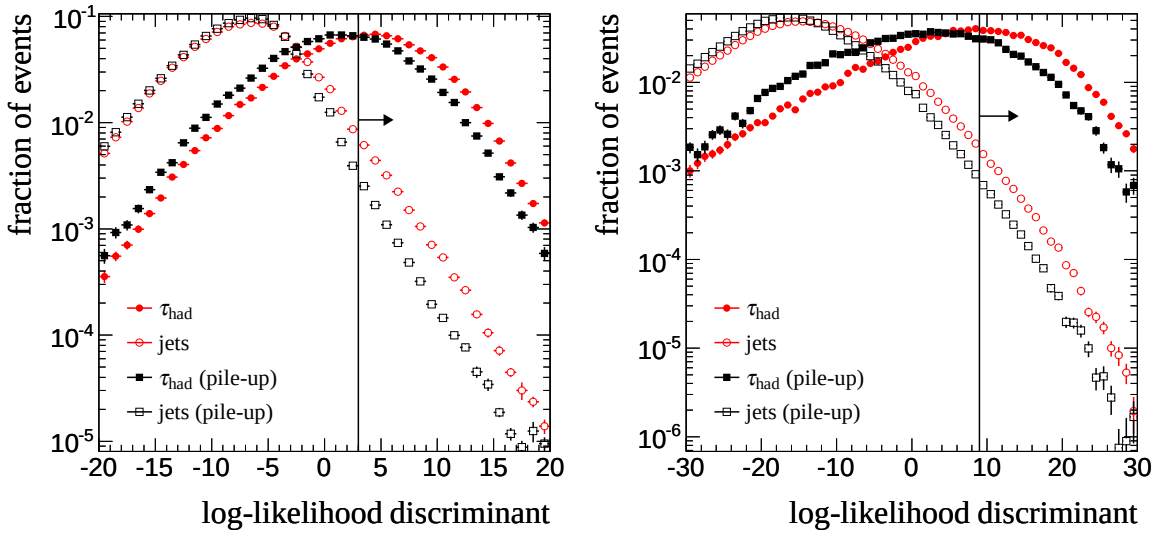
(a) efficiency vs p_T **(b)** efficiency vs η

Figure 8.5.: Reconstruction, identification and isolation efficiency of electrons from Z decays as a function of p_T and η.



(a) 1-prong candidates **(b)** 3-prong candidates

Figure 8.6.: Performance of the τ_{had} ID log-likelihood discriminant for candidates with $p_T > 30\,\text{GeV}$. Blue markers represent the working point for non-pile-up.



(a) 1-prong **(b)** 3-prong

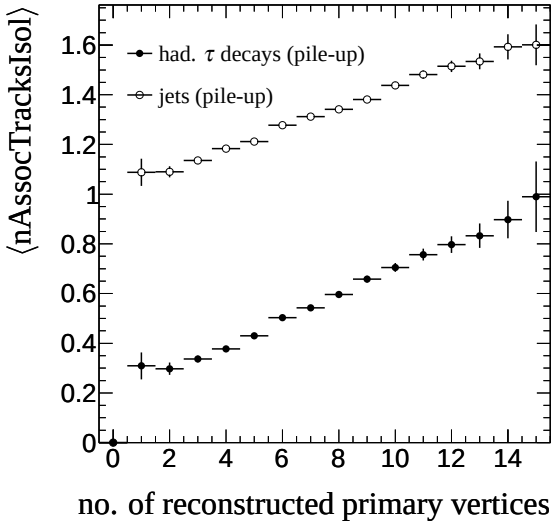
Figure 8.7.: τ_{had} identification log-likelihood discriminant. Vertical lines mark the threshold for candidates in the non-pile-up case.

the minimum ΔR with growing probability to find a track with increasing ΔR. The latter is much more pronounced with pile-up.

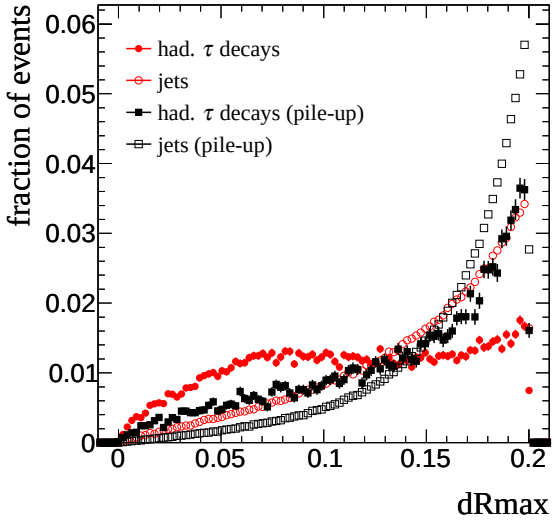
The additional particles deposit energy further off the barycentre. Thus, the τ_{had} candidate appears less well collimated in the calorimeter. This effect can be seen in Figure 8.8c, which shows the EM radius. The centrality fraction, being defined as E_T in $0 < \Delta R < 0.1$ over E_T in $0 < \Delta R < 0.1$, shows the same tendency (Fig. 8.8d).

As a result, real hadronic τ decays look more jet-like in the presence of pile-up. Hence, pile-up leads to a change in the working points accessible by the algorithm. Fig. 8.6 shows the possible values of signal efficiency vs rejection adjustable by a cut on the log-likelihood discriminant. As discussed before, staying with the same cuts leads to a decrease of signal efficiency but also some increase in the rejection of QCD jets. It is possible to adjust the cuts such that the efficiency of the non-pile-up case is recovered. This, however, leads to a severe decrease of rejection, especially for three-prong candidates. Figures 8.7a and 8.7b depict the distributions of the log-likelihood discriminant. For real hadronic τ decays the distributions are clearly shifted to left. This is also true for QCD jets but less visible because the distributions here are broader.

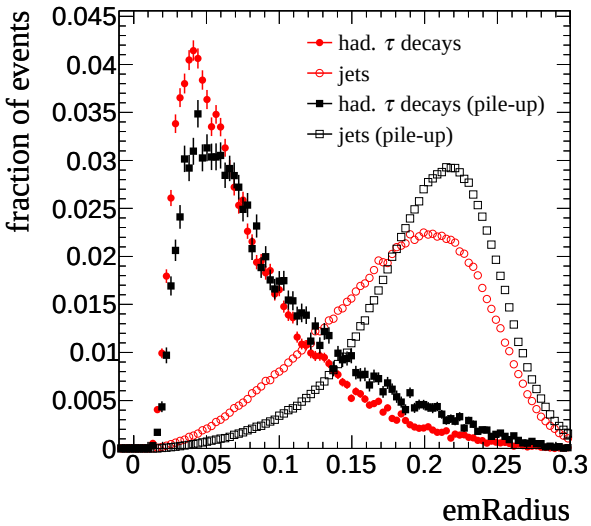
The τ_{had} identification observables indicate that the separation power does not deteriorate as much as the shift in the log-likelihood discriminant suggests. The distributions for signal and background are usually shifted into the same direction. The performance of the LLH-discriminant using an updated set of reference histograms is shown in Figure 8.6. Indeed, for 1-prong candidates the full separation power can be recovered. For 3-prong candidates, however, new reference histograms yield no improvement.



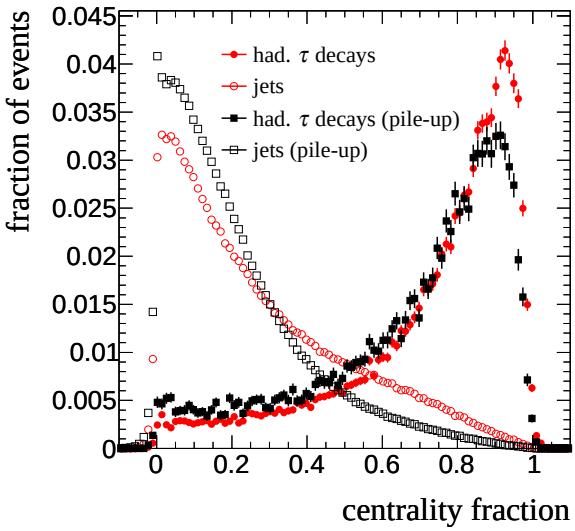
(a) Mean no. of low-quality tracks in $0.1 < \Delta R < 0.2$ around 1-prong τ_{had} candidates rec. by both algorithms as a function of the no. of reconstructed primary vertices.



(b) Maximum ΔR between barycentre and tracks within $\Delta R < 0.2$ of calorimeter-seeded 1-prong candidates.



(c) emRadius of calorimeter-seeded 1-prong candidates.



(d) Centrality fraction of calorimeter-seeded 1-prong candidates.

Figure 8.8.: Impact of pile-up on τ_{had} variables.

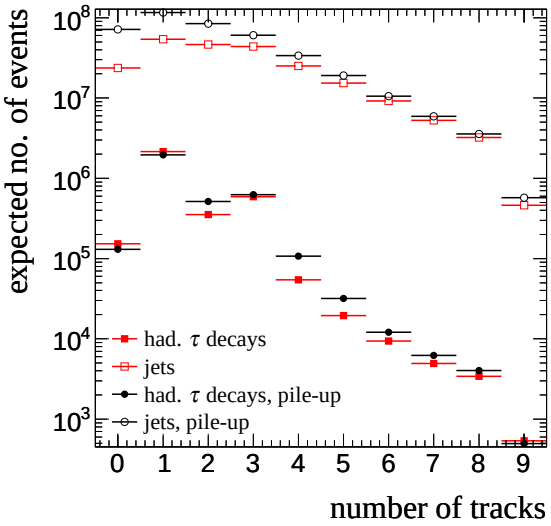


Figure 8.9.: Number of τ_{had} candidates per number-of-tracks bin.

8.3. Jet Veto

Hadronic activity in the central region is suppressed for the signal process. This restriction, however, does not apply to additional proton-proton interactions due to pile-up. Pile-up can trigger the jet veto in two ways: directly, by producing jets in the central region with sufficient p_{T}. And indirectly, by adding energy to jets with lower p_{T} and thus "pushing" them over the threshold. Those can be either other pile-up jets or jets from the hard signal process.

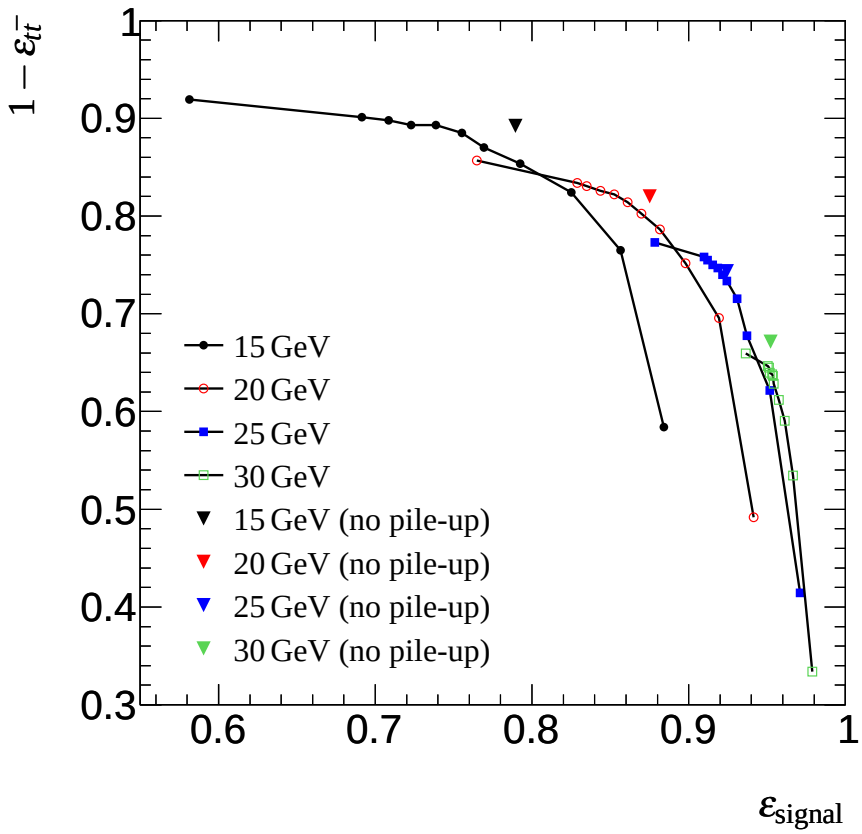
A *jet-vertex association* [70] (JVA) can be used to mitigate this effect within the tracking region ($|\eta| < 2.5$). The purpose of the jet veto is to discriminate against hard processes which do not have a rapidity gap. That means the jet veto should only be triggered by jets originating from the hard process, i.e. jets from the primary vertex of that hard process. A jet can be associated with a reconstructed vertex by identifying the vertex its tracks point to. The primary vertex of the hard scattering is selected as the vertex to which the electron or muon points to. In reality, pile-up does not produce separate jets that can be uniquely assigned to one vertex. Rather, jets can have an admixture of tracks from a number of different vertices. Therefore the procedure is as follows: first, the primary vertex is identified as the vertex to which the electron or muon required for the analysis points to. Second, for each jet the primary vertex p_{T} fraction (PVF) defined as

$$\text{PVF} = \frac{\sum_{\text{tracks in jet from pv}} p_{\text{T}}}{\sum_{\text{tracks in jet}} p_{\text{T}}} \tag{8.1}$$

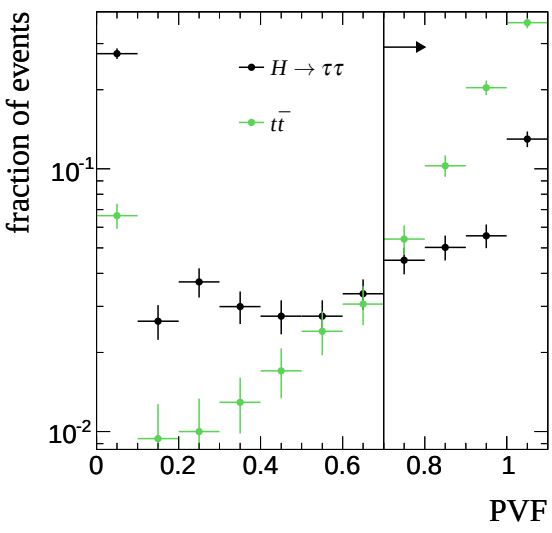
is calculated. In the last step, the veto is triggered only if a jet within $|\eta| < 2.5$ has a PVF greater than a certain threshold.

The PVF and the minimum p_{T} of jets are adjustable parameters of this method. The accessible working points are shown in Fig. 8.10a. Increasing either of the two parameters leads to a higher acceptance of both signal and background because less jets can trigger the jet veto. For large values of the p_{T} threshold, the signal efficiency is virtually constant and only the background rejection is varied by the PVF. In principle, a full optimisation of the parameters with respect to the signal significance would be preferable. But this is not feasible due to the limited size of the Monte-Carlo sample. Instead, the parameters are chosen such that the working point best reflects the non-pile-up case. The p_{T} threshold remains at 20 GeV. For PVF > 0.7 the signal efficiency of the no-pile-up scenario is recovered, while the rejection decreases only by a few percent. This choice coincides with the point where the fraction of $t\bar{t}$ events per bin is higher than the fraction of signal events (cf. Fig. 8.10b). This working point, however, is not unique: e.g. a higher threshold of $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and PVF > 0 (no jet-vertex association) yields a similar result. As mentioned at the beginning of this section, JVA needs tracking information and therefore works only within the narrower tracking region of $|\eta| < 2.5$ instead of the larger $|\eta| < 3.2$ that was used in the no-pile-up scenario. A hybrid solution is technically possible where only jets above a certain PVF within the tracking region and all additional jets in $2.5 < |\eta| < 3.2$ trigger the jet veto. This approach leads to a gain in rejection of $t\bar{t}$ events but comes at the cost of a significantly lower signal acceptance and therefore is not used in the following.

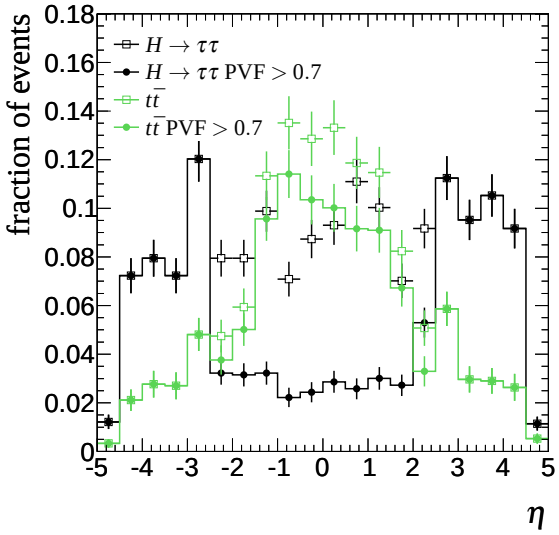
Fig. 8.10c shows the η distribution of jets which trigger the veto using jet-vertex association. Without JVA this distribution is almost flat for the signal process. JVA has the desired effect: In signal events, it rejects most of the jets within the tracking region as pile-up jets that will not trigger the veto. In $t\bar{t}$ events, with many genuine jets from the hard process, the majority of jets survives causing the event to be discarded.



(a) Performance of the jet veto for different p_{T} thresholds with jet-vertex association. Lines connect settings of the PVF threshold between 0.0 (leftmost) and 1.0 (rightmost) with a step size of 0.1. No cuts except for loose jet cuts where applied to the samples.

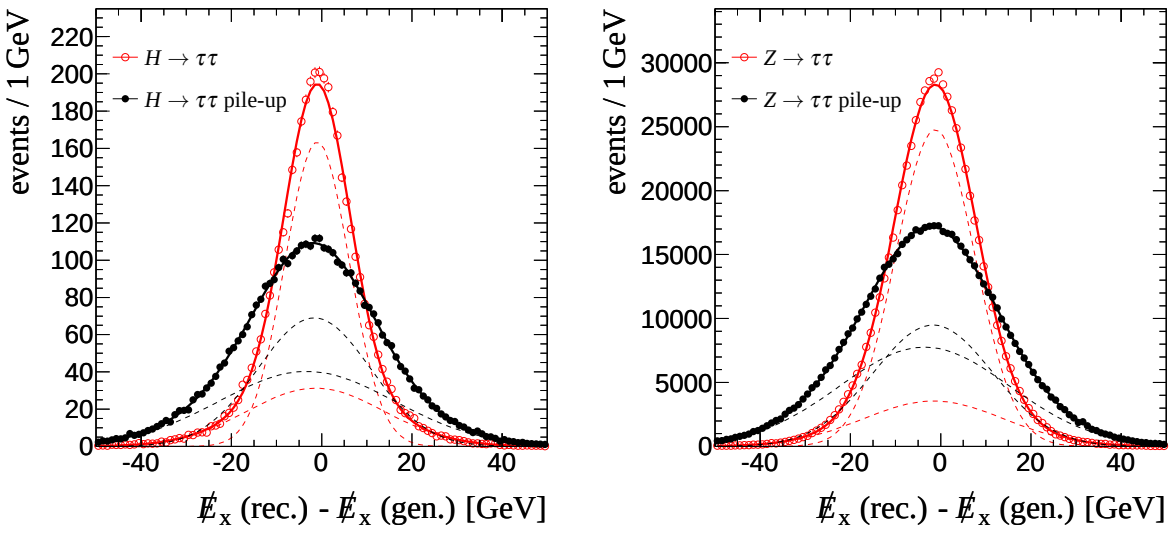


(b) PVF of jets within $|\eta| < 2.5$.



(c) η distribution of additional jets.

Figure 8.10.: Comparison of the jet veto performance for $H \to \tau\tau$ and $t\bar{t}$. Distributions with PVF > 0.7 are normalised to the corresponding distribution without the requirement for $|\eta| > 2.5$.



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

Figure 8.11.: Difference between reconstructed (RefFinal) and generated $\not{E}_x$. No cuts are applied. A double Gaussian function is used to fit the distributions (solid lines). The two Gaussian functions of each composite function are drawn as dashed lines.

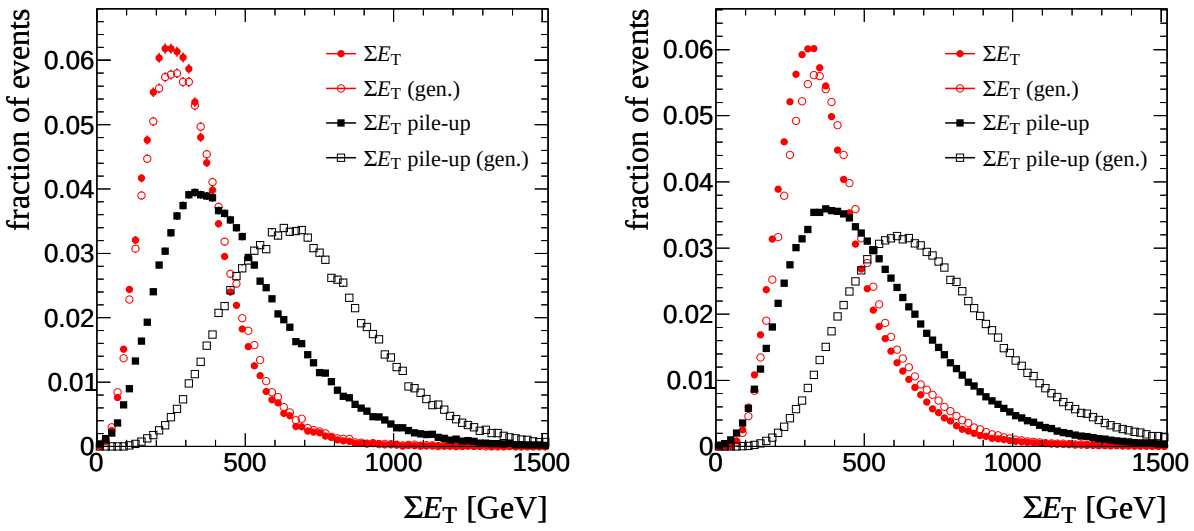
8.4. $\not{E}_T$ Resolution

Although the acceptance of the $\not{E}_T$ requirement is virtually unchanged, pile-up has a significant impact on $\not{E}_T$ resolution. As is apparent from Figure 8.11, the distribution of the difference between the reconstructed and the generated x-component of $\not{E}_T$ ($\not{E}_x$) becomes much broader in the presence of pile-up. In both luminosity scenarios, a single Gaussian function is insufficient to correctly describe the tails. The corresponding fit parameters are shown in Table 8.3. Especially the width of the inner Gaussian increases relative to the non-pile-up case: From (7.1 ± 0.3) GeV to (12 ± 2) GeV in $H \to \tau\tau$, (8.64 ± 0.03) GeV to (13.2 ± 0.3) GeV in $Z \to \tau\tau$. Note that the distributions are not exactly centred around zero but are slightly biased such that the reconstructed $\not{E}_x$ is on average smaller than the generated. This effect is caused by a displacement of the beam spot that was simulated in all Monte-Carlo samples used in this thesis. The bias can be corrected for, but this does not significantly improve the $\not{E}_T$ resolution. A more detailed discussion can be found in [71].

According to Eq. 3.14, $\not{E}_T$ resolution depends on the scalar sum of E_T in the calorimeter (ΣE_T). Pile-up affects the resolution in two ways: Proton-proton interactions from in-time pile-up deposit additional energy in the calorimeter and thus increase ΣE_T (Fig. 8.12). Furthermore, it adds noise from out-of-time pile-up. Fig. 8.13a shows the resolution of $\not{E}_x$ as a function of ΣE_T. Often only the statistical term, proportional to ΣE_T, is used to fit this distribution. This approximation works well without pile-up. Both $H \to \tau\tau$ and $Z \to \tau\tau$ show the expected behaviour with a coefficient of 0.51 (cf. Table 8.4), which agrees with 0.52 as stated

sample	σ_1 [GeV]	σ_2 [GeV]
$H \to \tau\tau$	7.1(3)	15.1(9)
$Z \to \tau\tau$	8.64(3)	15.2(1)
$H \to \tau\tau$ (pile-up)	12(2)	19(3)
$Z \to \tau\tau$ (pile-up)	13.2(3)	18.9(3)

Table 8.3.: Standard deviations of double Gaussian functions fit $\not{E}_x$ resolution in Fig. 8.11

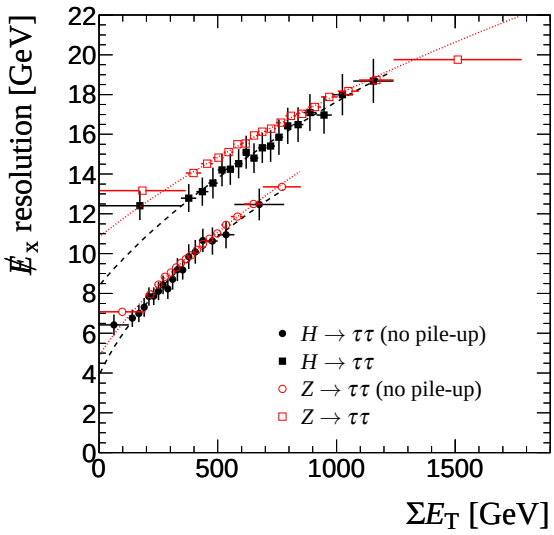


(a) $H \to \tau\tau$ **(b)** $Z \to \tau\tau$

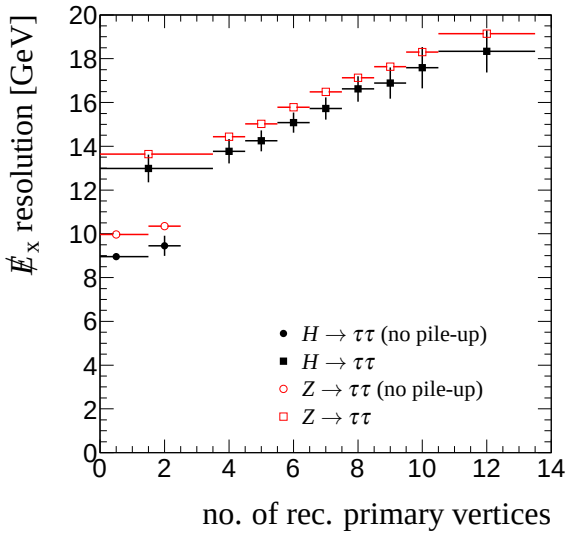
Figure 8.12.: ΣE_T with and without pile-up on generator level and reconstructed.

in [39]. It should be noted, though, that a slightly better fit can be obtained by adding an additional offset as in Fig. 8.13a. Pile-up not only adds additional transverse energy, increasing ΣE_T. This would result in the same curve. The worse $\not{E}_T$ resolution would then solely be a consequence of a higher average ΣE_T. But this is clearly not the case. Both, signal process and $Z \to \tau\tau$, have worse resolution for a given value of ΣE_T. The signal sample has an identical slope in the case of pile-up but the offset is approximately doubled. The higher offset can be attributed to additional noise from out-of-time pile-up. The different behaviour of signal and $Z \to \tau\tau$ can be explained by Figures 8.13b and 8.13c. ΣE_T is correlated to the number of primary vertices because this number directly translates into the number of proton-proton collisions.

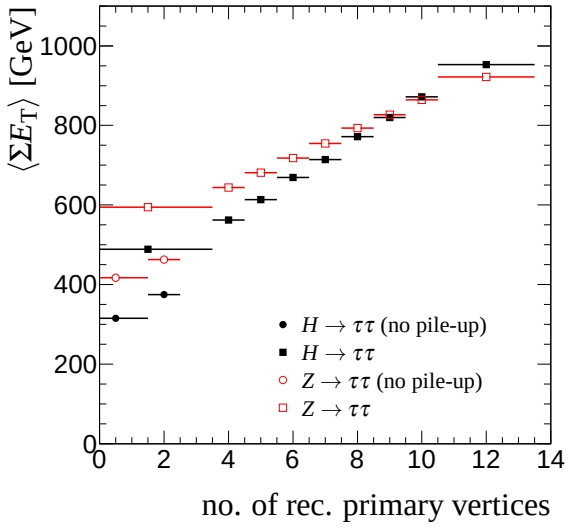
The deterioration of $\not{E}_T$ resolution has only a small influence on the distribution of $\not{E}_T$ (Fig. 8.14) itself except for the first bins. For smaller values of $\not{E}_T$, the bias seen in the difference between true and reconstructed $\not{E}_T$ becomes important. Thus, for pile-up the first bins of the distribution are more populated. For larger values, the bias can be seen as relatively small,



(a) $\not{E}_x$ resolution vs ΣE_T.



(b) $\not{E}_x$ resolution vs number of reconstructed primary vertices.



(c) Mean ΣE_T vs number of reconstructed primary vertices.

Figure 8.13.: Interdependence of $\not{E}_x$ resolution, ΣE_T and number of reconstructed primary vertices.

sample	a	b	c	a (stat.)
$H \to \tau\tau$ (no pile-up)	0.4(2)	5(2)	0	0.508(8)
$Z \to \tau\tau$ (no pile-up)	0.43(7)	5(1)	0	0.507(2)
$H \to \tau\tau$	0.5(3)	9(4)	0	0.590(8)
$Z \to \tau\tau$	0.453(9)	10.8(3)	0	0.611(3)

Table 8.4.: Fit parameters of $\not{E}_x$ resolution as a function of ΣE_T (Fig. 8.13a). c was in all cases compatible with zero within errors. The last column, $a(stat.)$ gives the coefficient of the statistical term assuming that $\sigma_{\not{E}_x} = a\sqrt{\Sigma E_T}$, only.

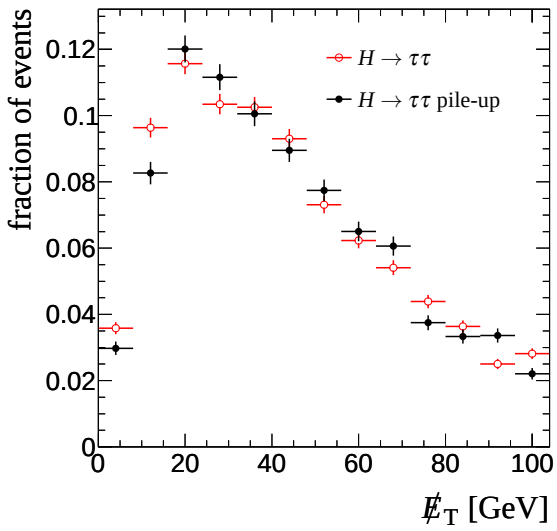
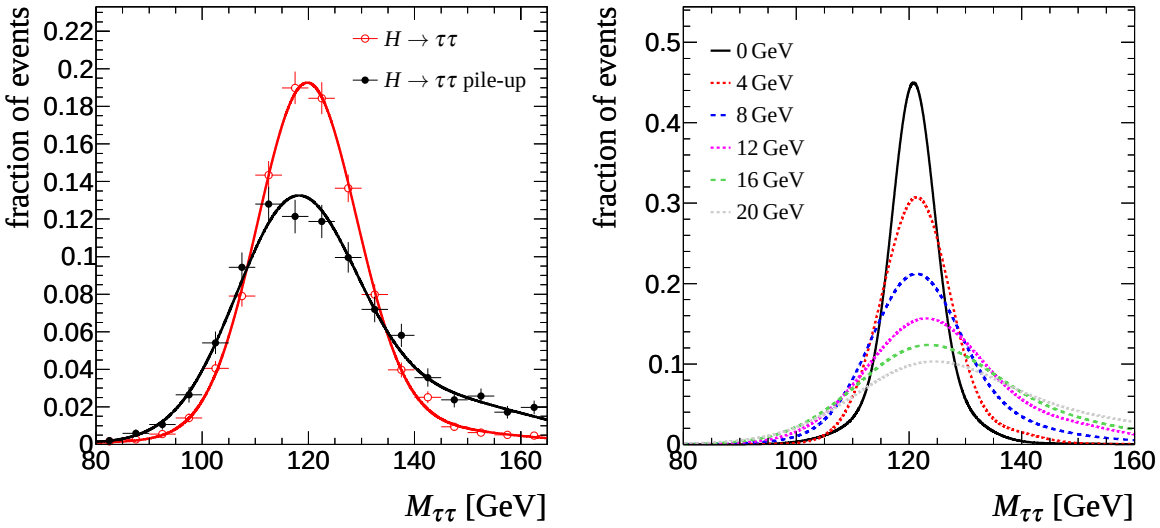


Figure 8.14.: Distribution of $\not{E}_T$ after pre-selection.

so the probability of an upwards fluctuation is approximately the same as for downwards. So, for a sufficiently high threshold, the acceptance is unchanged.



(a) $M_{\tau\tau}$ with and without pile-up.

(b) Fit to $M_{\tau\tau}$ calculated with generated $\not{E}_{\mathrm{T}}$ in $H \to \tau\tau$ without pile-up.

Figure 8.15.: $M_{\tau\tau}$ distributions after pre-selection and cuts on x_ℓ/x_h. A triple Gaussian function was used to fit the distributions. **(b)** shows fits to the non-pile-up sample using generated $\not{E}_{\mathrm{T}}$ instead of reconstructed. Gaussian smearing with a width between 0 and 20 GeV was applied to the generated $\not{E}_{\mathrm{T}}$.

8.5. $M_{\tau\tau}$ Resolution

The deterioration of $\not{E}_{\mathrm{T}}$ resolution has no significant impact on the acceptance of the $\not{E}_{\mathrm{T}}$ requirement in the signal sample. However, it directly affects the $M_{\tau\tau}$ resolution. Figure 8.15a shows the $M_{\tau\tau}$ distribution with and without pile-up. The presence of pile-up makes the distribution broader and more asymmetric. The width (FWHM[1]) increases from 22.5 GeV to 29.1 GeV ($m_H = 120\,\mathrm{GeV}$).

The collinear approximation assumes that $\not{E}_{\mathrm{T}}$ results only from neutrinos from the τ lepton decays. Deviation from this assumption directly results in a deterioration of the $M_{\tau\tau}$ resolution. This is shown in Figure 8.15b. Here, the measured $\not{E}_{\mathrm{T}}$ is substituted by the generated $\not{E}_{\mathrm{T}}$ which is equal to the sum of neutrino momenta, except for considering also particles with $|\eta| > 5$ as non interacting. The distributions show different levels of $\not{E}_{\mathrm{T}}$ resolution simulated by Gaussian smearing. Worse resolution leads to a broader peak but also makes the distribution more asymmetric. As can be seen from Figure 8.16, the resulting dependence of the width (FWHM) of $M_{\tau\tau}$ as a function of the width of the Gaussian smearing of $\not{E}_{\mathrm{x}}$ is linear in this simplified model. However, the additional points show that $H \to \tau\tau$ with real $\not{E}_{\mathrm{T}}$ is well described by this model. This is a strong indication that the broader $M_{\tau\tau}$ distribution in the presence of pile-up is caused dominantly by the $\not{E}_{\mathrm{T}}$ resolution.

[1] Since the $M_{\tau\tau}$ distribution is asymmetric, the full width at half maximum, FWHM, is a more suitable quantity than the σ of a Gaussian function.

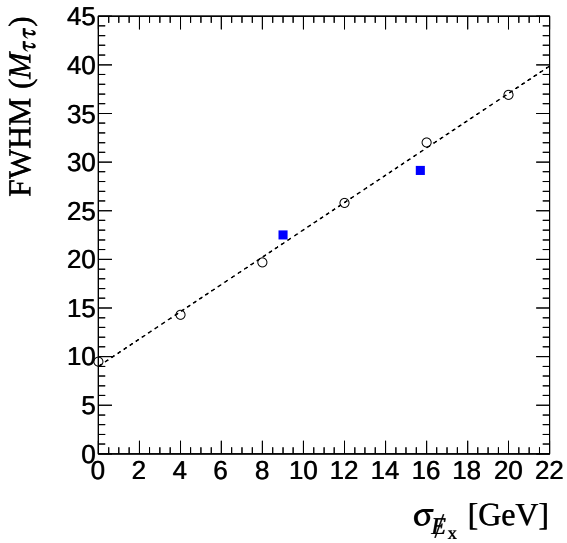
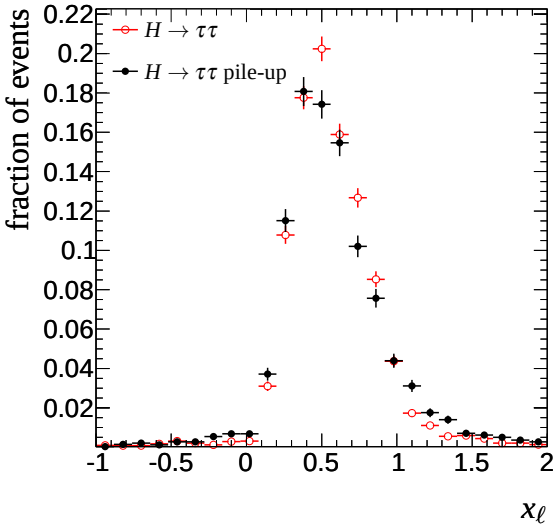


Figure 8.16.: Full width at half maximum of $M_{\tau\tau}$ for $H \to \tau\tau$ with generated $\not{E}_{\mathrm{T}}$ and different settings of Gaussian smearing (black open circles) (Fig. 8.15b). Blue squares mark $H \to \tau\tau$ with reconstructed $\not{E}_{\mathrm{T}}$ without (left) or with pile-up (right). Here, $\sigma_{\not{E}_x}$ is taken from a simple Gaussian function fit to the distribution. The dashed line is a fit to the black circles.

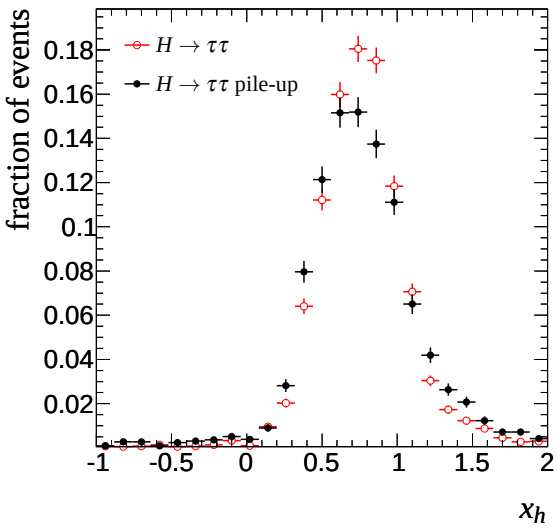
It stands to reason that this effect is not unique to the signal sample. The $Z \to \tau\tau$ resonance also becomes broader (see. Sec. 8.7). Pile-up causes both peaks, which are well separated in the no-pile-up scenario, to partially merge. Since one of the key features of $H \to \tau\tau$ search in VBF is mass reconstruction, this drastically affects the expected signal significance as will be further discussed in Chapter 10.

8.6. Other Effects

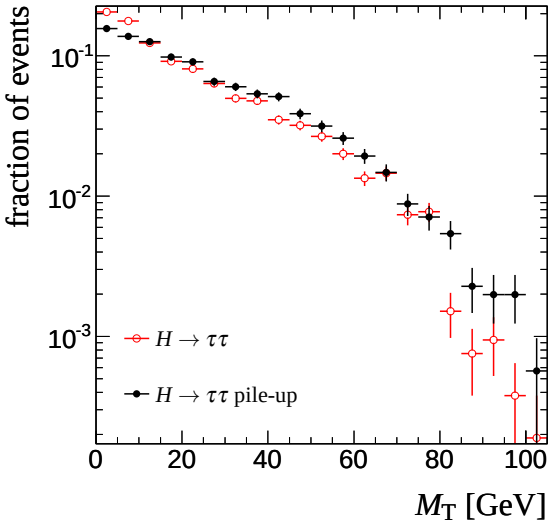
Looking at Table 8.1 reveals several minor deviations in acceptance caused by pile-up. The probability to find two jets with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ increases, since pile-up creates new jets or adds energy to jets from the hard process. Cuts on both x_ℓ and x_h have a reduced acceptance. This is, of course, closely related to the broader $M_{\tau\tau}$ distribution due to worse $\not{E}_{\mathrm{T}}$ resolution. In both cases, the distributions become broader (cf. Figs. 8.17a and 8.17b). As restricting the x-values to $0 < x < 1$ is only meaningful if $\not{E}_{\mathrm{T}}$ is equal to the sum of neutrino momenta from the τ decay, it makes sense to relax the requirement (see next section). The lower limit on M_{T} also has a lower acceptance. Again, this is an observable related to $\not{E}_{\mathrm{T}}$. Figure 8.17c reveals that the distribution is shifted towards higher values. Fig. 8.17c shows that also this effect can be explained by $\not{E}_{\mathrm{T}}$ resolution. Again, the reconstructed $\not{E}_{\mathrm{T}}$ is substituted by generated $\not{E}_{\mathrm{T}}$, leading to a similar effect.



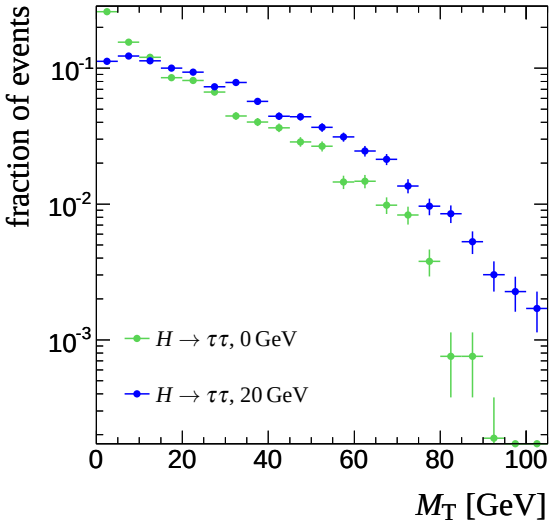
(a) x_ℓ.



(b) x_h.



(c) M_T.



(d) M_T in a sample with generated $\not{E}_T$ (green) and generated $\not{E}_T$ with Gaussian smearing of 20 GeV.

Figure 8.17.: Observables (after pre-selection) affected by deterioration of $\not{E}_T$ resolution.

8.7. Conclusions

To mitigate the impact pile-up has on the analysis, the selection is modified in the following way (the requirement in the no-pile-up scenario is given in brackets):

- $\not{E}_{\mathrm{T}} > 40(30)\,\mathrm{GeV}$
- $0 < x_\ell < 0.85(0.75)$
- $0 < x_h < 1.1(1.0)$
- $M_{\mathrm{T}} < 40(30)\,\mathrm{GeV}$
- jet veto: $|\eta| < 2.5(3.2)$, PVF > 0.7
- τ_{had} identification: new LLH, 1p: > 2, 3p: > 6.5

From looking only at the signal sample, the higher $\not{E}_{\mathrm{T}}$ threshold is not necessary. Since, however, the $\not{E}_{\mathrm{T}}$ resolution deteriorates in the presence of pile-up, and QCD background with pile-up is not available for this analysis, the threshold is set to a higher value. Table 8.5 shows the acceptance of the updated selection. The loss in acceptance due to pile-up can be reduced from about 50% to 15%. The $M_{\tau\tau}$ resolution on the other hand is not significantly affected by these modifications. Figure 8.18 shows that both the Higgs and Z boson resonance are broader in the presence of pile-up. As can be seen in Table 8.6, the expected number of non-resonant background is approximately equal to the number obtained in the no-pile-up scenario. The number of expected $Z \to \tau\tau$ events is reduced by about the same factor as the number of $H \to \tau\tau$ events. The final $M_{\tau\tau}$ distribution can be seen in Figure 8.19. The $H \to \tau\tau$ peak can barely be recognised as a separated resonance. The current luminosity delivered by the LHC already surpasses the simulation, with even more severe pile-up conditions. Nevertheless, more recent improvements in the description of pile-up in simulation as well as its handling in reconstruction and analysis could not be considered here. It seems not unreasonable that with a better understanding of pile-up, future developments can improve the situation. E.g. a new technique to reconstruct $M_{\tau\tau}$ [72] without relying on the collinear approximation seems a promising candidate.

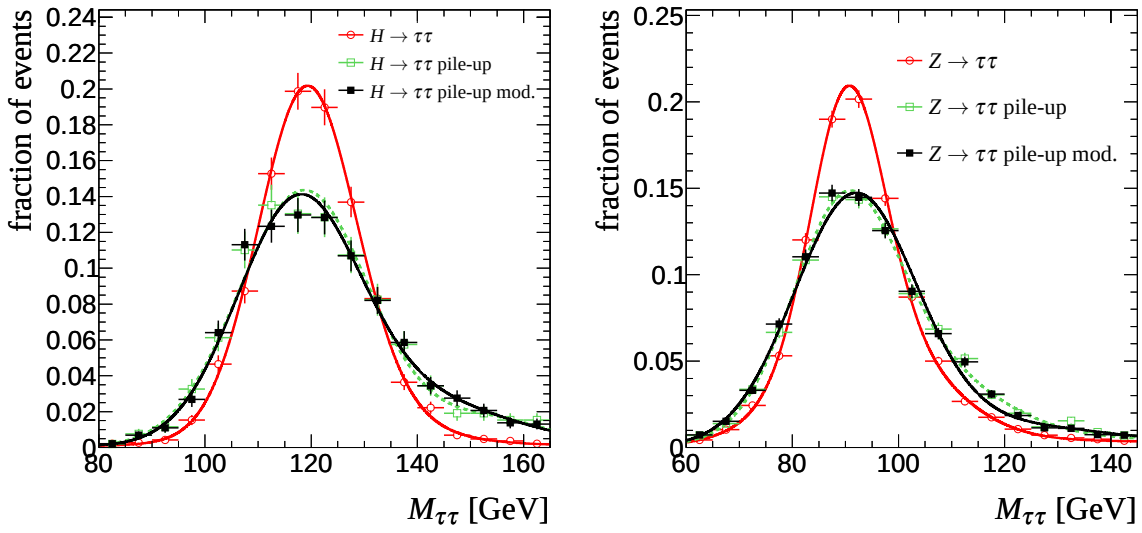
Clearly, the reduced separation of $H \to \tau\tau$ and $Z \to \tau\tau$ puts greater emphasis on understanding the background processes. As this chapter showed, correctly modelling $\not{E}_{\mathrm{T}}$ is crucial to this understanding. As modelling $\not{E}_{\mathrm{T}}$ in the simulation is difficult – all higher level objects enter the calculation and $\not{E}_{\mathrm{T}}$ is sensitive to the detector condition and LHC parameters – the next chapter will present methods to estimate the dominant background processes from data.

cut	$H \to \tau\tau \to \ell h$ (pile-up)	
total	3100(10)	-
trigger	1007(7)	0.33
trigger lepton	874(6)	0.87
no. of leptons	868(6)	0.99
no. of τ_{had}	239(3)	0.28
no. of jets	197(3)	0.82
charge product	192(3)	0.98
$\not{E}_{\text{T}}$	107(2)	0.55
x_ℓ	94(2)	0.88
x_h	77(2)	0.83
$\cos\Delta\phi(\ell, \tau_{\text{had}})$	75(2)	0.97
M_{T}	63(2)	0.84
$\text{sign}(\eta_{j1}\eta_{j2})$	52(1)	0.83
p_{T} (lead. jet)	52(1)	0.99
$\Delta\eta_{jj}$	29(1)	0.56
ℓ-τ_{had} centrality	28(1)	0.98
M_{jj}	26(1)	0.91
jet veto	23(1)	0.88

Table 8.5.: Comparison of cut acceptances in the signal sample for with and without and pile-up. Samples are filtered on generator level to contain only the ℓh channel.

	$H \to \tau\tau$	$Z \to \tau\tau$	$W+\text{jets}$	$t\bar{t}$	$Z \to \ell\ell$	di-boson
after cuts	23(1)	60(7)	12(5)	17(4)	-	-

Table 8.6.: Expected number of events after cuts. $\mathscr{L}_{\text{int}} = 30\ \text{fb}^{-1}$.



(a) $H \to \tau\tau \to \ell h$ **(b)** $Z \to \tau\tau \to \ell h$

Figure 8.18.: Impact of modifications of the analysis on $M_{\tau\tau}$ after pre-selection and all lepton and $\not{E}_{\mathrm{T}}$ cuts.

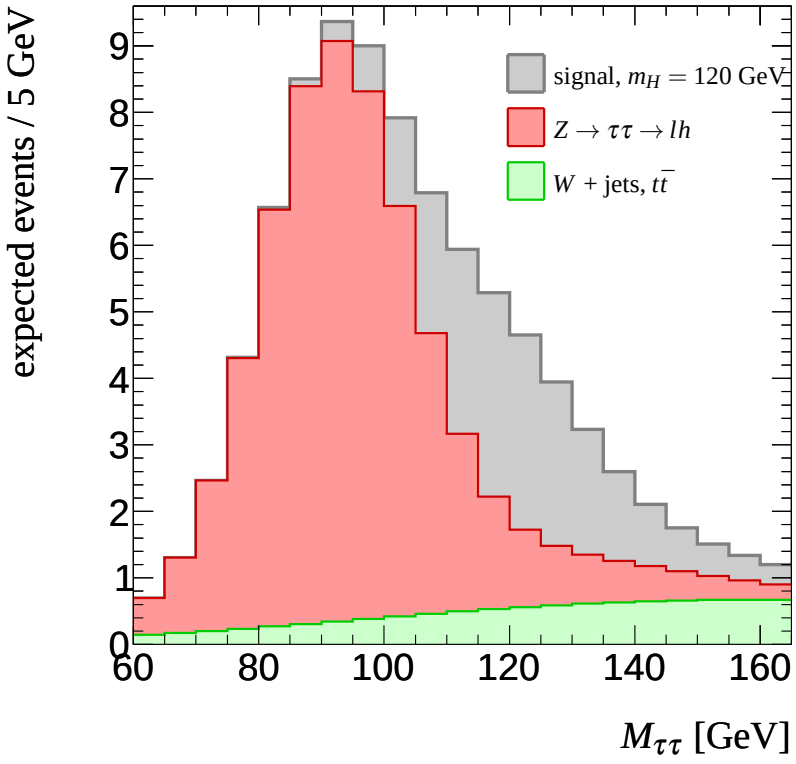


Figure 8.19.: Parametrised $M_{\tau\tau}$ distribution in the presence of pile-up for an integrated luminosity of $30\,\mathrm{fb}^{-1}$ after the modified full selection.

9

Estimation of Background Processes

The previous chapter emphasised that a precise knowledge of the background and especially the $M_{\tau\tau}$ shape is essential in order to be able to extract the signal process. $M_{\tau\tau}$ depends on $\not{E}_{\mathrm{T}}$, which, as was demonstrated, is strongly influenced by pile-up and also detector conditions. It is therefore not always reliably modelled by simulation. Moreover, the dominant background processes include the production of heavy vector bosons associated with several jets. The vector boson recoils against the jets, which thus determine the boost of the decay products. As was explained in Chapter 3, the jet kinematics are difficult to describe with standard leading order Monte-Carlo programs. Furthermore, it is doubtful whether it will be at all possible to produce a sufficient amount of MC events for processes like $W+\mathrm{jets}$. As a result, it is preferable to estimate the dominant background processes directly from data. In this chapter, I will present two such methods.

9.1. An Embedding Technique to Estimate $Z \to \tau\tau$

$Z \to \tau\tau$ is an irreducible background to $H \to \tau\tau$, i.e. it can have an identical final state. It can be separated from the signal process by reconstructing $M_{\tau\tau}$. However, both peaks overlap: The signal peak lies on top of the steeply falling flank of the Z peak. It is therefore impossible to select a $Z \to \tau\tau$ control sample from data in the signal mass region without signal contamination. The approach presented here makes use of as much data as possible while keeping input from simulation at a minimum.

Neglecting the difference between m_τ and m_μ, both of which are small compared to m_Z, $Z \to \mu\mu$ production is kinematically identical to $Z \to \tau\tau$, especially when considering the associated jets. In contrast to the latter process, it can be selected free of signal by selecting two muons with an invariant mass near the Z peak. If one replaces these two muons of a collision data event with simulated τ decay products, where the τ leptons had the same four-vectors as the muons, the result will be an almost perfect $Z \to \tau\tau$ event.

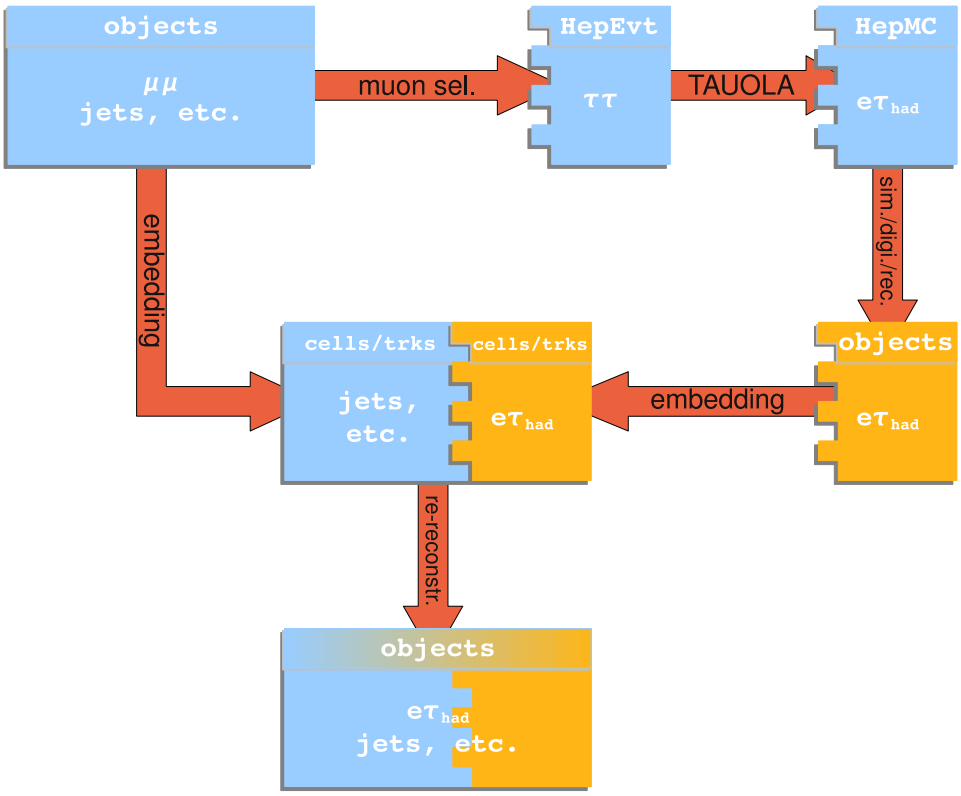


Figure 9.1.: Flowchart of the embedding procedure.

Several similar methods already exist and have been demonstrated. These methods, see e.g. [73], focus on replacing the original muons with the visible (reconstructed) tau decay products. This is sufficient to reproduce the kinematics of the τ decay products. However, $\not{E}_{\mathrm{T}}$ has to be corrected manually. In ATLAS, $\not{E}_{\mathrm{T}}$ is calculated from calorimeter cell depositions (cf. Sec. 3.4.6). Starting from this premise, the most "natural" way to manipulate the original event is to replace energy depositions in calorimeter cells instead of reconstructed objects and re-run the standard reconstruction algorithms. This in turn produces all the required final reconstructed objects and provides $\not{E}_{\mathrm{T}}$ without having to correct.

The advantages which this method offers come at the price of increased complexity compared to other methods. The whole procedure, called *embedding* in the following, runs in several steps. Fig. 9.1 gives an overview.

9.1.1. The Embedding Procedure

Muon Selection

The first step of the embedding procedure is to select the muons in a collision data event which will be replaced by τ decay products. The proper selection of a clean and unbiased control sample will be discussed in Sec. 9.1.5. The embedding procedure itself should be independent

of the actual selection. For the purpose of validating the method (Sec. 9.1.4), an event is required to contain two muons with:

- opposite charge
- $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ each
- the same reconstructed vertex as origin

In case there is more than one valid pair of muons in an event, the pair with the highest scalar p_{T} sum is selected. The $Z \to \mu\mu$ decay vertex is written out in HEPEVT [74] format. The Z boson is taken as four-vector sum of the muons. The reconstructed vertex to which the muons point is used as vertex position. Both muons are declared to be τ leptons, i.e. the identifier in HEPEVT is set to that of a τ and the four-vector is rescaled to obtain m_τ. An example of a HEPEVT record is shown in Appendix E.

Simulation of the $\tau\tau$ Decay

The new $Z \to \tau\tau$ "event" is processed by TAUOLA, which simulates the τ decays, taking care of spin correlations and the desired decay mode. The result is then fed into the full detector simulation, digitisation and reconstruction. Calorimeter noise simulation, as well as vertex smearing are switched off. The final output of this is an ESD (see Sec. "Data Format").

Merging Original Event and Simulated $\tau\tau$ Decay

The merging step is the core of the embedding procedure. Both the original ESD and the new one containing the simulated τ decays are processed in parallel. The result is an ESD containing hybrid events consisting of the original event in which the muons are replaced by simulated τ decay products. The merging is done in three separate sub-tasks. First, the original muons, which were used as τ leptons, have to be identified and associated with the τ decays. The reason for this is a detail of the implementation but not entirely trivial: muon selection and merging are run in separate steps. τ leptons in the simulated $\tau\tau$ decays are matched to original muons within a cone around the generated τ. The direction of a τ lepton in the simulated $\tau\tau$ decay and the original muon are not always identical because TAUOLA sometimes creates additional photon radiation which changes the four-vectors of the τ to ensure energy and momentum conservation in the decay process. Possible effects will be discussed in more detail in Sec. 9.1.8.

The second step is track embedding. The embedding procedure works on the most basic entities in the calorimeter, i.e. energy depositions in cells. But instead of replacing hits or drift circles in the tracking sub-detectors, reconstructed tracks are replaced. First of all, ESDs do not contain all hits. In addition, the procedure would be much more sensitive to misalignment. Also, $\not{E}_{\mathrm{T}}$ reconstruction runs on tracks not hits. So, while embedding on hit level might be more accurate, it is not strictly necessary. Track embedding removes the tracks of the original muons in the inner detector and muon spectrometer. All tracks in the simulated $\tau\tau$ decay are

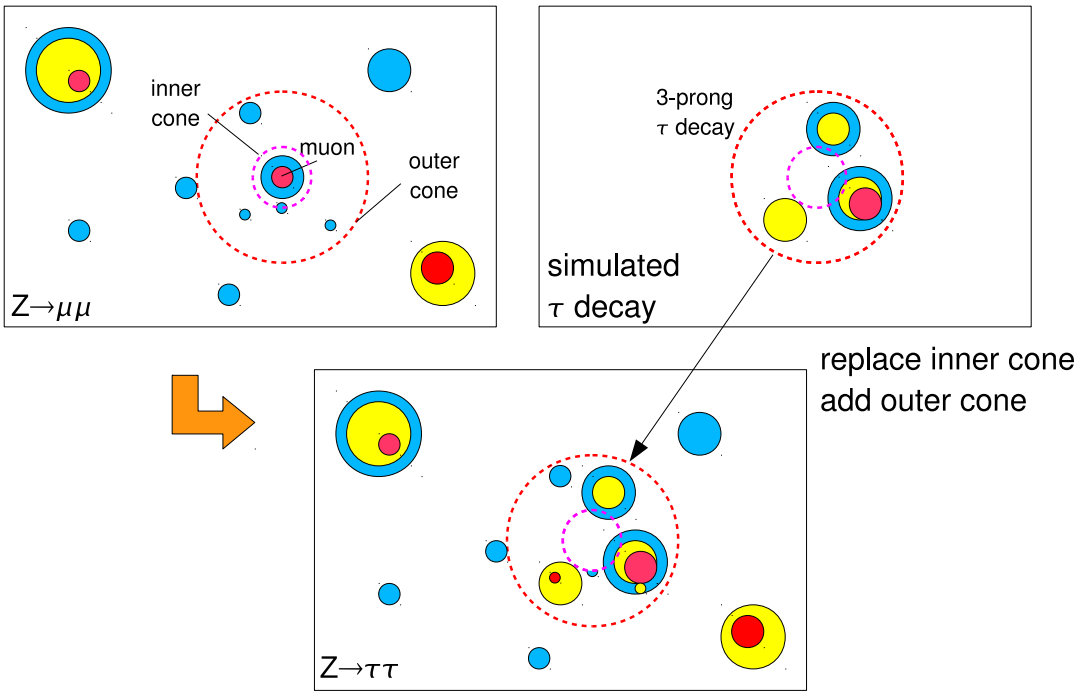


Figure 9.2.: Stylised event display: Rectangular areas represent the calorimeter in the η-ϕ plane. Filled circles depict energy depositions, where the colour coding stands for an increasing amount of deposited energy from blue to red.

copied to the new event. All track segments in the muon spectrometer in a cone around the original muons are removed and replaced by the simulated track segments in the same cone. This is necessary in order to completely remove the original muons. Muon tracks not always make use of all the track segments and leftover segments may lead to the reconstruction of additional muons by MuTag (cf. Sec. 3.4.3).

Calorimeter embedding is done in the third step. See Figure 9.2 for an example. The original embedding method described in [75] works but introduces small biases. A more sophisticated method (cf. [76]) is used instead. To remove the energy the original muon deposited in the calorimeter, all depositions within a small cone (C_{in}) around that muon are removed. In a larger cone (C_{out}) around the muon all cell entries in the simulated $\tau\tau$ decay are added to those of the original event.

Re-Reconstruction

The new hybrid event now consists of tracks and calorimeter cell information. To obtain higher-level objects all available standard reconstruction algorithms are run on the event. This slightly limits the method: Reconstruction on ESDs is in principle foreseen in the ATLAS software but has to be run mostly manually. Thus, only muons from the STACO algorithm are available. Fortunately, this is irrelevant for many use cases. In the context of this analysis muons are defined as being reconstructed by STACO, and also $\not{E}_{\text{T}}$ makes use of STACO muons.

Data Format

The ATLAS analysis data format contains only limited calorimeter cell information. Energy depositions in cells are stored if the cell lies within a cone around a reconstructed electron, muon or τ_{had}. This is insufficient for the embedding procedure as it is implemented: On the one hand, the reconstruction of $\not{E}_{\text{T}}$ requires full calorimeter information. On the other hand, not all jets are reconstructed: Although τ_{had} candidates are reconstructed from seed jets, not every jet is reconstructed as τ_{had} candidate, especially in the low-p_{T} regime – $\not{E}_{\text{T}}$ uses jets with $p_{\text{T}} > 5\,\text{GeV}$. Complete calorimeter information is available in raw data but access to this is restricted. The choice to use the ESD[1] format is therefore a trade-off between contained information and accessibility of data. ESDs contain full, albeit compressed (lossy) [77], calorimeter information and is accessible to all users via special skims.

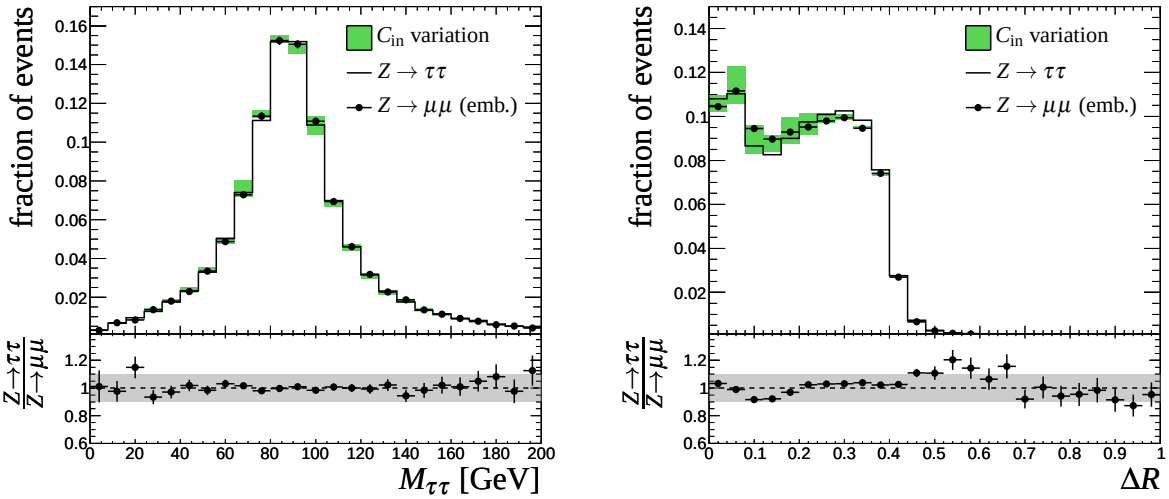
9.1.2. Cone Optimisation

Cone sizes are the main tunable parameters of the embedding procedure. The size of the inner cone C_{in} is a trade-off between removing the entire energy deposited by the original muon and keeping as much of the original event as possible. While calorimeter depositions of the simulated $Z \to \tau\tau$ are added to the original event independent of C_{in}, the hybrid event will contain too much energy in the vicinity of the τ decay product if C_{in} is too small. Making C_{in} too large, removes noise, out-of-time pile-up or even real objects from the collision data event. The outer cone C_{out} is a detail of the implementation used in this study: noise is no longer simulated for the $Z \to \tau\tau$ decay, which means there is no need to restrict the area in which energy is added. In the latest releases of the package all energy depositions in the calorimeter are added to the event. Restricting C_{out} in the version used for this study is necessary in order to prevent significant double counting of energy depositions as a cell is added each time it is situated with an outer cone around an original muon. Note also that in recent releases C_{out} is centred around the simulated τ decay products instead of the original muon. This is more reasonable because C_{out} is used to implant the τ decay products into the new event.

For the optimisation, C_{in} is varied between 0.05 and 0.12 in steps of 0.01. To limit CPU time consumption the study is carried out on the ALPGEN two-parton samples (cf. Chapter 5), only. The simulation of the $Z \to \tau\tau$ decays involves randomisation of e.g. the τ decays and noise. To avoid statistical effects due to this randomisation on top of the variation, the same simulated $Z \to \tau\tau$ decays are used for all values of C_{in}. A Kolmogorov test of the $Z \to \tau\tau$ distribution and the respective $Z \to \mu\mu$ distribution after embedding serves as a figure of merit. In general, calorimeter related observables involving the τ decay products are suited for the optimisation. It turns out that the optimal value of C_{in} depends on the observable in question. Hence, $M_{\tau\tau}$ is the best choice for this problem.

Fig. 9.3a shows $M_{\tau\tau}$ calculated from $Z \to \tau\tau$ decay products matched to generated objects and reconstructed $\not{E}_{\text{T}}$. Variation of C_{in} over the full range leads to changes which are well below the statistical uncertainty of the expected number of events after cuts. Nevertheless, it

[1] Event Summary Data



(a) $M_{\tau\tau}$.

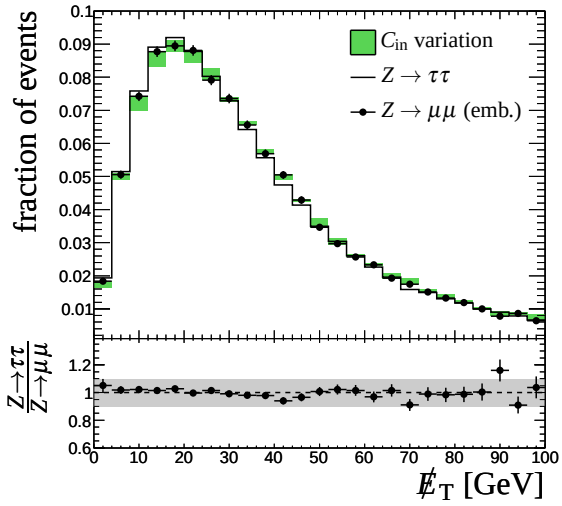
(b) ΔR between barycentre and subclusters of τ_{had} candidates.

Figure 9.3.: Effect of varying the the size of the inner cone C_{in} between 0.05 and 0.12. The solid green areas range between the minimum and maximum central value per bin of the modified histograms. The sample giving the best agreement with $Z \to \tau\tau$ in a Kolmogorov-Smirnov test is used for the ratio.

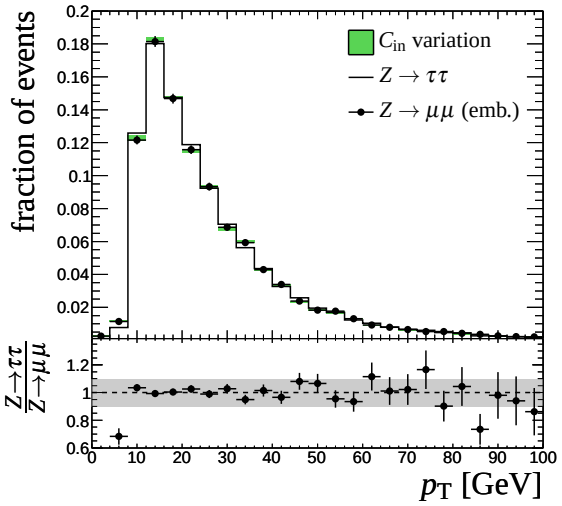
should be noted that the bin entries vary in a correlated way, leading to a systematic shift of roughly 1 GeV in the peak. The optimal value for C_{in} is 0.1.

As tracks are copied independent of the cone sizes, muons are not affected. Some other relevant observables are shown in Figures 9.3b, 9.4 and 9.5. $\not{E}_{\text{T}}$ and electron p_{T}, although not much affected by the cone variations, both prefer small cones. $C_{\text{in}} = 0.05$ is the optimal value within the tested range. Relative isolation on the other hand agrees best with a large cone. τ_{had} related observables consistently show an optimal C_{in} of 0.09. The most sensitive observable for τ_{had} candidates is the p_{T} resolution. Cone size variation again leads to a shift of the mean value. The jet algorithm which provides a seed for the τ_{had} reconstruction algorithm gathers more than just energy depositions of the hadronic τ decay itself. This leads to an average reconstructed p_{T} of the τ_{had} that is greater than the true p_{T}. A large cone therefore cuts away too much energy around the τ_{had} making the result "too good". A small cone leaves too much energy deposited by the original muon, especially in the hadronic calorimeter.

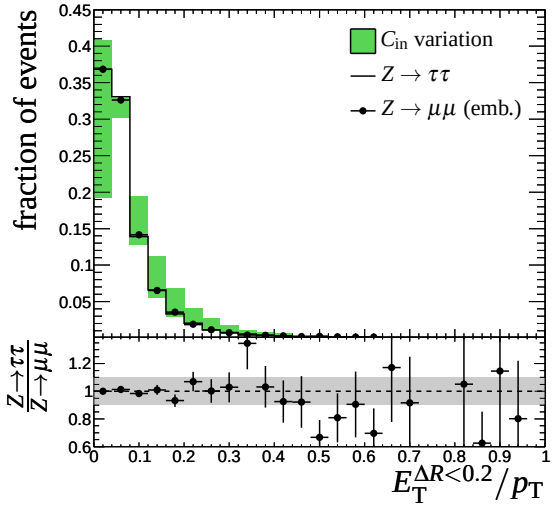
The optimal value for $M_{\tau\tau}$ (0.98 in a Kolmogorov-Smirnov test), $C_{\text{in}} = 0.1$, is used in the following. All of the above tests can also be performed on data. The information on generator level comes solely from the simulated $\tau\tau$ decay and is fully available also when using $Z \to \mu\mu$ events form collision data. Thus, assuming well understood MC by the time an integrated luminosity of $30\,\text{fb}^{-1}$ is reached, the same embedding vs. MC comparison can be done. Systematic uncertainties arising from badly adjusted cone sizes can be expected to be well under control.



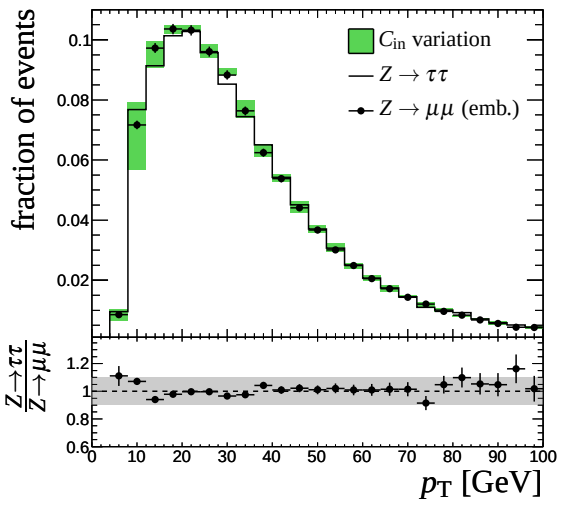
(a) MET RefFinal.



(b) p_{T} of electrons from the $Z \to \tau\tau$ decay (matched to generated electrons).

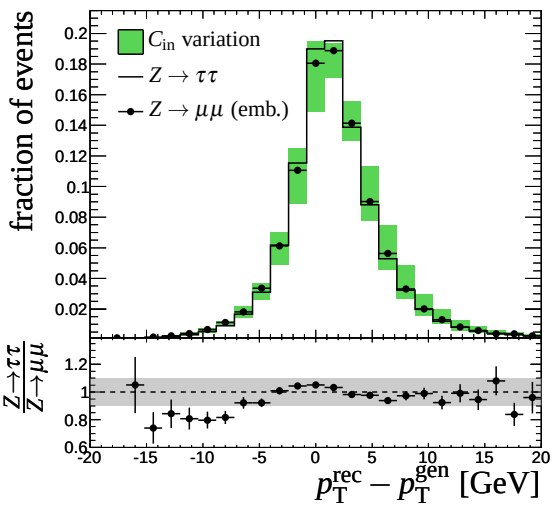


(c) Relative isolation of electrons from the $Z \to \tau\tau$ decay (matched to generated electrons).

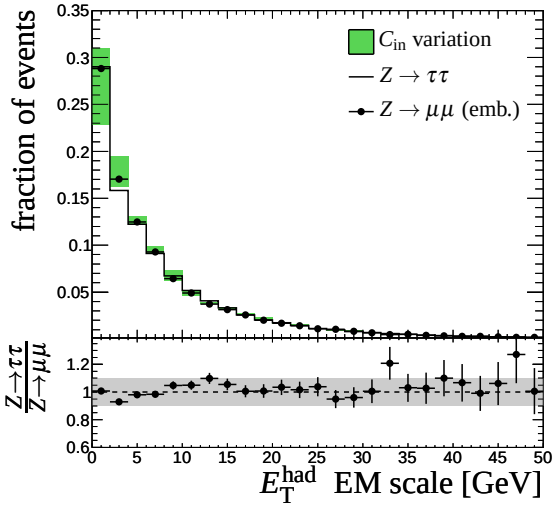


(d) p_{T} of τ_{had} candidates from the $Z \to \tau\tau$ decay (matched to the visible hadronic τ decay product on generator level).

Figure 9.4.: Effect of varying the the size of the inner cone C_{in} between 0.05 and 0.12. The solid green areas range between the minimum and maximum central value per bin of the modified histograms. The sample giving the best agreement with $Z \to \tau\tau$ in a Kolmogorov-Smirnov test is used for the ratio.



(a) p_T resolution of τ_{had} candidates from the $Z \to \tau\tau$ decay (matched to the visible hadronic τ decay product on generator level).



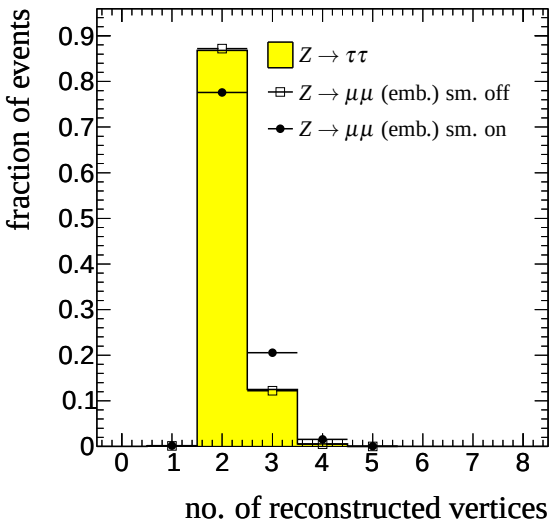
(b) E_T in the HCAL of track-seeded τ_{had} candidates from the $Z \to \tau\tau$ decay (matched to the visible hadronic τ decay product on generator level).

Figure 9.5.: Effect of varying the the size of the inner cone C_{in} between 0.05 and 0.12. The solid green areas range between the minimum and maximum central value per bin of the modified histograms. The sample giving the best agreement with $Z \to \tau\tau$ in a Kolmogorov-Smirnov test is used for the ratio.

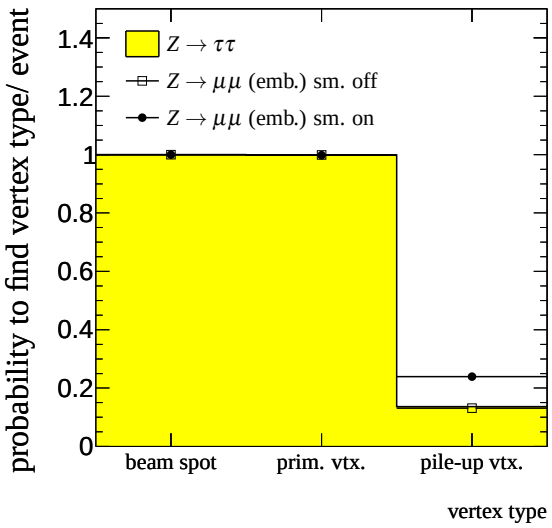
9.1.3. Vertex Reconstruction

Proton-proton collisions do not happen exactly at the centre of the detector. They are spread around the beam spot, especially along the z-axis. In addition, the actual beam spot can be displaced from the nominal one. In the simulation, this is realised by going through the list of generated vertices, possibly setting the beam spot to a new location and performing a Gaussian smearing of the coordinates. In the context of the embedding procedure, this leads to a problem: The reconstructed vertex is already smeared with respect to the true collision. The Z production vertex of the simulated $Z \to \tau\tau$ decay is set to the vertex to which the two muons point. By default this is then smeared according to the parameters currently employed by the simulation. This obviously means another very substantial smearing of the vertex from which the τs of the $Z \to \tau\tau$ decay originate.

Any track that cannot be associated with a reconstructed vertex is attached to the beam spot. Thus, an ATLAS event usually has at least two vertices: the beam spot and the primary vertex. If more than one primary vertex is reconstructed, the one with the highest scalar sum of p_T^2 of tracks is called primary, the others are classified as pile-up vertices. Hence, an event without additional pile-up interactions should have two vertices (beam spot and primary) but no pile-up vertices. The number of vertices per event can be seen in Fig. 9.6a. Due to the default vertex smearing in the simulated $\tau\tau$ decay, a $Z \to \mu\mu$ sample has on average more vertices than a real $Z \to \tau\tau$ event. Especially in z-direction, tracks of the τ decay products are displaced with respect to the original primary vertex. The vertex reconstruction algorithm



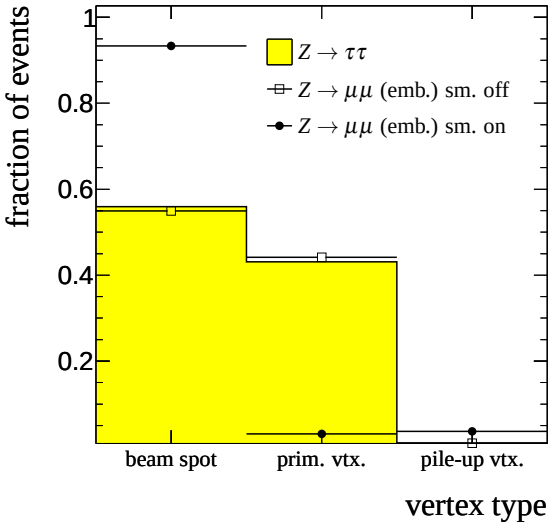
(a) Number of reconstructed primary vertices.



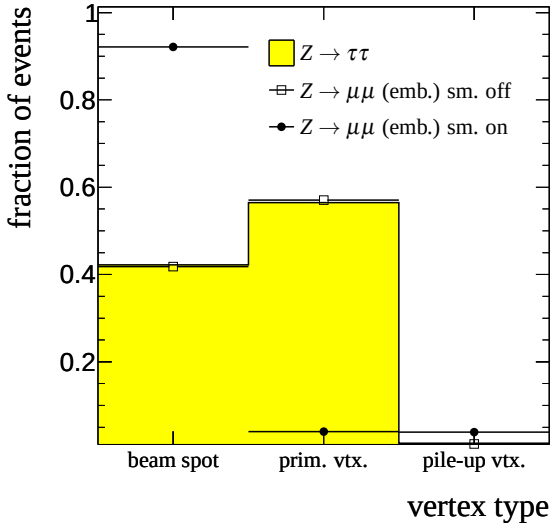
(b) Probability to find the different types of vertices per event.

Figure 9.6.: Effect of vertex smearing on the primary vertex reconstruction.

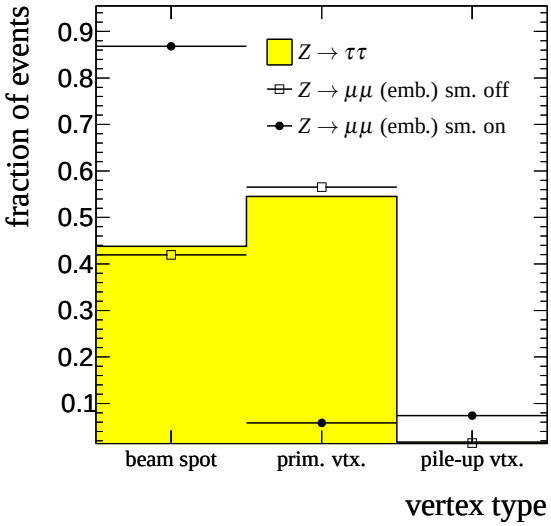
does not consider them to be coming from the primary vertex. The tracks are then either attached to the beam spot, or, in case of e.g. a 3-prong τ_{had}, a new pile-up vertex is created. In the former case, tracks can have a large $z0$, which means that they fail track quality cuts of object reconstruction algorithms like the τ_{had} reconstruction (cf. e.g. Fig. 3.4.5). This not only leads to an inefficiency of the reconstruction but also to a wrong ratio of 1- to 3-prong τ_{had} candidates. When the default vertex smearing is switched off, the distributions are in good agreement. The effect of vertex smearing on the assignment of tracks belonging to reconstructed objects to the different types of primary vertices is shown in Figure 9.7. In the presence of pile-up, the track multiplicity and the number of vertices are much higher than without pile-up. Vertex smearing in the simulated $\tau\tau$ decay has therefore less impact on the general vertex reconstruction. With smearing switched off, the number of reconstructed primary vertices is in good agreement (Fig. 9.8a). The number of vertices is on average 0.6% lower in the embedding sample. Although visible, this effect can probably be safely neglected. Without smearing, tracks point more often to the primary vertex than in real $Z \to \tau\tau$ MC. This effect is presumably caused by the much cleaner environment in which track reconstruction in the simulated $\tau\tau$ decay acts, especially when considering pile-up. Consequently, the effect is more visible in the presence of pile-up (Fig. 9.8b). Also, requiring the two original muons to come from the same reconstructed vertex probably biases the distribution towards real primary vertices.



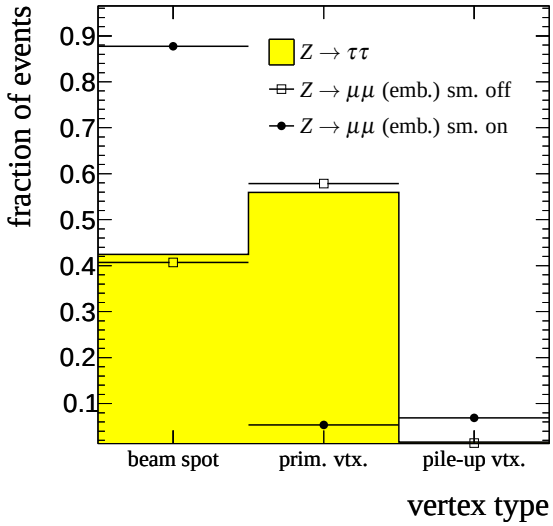
(a) Electrons.



(b) Muons.

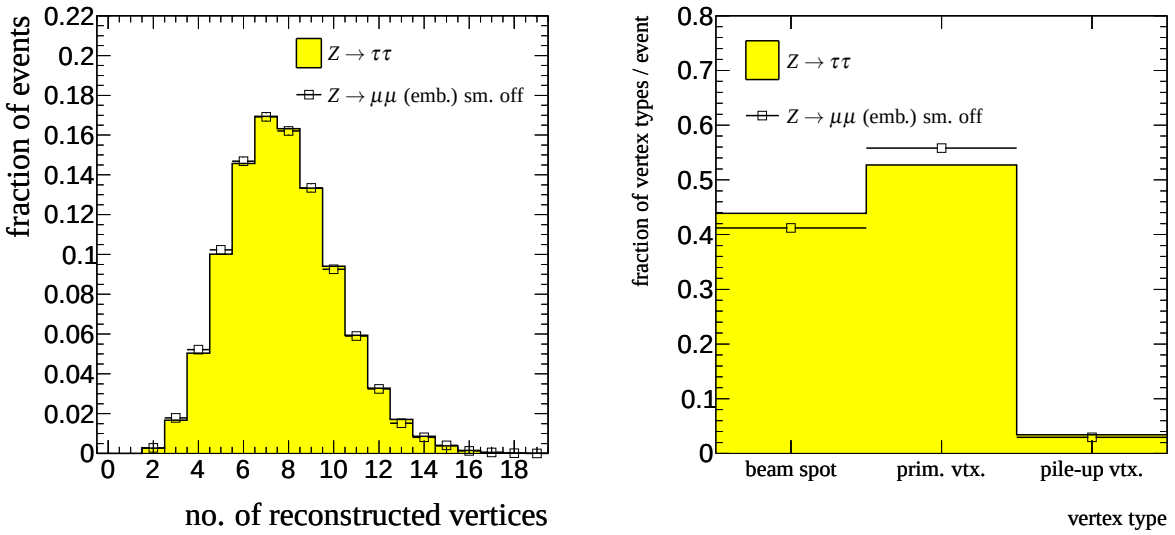


(c) Calorimeter-seeded τ_{had} candidates (leading track).



(d) Track-seeded τ_{had} candidates (leading track).

Figure 9.7.: Types of reconstructed primary vertices reconstructed objects from the $Z \to \tau\tau$ decay are associated with. Reconstructed objects are matched to the corresponding generator-level object. The embedded $Z \to \mu\mu$ sample is shown with vertex smearing on and off.



(a) Number of reconstructed primary vertices (pile-up).

(b) Types of reconstructed primary vertices to which track-seeded τ_{had} candidates point (pile-up).

Figure 9.8.: Primary vertex reconstruction in the presence of pile-up without vertex smearing.

9.1.4. Validation of the Method

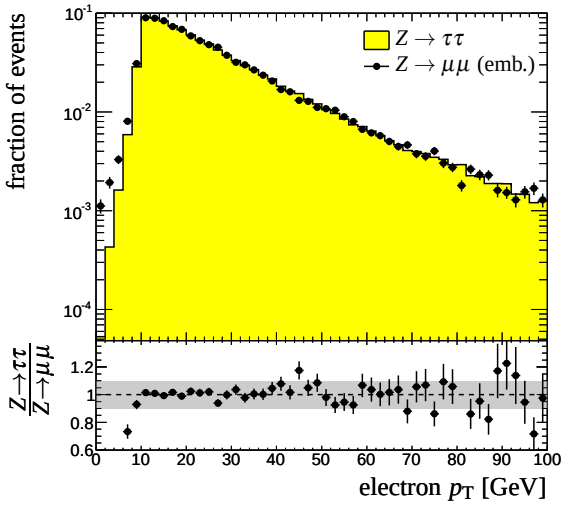
The discriminating variable for this analysis is $M_{\tau\tau}$. Accordingly, any method to obtain a control sample has to correctly reproduce the quantities from which $M_{\tau\tau}$ is calculated. These are (cf. Eqns. 4.5, 4.6 and 4.9):

- lepton/τ_{had} momenta: Figs. 9.9a-9.9c show p_{T} distributions of reconstructed τ decay products from the $Z \to \tau\tau$ decay. Reconstructed objects are matched to generated particles from the Z decay. No additional cuts are applied. The agreement between $Z \to \tau\tau$ MC and $Z \to \mu\mu$ after embedding is generally better than 10%. Low-p_{T} electrons have a higher efficiency in the embedded sample. As they are below the threshold for the analysis (cf. Sec. 7.1), the effect can be neglected.

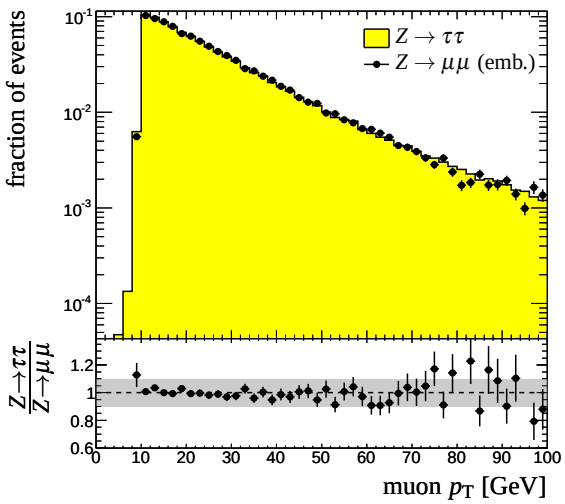
- missing transverse energy: Although the $\not{E}_{\text{T}}$-distribution of the embedded sample in Fig. 9.9d is slightly shifted, the deviation is below 10%.

- angles between above objects:
  Angles are reproduced correctly with the embedding procedure (Figs. 9.10a-9.10c). The agreement is, again, better than 10%, except for small angles between muon and τ_{had}. A deviation is visible for angles below 0.5.

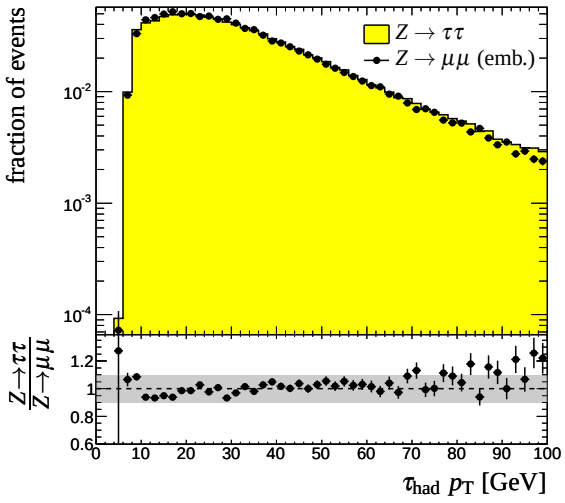
All observables which enter the $M_{\tau\tau}$ calculation are reproduced by the embedding procedure. The resulting $M_{\tau\tau}$ distribution, Fig. 9.10d, matches the one obtained from $Z \to \tau\tau$ with high precision.



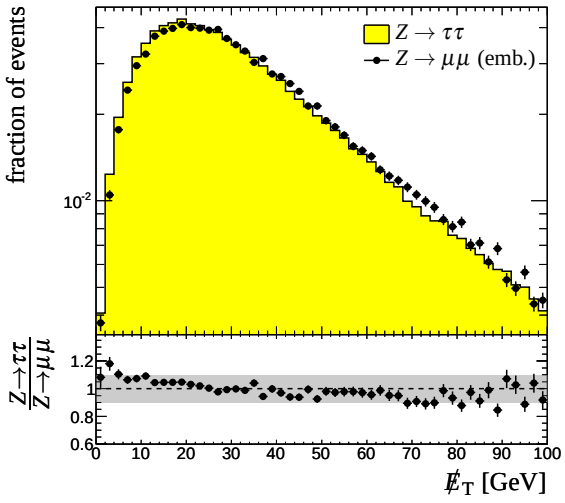
(a) p_T of reconstructed electrons from the $Z \to \tau\tau$ decay (matched to generated electrons).



(b) p_T of reconstructed muons from the $Z \to \tau\tau$ decay (matched to generated muons).

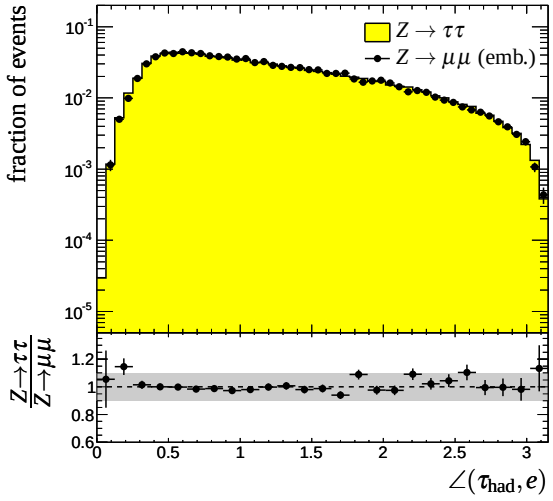


(c) p_T of reconstructed τ_{had} candidates, calorimeter-seeded, from the $Z \to \tau\tau$ decay (matched to the visible hadronic τ decay product on generator level)

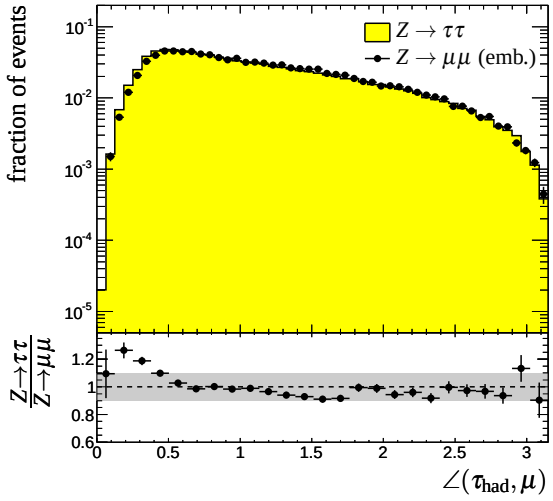


(d) MET RefFinal

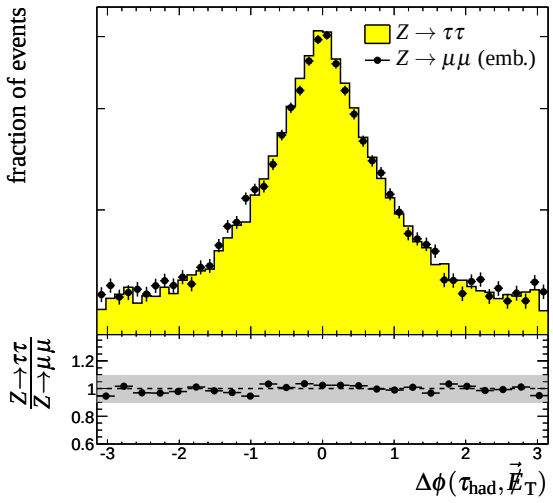
Figure 9.9.: Comparison of quantities which enter the calculation of $M_{\tau\tau}$.



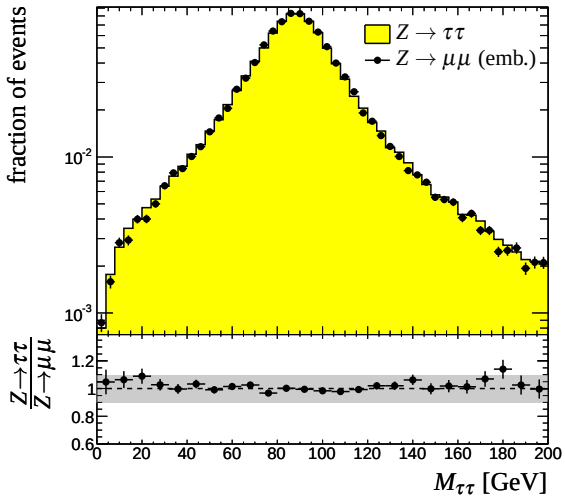
(a) Angle between reconstructed τ_{had} candidates and electrons from the $Z \to \tau\tau$ decay (objects matched to the corresponding generator-level objects).



(b) Angle between reconstructed τ_{had} candidates and muons from the $Z \to \tau\tau$ decay (objects matched to the corresponding generator-level objects).



(c) $\Delta\phi$ between reconstructed τ_{had} candidates from the $Z \to \tau\tau$ decay (matched to the visible hadronic τ decay product on generator level) and $\vec{E}_{\text{T}}$



(d) $M_{\tau\tau}$ calculated from reconstructed objects from the $Z \to \tau\tau$ decay (objects matched to the corresponding generator-level objects).

Figure 9.10.: Comparison of quantities which enter the calculation of $M_{\tau\tau}$. **(d)** shows a comparison of the resulting $M_{\tau\tau}$ distribution in embedded $Z \to \mu\mu$ events to $Z \to \tau\tau$ events.

cut	standard	pile-up
trigger	EF mu20	
muon charge	charge(μ_1)·charge(μ_2) = -1	
muon vertex	same vertex	
muon p_T	> 20 GeV	
muon $E_T^{\text{isol.}}$	< 3 GeV	< 4 GeV
$M_{\mu\mu}$	> 60 GeV	
$\not{E}_{\text{T}}$	< 30 GeV	< 50 GeV

Table 9.1.: Basic $Z \to \mu\mu$ selection as input for the embedding method.

9.1.5. $Z \to \mu\mu$ selection

Above sections show that, given a pure $Z \to \mu\mu$-sample, $Z \to \tau\tau$ distributions can be reproduced with sufficient accuracy. In a real experiment, the input sample has to be selected from data. The selection has to be:

- pure
- unbiased
- efficient

A high purity is mandatory because a significant admixture of other processes can be expected to change the resulting shapes. Special care has to be taken in order to avoid biases. The selected muons have to represent the true τ leptons. On the other hand, muon observables are the main tool to separate $Z \to \mu\mu$ from the background. Efficiency is important to ensure a sufficiently large control sample. In general, the size of a control sample should be larger than the expected number of events for the process in question. Less events result in larger statistical fluctuations which ultimately reduce the significance with which a signal process can be detected.

It should be noted that a full-fledged $Z \to \mu\mu$ analysis is beyond the scope of this thesis. The following selection sketches the basic parts of a proper $Z \to \mu\mu$ selection with respect to the processes studied here and its impact on embedding. The selection criteria employed in the $Z \to \mu\mu$ selection are summarised in Table 9.1 and further discussed in the following. The relevant distributions are shown in Appendix F.

Muons

$Z \to \mu\mu$ produces two high-energetic muons with opposite charge. For the embedding procedure, a cut on the p_{T} of muons from $Z \to \mu\mu$ decays directly defines the minimum p_{T} of τs in the $Z \to \tau\tau$ control sample. In a real $Z \to \tau\tau$ process there is no such direct threshold for the τs. Assuming the collinear approximation is valid, a cut on the p_{T} of the τ decay products poses an indirect cut on the τ p_{T}. The thresholds for muons, electrons and τ_{had} candidates in the $H \to \tau\tau$ analysis are 20, 25 and 30 GeV respectively (cf. Sec. 7.3). To avoid biases, a safe requirement for muons from $Z \to \mu\mu$ is therefore 20 GeV. Although muons (electrons) with $p_{\mathrm{T}} > 10(15)$ GeV are used in the analysis to remove overlapping reconstructed jets, the effect can be considered small since events with more than one identified muon or electron are discarded. The $\not{E}_{\mathrm{T}}$ requirement in the $H \to \tau\tau$ analysis effectively further increases the minimal τ p_{T}. Hence, if a harder cut is necessary, it will probably not have a large effect.

It is natural to require isolated muons. Several effects have to be taken into account. First of all, an isolation requirement can bias the p_{T} spectrum of the muons. This happens because the energy deposited in a cone around the muon is a function of the muon energy itself (see Sec. 7.1 for a more detailed discussion). Neither the absolute nor relative (i.e. E_{T} over p_{T} of the muon) E_{T} around the muon are flat in muon p_{T}. In addition, there is a more or less subtle interplay between isolation, the embedding procedure and the resulting control sample. Again, there is no direct restriction on the energy around a real τ in a real $Z \to \tau\tau$ process. There is an indirect one, manifest in the isolation requirements for the τ decay products. These however, are not identical. While for muons and electrons from a τ decay a cut on E_{T} around the particle is applied, the energy around a τ_{had} candidate enters the likelihood discriminant. Also, in a small cone around the original muon, the energy depositions in the calorimeter are replaced entirely. Thus, isolation is influenced by the size of this cone. On the other hand, isolation has a positive effect apart from the suppression of background: Muons can lose energy due to photon radiation. Hence, using reconstructed muons as τ leptons broadens the Z peak. Using isolated muons mitigates this effect.

As both muons come from the same mother particle it, makes sense to require them to be associated with the same reconstructed vertex - especially in the presence of pile-up. In addition, for TAUOLA to work correctly, a decay vertex needs to be assigned to the Z boson in each HEPEVT entry. This decay vertex should be identical for both decay products.

Z peak

The main discriminating observable for $Z \to \mu\mu$ is the invariant mass of the muon pair. Using a narrow window around the mass of the Z boson would discard most of the background. Unfortunately, cutting into the Z peak directly means changing the $M_{\tau\tau}$ distribution of the control sample. Therefore, the cut on $M_{\mu\mu}$ must be relatively loose. $M_{\mu\mu} > 60$ GeV keeps almost all $Z \to \mu\mu$ events, while still rejecting the vast majority of background processes. Background processes considered here do not contain events far above the Z peak, hence an

upper cut is not necessary. In reality the situation may be different, in which case $M_{\mu\mu} < 120\,\text{GeV}$ would be an adequate upper limit.

Missing energy

$\not{E}_{\text{T}}$ in $Z \to \mu\mu$ is almost exclusively an effect of the finite resolution of the $\not{E}_{\text{T}}$ reconstruction itself and, most prominent, of jets. A cut on $\not{E}_{\text{T}}$ suppresses $H \to \tau\tau$, $Z \to \tau\tau$ and $t\bar{t}$. But since jets affect $\not{E}_{\text{T}}$, cutting hard on $\not{E}_{\text{T}}$ will favour events with well-measured jets and $\not{E}_{\text{T}}$. One of the key advantages of the embedding procedure is that $\not{E}_{\text{T}}$ comes directly from the re-reconstruction and does not need to be corrected for. The $H \to \tau\tau$ analysis requires a minimum missing transverse energy and cannot reject events where jets contribute significantly to the $\not{E}_{\text{T}}$ measurement. A hard cut would therefore bias $\not{E}_{\text{T}}$ and destroy one of the main features of the embedding technique. $\not{E}_{\text{T}} < 30\,\text{GeV}$ is used for the $Z \to \mu\mu$ selection without pile-up, 50 GeV with pile-up.

Jets (VBF Signature)

Since the VBF signature has great separation-power and the cuts are applied in the $H \to \tau\tau$ selection anyway it is tempting to exploit the signature already during the $Z \to \mu\mu$ selection. The MC samples are filtered already on generator level (cf. Chapter 5). Beyond this, no further cuts are applied here. Harder cuts would very much reduce the number of events available for validation. In addition, cutting on jets in $Z \to \mu\mu$ would already decide at this stage which of the jets are going to be used as tagging jets in the $H \to \tau\tau$ selection. While this decision will be correct in most of the cases, it suppresses configurations in which a hadronic τ decay replaces one of the tagging jets and the actual τ_{had} candidate is a misidentified QCD jet. Given a proper τ_{had} identification the probability for this to happen is small. Nevertheless, such events occur (see Sec. 5.1), and the embedding technique should be able to reproduce them.

Composition of the Input Sample

The resulting composition of the selected $Z \to \mu\mu$ sample is shown in Table 9.2. With a not overly sophisticated selection a purity of above 90% is possible. The purity suffers from pile-up because some cuts had be loosened. The background mostly consists of $t\bar{t}$. Although 6.1% may seem problematic, it should be noted that – in contrast to all other non-vanishing background processes – the $t\bar{t}$ sample has no VBF filter applied. This means that the numbers in Table 9.2 are largely overestimated.

Unfortunately, the effects of such an impurity cannot be studied properly: Taking into account events with negative weight, the $t\bar{t}$ sample has roughly 700,000 effective events (twice that without pile-up). Only 10% of the sample is available in ESD format. Given the selection efficiency of 6.1%, this leaves roughly 4,000 events, most of which will fail proper VBF cuts. Hence, the control sample has to be assumed pure in the following.

process	no. of events	fraction
$Z \to \mu\mu$	$6.29(1) \times 10^5$	97.24%
$t\bar{t}$	$1.73(5) \times 10^4$	2.68%
$Z \to \tau\tau$	470(20)	0.07%
$W \to \mu\nu$	20(20)	$< 0.01\%$
$H \to \tau\tau$	7.0(5)	$< 0.01\%$

process	no. of events	fraction
$Z \to \mu\mu$	$6.17(1) \times 10^5$	93.85%
$t\bar{t}$	$4.0(1) \times 10^4$	6.05%
$Z \to \tau\tau$	500(20)	0.08%
$W \to \mu\nu$	110(60)	0.02%
$H \to \tau\tau$	10.3(7)	$< 0.01\%$

Table 9.2.: Expected number of events for an integrated luminosity of $30\,\text{fb}^{-1}$ after $Z \to \mu\mu$ selection without pile-up (left) and with pile-up (right).

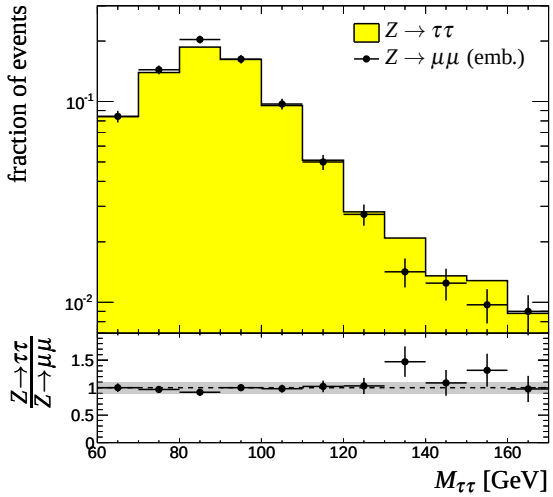
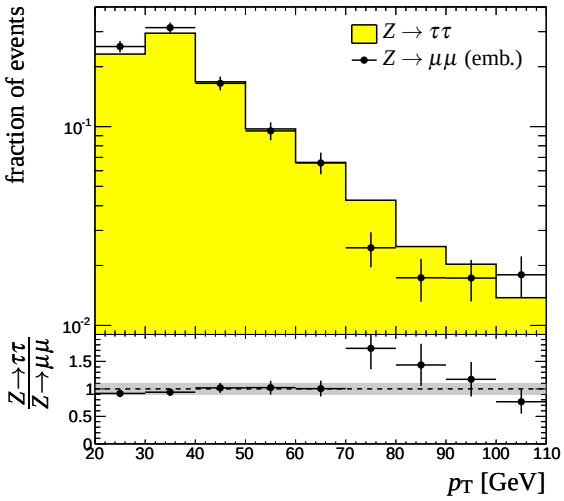


Figure 9.11.: Comparison of $M_{\tau\tau}$ in an embedded $Z \to \mu\mu$ sample (with muon selection cuts applied) to $M_{\tau\tau}$ in $Z \to \tau\tau$ events. The distributions are shown after the pre-selection. Both samples include pile-up.

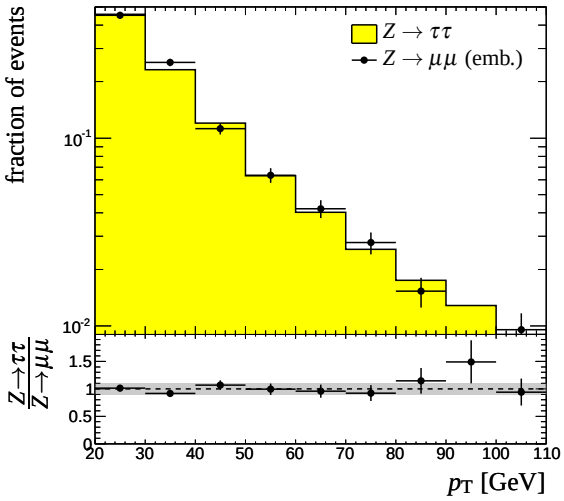
9.1.6. Validation of Distributions after $Z \to \mu\mu$ Selection

Since the purpose of this control sample is to estimate the $M_{\tau\tau}$ shape, the correct validation would be to compare this distribution after all cuts. Again, this is not possible due to the limited size of the $Z \to \mu\mu$ MC sample. However, $M_{\tau\tau}$ distributions after the pre-selection are in good agreement (Fig. 9.11).

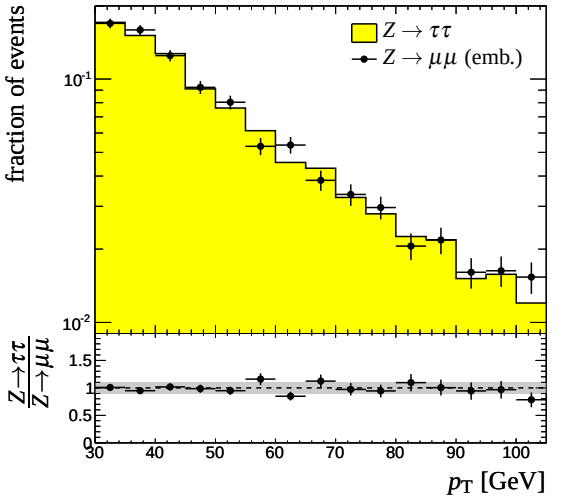
Figures 9.12a-9.16 show variables relevant for the cut analysis. For the most part, there is good agreement between the embedded $Z \to \mu\mu$ sample and $Z \to \tau\tau$. Most notably, some jet related observables (e.g. Figs. 9.14d and 9.15a) and the angle between lepton and τ_{had} (Fig. 9.14a) show disagreement. All of the mentioned observables depend on the number of jets in an event. Since the $Z \to \mu\mu$ events are produced in separate sub-samples depending on the number of jets (cf. Sec. 5.3), the reason might be a wrong weighting of the sub-samples after the embedding procedure.



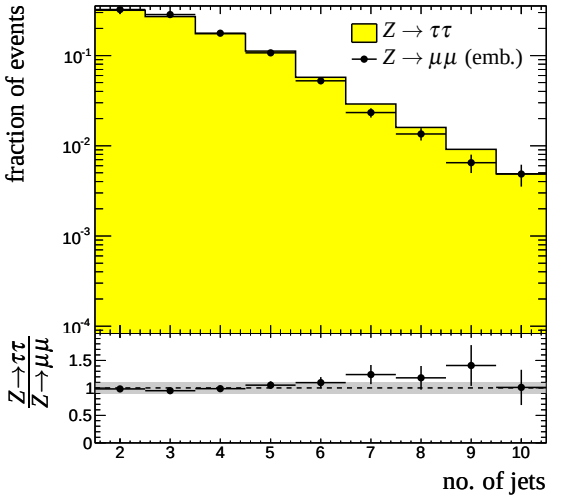
(a) p_T of reconstructed electrons.



(b) p_T of reconstructed muons.

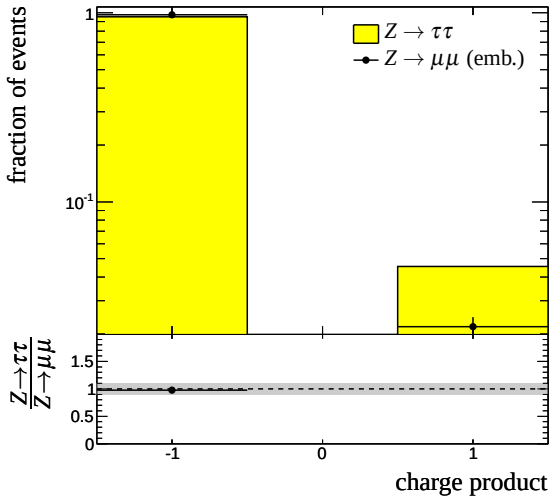


(c) p_T of τ_{had} candidates.

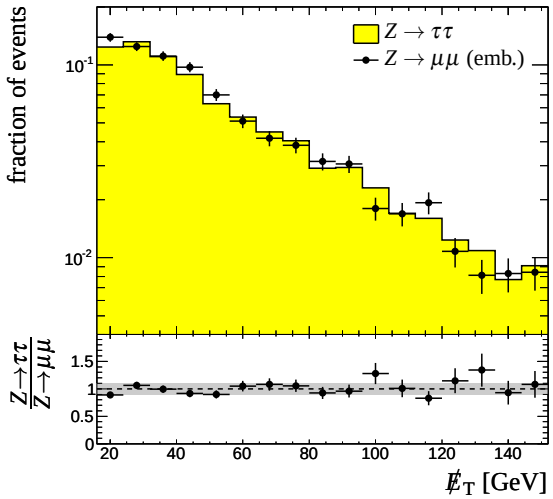


(d) Number of reconstructed jets.

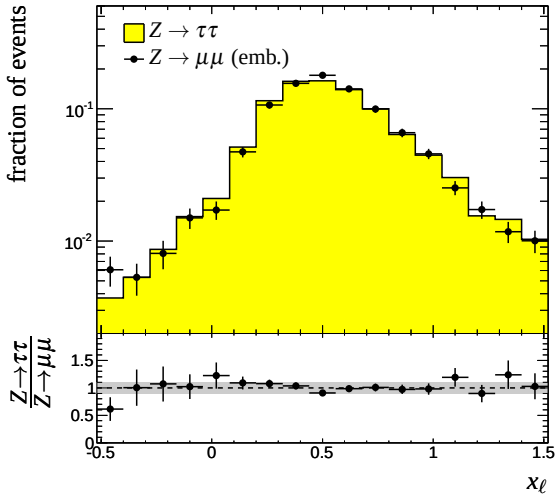
Figure 9.12.: Observables used in the $H \to \tau\tau \to \ell h$ cut selection. Both samples include pile-up. Distributions are shown after the pre-selection.



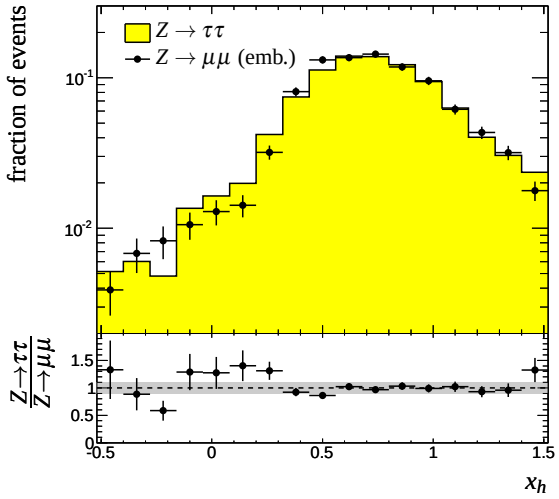
(a) Lepton-τ_{had} charge product.



(b) MET RefFinal.

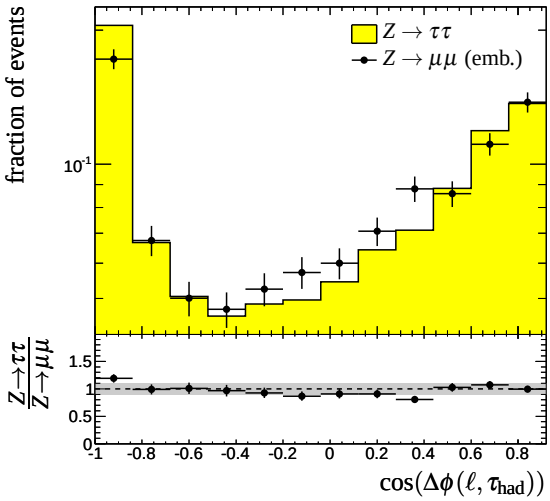


(c) x_ℓ.

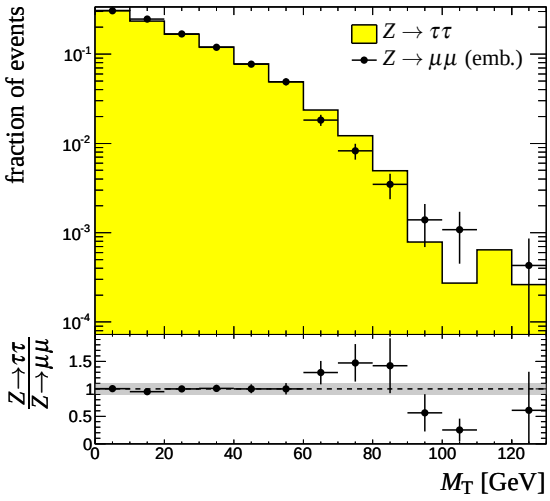


(d) x_h.

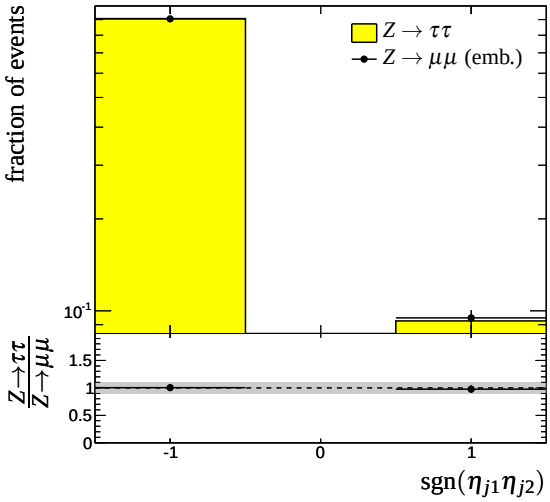
Figure 9.13.: Observables used in the $H \to \tau\tau \to \ell h$ cut selection. Both samples include pile-up. Distributions are shown after the pre-selection.



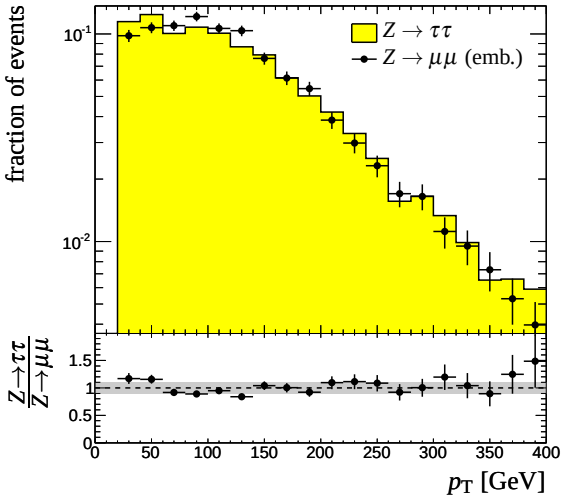
(a) $\cos(\Delta\phi(\ell, \tau_{\text{had}}))$.



(b) M_{T} after pre-selection.

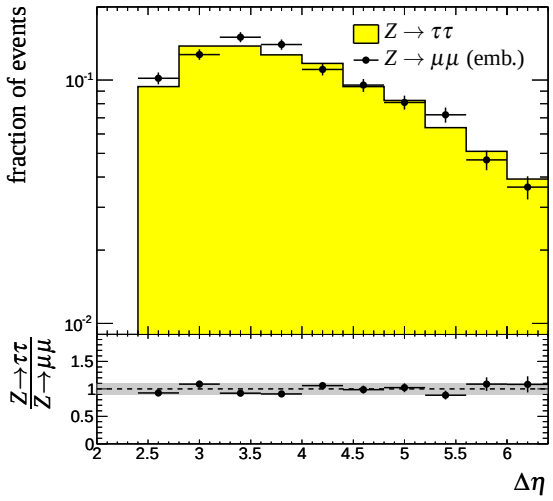


(c) Sign of η product of tagging jets.

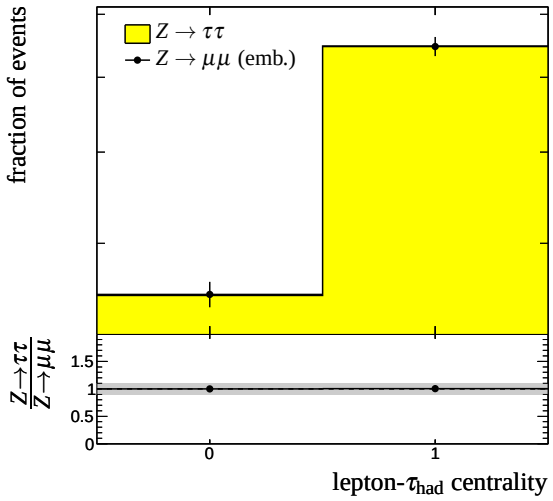


(d) p_{T} of leading jet.

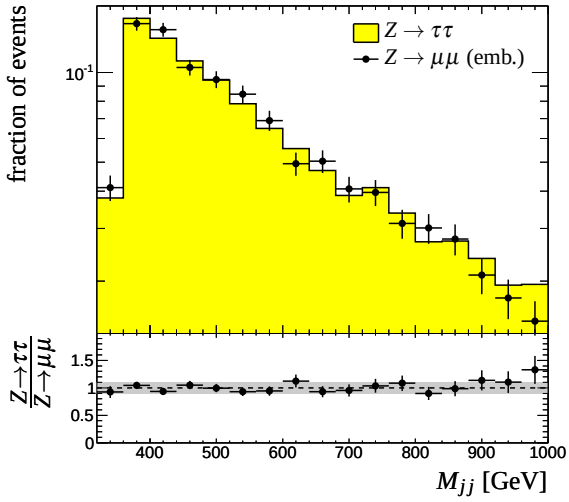
Figure 9.14.: Observables used in the $H \to \tau\tau \to \ell h$ cut selection. Both samples include pile-up. Distributions are shown after the pre-selection.



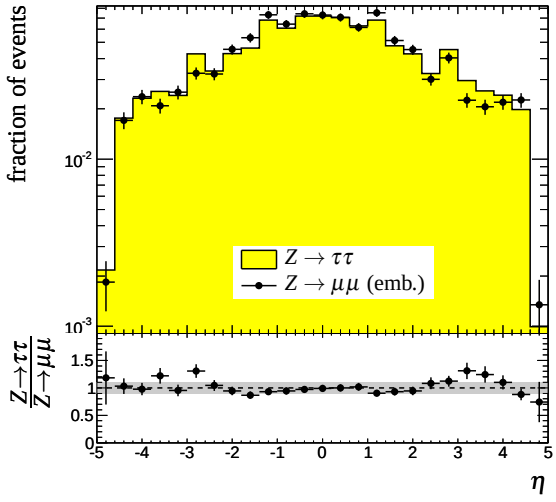
(a) $\Delta\eta$ between tagging jets.



(b) Lepton-τ_{had} centrality.



(c) Invariant mass of tagging jets.



(d) η distribution of additional jets after pre-selection.

Figure 9.15.: Observables used in the $H \to \tau\tau \to \ell h$ cut selection. Both samples include pile-up. Distributions are shown after the pre-selection.

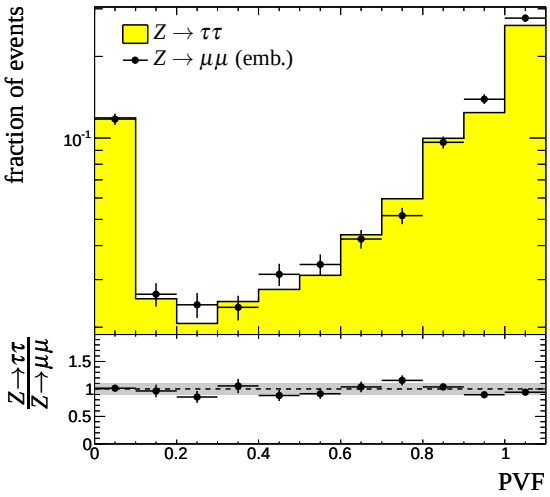


Figure 9.16.: Primary vertex p_T fraction of additional jets. Both samples include pile-up. Distributions are shown after the pre-selection.

9.1.7. Normalisation

The ratio of events in the $Z \to \tau\tau$ control sample to $Z \to \tau\tau$ events in the signal region, R_{emb}, has an upper limit which depends on the $\tau\tau$ branching ratio into final state in question. The total production cross sections $\sigma_{Z\to\mu\mu}$ and $\sigma_{Z\to\tau\tau}$ are approximately equal. Therefore the maximum value of R_{emb} in the ℓh channel this is given by

$$R_{\text{emb}}^{\text{max}} = \frac{\sigma_{Z\to\mu\mu}}{\sigma_{Z\to\tau\tau}\,\text{BR}(\tau\tau\to\ell h)} = \frac{1}{\text{BR}(\tau\tau\to\ell h)} \approx 2.2. \tag{9.1}$$

In general this maximum will not be reached because additional factors modify R_{emb}. ε_{sel} is the selection efficiency of $Z \to \mu\mu$. The selection reduces the number of available events to embed. ε_{cut} and $\varepsilon_{\text{cut,emb}}$ are the acceptances of the cut selection in $Z \to \tau\tau$ and the embedded $Z \to \mu\mu$ sample, respectively. They are usually not equal: $\varepsilon_{\text{cuts,emb}}$ is not independent of the $Z \to \mu\mu$ selection. A higher p_T threshold for muons in $Z \to \mu\mu$, for instance, leads to a higher average p_T of the decay products after embedding. For loose selection cuts the embedding sample should give the same cut efficiencies, except for one effect: Fig. 9.17 shows the expected number of events after each cut. As the ratio between the samples changes before cut number 5 (requiring exactly one identified τ_{had}) the efficiencies are obviously not equal. The $Z \to \mu\mu$ selection requires two reconstructed combined muons. These can only be reconstructed in the tracking region. As a result, the τ leptons in the embedded $Z \to \mu\mu$ sample are already within $|\eta| < 2.5$. This is not true for $Z \to \tau\tau$ (cf. Fig. 9.18). In both cases, the τ decay products have to be within $|\eta| < 2.5$, thus the lepton/τ_{had} pre-selection is more efficient in the embedded sample. So, part of the loss in statistics due to the $Z \to \mu\mu$ selection is recovered by a higher acceptance of the pre-selection. Also, reconstruction and identification efficiencies of the τ decay products are not necessarily equal since e.g. tracks are reconstructed in a cleaner environment and copied to the new hybrid event. Finally, ε_{emb} gives the efficiency

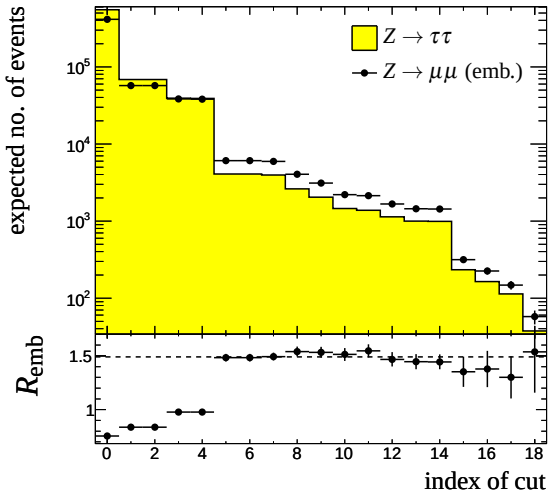


Figure 9.17.: Expected number of events after each stage of the cut selection for samples including pile-up using $C_{\text{in}} = 0.08$.

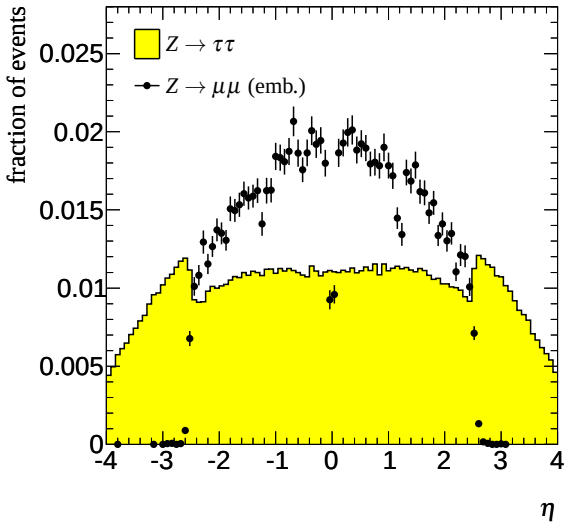


Figure 9.18.: η distribution of visible hadronic τ decay products on generator level.

of the embedding procedure, which can be smaller than one. If no generator-level τ lepton in the simulated $Z \to \tau\tau$ decay could be matched to an original muon (cf. Sec. 9.1.1), that event gets rejected. Using a sufficiently large matching cone, this can be neglected. In addition, in the version of the code used for this analysis, the last event per file cannot be processed. Given the usual number of events per file this effect is below the percent level. A consequence of the generator level filters used to produce the ALPGEN samples is that additional cuts have to be applied to the generated τ leptons and τ decay products of the simulated $\tau\tau$ decay to ensure equal starting conditions for both samples and obtain comparable distributions. The cuts lead to $\varepsilon_{\text{emb}} \approx 0.6$, but this factor gets entirely absorbed by a higher pre-selection acceptance. To summarise, R_{emb} is given by

$$R_{\text{emb}} = \frac{\sigma_{Z\to\mu\mu}}{\sigma_{Z\to\tau\tau}\,\text{BR}(\tau\tau\to\ell h)}\,\frac{\varepsilon_{\text{cuts,emb}}}{\varepsilon_{\text{cuts}}}\,\varepsilon_{\text{sel}}\,\varepsilon_{\text{emb}}. \tag{9.2}$$

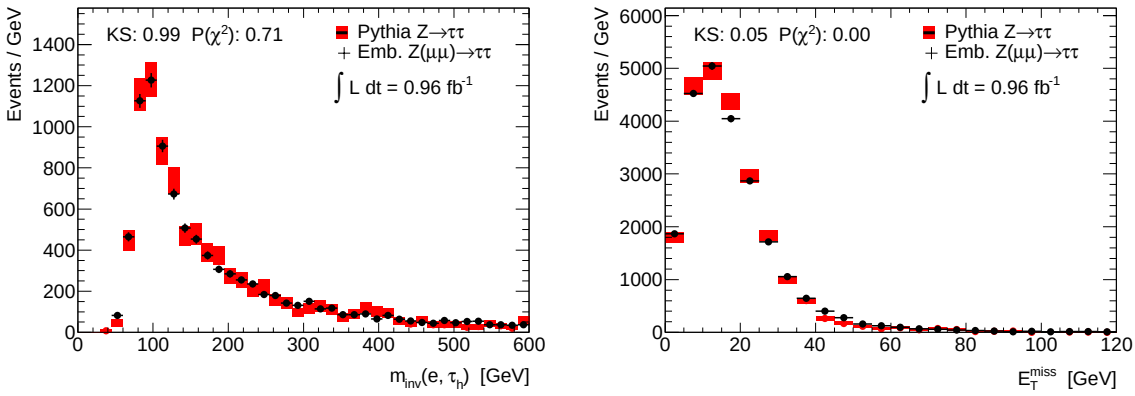
While ε_{emb} is a known quantity, $\varepsilon_{\text{cuts,emb}}$, $\varepsilon_{\text{cuts}}$ and ε_{sel} have to be estimated using MC, giving rise to a number of systematic uncertainties. Although prior knowledge of R_{emb} could be used to constrain the fit, for the signal extraction in this analysis it is only necessary to know the shape of $M_{\tau\tau}$ not the normalisation. It enters the significance calculation as an additional free parameter in the fit (cf. Chapter 10). With $\varepsilon_{\text{sel}} = 0.55$, $\varepsilon_{\text{cuts,emb}} = 1.43 \times 10^{-4}$ and $\varepsilon_{\text{cuts}} = 6.8 \times 10^{-5}$, a value of $R_{\text{emb}} \approx 1.5$ is obtained for pile-up. This factor is assumed for both luminosity scenarios in the significance calculation.

In principle a gain in statistics of another factor of 2 would be possible by producing eh and μh final states separately, using the same selected $Z \to \mu\mu$ events. Nevertheless, this gain is rather questionable because only the τ decay products would be different in both samples. In all other aspects the events would be correlated.

9.1.8. Conclusions

The embedding method to estimate $Z \to \tau\tau$ from $Z \to \mu\mu$ data is able to reproduce most relevant distributions with high precision. It is already successfully used in real-data analyses, e.g. [68]. On the technical side the framework is flexible enough to be employed in other use cases like the estimation of $t\bar{t}$ background in searches for charged Higgs bosons [78] or $W \to \tau\nu$ from $W \to \mu\nu$ [79]. As an outlook, Figure 9.19 shows to distributions from data collected in 2011.

Although the embedding method produces events which are very similar to real $Z \to \tau\tau$ events, these events - by construction - cannot be fully identical. One of the main reason is the fact that reconstructed muons are used as input objects to the simulation. The reconstruction efficiency is not independent of p_{T} and η. Most prominent is the gap in the muon spectrometer around $\eta = 0$, which leads to a dip in the η spectrum of τ decay products in the embedded sample (cf. Fig 9.18). As muon efficiencies can be estimated using tag and probe methods they do not need input from MC and could therefore be used to correct e.g. the η distribution. Reweighting events in the gap region, however, might lead to additional unwanted effects, because it is doubtful how trustworthy muons in this region really are. Also, muons are subject



(a) $M_{\tau\tau}$. **(b)** $\not{E}_{\mathrm{T}}$.

Figure 9.19.: Comparison of an embedded $Z \to \mu\mu$ sample selected in data from 2011 to simulated $Z \to \tau\tau$ events. By courtesy of Thomas Schwindt. The $M_{\tau\tau}$ distributions are in excellent agreement. Slight deviations can be observed in the $\not{E}_{\mathrm{T}}$ distribution, which might point to an insufficient modelling of $\not{E}_{\mathrm{T}}$ in the simulation.

to photon radiation, changing the original four-vector of the muon. While this can happen also in $Z \to \tau\tau$ events, double counting can occur if TAUOLA/PHOTOS produce an extra photon. Another source of deviation is the treatment of noise. The simulation of calorimeter noise is switched off for the $\tau\tau$ decay because otherwise it would be added on top of the already existing noise in the data event. On the other hand, this leads to a lack of noise within C_{in}, the small cone around the original muons. Finally, the resolution of vertex reconstruction, especially in z-direction, leads to imprecise vertex positions which are then used as true vertices for the $\tau\tau$ simulation.

Using simulated τ decay products to replace muons from collision data is a possible source of systematic uncertainties if the simulation does not sufficiently model τ lepton decays and the response of the detector to the decay products. This issue is discussed in Sec. 10.3.

9.2. Estimation of the Non-Resonant Background

The usual and well established – see e.g. [80] – way to select W boson events with high purity involves requiring a large enough transverse mass (cf. Sec. 5.2) and a good electron or muon. M_{T} is a powerful handle to suppress W background in the signal region and would be an ideal observable to construct a W control region. However, M_{T} is correlated to $M_{\tau\tau}$ as can be seen for instance by reformulating x_h (Eq. 4.6), which enters the calculation of $M_{\tau\tau}$ (Eq. 4.9):

$$x_h = \frac{1}{1 + \dfrac{\not{E}_{\mathrm{T}} \sin\Delta\phi(\vec{p}_{\mathrm{T},\ell}, \vec{\not{E}}_{\mathrm{T}})}{p_{\mathrm{T},h} \sin\Delta\phi(\vec{p}_{\mathrm{T},\ell}, \vec{p}_{\mathrm{T},h})}} = \frac{1}{1 + \dfrac{M_{\mathrm{T}}^2 \cot[\Delta\phi(\vec{p}_{\mathrm{T},\ell}, \vec{\not{E}}_{\mathrm{T}})/2]}{2 p_{\mathrm{T},h}\, p_{\mathrm{T},\ell} \sin\Delta\phi(\vec{p}_{\mathrm{T},\ell}, \vec{p}_{\mathrm{T},h})}} \tag{9.3}$$

	$H \to \tau\tau$	$Z \to \tau\tau$	$W+$jets	$t\bar{t}$
without pile-up				
di-jet mass	0.13(7)	4(2)	10(2)	10(2)
jet veto	0.13(7)	3(1)	3.6(9)	0.4(2)
with pile-up				
di-jet mass	0.2(1)	2(1)	22(6)	16(5)
jet veto	0.09(6)	0.5(5)	13(5)	4(2)

Table 9.3.: Expected number of events after cuts for an integrated luminosity of $30\,\text{fb}^{-1}$.

Unfortunately, this renders inverting the M_T requirement unsuitable for creating a control region in which to estimate $M_{\tau\tau}$. Instead, one can make use of the fact that in $W+$jets events all τ_had candidates are actually misidentified quark or gluon jets. The probability to mismeasure the charge of a real τ_had is expected to be small: 1.7%(3.6%) for 1(3)-prong candidates [1]. In an $H/Z \to \tau\tau$ event, lepton and τ_had will almost always have opposite charge. A simple example of a background event where this is not the case is the production of a W boson with one associated gluon jet. A τ_had candidate originating from that gluon will have a random charge, because the gluon itself has a charge of zero. Since the charge is random, it is not correlated to the kinematics of such a candidate. Hence, a control sample can be selected by requiring the charge of lepton and τ_had candidate to be equal. Such events are often referred to as *same-sign* events, denoted SS in the following. Events with oppositely charged lepton and τ_had are called opposite-sign events (OS).

As Table 9.3 shows, selecting SS events creates a sample that is virtually free of signal. However, a W boson with one additional gluon jet is an overly simplified example: The sample contains $t\bar{t}$ production in addition to $W+$jets as well as a smaller admixture of $Z \to \tau\tau$. Moreover, a W event needs to contain at least three high-p_T jets to be a valid background event. These jets consist of quark as well as gluon jets. Quarks have an electrical charge, so the naïve expectation is that jets originating from up quarks are preferentially reconstructed as τ^+_had candidates, while jets from down quarks generate τ^-_had candidates. Fig. 9.20 shows the distribution of initial partons associated with τ_had candidates in $\ell^+\tau_\text{had}$ events. The exact numbers should be regarded with some caution due the ambiguities inherent to matching generated partons to reconstructed objects. Nevertheless, the figure clearly confirms the simple picture in which the charge of the quark determines the charge of the τ_had candidate, whereas gluon jets induce no preference. In the example of W production with one associated jet, this leads to a correlation between the charge of the W boson and the charge of the τ_had candidate. With the requirement of one jet with $p_\text{T} > 20\,\text{GeV}$ and one lepton with $p_\text{T} > 10\,\text{GeV}$, both within $|\eta| < 2.5$, MCFM [81] predicts at LO the relative abundances of quarks and gluons in the final state of $W^+ +$jet events to be: 24.7% gluon, 22.6% $\bar{d}$ quark and 52.8% u quark. An excess of OS events with respect to SS events is the consequence.

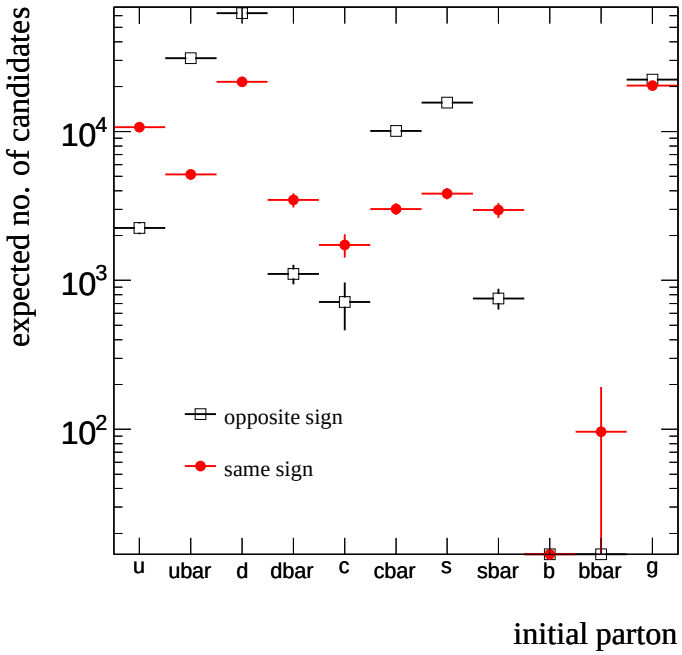
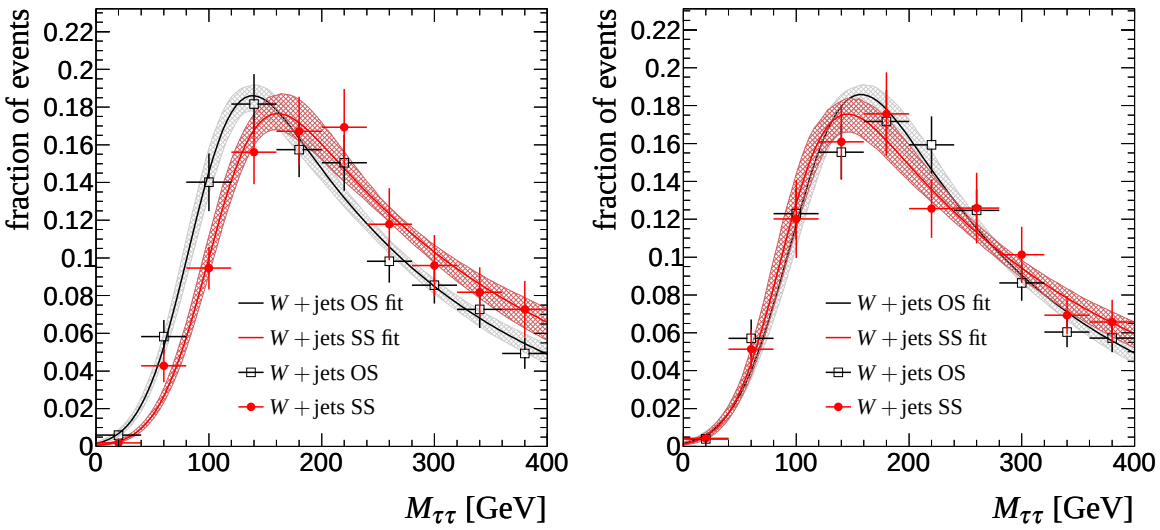


Figure 9.20.: Partons matched to τ_{had} candidates in $\ell^+\tau_{\text{had}}$ events after the pre-selection.

9.2.1. Validation

For the method to work, all processes of which the $M_{\tau\tau}$ distribution is to be estimated from the control region have to fulfil individually $M_{\tau\tau}$ (OS) = $M_{\tau\tau}$ (SS). Furthermore, they have to have either identical $M_{\tau\tau}$ shapes in the SS region or have the same ratio of OS to SS events. Figure 9.21 shows a comparison of the $M_{\tau\tau}$ distribution of $W+$jets in the signal and the SS control region. All distributions are well described by an exponential decay function convoluted with a Gaussian function (see Sec. 10.2. for details). In the no-pile-up scenario, however, $M_{\tau\tau}$ in the control region is shifted with respect to the signal region. It should be noted that the distributions are obtained by employing τ_{had} ID factorisation. This method does not perfectly model the lepton-τ_{had} charge product (cf. Sec. 6.3.1). Further investigation using a larger sample or, preferably, real data is necessary in order to establish if this is a real effect. It could be caused e.g. by the charge correlation between the W boson and additional quarks in the final state of the process. As described above, selecting SS events favours gluon jets. If gluon jets have different kinematics than quark jets in $W+$jets events, this can bias the $M_{\tau\tau}$ distribution. This hypothesis is supported by the absence of the effect in the pile-up scenario. The charge of additional jets or tracks produced by pile-up is not related to the charge of the W boson.

As stated above, the probability to mismeasure the charge of a real hadronic τ decay candidate is small. Hence, $t\bar{t}$ events passing the SS selection can be considered to contain almost only misidentified τ_{had} candidates. As shown in Chapters 7 and 8, this is not true in the signal region. Nevertheless, SS $t\bar{t}$ events can be used to estimate the $M_{\tau\tau}$ shape of $t\bar{t}$ production. In the signal region the $M_{\tau\tau}$ distribution of misidentified τ_{had} candidates and the full $M_{\tau\tau}$ distribution are in acceptable agreement given the statistical uncertainties (Fig. 9.22). Figure 9.23



(a) Without pile-up. **(b)** With pile-up.

Figure 9.21.: $M_{\tau\tau}$ shapes of $W+$jets in the SS control region and the signal region after pre-selection and lepton/$\not\!\!E_{\mathrm{T}}$ requirements. The shaded areas represent the 1σ uncertainties of the fit function. τ_{had} ID factorisation is used to increase the number of available events.

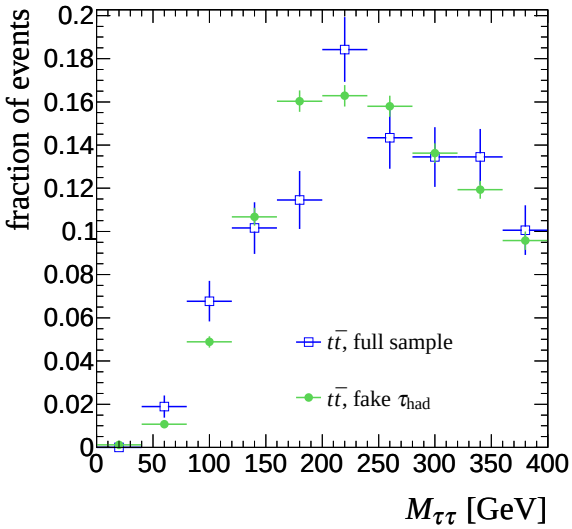
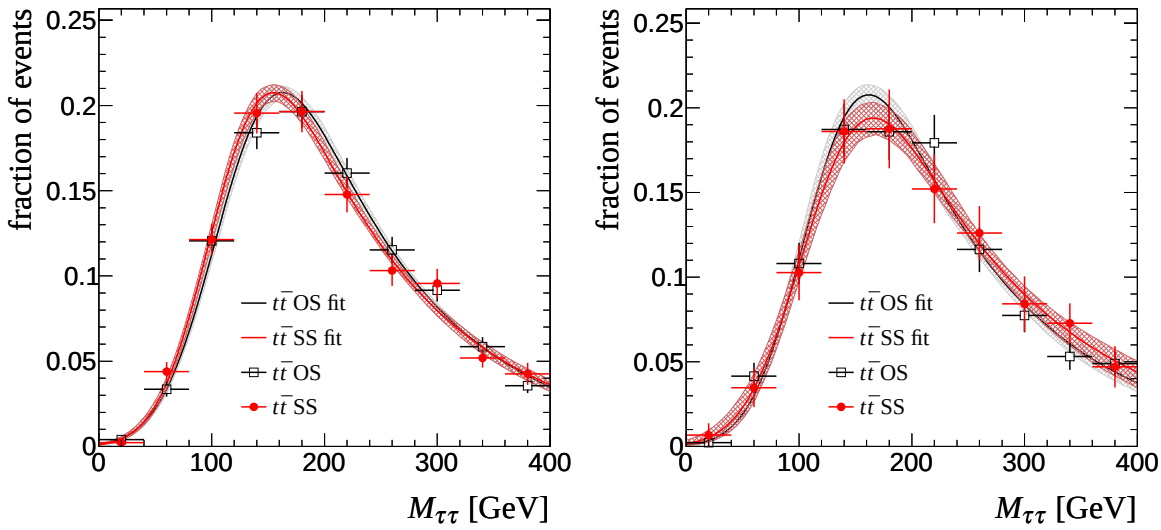


Figure 9.22.: $M_{\tau\tau}$ distributions of $t\bar{t}$ events in the signal region. The full $t\bar{t}$ sample is compared to a sample containing only misidentified τ_{had} candidates. τ_{had} ID factorisation is applied to the latter sample.

shows a comparison of $M_{\tau\tau}$ shapes in the signal and the control region for events with misidentified τ_{had} candidates. In both luminosity scenarios the distributions are in good agreement.



(a) Without pile-up. **(b)** With pile-up.

Figure 9.23.: $M_{\tau\tau}$ shapes of $t\bar{t}$ events with misidentified τ_{had} candidates in the SS control region and the signal region after pre-selection and lepton/$\not\!\!E_{\text{T}}$ requirements. The shaded areas represent the 1σ uncertainties of the fit function. τ_{had} ID factorisation is used to increase the number of available events.

The $M_{\tau\tau}$ distribution of $Z \to \tau\tau$ events in the SS control region is shown in Figure 9.24. Apart from the charge product requirement, the acceptance of the selection is higher for $Z \to \tau\tau$ events than for $t\bar{t}$ production or $W+\text{jets}$. This leads to a significant contribution of $Z \to \tau\tau$ events with real τ_{had} candidates in the control sample. The resulting $M_{\tau\tau}$ shape is a superposition of the Z peak and a non-resonant part similar to that of $W+\text{jets}$ or $t\bar{t}$ events. Fig. 9.24 shows that this superposition can be modelled by adding the $Z \to \tau\tau$ shape obtained from the signal region on top of the $M_{\tau\tau}$ distribution of $Z \to \tau\tau$ events with misidentified τ_{had} candidates in the control region. The former is provided by the embedding method. A comparison of the latter to a combination of $W+\text{jets}$ and $t\bar{t}$ events is shown in Figure 9.25. $Z \to \tau\tau$ is clearly shifted with respect to $W+\text{jets}$ and $t\bar{t}$ production. Nevertheless, the total number of expected $Z \to \tau\tau$ events and their relative contribution to the control sample is small enough that this can be considered a reasonable approximation.

The combination of SS events from $t\bar{t}$ production, $W+\text{jets}$ and the non-resonant part of $Z \to \tau\tau$ to $t\bar{t}$ and $W+\text{jets}$ in the signal region is shown in Figure 9.26. Despite the discrepancies discussed above, these are in good agreement for both luminosity scenarios.

A possible contamination of the SS region with QCD multi-jet production cannot be investigated in the scope of this analysis. If the charge of lepton and τ_{had} candidates in multi-jet production is totally random, one can expect a ratio of OS to SS events of one. Thus, a significant amount of multi-jet events in the control region implies that multi-jet production is also a significant background in the signal region. The $M_{\tau\tau}$ shape of non-resonant background processes is presumably a result of an exponential decay cut off at lower values of $M_{\tau\tau}$ due to

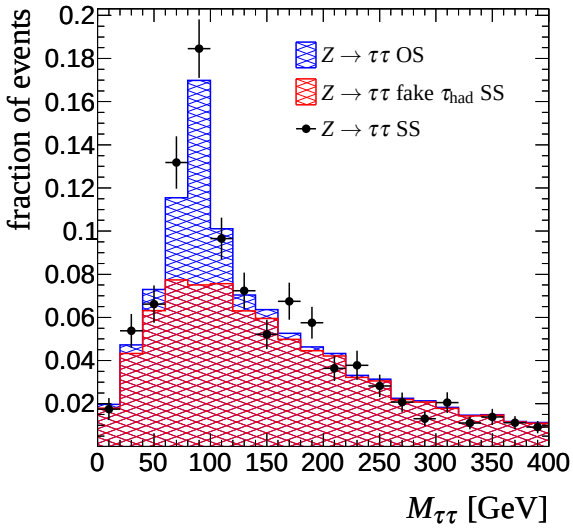
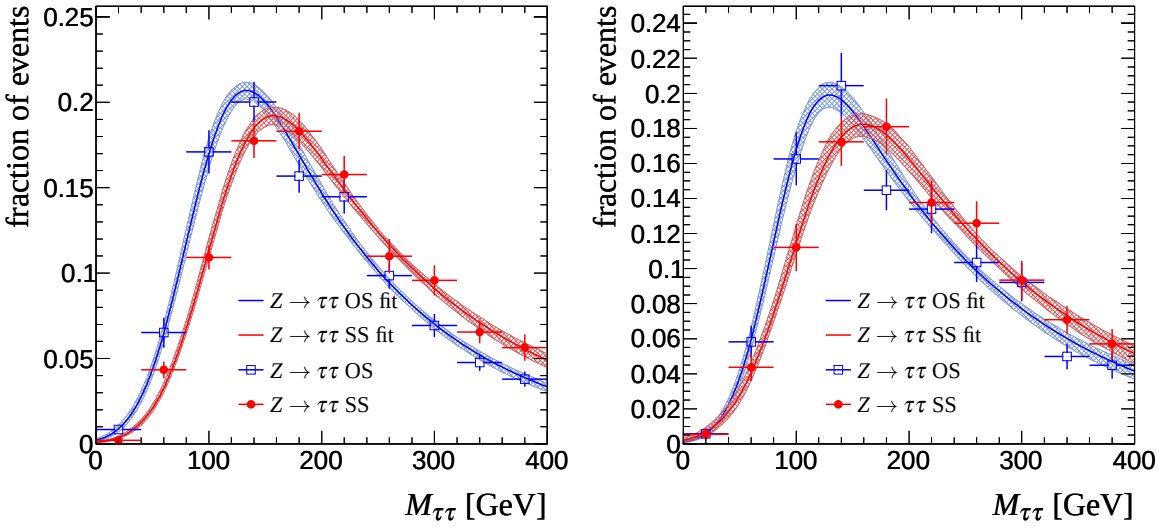
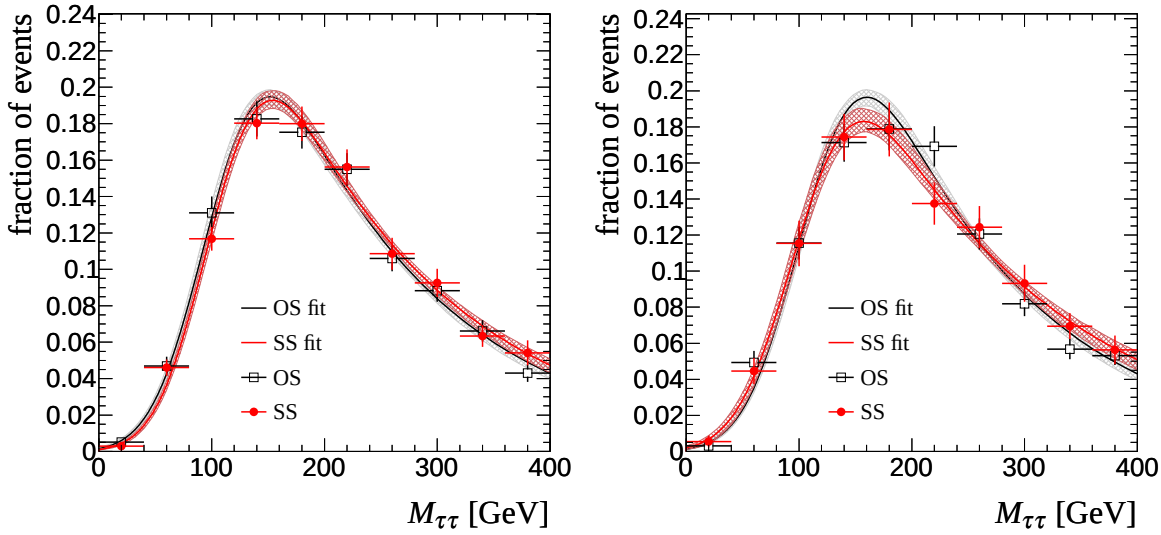


Figure 9.24.: The $M_{\tau\tau}$ shape of $Z \to \tau\tau$ events in the SS control region. The shape can be modelled by adding the $Z \to \tau\tau$ OS events to the shape obtained from $Z \to \tau\tau$ events in which the τ_{had} candidate is a misidentified jet.



(a) Without pile-up. **(b)** With pile-up.

Figure 9.25.: $M_{\tau\tau}$ shapes of $Z \to \tau\tau$ events with misidentified τ_{had} candidates in the SS control region and the signal region after pre-selection and lepton/$\not{E}_T$ requirements. The shaded areas represent the 1σ uncertainties of the fit function. τ_{had} ID factorisation is used to increase the number of available events.



(a) Without pile-up. **(b)** With pile-up.

Figure 9.26.: Combined $M_{\tau\tau}$ shapes in the SS control region ($t\bar{t}$, $W+$jets, $Z \to \tau\tau$) and the signal region ($t\bar{t}$ and $W+$jets) for events with misidentified τ_{had} candidates. The shaded areas represent the 1σ uncertainties of the fit function. τ_{had} ID factorisation is used to increase the number of available events.

the selection. Therefore, multi-jet production can be expected to have a similar $M_{\tau\tau}$ shape as the processes discussed above. Hence, it can be integrated into the method. This, in fact, is demonstrated in [68].

9.2.2. Normalisation

Similar to the embedding method to estimate $Z \to \tau\tau$ from data, this method does not provide an absolute normalisation. As discussed above, there is a charge correlation between the W boson and quark jets, leading to an excess of OS events. The exact ratio of OS events to SS events depends on details of the τ_{had} identification. The identification efficiency is different for quark and gluon jets. Thus, changing the working point of the τ_{had} identification can change the ratio of OS events to SS events. Furthermore, the efficiency is process dependent. A change in the relative abundance of $t\bar{t}$ and $W+$jets events in the control region – e.g. by using a different selection – will have an impact the OS to SS ratio. Finally, the relative survival probability of events with different jet multiplicity has an effect on the ratio. $W+$jets events with a higher number of jets have a different admixture of quark and gluon jets which can be identified as τ_{had} candidates. Changing, e.g., the configuration of the jet veto would favour lower jet multiplicities. A method to obtain the ratio from data and additional input from MC is demonstrated in [68].

In this analysis the normalisation of the SS control samples is an additional free parameter of the fit to calculate the signal significance. See Chapter 10 for details. Since the number

of expected events in the control region after the full selection is expected to be small, the jet veto is omitted in the selection of SS events.

9.2.3. Conclusion

A control sample constructed by selecting events in which lepton and τ_{had} have the same charge can produce an $M_{\tau\tau}$ distribution which is in good agreement with that of a combination of $W+$jets and $t\bar{t}$ events in the signal region. In conjunction with the embedding technique to estimate $Z \to \tau\tau$ all major background processes to VBF $H \to \tau\tau \to \ell h$ can be estimated from data.

10

Significance Calculation

10.1. Profile Likelihood Method

A proton-proton collision measured by a detector can never be unambiguously identified as a specific process. Therefore, establishing a signal in particle physics is a statistical statement: a signal is found if the observed data is statistically incompatible with the *background-only hypothesis* H_0. The threshold beyond which data is incompatible to H_0 is in principle an arbitrary value. It is quantified by the so-called *p-value* which gives the probability to find data, e.g. in a repeated experiment, that is equally or less compatible with H_0. A simple interpretation of the *p*-value is the one-sided tail of a Gaussian distribution: the *p*-value is the probability with which a Gaussian distributed random variable, e.g. the expected number of background events, fluctuates such that it is found in the tail area equal to the *p*-value. Hence, in particle physics the *signal significance Z* is the *p*-value measured in units of standard deviations of a Gaussian function [60]:

$$Z = \Phi^{-1}(1-p) \tag{10.1}$$

Here Φ^{-1} is the inverse cumulative distribution of the Gaussian function. It is commonly agreed to claim a discovery at a *Z*-value of $Z \geq 5$ or $p \leq 2.87 \times 10^{-7}$.

In this analysis the actual hypothesis test is done using the *profile likelihood method* [60]. The so-called profile likelihood ratio λ is defined as:

$$\lambda_0 = \frac{L(f_s = 0, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{f}_s, \hat{\boldsymbol{\theta}})}. \tag{10.2}$$

The likelihood function L reflects the conditional probability to observe the actual data given that the parameters $\boldsymbol{\theta}$ are realised in nature. f_s is the signal strength parameter, which here equals the ratio of signal events to the total number of observed events. $\boldsymbol{\theta}$ denotes the other

parameters of the probability density function (PDF) taken to describe the data (cf. Sec. 10.2). They are called *nuisance parameters* because they are considered to be of no direct interest for the measurement. $\hat{f}_s$ and $\hat{\boldsymbol{\theta}}$ are the maximum likelihood estimators, i.e. the set of parameters that maximise the full likelihood. The parameters $\hat{\hat{\boldsymbol{\theta}}}$ maximises the likelihood when assuming the data to consist of background only ($f_s = 0$).

Given an estimator q, the p-value of the background-only hypothesis for a measurement with an observed q_{obs} can be calculated as

$$p = \int_{q_{\text{obs}}}^{\infty} f(q|0)\,\mathrm{d}q, \tag{10.3}$$

where $f(q|0)$ is the *sampling distribution* of q, i.e. the distribution of q for repeated experiments given that H_0 is true. The sampling distribution can be a simple Poissonian distribution. In general, however, $f(q|0)$ can be more complicated and has to be obtained, e.g., by performing pseudo experiments. Unfortunately, the number of pseudo experiments needed for a discovery experiment is very large and thus their generation is very (CPU-)time consuming. An advantage of the profile likelihood method is the fact that, for a sufficiently large sample size, the sampling distribution is known beforehand. The estimator q_0 is defined as:

$$q_0 = \begin{cases} -2\ln\lambda_0 & \hat{f}_s \geq 0 \\ 0 & \hat{f}_s < 0 \end{cases}. \tag{10.4}$$

While in general a deviation from H_0 with $\hat{f}_s < 0$ could be interpreted as the presence of a signal which destructively interferes with the background, this definition reflects the assumption that the presence of a Higgs boson always leads to an excess of events. Using Wilks' theorem [82], the sampling distribution of q_0 can be approximated by the limit distribution of q_0, a χ^2 distribution with one degree of freedom plus a δ function:

$$f(q_0|0) = \frac{1}{2}\delta(q_0) + \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{q_0}}\exp(-q_0/2). \tag{10.5}$$

The δ function is a consequence of the definition of q_0: measurements with $\hat{f}_s < 0$ are set to zero. Since in a background-only experiment upward fluctuations of $\hat{f}_s$ are equally likely as downward fluctuations, the δ function is assigned the same weight as the χ^2 distribution. The sampling distribution is thus known. Moreover, the cumulative distribution of $f(q_0|0)$ is simply the cumulative distribution of a Gaussian function of $\sqrt{q_0}$. Hence, according to Eq. 10.3 the p-value is

$$p_0 = 1 - \Phi(\sqrt{q_0}), \tag{10.6}$$

corresponding to a signal significance

$$Z = \sqrt{q_0}. \tag{10.7}$$

The sensitivity of an experiment can be characterised by the expected signal significance and the statistical variation of actual measurements around the expected significance. Thus, one is interested not only in the distribution of q_0 in background-only experiments, $f(q_0|0)$, but also in the distribution of q_0 in experiments with a true signal strength parameter f_s' different from zero, $f(q_0|f_s')$. In the limit of an infinite sample size the estimator q_0 is related to the measured signal strength parameter in a simple way [83],

$$q_0 = \frac{\hat{f}_s^2}{\sigma^2}, \tag{10.8}$$

where $\hat{f}_s$ follows a Gaussian distribution with mean f_s' and standard deviation σ. Hence, $\sqrt{q_0}$ follows a Gaussian distribution with a unit standard deviation. According to Eq. 10.7, the significance becomes:

$$Z = \sqrt{q_0} = \begin{cases} \dfrac{\hat{f}_s}{\sigma} & \hat{f}_s \geq 0 \\ 0 & \hat{f}_s < 0 \end{cases}. \tag{10.9}$$

The expected or median significance, $\text{med}[Z|f_s'] = f_s'/\sigma$, can be obtained from the so-called *Asimov sample* [60]. Asimov data corresponds to the signal plus background hypothesis with infinite statistics and thus is equal to the expected number of events per bin. The measured signal strength parameter $\hat{f}_s$ is then identical to the true value f_s'. In this thesis Asimov data is approximated by using binned data with bin entries obtained from the parametrisation described in Sec. 10.2. With the estimator q_0 obtained from Asimov data $q_{0,\text{A}}$, the median significance and the error bands corresponding to a variation of $\hat{f}_s$ of $\pm N\sigma$ are (Eq. 10.9):

$$Z(f_s') = \sqrt{q_{0,\text{A}}} \tag{10.10}$$

$$Z(f_s' + N\sigma) = \sqrt{q_{0,\text{A}}} + N \tag{10.11}$$

$$Z(f_s' - N\sigma) = \max[\sqrt{q_{0,\text{A}}} - N, 0] \tag{10.12}$$

10.2. Modelling the Input Distributions

To build the likelihood function the data PDF must be modelled. Such a parametrisation is by necessity a trade-off between an accurate description of the distribution and a limitation of the number of free parameters. Since given a sufficient number of free parameters it is always possible to fit any distribution, too many free parameters reduce the sensitivity to the signal.

Signal: The central part of the signal $M_{\tau\tau}$ shape is nearly Gaussian, especially without pile-up. A single Gaussian function, however, does not sufficiently describe the tails and the asymmetry (cf. Fig. 8.15a). Hence, a double "asymmetric" Gaussian function (*AG*) is

used to parametrise the signal process (Fig. 10.1a),

$$\mathrm{AG}(x,\vec{p}) = \frac{1}{2}+\frac{1}{2}\cdot\mathrm{Erf}\left(\frac{x-p_1}{\sqrt{2}p_2}\right)\cdot\exp\left(-\frac{1}{2}\frac{(x-p_3)^2}{p_4^2}\right) \tag{10.13}$$

$$\mathrm{PDF}_{H\to\tau\tau}(\boldsymbol{\theta}_H) = p_6\,\mathrm{AG}(x,p_1,\ldots,p_4)+(1-p_6)\,\mathrm{AG}(x,p_1,\ldots,p_3,p_5), \tag{10.14}$$

where $\boldsymbol{\theta}_H=(p_1,\ldots,p_6)$. Both asymmetric Gaussian functions share a common mean, p_3, and the error function (p_1 and p_2). Only the widths of the two Gaussian functions, p_4 and p_5, are different, leading to six free parameters in total.

$Z\to\tau\tau$: The $M_{\tau\tau}$ shape of $Z\to\tau\tau$ is parametrised by the same function as the signal process (Fig. 10.1b), with $\boldsymbol{\theta}_Z=(p_7,\ldots,p_{12})$. The same PDF is used to describe the $Z\to\tau\tau$ control sample.

$$\mathrm{PDF}_{Z\to\tau\tau}(\boldsymbol{\theta}_Z) := \mathrm{PDF}_{H\to\tau\tau}(\boldsymbol{\theta}_Z) \tag{10.15}$$

Non-resonant: The $M_{\tau\tau}$ shape of the non-resonant background is actually determined by an exponentially falling distribution of the visible mass $M_{\ell\tau_{\mathrm{had}}}$, which is made to peak due to p_{T} thresholds for lepton and τ_{had}, as well as by a lower limit on $\not{E}_{\mathrm{T}}$ and by cuts on x_ℓ and x_h. To parametrise this distribution an exponential decay function is used, convoluted with a Gaussian resolution function (Fig. 10.1c):

$$\mathrm{PDF}_{\text{non-res}}(\boldsymbol{\theta}_{\mathrm{NR}}) = \exp\left(-\frac{x}{p_{13}}\right)\otimes\exp\left(-\frac{1}{2}\frac{(x-p_{14})^2}{p_{15}^2}\right). \tag{10.16}$$

The SS control sample to estimate the non-resonant background additionally includes $Z\to\tau\tau$. While $Z\to\tau\tau$ events with misidentified jets as τ_{had} candidates can be described by the same shape as $W+$jets and $t\bar{t}$, $Z\to\tau\tau$ events with a real τ_{had} but mismeasured charge follow the same distribution as OS $Z\to\tau\tau$ events. The parametrisation for the control sample therefore has an additional $Z\to\tau\tau$ component (Fig. 10.1d):

$$\mathrm{PDF}_{\text{non-res,SS}}(\boldsymbol{\theta}_{\mathrm{NR}},\boldsymbol{\theta}_Z) = (1-f_Z)\,\mathrm{PDF}_{\text{non-res}}+f_Z\,\mathrm{PDF}_{Z\to\tau\tau} \tag{10.17}$$

The parameters $\boldsymbol{\theta}_Z$ are the same as in the $Z\to\tau\tau$ parametrisation. f_Z represents the relative contamination of the SS control sample with resonant $Z\to\tau\tau$ background.

The combined data PDF is the sum of $H\to\tau\tau$, $Z\to\tau\tau$ and the non-resonant component,

$$\mathrm{PDF}_{\mathrm{data}} = (1-f_s)\,[f_R\,\mathrm{PDF}_{Z\to\tau\tau}+(1-f_R)\,\mathrm{PDF}_{\text{non-res}}]+f_s\,\mathrm{PDF}_{H\to\tau\tau}, \tag{10.18}$$

where f_R is the ratio of the resonant $Z\to\tau\tau$ background over the total background. The signal parameters $\boldsymbol{\theta}_H$ are fixed in order to avoid the *look-elsewhere effect* (see e.g. [84]). The data PDF has $6+3+2=11$ free parameters, six from the $Z\to\tau\tau$ PDF, three from the non-resonant PDF and another two parameters, f_s and f_R for the normalisation. The input distributions after all cuts are very much determined by statistical fluctuations due to the limited sample sizes. For the signal and the $Z\to\tau\tau$ distribution the lepton-τ_{had} centrality, $\eta_{j1}\times\eta_{j2}<0$ for the

tagging jets, the upper limit on M_{T} and the jet veto are omitted from the selection criteria. These have only a minimal correlation with $M_{\tau\tau}$ in both samples. For all samples contributing to the non-resonant distribution all requirements after the upper limit on M_{T} are omitted. The signal significance is calculated assuming that, apart from statistical fluctuations, the shape of the $M_{\tau\tau}$ distributions of $Z \to \tau\tau$ and the embedding control sample, as well as the non-resonant component of the signal and the SS control region are identical.

The likelihood L is a product of three components, one for the signal region obtained by the selection described in Chapters 7 and 8 and two for the background control regions described in Chapter 9. It depends on the $M_{\tau\tau}$ measurements in the three regions and the parameters of the PDFs.

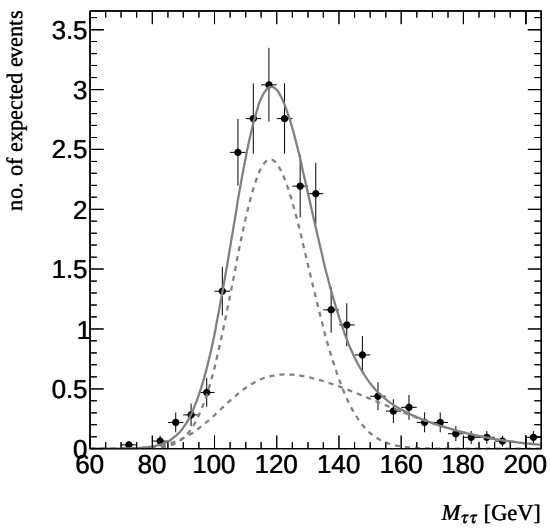
$$\begin{aligned} L(M_{\tau\tau}|\, f_s, \boldsymbol{\theta}) &= L_{\text{signal region}}\,(M_{\tau\tau}(\text{signal region})|\boldsymbol{\theta}_H, \boldsymbol{\theta}_Z, \boldsymbol{\theta}_{\text{NR}}, f_s, f_R) \\ &\times\ L_{Z\to\tau\tau}\,(M_{\tau\tau}(\text{embedding})|\boldsymbol{\theta}_Z) \\ &\times\ L_{\text{non-resonant}}\,(M_{\tau\tau}(\text{SS})|\boldsymbol{\theta}_Z, \boldsymbol{\theta}_{\text{NR}}, f_Z) \end{aligned} \tag{10.19}$$

The likelihood components of the two control regions share parameters of the PDFs used in the signal region. The control regions thus act as additional measurements which constrain the parameters of $\text{PDF}_{Z\to\tau\tau}$ and the non-resonant PDF, $\boldsymbol{\theta}_Z$ and $\boldsymbol{\theta}_{\text{NR}}$. The way the likelihood is built reveals another feature of the profile likelihood method that makes it very suited for this analysis: The two methods for background estimation provide only shapes. But more knowledge is not necessary as PDFs are normalised to unity by definition and the normalisation is a free parameter of the fit.

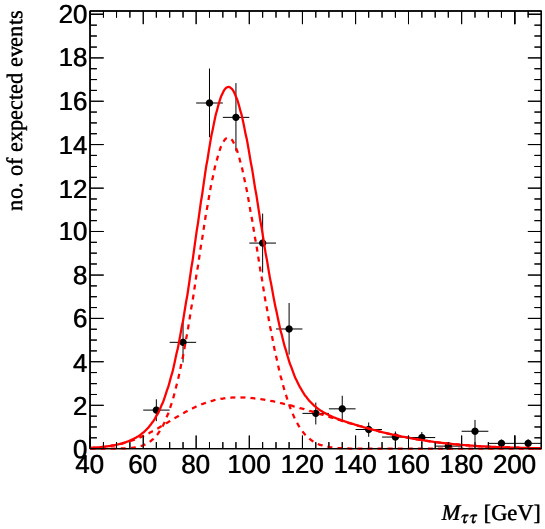
10.3. Systematic Uncertainties

The prediction of the expected signal significance differs from the calculation of the significance of an actual measurement in that different uncertainties have to be considered. Systematic uncertainties here have much the same meaning as in real experiments. They are the consequence of an imprecise knowledge of theoretical or experimental parameters which does not vanish in the limit of infinite data. They are often introduced if Monte-Carlo simulation is used to estimate aspects of the experiment. Common examples are the energy measurements of jets or the theoretical uncertainty in the cross section of a process. Another example is the estimation of processes from data which has to be validated by using Monte-Carlo simulation. The lack of knowledge has to be taken into account when calculating the significance and, in contrast to statistical uncertainties, can only decrease the significance.

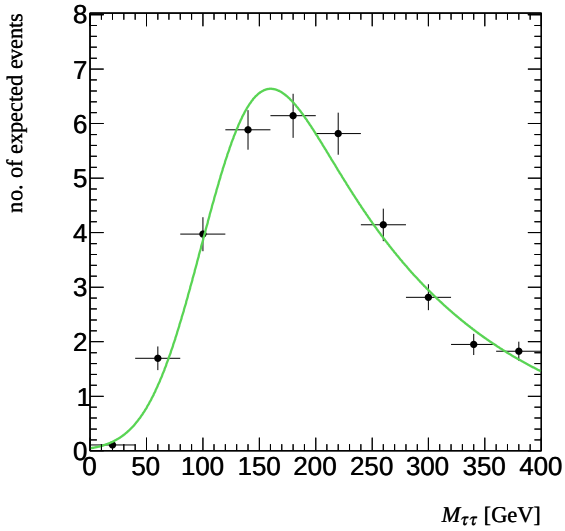
In addition to systematic uncertainties, there are uncertainties in various predictions: This kind of uncertainties is similar to real systematic uncertainties but does not occur when carrying out a real experiment. The uncertainty of the absolute normalisation of background in this analysis is an example. The imprecise knowledge comes from the theoretical cross section, from the limited amount of MC events or from using e.g. cut factorisation. If the normalisation estimated from simulation does not give the correct expectation value the predicted signal significance will be inaccurate. But these uncertainties will not be present in a real ex-



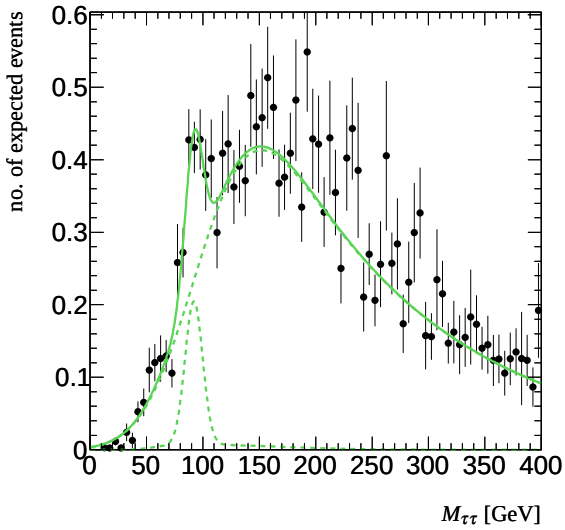
(a) $H \to \tau\tau$.



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



(c) Non-resonant.



(d) Non-resonant (SS). Empty circles show the contribution $Z \to \tau\tau$ with real τ_{had}.

Figure 10.1.: Parametrisation of input samples for significance calculation.

periment because the number of observed events is precisely known. Effects of uncertainties of the prediction can be estimated in principle. E.g. the normalisation can be varied within the theoretical errors of the cross section. Other sources, however, like uncertainties of the shower model of the MC generator or PDF uncertainties need additional MC samples which are not available for this analysis. Statistical uncertainties are estimated using toy experiments as described in the next section.

The likelihood function, Eq. 10.22, takes as input the observed $M_{\tau\tau}$ distribution from the signal selection and the two $M_{\tau\tau}$ distributions from the background estimation methods. The selected data in the signal region is, by definition, free of systematic uncertainties. The normalisation parameters are determined by the fit. Only the background estimation methods are left as sources of systematic effects. The ratio of events in a control region to events in the signal region determine how strongly the shape parameters are constrained, which directly influences the signal significance. This however is not a systematic uncertainty but only an uncertain prediction: Once the experiment is carried out the number of events will be known with absolute certainty. Thus, systematic effects need to be taken into account only if they lead to a discrepancy between the $M_{\tau\tau}$ shape in the signal and in the control region.

The embedding method to estimate $Z \to \tau\tau$ relies on simulation to replace the muons with simulated τ decay products. To estimate the effect on the $M_{\tau\tau}$ shape, variations as suggested in [61] have been applied to electrons, muons and τ_{had} candidates which replace the original muons in an embedded $Z \to \mu\mu$ sample including pile-up (cf. Table 10.1). For all three kinds of decay products the four-vectors are scaled up and downwards. The E_{T} or p_{T} resolutions are varied by Gaussian smearing with the given parameters. Variation of the reconstruction and identification efficiency is simulated by randomly deleting a fraction of candidates. This is obviously possible for variations towards lower values only, since particles can be deleted but not (easily) created. The impact of these variations on the $M_{\tau\tau}$ shape is shown in Figure 10.2. The effect is below the statistical error bars per bin and is thus neglected in the following. Effects of cone sizes as discussed in Sec. 9.1.2 are not considered here as it is assumed that a proper optimisation can be performed on e.g. $Z \to ee$ and $Z \to \mu\mu$.

In Chapter 9 it has been demonstrated that the $M_{\tau\tau}$ distributions obtained from the control regions are generally in good agreement with corresponding distributions in the signal region. Residual deviations come presumably from statistical fluctuations. As the statistical uncertainties of the distributions are rather large due to the limited number of available MC events these might mask true shape deviations. Hence, a conservative approach is to take the statistical uncertainties of the fit parameters obtained from the background component in the signal region as systematic uncertainties. Systematic uncertainties of this kind can be implemented in the profile likelihood method in the form of additional nuisance parameters. Each parameter p' of the PDFs used to fit the control region for $Z \to \tau\tau$ and the non-resonant background is shifted with respect to the parameter p in the signal region:

$$p' = p + \alpha. \tag{10.20}$$

The systematic uncertainties are assumed to be distributed according to a Gaussian probability. For each parameter the Likelihood function is multiplied by a Gaussian function with the

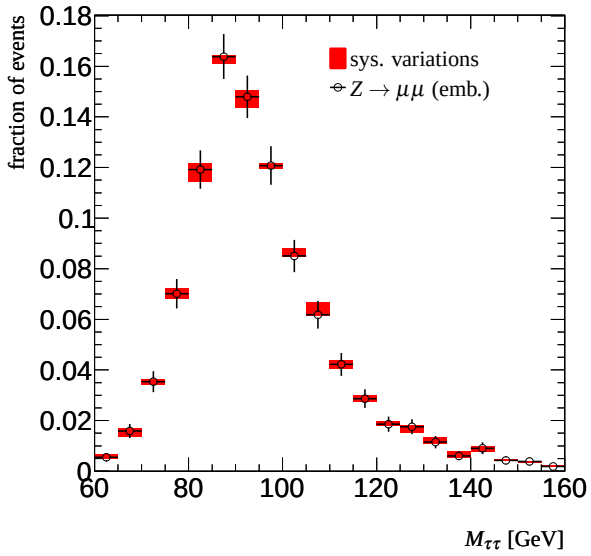


Figure 10.2.: Effect of variations to estimate systematic uncertainties. All variations are performed independently. The solid red areas range between the minimum and maximum central value per bin found in all variations and the unmodified distribution. The $M_{\tau\tau}$ distribution is shown after pre-selection and all $\not{E}_{\mathrm{T}}$/lepton-requirements.

variable α centred around zero.

$$G(\alpha, 0, \sigma_\alpha) = \exp\left(-\frac{\alpha^2}{2\sigma_\alpha}\right) \tag{10.21}$$

The width of the Gaussian function, σ_α, is set to the statistical uncertainty of parameter p obtained from fitting the PDF to the signal region. Since $\mathrm{PDF}_{Z\to\tau\tau}$ and $\mathrm{PDF}_{\text{non-res}}$ contain six and three free parameters, respectively, a total of nine additional nuisance parameters is added to the likelihood function. The values of the σ_{α_i} are shown in Table 10.2. The likelihood function including systematic uncertainties now reads:

$$\begin{aligned} L(M_{\tau\tau}|f_s, \boldsymbol{\theta}) &= L_{\text{signal}}\left(M_{\tau\tau}(\text{signal region})|\boldsymbol{\theta}_H, \boldsymbol{\theta}_Z, \boldsymbol{\theta}_{\text{NR}}, f_s, f_R\right) \\ &\times L_{Z\to\tau\tau}\left(M_{\tau\tau}(\text{embedding})|\boldsymbol{\theta}'_Z\right) \times \prod_{\alpha_i \text{ in } \theta'_Z} G(\alpha_i, 0, \sigma_{\alpha_i}) \\ &\times L_{\text{non-resonant}}\left(M_{\tau\tau}(\text{SS})|\boldsymbol{\theta}_Z, \boldsymbol{\theta}'_{NR}, f_Z\right) \times \prod_{\alpha_j \text{ in } \theta'_{NR}} G(\alpha_j, 0, \sigma_{\alpha_j}) \end{aligned} \tag{10.22}$$

As a consequence of the additional parameters both background PDFs are less constrained by the corresponding control region. This lowers the sensitivity to a signal and therefore reduces the signal significance for a given observation.

parameter	variation
electrons:	
E scale	$\pm 0.5\%$
E_{T} resolution	$0.0073 \cdot E_{\mathrm{T}}$
rec./ID efficiency	-0.2%
muons:	
E scale	$\pm 1\%$
$1/p_{\mathrm{T}}$ resolution	$0.011/p_{\mathrm{T}}[\mathrm{GeV}] \oplus 0.00017\,\mathrm{GeV}$
rec./ID efficiency	-1%
τ_{had} :	
E scale	$\pm 3\%$
E resolution	$0.45\sqrt{E[\mathrm{GeV}]}$
rec./ID efficiency	-5%

Table 10.1.: Variations applied to the embedded $Z \to \mu\mu$ sample to estimate the impact of systematic uncertainties.

parameter		σ_{α_i}	
		w/o pile-up	w/ pile-up
$Z \to \tau\tau$	p_7	11.31 GeV	14.77 GeV
	p_8	15.46 GeV	10.79 GeV
	p_9	0.69 GeV	3.34 GeV
	p_{10}	0.48 GeV	7.97 GeV
	p_{11}	38.26 GeV	2.46 GeV
	p_{12}	0.016	0.072
non-resonant	p_{13}	10.92 GeV	16.15 GeV
	p_{14}	3.75 GeV	5.80 GeV
	p_{15}	2.33 GeV	3.20 GeV

Table 10.2.: Statistical uncertainties of the fit parameters used as systematic uncertainties of the $M_{\tau\tau}$ shapes estimated from the control samples.

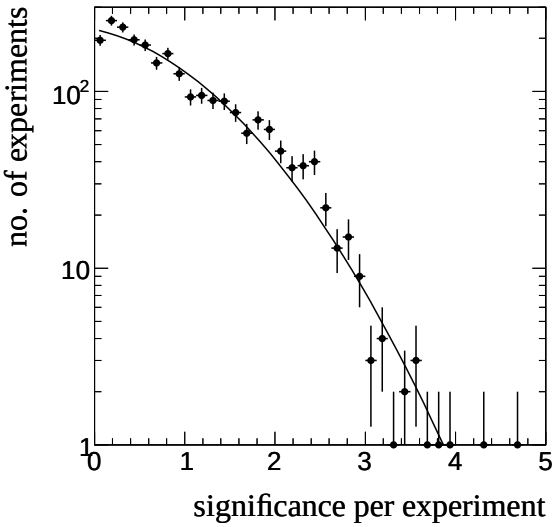


Figure 10.3.: Signal significance in background-only pseudo experiments probing $m_H = 115\,\text{GeV}$. The solid line is a Gaussian function fit to the data points.

10.4. Pseudo Experiments

Pseudo experiments can be generated using the PDFs for $M_{\tau\tau}$ to obtain random $M_{\tau\tau}$ distributions. The number of events is a Poissonian random variable with the nominal normalisation as mean. These pseudo experiments can be used to test the method itself and the validity of some of the assumptions made above.

According to Equation 10.5, the distribution of q_0 in background-only experiments should follow an exponential distribution for values of $q_0 > 0$. Equivalently the signal significance is supposed to follow a half Gaussian distribution with a mean at zero. Fig. 10.3 shows that the significance is indeed well described by a Gaussian function. Without pile-up, however, the mean of the Gaussian function is smaller than zero. This effectively subtracts the shift from the real significance, decreasing the significance obtained from pseudo experiments. For the final results shown in the next section this offset is corrected for.

The distribution of the signal fraction, $\hat{f}_s$, and, as a result, the significance in experiments including signal are supposed to be Gaussian distributed, as well. Both assumptions can be verified using the pseudo experiments. The outcome is shown in Figures 10.4 and 10.5. In both cases the distribution is in good agreement with a Gaussian function. Moreover, Eqs. 10.11 and 10.12 state that the standard deviation of the significance distribution is equal to one. As can be seen in Fig. 10.5, also this assumption is well justified.

In addition, the pseudo experiments give information about the accuracy of the normalisation obtained from the fit. Figure 10.6 shows the mean of a Gaussian function fit to the relative difference between the estimated and true number of signal events as a function of m_H. In both luminosity scenarios there is a bias towards a smaller signal strength. This bias is probably related to the shift in the distribution of $\sqrt{q_0}$ observed above. Since with the latter

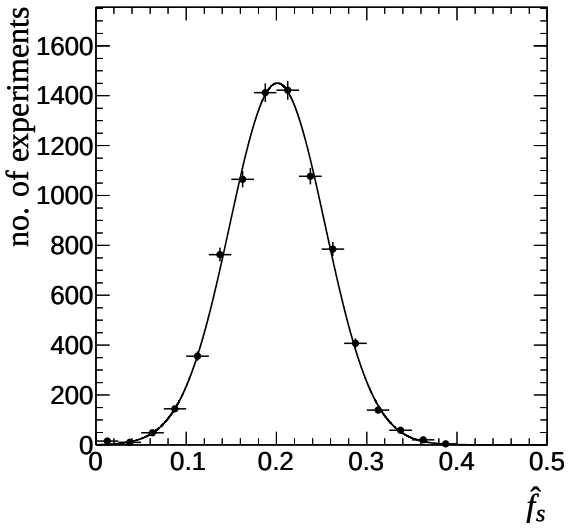


Figure 10.4.: The measured signal strength parameter in pseudo experiments containing the signal with $m_H = 120\,\text{GeV}$. The solid line is a Gaussian function fit to the data points.

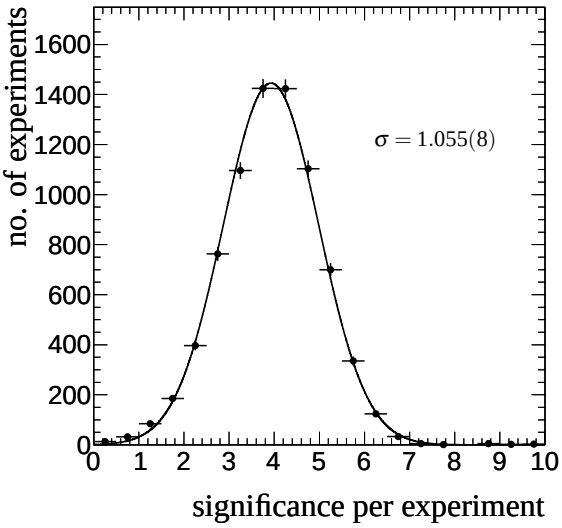
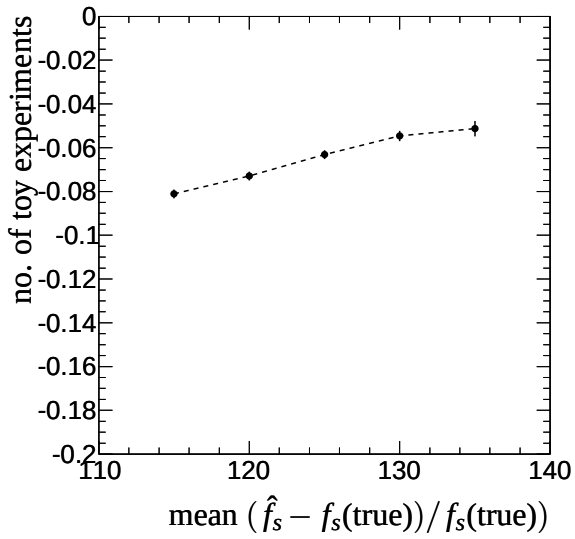
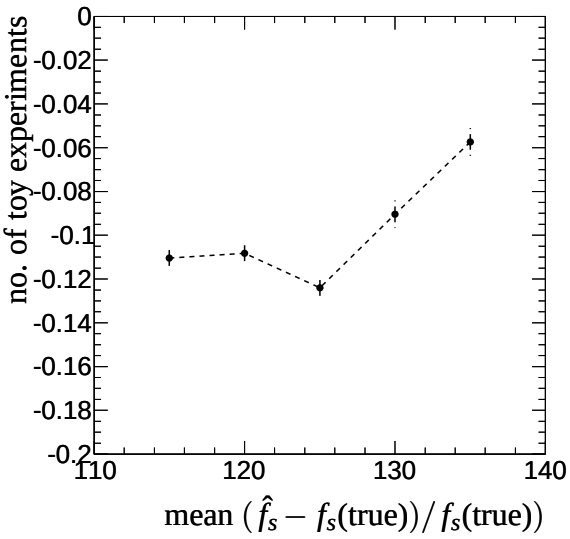


Figure 10.5.: Signal significance in signal plus background pseudo experiments containing signal with $m_H = 120\,\text{GeV}$. The solid line is a Gaussian function fit to the data points.



(a) Without pile-up



(b) With pile-up.

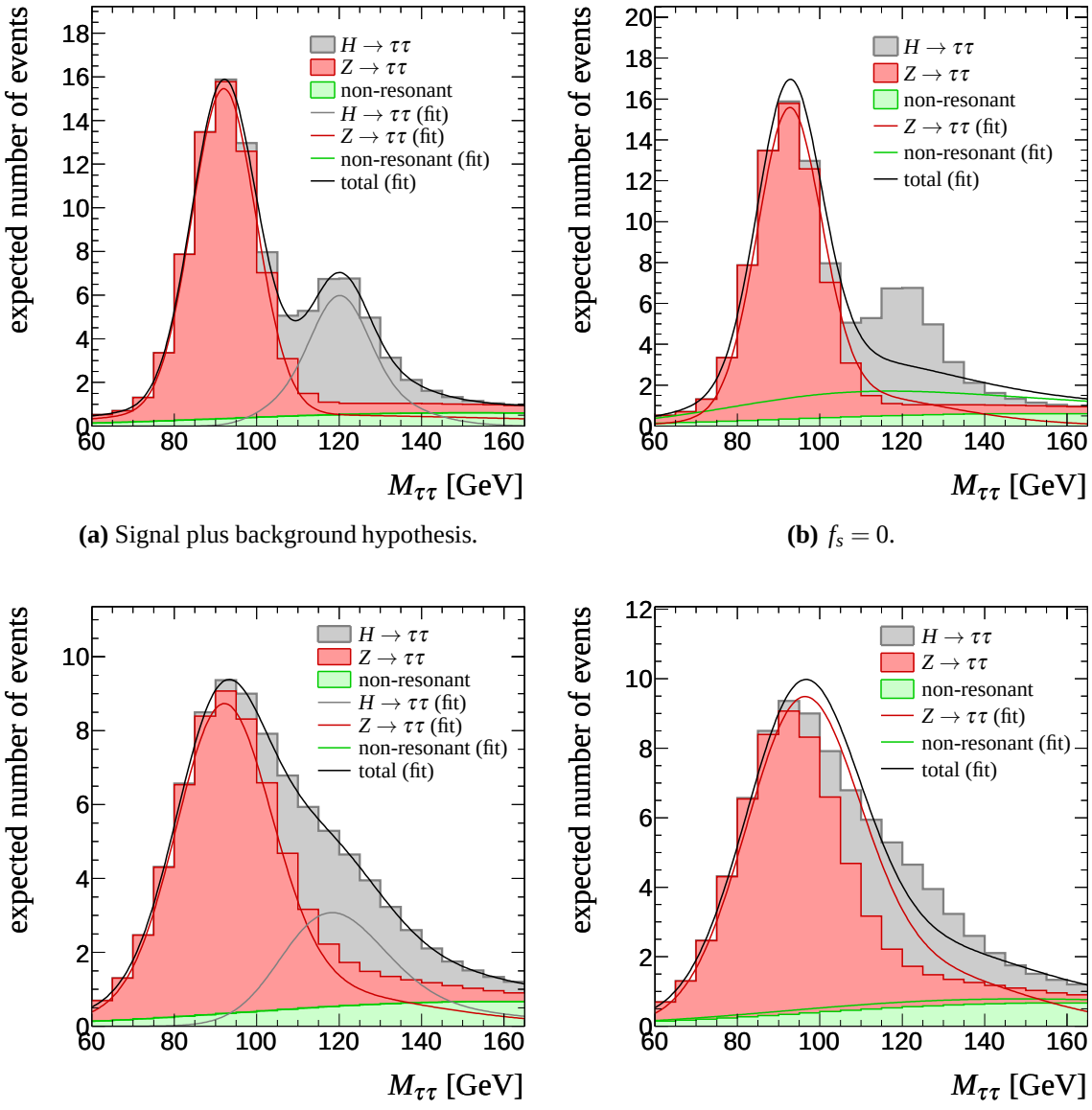
Figure 10.6.: Mean difference between $\hat{f}_s$ and the true value f_s' as a function of m_H.

effect corrected for the signal significance obtained from the Asimov data and from the pseudo experiments are in acceptable agreement, there is no need to take additional measures against the bias within the context of this thesis.

10.5. Expected Signal Significance

The likelihood fit to Asimov data for an assumed Higgs boson mass of $m_H = 120\,\text{GeV}$ is shown in Fig. 10.7 for the signal region and in Figures 10.8 and 10.9 for the two background control regions. Due to the way the Asimov data is constructed, the likelihood fit of the signal plus background hypothesis exactly reproduces the data points in the signal region and the background control regions. For the background-only hypothesis the data PDF is less consistent with the data points.

The resulting expected signal significance for an integrated luminosity of $30\,\text{fb}^{-1}$ is shown in Figure 10.10 for both luminosity scenarios as a function of the hypothetical Higgs boson mass. Systematic uncertainties have a major effect only on smaller Higgs boson masses: The systematic uncertainties reflect uncertainties in the $M_{\tau\tau}$ shape of either $Z \to \tau\tau$ or the non-resonant background. Since the contribution of non-resonant background in the signal mass region is relatively small, the further the signal mass peak moves away from the Z peak the smaller the effect of the shape variations becomes. The central value of the significance obtained from pseudo experiments is in acceptable agreement with the value obtained from the Asimov sample. The expected signal significance without pile-up ranges between 3.0σ for $m_H = 135\,\text{GeV}$ and 4.4σ for $m_H = 125\,\text{GeV}$. The behaviour of the significance as a function of m_H is an effect of two competing effects: on the one hand, the cross section of VBF $H \to \tau\tau$ and thus the size of the signal peak monotonously decreases for increasing values of m_H. On the other hand, the signal peak moves farther away from the Z peak for increasing values of m_H. Therefore, the signal significance has a maximum in the vicinity of $m_H = 125\,\text{GeV}$. In the presence of pile-up a substantial decrease of the signal significance with respect to the no-pile-up scenario is observed: With between 1.6σ for $m_H = 115\,\text{GeV}$ and 2.0σ for $125\,\text{GeV}$ a data sample corresponding to an integrated luminosity of $30\,\text{fb}^{-1}$ is not sufficient for a sensitive measurement of VBF $H \to \tau\tau \to \ell h$.



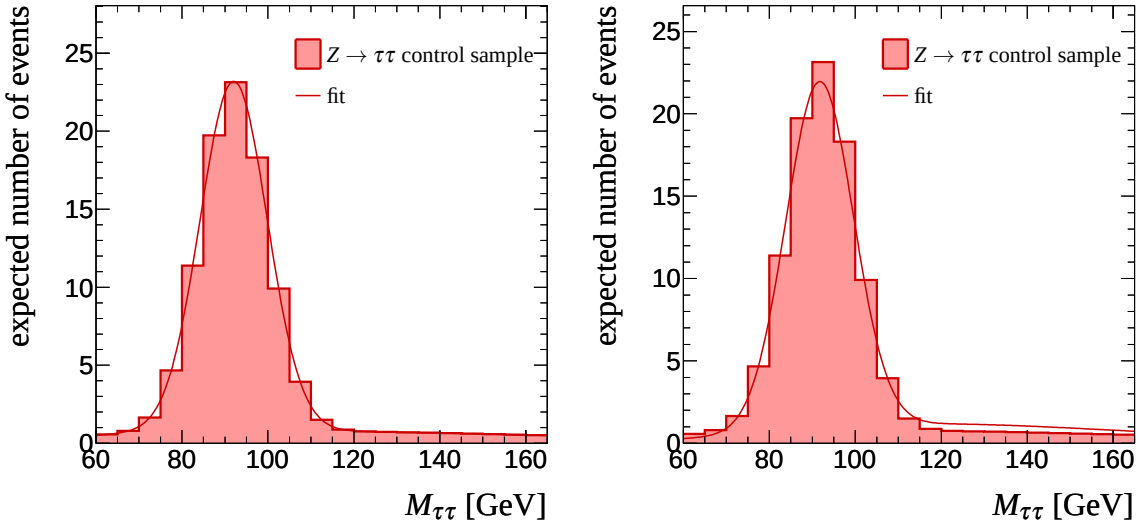
(a) Signal plus background hypothesis.

(b) $f_s = 0$.

(c) Signal plus background hypothesis, pile-up

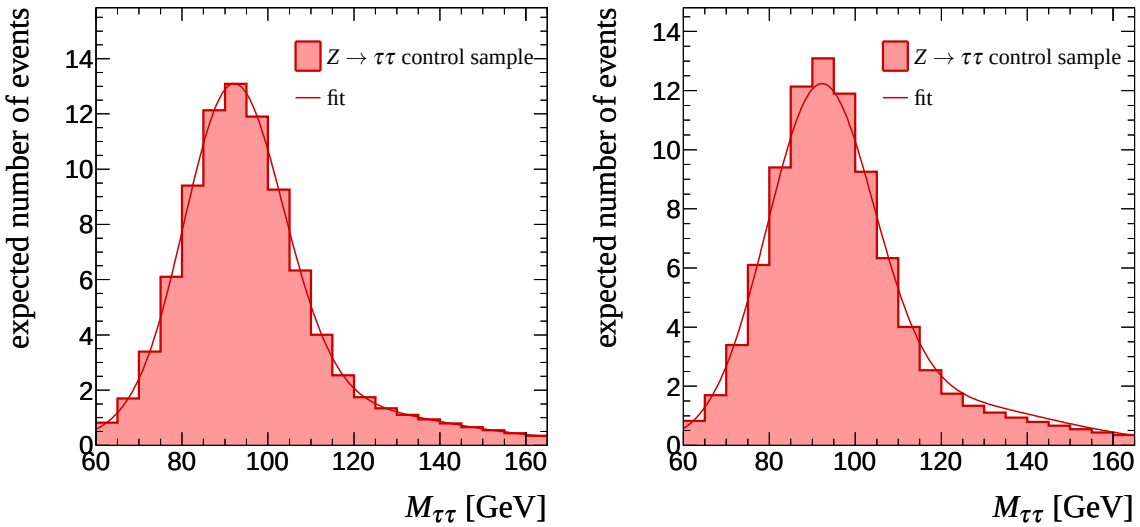
(d) $f_s = 0$, pile-up

Figure 10.7.: Likelihood fit to Asimov data in the signal region corresponding to the expectation for an integrated luminosity of $30\,\text{fb}^{-1}$ and an assumed Higgs boson mass of $m_H = 120\,\text{GeV}$. The unconstrained fits are shown on the left side, fits with $f_s = 0$ on the right side.



(a) Signal plus background hypothesis.

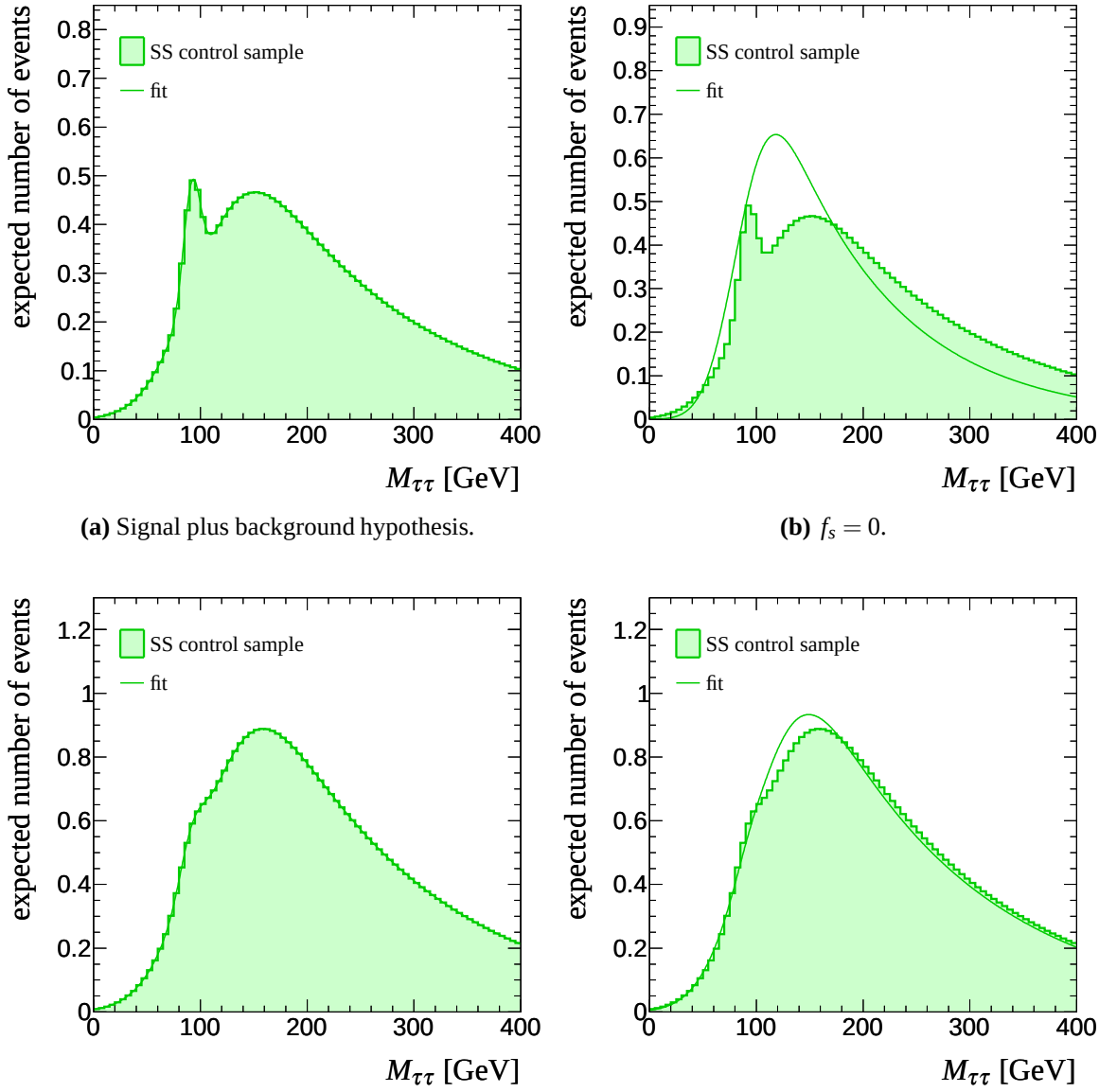
(b) $f_s = 0$.



(c) Signal plus background hypothesis, pile-up

(d) $f_s = 0$, pile-up

Figure 10.8.: Likelihood fit to Asimov data in the $Z \to \tau\tau$ control region corresponding to the expectation for an integrated luminosity of $30\,\text{fb}^{-1}$ and an assumed Higgs boson mass of $m_H = 120\,\text{GeV}$. The unconstrained fits are shown on the left side, fits with $f_s = 0$ on the right side.



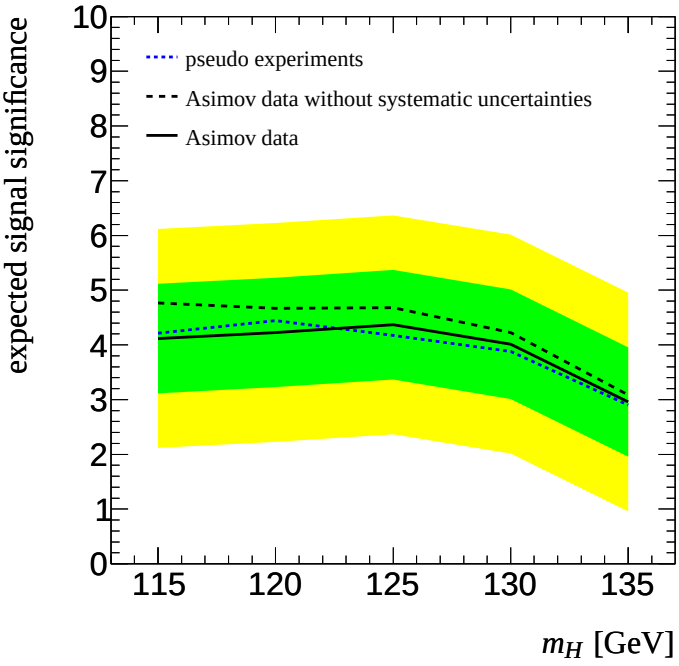
(a) Signal plus background hypothesis.

(b) $f_s = 0$.

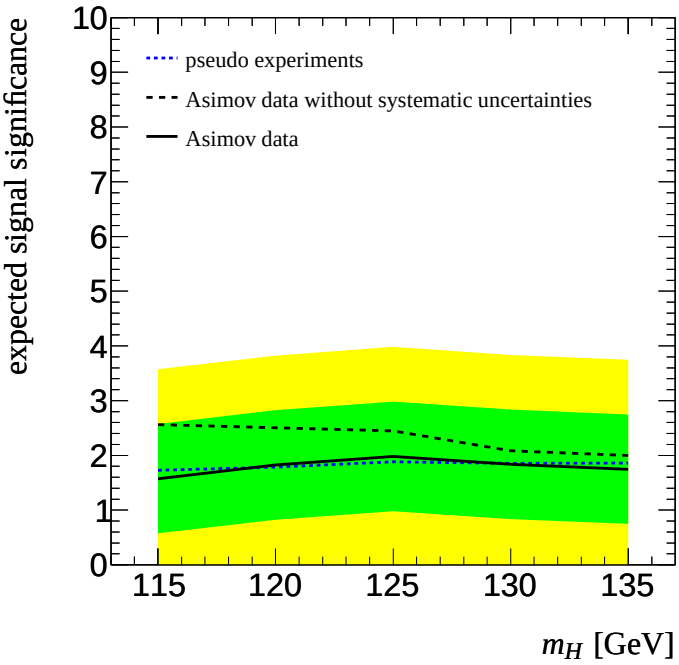
(c) Signal plus background hypothesis, pile-up

(d) $f_s = 0$, pile-up

Figure 10.9.: Likelihood fit to Asimov data in the SS control region corresponding to the expectation for an integrated luminosity of $30\,\text{fb}^{-1}$ and an assumed Higgs boson mass of $m_H = 120\,\text{GeV}$. The unconstrained fits are shown on the left side, fits with $f_s = 0$ on the right side.



(a) Without pile-up.



(b) With pile-up.

Figure 10.10.: Expected signal significance for data corresponding to an integrated luminosity of $30\,\mathrm{fb}^{-1}$. Error bands correspond to variations of $\hat{f}_s$ of $\pm 1(2)\sigma$.

11
Conclusions

At the time of writing this thesis ATLAS has collected a data sample corresponding to an integrated luminosity of $2.69\,\text{fb}^{-1}$ [69]. With this data, Higgs boson searches at ATLAS are able to exclude a Standard Model Higgs boson in most of the range between $m_H = 146\,\text{GeV}$ and 466 GeV [16]. In order to discover the Higgs boson, especially with a mass close to the LEP exclusion limit of 114.4 GeV, a larger sample, preferably at a higher centre-of-mass energy, is required. In this mass regime, $H \to \tau\tau$ is an important search channel. In order to prepare for future Higgs boson searches, this thesis presents a study of the sensitivity of the ATLAS detector for a Standard Model Higgs boson in the channel VBF $H \to \tau\tau \to \ell h + 3\nu$. For this study a data sample corresponding to $30\,\text{fb}^{-1}$ at a centre-of-mass energy of 14 TeV is assumed. The analysis updates earlier results in the light of more recent developments and significantly extends previous studies in two important aspects: For the first time in this search channel, pile-up is completely taken into account. Secondly, novel methods have been developed which now allow for estimating all dominant background processes from data.

In this thesis a pile-up scenario with a luminosity of $\mathscr{L} = 10^{33}\,\text{s}^{-1}\text{cm}^{-2}$, a proton bunch spacing of 75 ns and 6.9 simultaneous minimum bias interactions has been investigated. In this scenario, the signal acceptance decreases by about 50% using the standard selection criteria. Three main sources have been identified: The trigger, the identification of hadronic τ decays and the veto on additional jets in the central region of the detector. The efficiency of the electron and muon triggers decreases on level-2. The performance of the τ_{had} identification suffers from additional energy deposited close to the candidate. The veto on additional central jets rejects events due to jets from pile-up. In addition to a lower signal efficiency, pile-up results in a worse $\not{E}_{\text{T}}$ resolution caused by the increased amount of energy deposited in the calorimeter and additional noise. As a consequence, the resolution of the reconstructed Higgs boson mass, $M_{\tau\tau}$, deteriorates. The width of the mass peak increases by about 30%. The $Z \to \tau\tau$ background is affected in the same way as the signal. Therefore, $M_{\tau\tau}$ as the discriminating observable loses much of its separation power. With updated selection criteria, a τ_{had} identification which includes the effects of pile-up and the using of jet-vertex associations for the veto on central jets, most of the loss in signal acceptance has been recovered.

The influence of pile-up on the $M_{\tau\tau}$ distribution puts even greater emphasis on the estimation of the dominant background processes from data. A method has been developed to estimate the $Z \to \tau\tau$ background from $Z \to \mu\mu$ data. Muons in a data event are replaced by simulated τ-lepton decays. The simulated decay products are integrated into the data event on the level of calorimeter cells and tracks. The hybrid events resulting from rerunning the standard reconstruction algorithms model $Z \to \tau\tau$ events to high precision. Possible systematic uncertainties related to the use of simulated τ-lepton decays have been studied and found to be negligible. A second control sample to estimate the dominant non-resonant background processes, $W+$jets and $t\bar{t}$ production, is constructed by selecting lepton-τ_{had} pairs with equal charge. The influence of quark and gluon jets and the resulting correlation between the charge of the lepton and the τ_{had} candidate have been investigated. The composition of the control sample and the effect of a contamination with resonant background have been studied. The mass shape extracted from this control sample is in good agreement with the $M_{\tau\tau}$ distribution of the non-resonant background in the signal region.

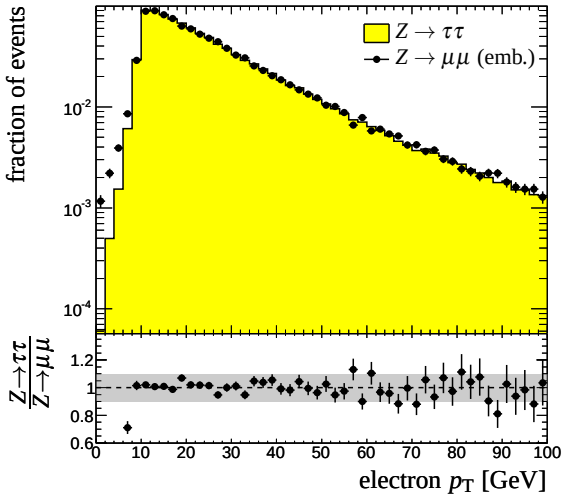
The profile likelihood method has been used to calculate the expected signal significance. Since all dominant background processes can be estimated from data, the only remaining source of systematic uncertainties is a possible discrepancy between the mass shapes in the control samples and in the signal region. Based on simulation, no such discrepancies have been found. Instead the statistical uncertainties of the parametrisation have been used as a conservative estimate of the systematics. Not considering pile-up, the expected signal significance ranges between 3.0σ for $m_H = 135\,\text{GeV}$ and 4.4σ for $m_H = 125\,\text{GeV}$. This confirms the outcome of previous studies. In the presence of pile-up the expected significance drops to between 1.6σ for $m_H = 115\,\text{GeV}$ and 2.0σ for $125\,\text{GeV}$. This substantial decrease in sensitivity is in accordance with a recent study of VBF $H \to \tau\tau \to \ell\ell$ [71].

With the increasing amount of data collected by ATLAS, many of the assumptions being made in this thesis have already been confirmed. The embedding technique to estimate $Z \to \tau\tau$ from $Z \to \mu\mu$ data has been successfully employed in recent analyses [68, 78, 79]. Although Monte Carlo simulation describes the data collected in 2010 and 2011 already reasonably well within the limited statistics, the novel methods for the background estimation developed in the context of this thesis will allow to perform a search for $H \to \tau\tau$ in VBF with almost no input from simulation. Future studies of VBF $H \to \tau\tau$ will have to take into account the effects of pile-up on various aspects of the analysis. The pile-up conditions encountered during data taking in 2011 is already similar to the scenario this study is based upon. Nonetheless, with the growing understanding of pile-up that will develop from the many different analyses being performed on current data it is reasonable to assume that VBF $H \to \tau\tau$ will profit from more sophisticated reconstruction algorithms, new tools and techniques which may mitigate the effects of pile-up and improve the sensitivity of the experiment.

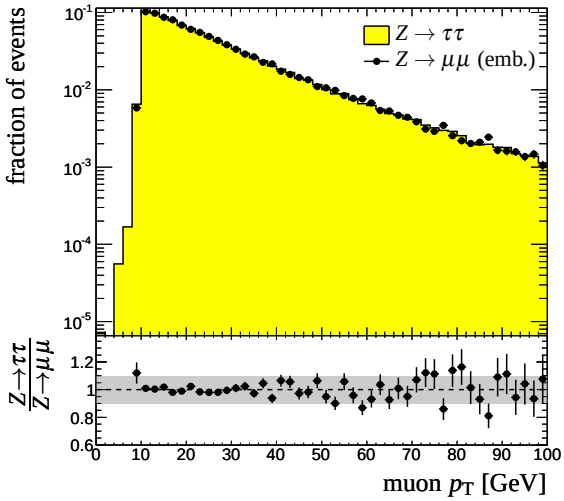
A

Validation of the Embedding Method with Pile-up

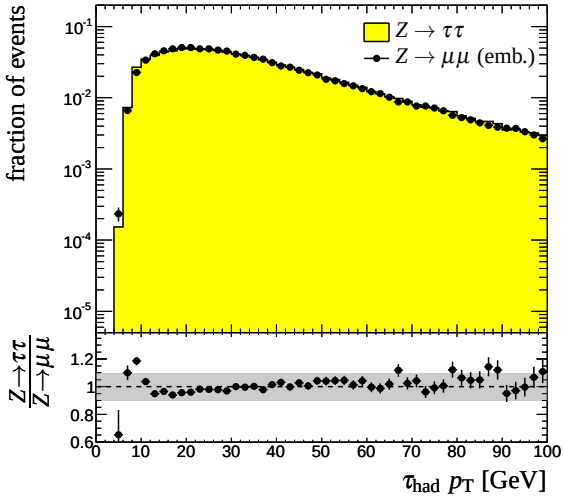
The validation of the embedding method shown in Sec. 9.1.4 is performed on a sample which does not contain pile-up. The corresponding distributions in the presence of pile-up are shown in the following. As in the non-pile-up case, only the ALPGEN 2, 3 and 4-parton samples are used in order to avoid high weights from samples with low statistical power. The conclusions made in Sec. 9.1.4 remain unchanged. Note, however, that the sample used for this comparison has been generated using an inner cone size, C_{in}, of 0.08. This fact explains e.g. the less visible shift in the distribution of $\not{E}_{\text{T}}$.



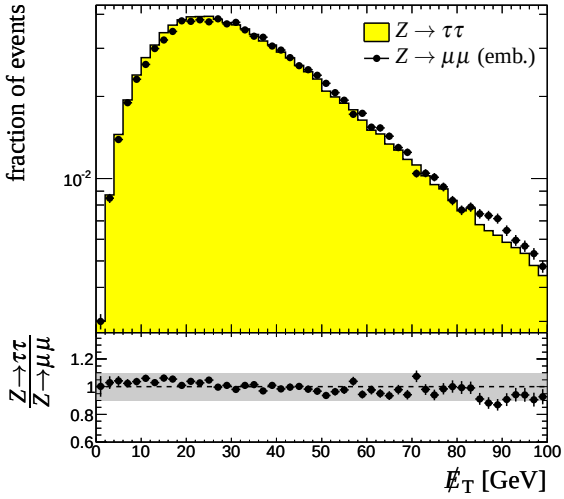
(a) p_T of reconstructed electrons from the $Z \to \tau\tau$ decay (matched to generated electrons).



(b) p_T of reconstructed muons from the $Z \to \tau\tau$ decay (matched to generated muons).

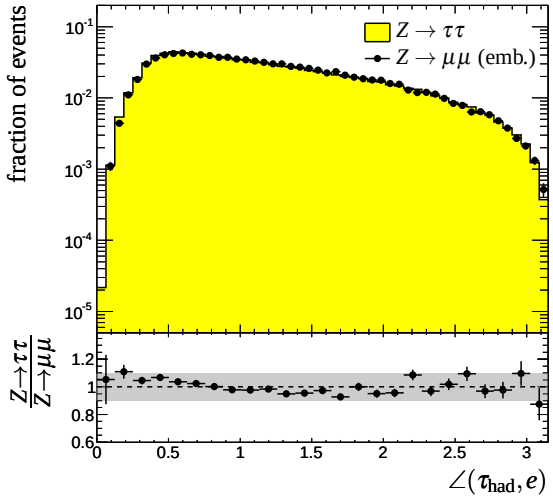


(c) p_T of reconstructed τ_{had} candidates, calorimeter-seeded, from the $Z \to \tau\tau$ decay (matched to the visible hadronic τ decay product on generator level)

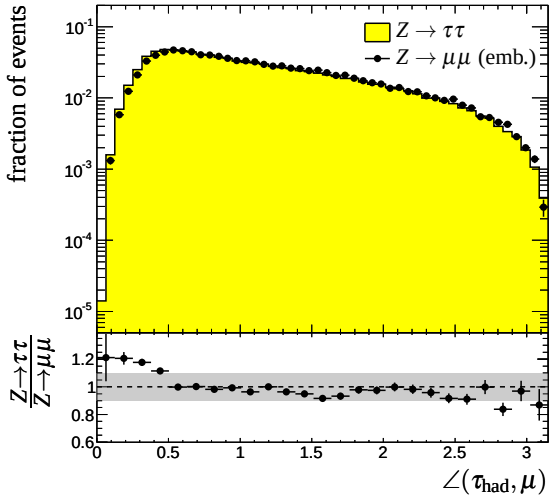


(d) MET RefFinal

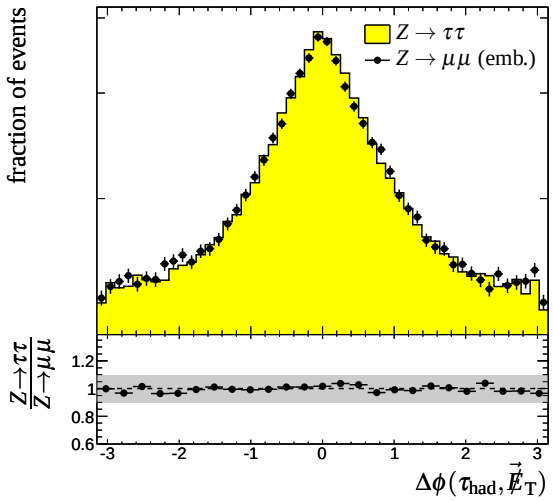
Figure A.1.: Comparison of quantities which enter the calculation of $M_{\tau\tau}$. ($C_{in} = 0.08$)



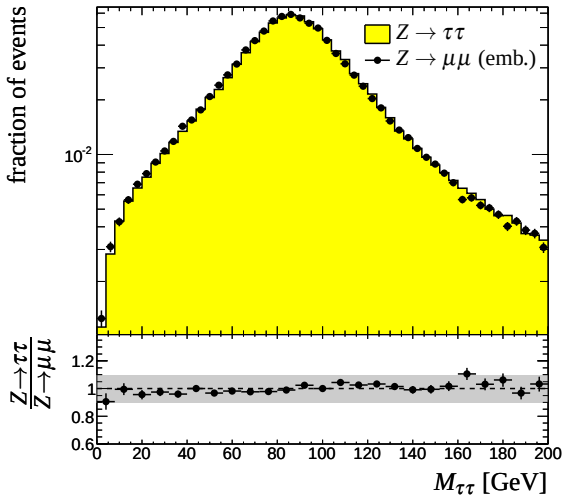
(a) Angle between reconstructed τ_{had} candidates and electrons from the $Z \to \tau\tau$ decay (objects matched to the corresponding generator-level objects).



(b) Angle between reconstructed τ_{had} candidates and muons from the $Z \to \tau\tau$ decay (objects matched to the corresponding generator-level objects).



(c) $\Delta\phi$ between reconstructed τ_{had} candidates from the $Z \to \tau\tau$ decay (matched to the visible hadronic τ decay product on generator level) and $\vec{E}_{\text{T}}$



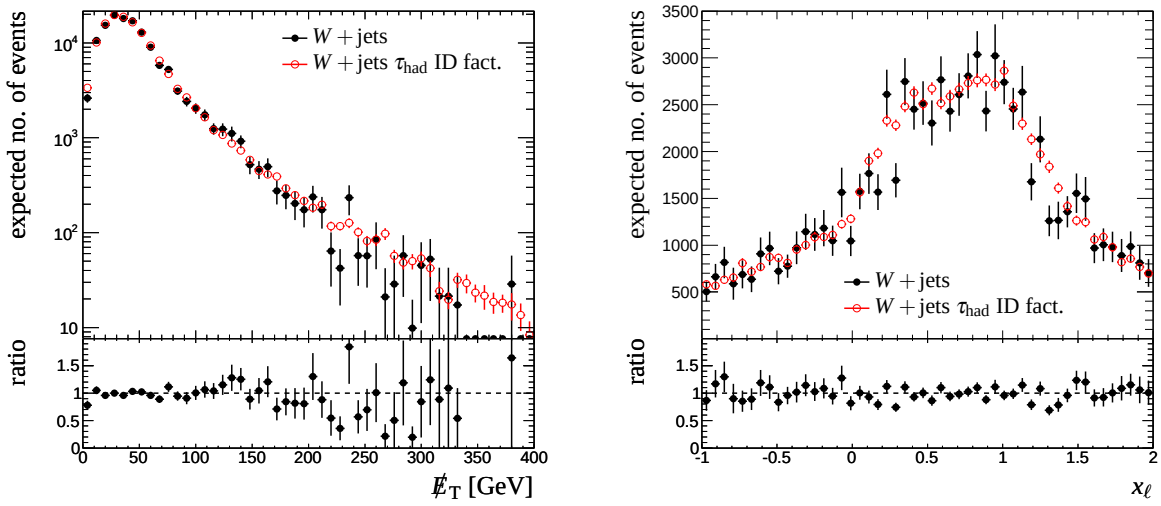
(d) $M_{\tau\tau}$ calculated from reconstructed objects from the $Z \to \tau\tau$ decay (objects matched to the corresponding generator-level objects).

Figure A.2.: Comparison of quantities which enter the calculation of $M_{\tau\tau}$. **(d)** shows a comparison of the resulting $M_{\tau\tau}$ distribution in embedded $Z \to \mu\mu$ events to $Z \to \tau\tau$ events. ($C_{\text{in}} = 0.08$)

B

Validation of the τ_{had} ID factorisation.

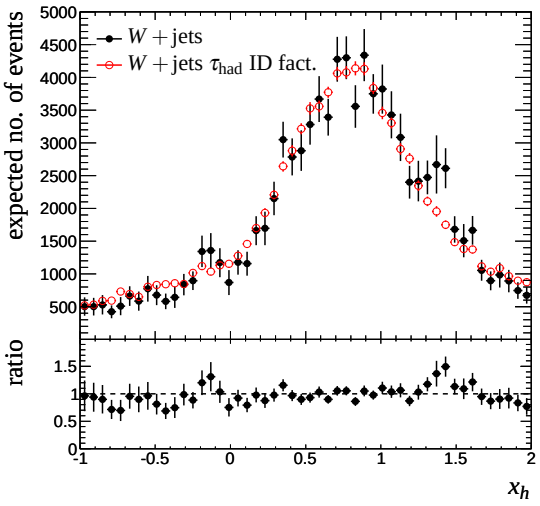
The following validation figures show all observables necessary for the cut selection presented in Chapter 7 after the pre-selection. A separate validation for $W+\text{jets}$ and $t\bar{t}$ production is shown. The $t\bar{t}$ sample is filtered to contain events in which the τ_{had} candidates are misidentified QCD jets. Note that the regular and factorised samples are not normalised to each other: the distributions show the expected number of events after the pre-selection.



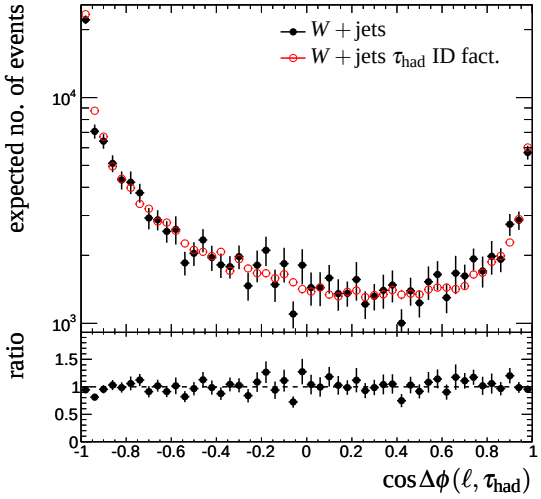
(a) Missing transverse energy.

(b) x_ℓ.

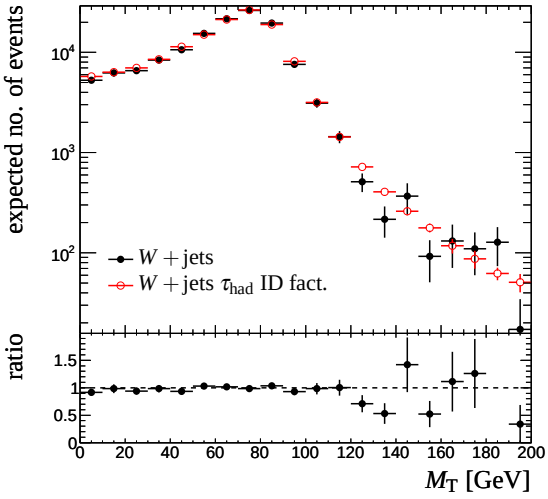
Figure B.1.: Observables from $W+\text{jets}$ samples used in the $H \to \tau\tau$ cut selection.



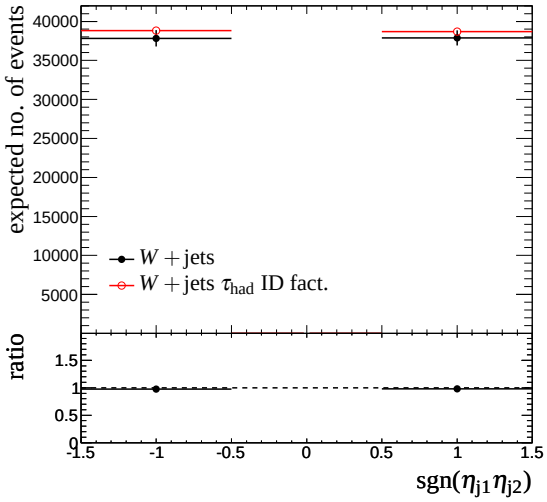
(a) x_h.



(b) $\Delta\phi$ between lepton and τ_{had} candidate.

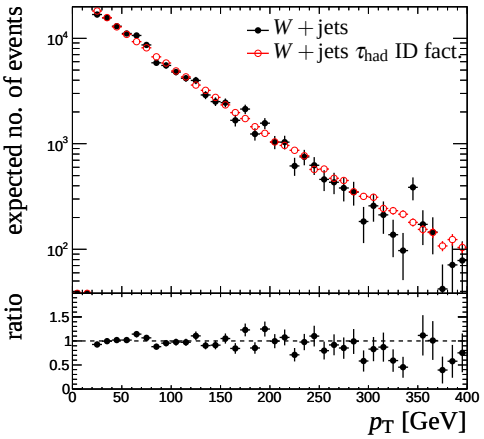


(c) Transverse mass of the lepton-$\not{E}_{\text{T}}$ system.

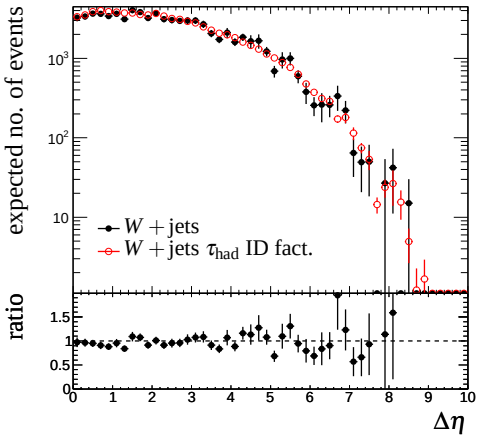


(d) Sign of the product of η of the tagging jets.

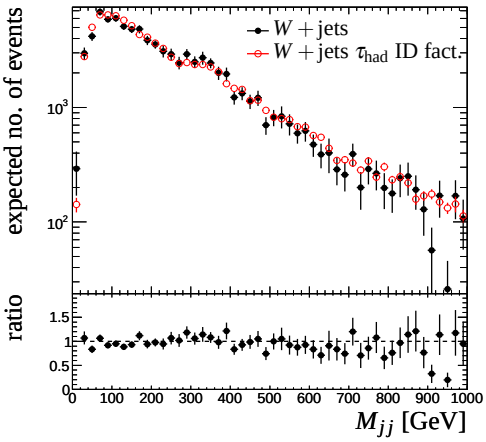
Figure B.2.: Observables from W + jets samples used in the $H \to \tau\tau$ cut selection.



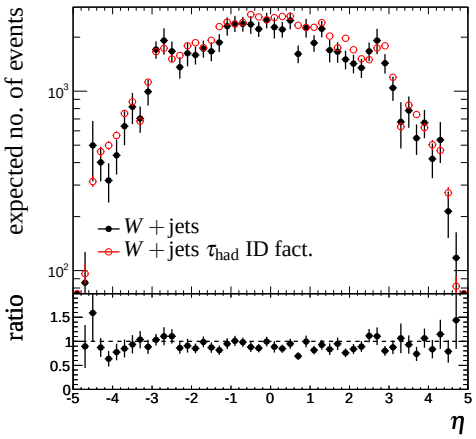
(a) p_T of the leading jet.



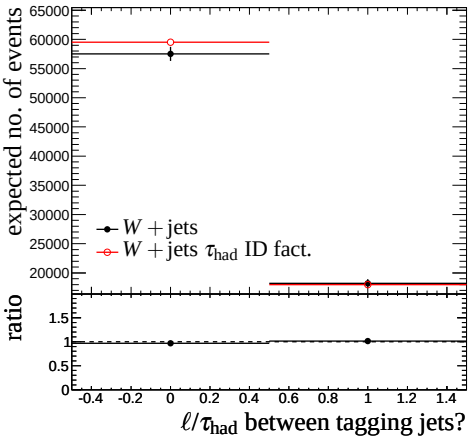
(b) $\Delta\eta$ between the tagging jets.



(c) Invariant mass of the tagging jets.

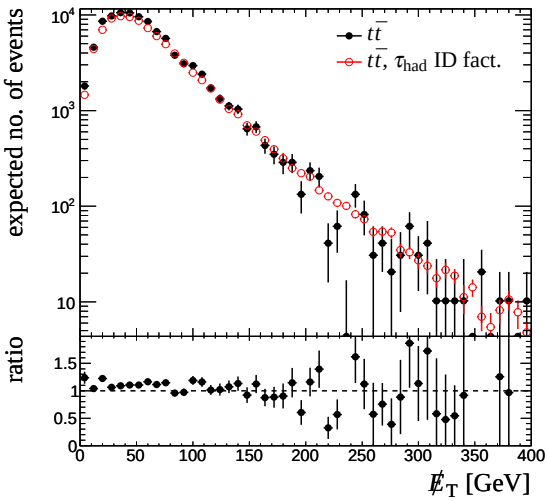


(d) η distribution of additional jets.

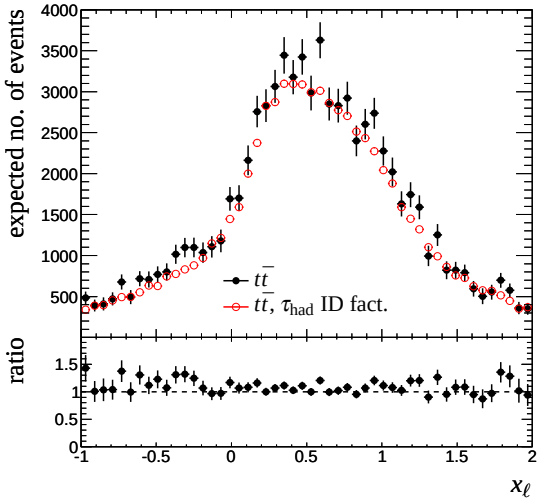


(e) Lepton-τ_{had} centrality.

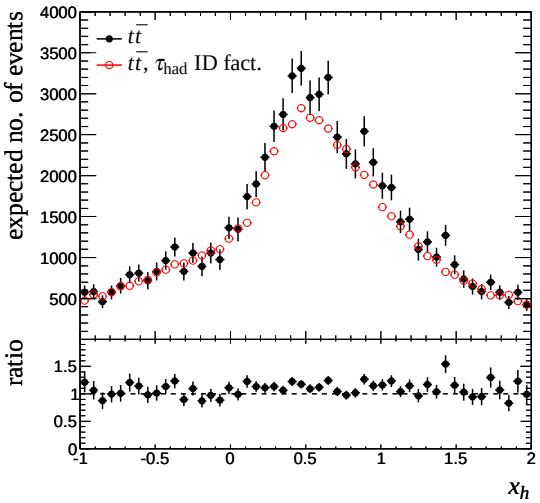
Figure B.3.: Observables from W + jets samples used in the $H \to \tau\tau$ cut selection.



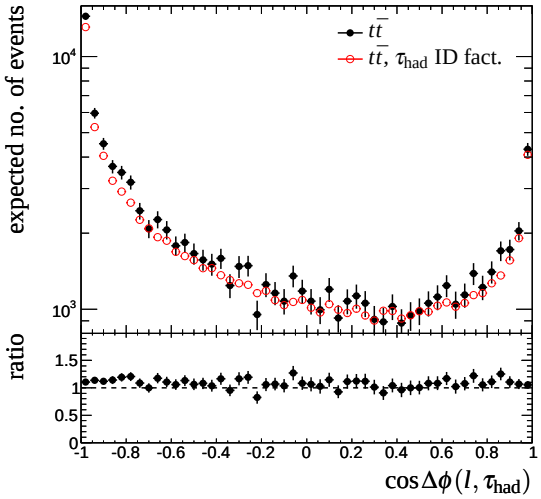
(a) Missing transverse energy.



(b) x_ℓ.

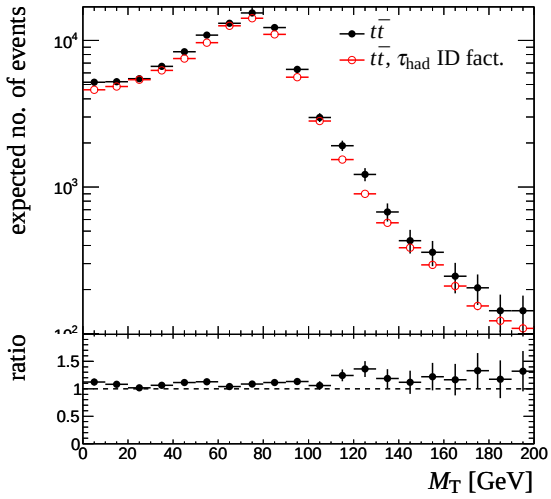


(c) x_h.

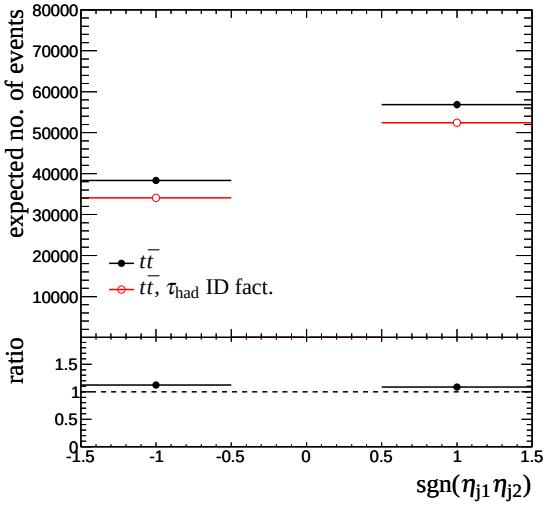


(d) $\Delta\phi$ between lepton and τ_{had} candidate.

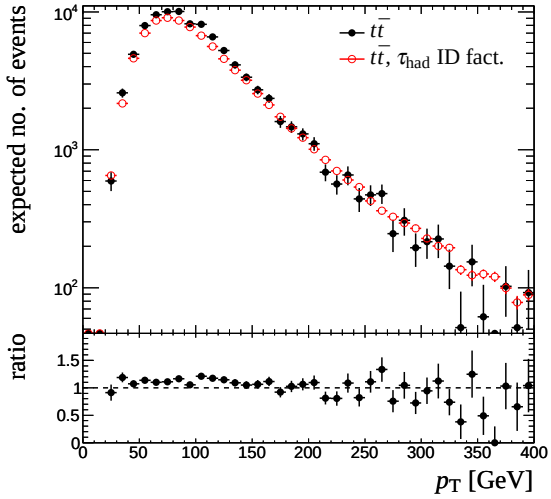
Figure B.4.: Observables from $t\bar{t}$ samples used in the $H \to \tau\tau$ cut selection.



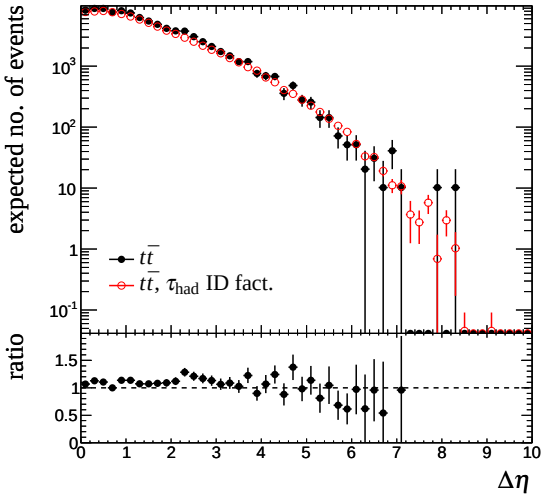
(a) Transverse mass of the lepton-$\not{E}_T$ system.



(b) Sign of the product of η of the tagging jets.

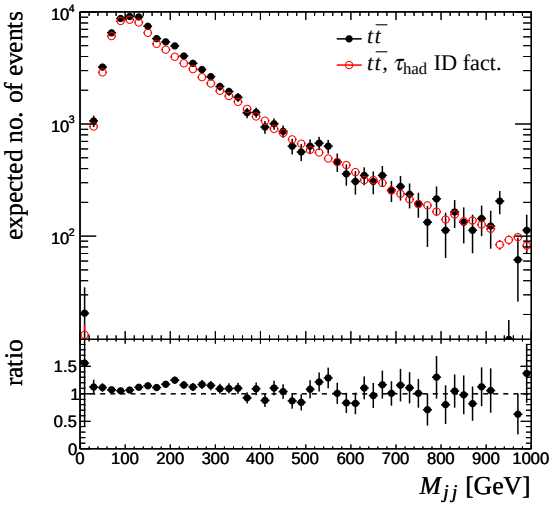


(c) p_T of the leading jet.

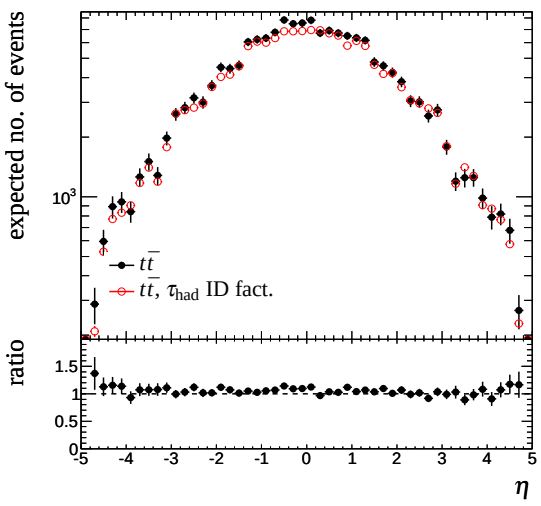


(d) $\Delta\eta$ between the tagging jets.

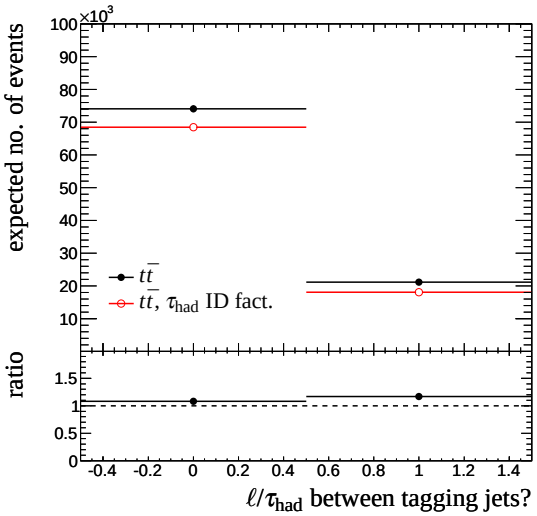
Figure B.5.: Observables from $t\bar{t}$ samples used in the $H \to \tau\tau$ cut selection.



(a) Invariant mass of the tagging jets.



(b) η distribution of additional jets.



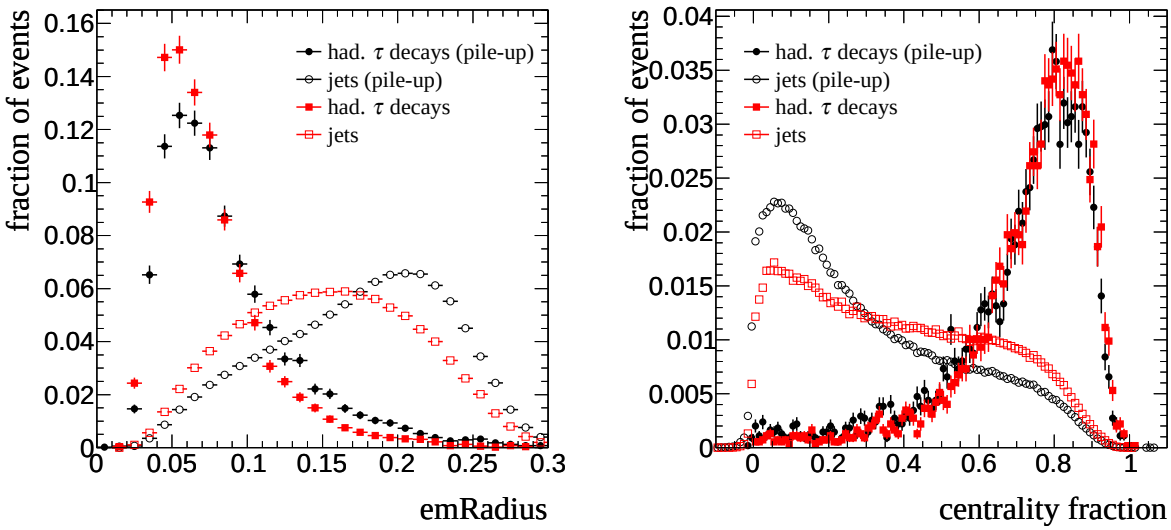
(c) Lepton-τ_{had} centrality.

Figure B.6.: Observables from $t\bar{t}$ samples used in the $H \to \tau\tau$ cut selection.

C

Additional τ_{had} ID Variables

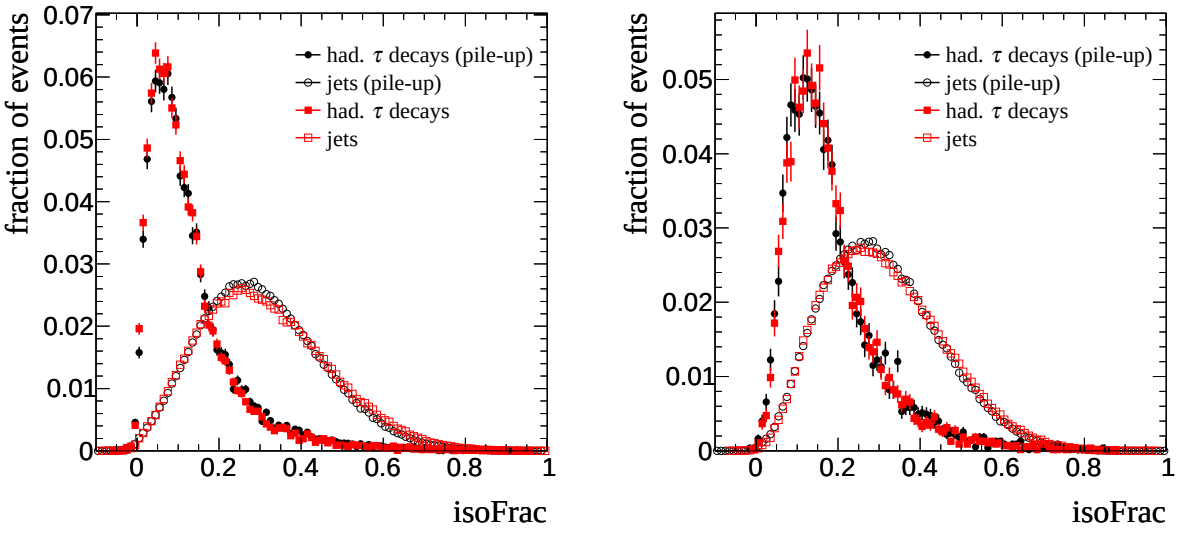
The following figures contain some τ_{had} variables in addition to those shown in Chapter 8.



(a) EM radius.

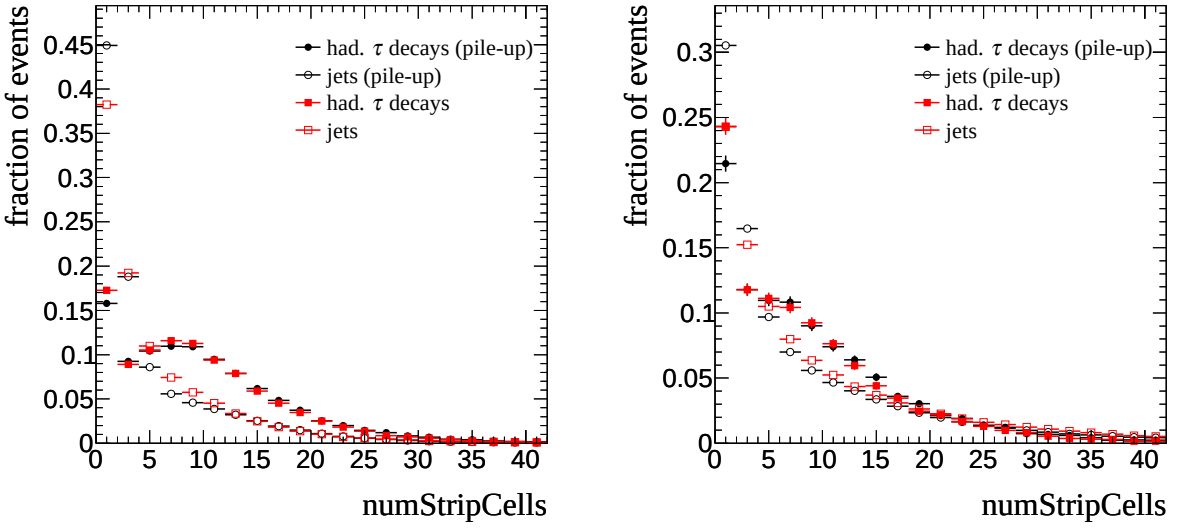
(b) Centrality fraction.

Figure C.1.: EM radius and centrality fraction of 3-prong τ_{had} candidates.



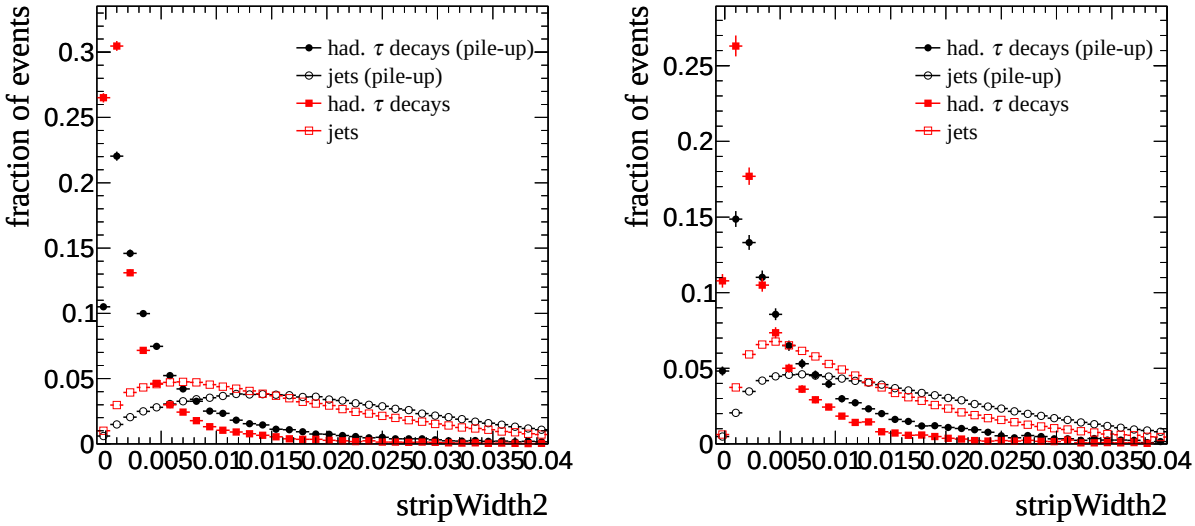
(a) 1-prong τ_{had} candidates **(b)** 3-prong τ_{had} candidates

Figure C.2.: Isolation fraction.



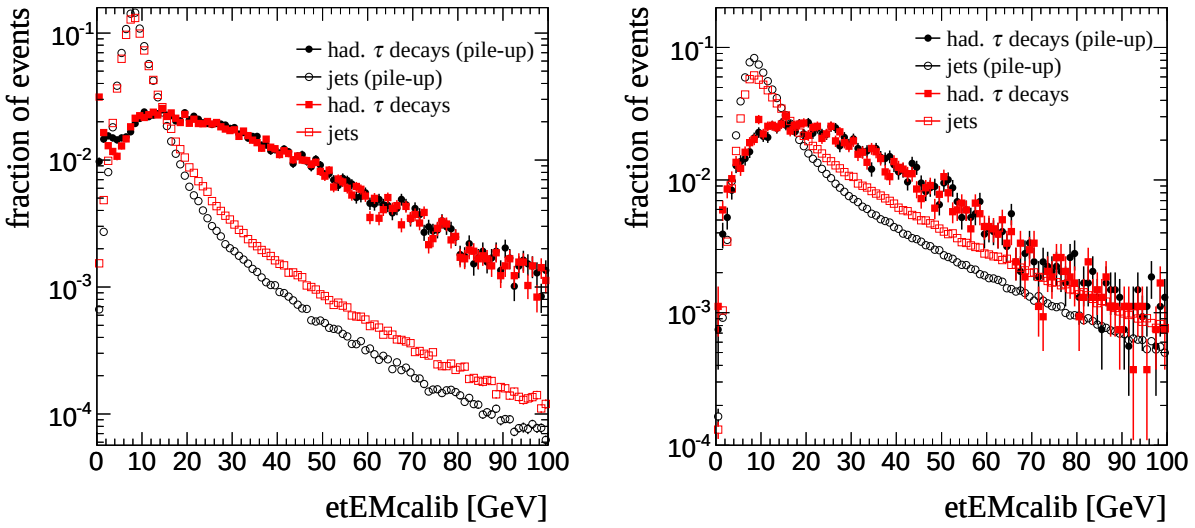
(a) 1-prong τ_{had} candidates **(b)** 3-prong τ_{had} candidates

Figure C.3.: Number of strip cells.



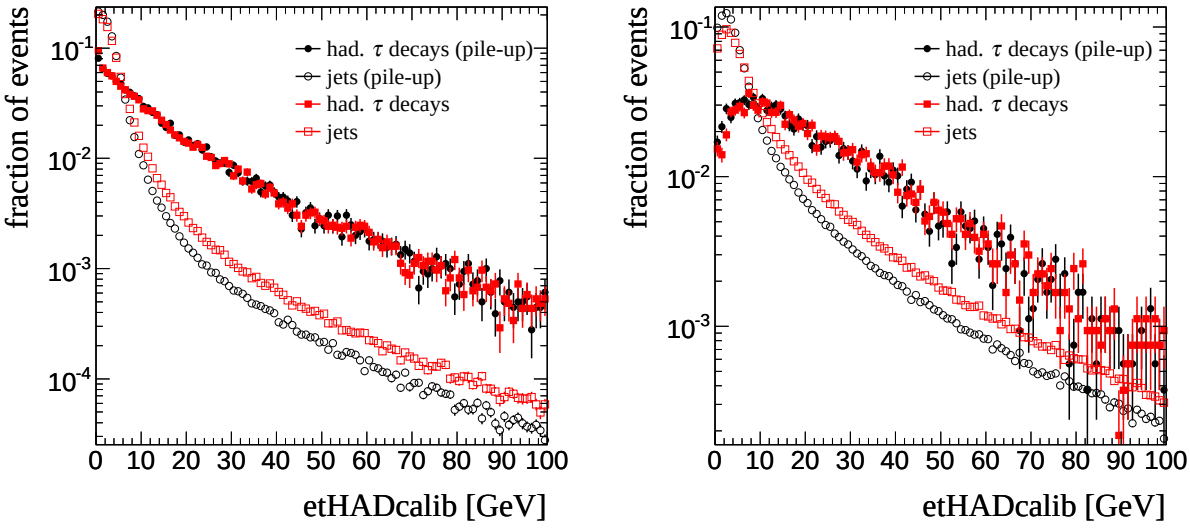
(a) 1-prong τ_{had} candidates **(b)** 3-prong τ_{had} candidates

Figure C.4.: Strip width.



(a) 1-prong τ_{had} candidates **(b)** 3-prong τ_{had} candidates

Figure C.5.: Calibrated E_T in the ECAL.



(a) 1-prong τ_{had} candidates

(b) 3-prong τ_{had} candidates

Figure C.6.: Calibrated E_T in the HCAL.

D

Loose Cuts Selection

The loos cuts selection comprises the standard selection cuts with modified cut values. For the pre-selection, electrons and muons are required to have $p_{\mathrm{T}} > 15\,\mathrm{GeV}$. τ_{had} candidates must have $p_{\mathrm{T}} > 20\,\mathrm{GeV}$. The following cuts are modified:

- $\not{E}_{\mathrm{T}} > 20\,\mathrm{GeV}$,
- $0 < x_{\ell} < 1$,
- $M_{\mathrm{T}} < 40\,\mathrm{GeV}$.
- $p_{T,j_1} > 20\,\mathrm{GeV}$,
- $\Delta\eta_{j_1,j_2} > 2.5$,
- $M_{jj} > 350\,\mathrm{GeV}$,

E

HEPEVT

The HEPEVT record of single $Z \to \tau\tau$ decay extracted from a $Z \to \mu\mu$ event comprises 10 lines. The first line contains an event number and the number of particles in this event. The other 9 lines contain information about the two τ leptons (lines 5-10) and the Z boson (lines 2-4). The first line of each particle entry contains an index, the status code (2 for unstable, 1 for stable final state particles), the pdg ID (23 for the Z boson and ± 15 for the τ leptons), two indices of mother particles and two indices of decay products. The next line is made up of the three momentum components, energy and mass. The last line contains the position of the production vertex. The two τ leptons have the 4-vectors of the original muons scaled such that the mass is equal to m_τ. The Z boson is the sum of these 4-vectors. The production vertex is the reconstructed vertex to which the muons point.

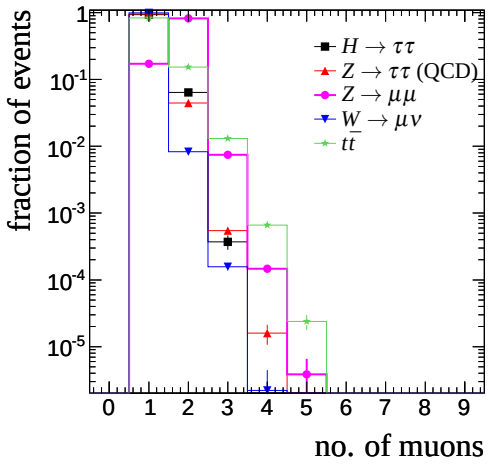
```
4951 3
1 2 23 0 0 2 3
13.19189882 -9.36381888 358.26952854 369.87071573 90.47421719
1.50303829 2.49593687 -4.44014549 0.00000000
2 1 -15 1 0 0 0
41.35693271 17.15949321 266.84691430 270.58321731 1.77700000
1.50303829 2.49593687 -4.44014549 0.00000000
3 1 15 1 0 0 0
-28.16503389 -26.52331210 91.42261424 99.28749842 1.77700000
1.50303829 2.49593687 -4.44014549 0.00000000
```

Figure E.1.: A HEPEVT record of a typical $Z \to \tau\tau$ decay in the embedding procedure.

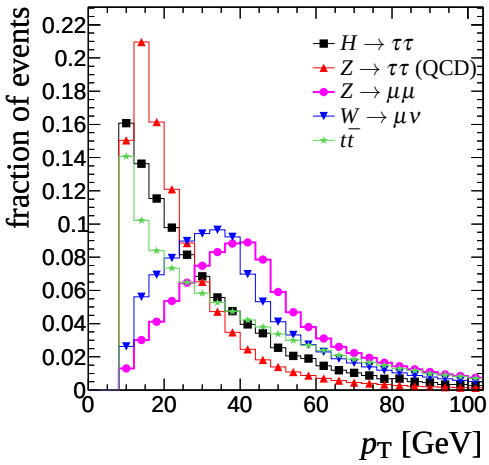
F

Selection of $Z \rightarrow \mu\mu$ Events for the Embedding Procedure

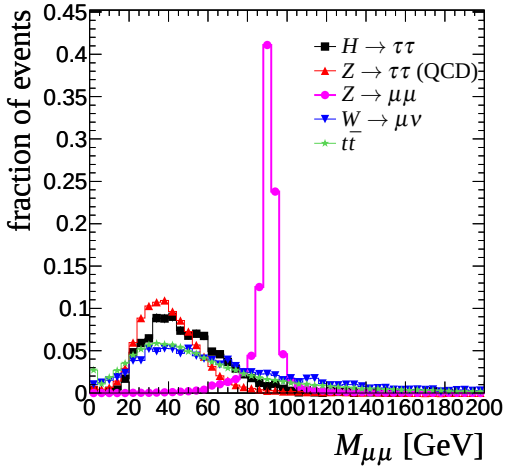
The figures in this chapter show the observables used by the $Z \rightarrow \mu\mu$ selection as input to the embedding procedure.



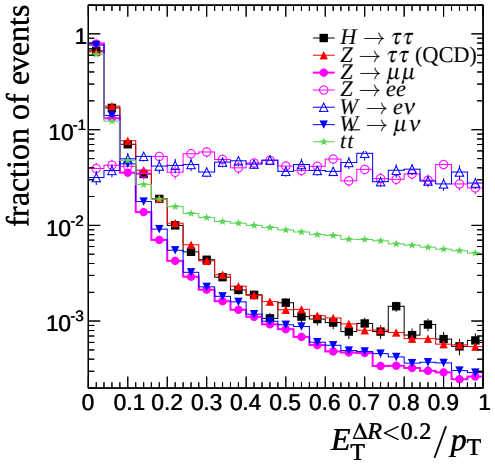
(a) Number of reconstructed muons per event.



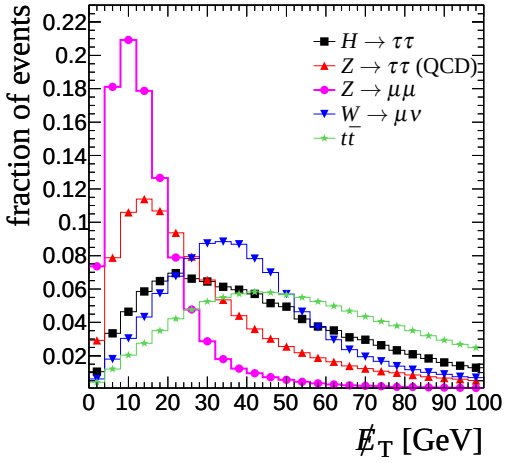
(b) Muon p_T.



(c) Di-muon mass.

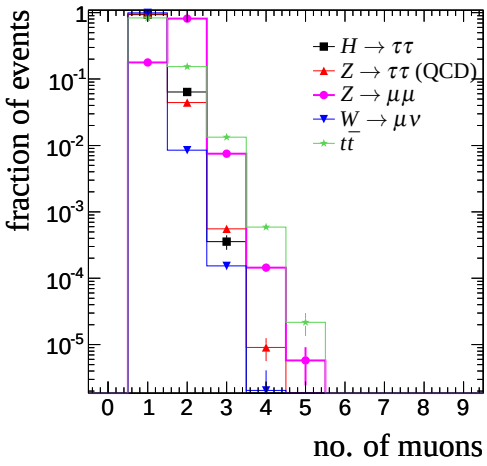


(d) Relative isolation of muons.

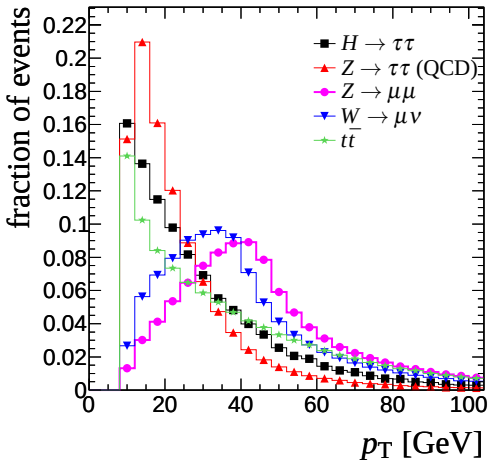


(e) Missing transverse energy.

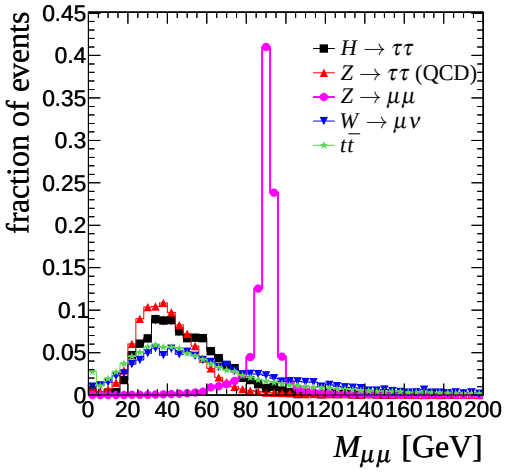
Figure F.1.: Observables used in the $Z \to \mu\mu$ selection (without pile-up).



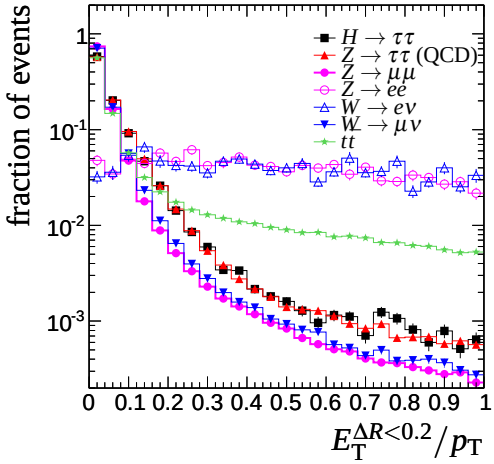
(a) Number of reconstructed muons per event.



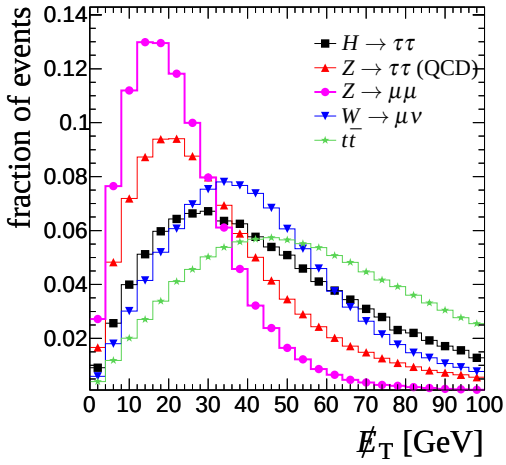
(b) Muon p_T.



(c) Di-muon mass.



(d) Relative isolation of muons.



(e) Missing transverse energy.

Figure F.2.: Observables used in the $Z \to \mu\mu$ selection (with pile-up).

Bibliography

[1] **The ATLAS** Collaboration, G. Aad, NM, *et al.*, "Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics," arXiv:0901.0512 [hep-ex].

[2] **Particle Data Group** Collaboration, K. Nakamura *et al.*, "Review of particle physics," *J. Phys.* **G 37** (2010) 075021

[3] F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. Wiley, 1984

[4] M. Warsinsky, M. Kobel, M. Schumacher, and F. Krauss, *Studien zur Produktion neutraler Higgs-Bosonen des Minimal Supersymmetrischen Standardmodells zusammen mit b-Quarks und deren Zerfall in Myonen am ATLAS Experiment. oai:cds.cern.ch:1312682*. PhD thesis, Dresden, TU Dresden, Dresden, 2008. Presented on 15 Sep 2008

[5] T. Nakano and K. Nishijima, "Charge Independence for V-particles," *Progress of Theoretical Physics* **10** no. 5, (1953) 581–582. http://ptp.ipap.jp/link?PTP/10/581/

[6] M. Gell-Mann, "The interpretation of the new particles as displaced charge multiplets," *Il Nuovo Cimento (1955-1965)* **4** (1956) 848–866. http://dx.doi.org/10.1007/BF02748000. 10.1007/BF02748000

[7] **LEP Electroweak Working Group** Collaboration, "Update Summer 2010,". http://lepewwg.web.cern.ch/LEPEWWG/

[8] G. 't Hooft, "Renormalization of Massless Yang-Mills Fields," *Nucl. Phys.* **B33** (1971) 173–199.

[9] G. 't Hooft, "Renormalizable Lagrangians for Massive Yang-Mills Fields," *Nucl. Phys.* **B35** (1971) 167–188.

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

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

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

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

[14] A. Djouadi, "The Anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model," *Phys. Rept.* **457** (2008) 1–216, arXiv:hep-ph/0503172.

[15] T. CDF, D. Collaborations, t. T. N. Phenomena, and H. W. Group, "Combined CDF and D0 Upper Limits on Standard Model Higgs Boson Production with up to 8.6 fb-1 of Data," arXiv:1107.5518 [hep-ex].

[16] **ATLAS** Collaboration, "Update of the Combination of Higgs Boson Searches in 1.0 to $2.3\,\text{fb}^{-1}$ of pp Collisions Data Taken at $\sqrt{s} = 7\,\text{TeV}$ with the ATLAS Experiment at the LHC," *ATLAS Note* **ATL-PHYS-PUB-2011-135** (2011)

[17] **CMS** Collaboration, "Combination of Higgs Searches," *CMS Note* **CMS-PAS-HIG-11-022** (2011)

[18] R. P. Feynman, "Very High-Energy Collisions of Hadrons," *Phys. Rev. Lett.* **23** no. 24, (Dec, 1969) 1415–1417

[19] *ATLAS detector and physics performance: Technical Design Report, 1*. Technical Design Report ATLAS. CERN, Geneva, 1999

[20] G. L. Bayatian *et al.*, *CMS Physics Technical Design Report Volume I: Detector Performance and Software*. Technical Design Report CMS. CERN, Geneva, 2006. There is an error on cover due to a technical problem for some items

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

[22] *LHCb : Technical Proposal*. Tech. Proposal. CERN, Geneva, 1998

[23] S. Agostinelli *et al.*, "G4–a simulation toolkit," *Nucl. Instrum. Methods A* **506** no. 3, (2003) 250 – 303

[24] J. Allison *et al.*, "Geant4 developments and applications," *IEEE Trans. Nucl. Sci.* **53** no. 1, (Feb., 2006) 270 –278

[25] CERN AC Team, "The four main LHC experiments." Jun, 1999

[26] T. S. Pettersson and P. Lefèvre, "The Large Hadron Collider: conceptual design. oai:cds.cern.ch:291782," Tech. Rep. CERN-AC-95-05 LHC, CERN, Geneva, Oct, 1995

[27] V. Gribov and L. Lipatov, "e+ e- pair annihilation and deep inelastic e p scattering in perturbation theory," *Sov.J.Nucl.Phys.* **15** (1972) 675–684

[28] Y. 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) 641–653

[29] G. Altarelli and G. Parisi, "Asymptotic Freedom in Parton Language," *Nucl.Phys.* **B126** (1977) 298

[30] V. V. Sudakov, "Vertex parts at very high-energies in quantum electrodynamics," *Sov.*

Phys. JETP **3** (1956) 65–71.

[31] T. Sjostrand, "Monte Carlo Generators," arXiv:hep-ph/0611247.

[32] T. Sjostrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual," *JHEP* **05** (2006) 026, arXiv:hep-ph/0603175.

[33] G. Corcella *et al.*, "HERWIG 6.5: an event generator for Hadron Emission Reactions With Interfering Gluons (including supersymmetric processes)," *JHEP* **01** (2001) 010, arXiv:hep-ph/0011363.

[34] G. P. Salam, "Elements of QCD for hadron colliders," arXiv:1011.5131 [hep-ph].

[35] **D0** Collaboration, V. M. Abazov *et al.*, "Measurements of differential cross sections of Z/γ^*+jets+X events in proton anti-proton collisions at $\sqrt{s}$=1.96 TeV," *Phys. Lett.* **B678** (2009) 45–54, arXiv:0903.1748 [hep-ex].

[36] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, "ALPGEN, a generator for hard multiparton processes in hadronic collisions," *JHEP* **07** (2003) 001, arXiv:hep-ph/0206293.

[37] T. Gleisberg *et al.*, "Event generation with SHERPA 1.1," *JHEP* **02** (2009) 007, arXiv:0811.4622 [hep-ph].

[38] J. Pequenao, "Computer generated image of the whole ATLAS detector." Mar, 2008

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

[40] J. Pequenao, "Computer generated image of the ATLAS inner detector." Mar, 2008

[41] J. Pequenao, "Computer Generated image of the ATLAS calorimeter." Mar, 2008

[42] J. Pequenao, "Computer generated image of the ATLAS Muons subsystem." Mar, 2008

[43] R. Nicolaidou *et al.*, "Muon identification procedure for the ATLAS detector at the LHC using Muonboy reconstruction package and tests of its performance using cosmic rays and single beam data," *J. Phys. Conf. Ser.* **219** (2010) 032052.

[44] **H1** Collaboration, I. Abt *et al.*, "The Tracking, calorimeter and muon detectors of the H1 experiment at HERA ," *Nucl. Instrum. Meth.* **A386** (1997) 348–396.

[45] E. Richter-Was and T. Szymocha, "Hadronic tau identification with track based approach: the $Z \to \tau\tau, W \to \tau\nu$ and dijet events from DC1 data samples," Tech. Rep. ATL-PHYS-PUB-2005-005, CERN, Geneva, 2005

[46] A. Kalinowski and K. Benslama, "Tau identification with the logarithmic likelihood method," Tech. Rep. ATL-PHYS-INT-2008-037. ATL-COM-PHYS-2008-125, CERN, Geneva, Sep, 2008

[47] H. Pi, P. Avery, D. Green, J. Rohlf, and C. Tully, "Measurement of Missing Transverse

Energy With the CMS Detector at the LHC," Tech. Rep. CMS-NOTE-2006-035, CERN, Geneva, Feb, 2006

[48] **LHC Higgs Cross Section Working Group** Collaboration, S. Dittmaier *et al.*, "Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables," arXiv:1101.0593 [hep-ph].

[49] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, "Jet substructure as a new Higgs search channel at the LHC," *Phys. Rev. Lett.* **100** (2008) 242001, arXiv:0802.2470 [hep-ph].

[50] A. Stahl, *Physics with Tau Leptons (Springer Tracts in Modern Physics)*. Springer, 2000

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

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

[53] S. Jadach, Z. Was, R. Decker, and J. H. Kuhn, "The tau decay library TAUOLA: Version 2.4," *Comput. Phys. Commun.* **76** (1993) 361–380

[54] E. Barberio and Z. Was, "PHOTOS - a universal Monte Carlo for QED radiative corrections: version 2.0," *Computer Physics Communications* **79** no. 2, (1994) 291 – 308

[55] N. Andari *et al.*, "Higgs Production Cross Sections and Decay Branching Ratios," Tech. Rep. ATL-COM-PHYS-2010-046, CERN, Geneva, Jan, 2010

[56] S. Asai *et al.*, "Cross sections for Standard Model processes to be used in the ATLAS CSC notes," Tech. Rep. ATL-PHYS-INT-2009-003. ATL-COM-PHYS-2008-077, CERN, Geneva, May, 2008

[57] K. Melnikov and F. Petriello, "The W boson production cross section at the LHC through O(alpha**2(s)," *Phys. Rev. Lett.* **96** (2006) 231803, arXiv:hep-ph/0603182.

[58] K. Melnikov and F. Petriello, "Electroweak gauge boson production at hadron colliders through O(alpha(s)**2)," *Phys. Rev.* **D74** (2006) 114017, arXiv:hep-ph/0609070.

[59] S. Moch and P. Uwer, "Theoretical status and prospects for top-quark pair production at hadron colliders," *Phys. Rev.* **D78** (2008) 034003, arXiv:0804.1476 [hep-ph].

[60] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics," *The European Physical Journal C - Particles and Fields* **71** (2011) 1–19. http://dx.doi.org/10.1140/epjc/s10052-011-1554-0. 10.1140/epjc/s10052-011-1554-0

[61] **ATLAS** Collaboration, "Discovery Potential of $A/H \to \tau^+\tau^- \to \ell h$ in ATLAS," Tech. Rep. ATL-PHYS-PUB-2010-011, CERN, Geneva, Jul, 2010

[62] E. T. Bell, "Exponential Numbers," *Amer. Math. Monthly* **41** (1934) 411–419

[63] D. L. Rainwater, D. Zeppenfeld, and K. Hagiwara, "Searching for $H \to \tau\tau$ in weak boson fusion at the LHC," *Phys. Rev.* **D59** (1998) 014037, `arXiv:hep-ph/9808468`.

[64] **ATLAS** Collaboration, "Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment," Tech. Rep. ATLAS-CONF-2011-077, CERN, Geneva, May, 2011

[65] **ATLAS** Collaboration, "Reconstruction and Identification of Hadronic τ Decays," in *[1]*, pp. 230–259. 2009. Detailed internal record: ATL-PHYS-INT-2008-003

[66] **ATLAS** Collaboration, "Search for the Standard Model Higgs Boson via Vector Boson Fusion Production Process in the Di-Tau Channels," in *[1]*, pp. 1271–1305. 2009. Detailed internal record: ATL-PHYS-INT-2009-053

[67] M. Röder, "Untersuchung verschiedener Monte-Carlo-Generatoren für die Analyse der Vektor-Boson-Fusions-Reaktion $qq \to H$ mit $H \to \tau\tau \to \ell\ell + 4\nu$ mit dem ATLAS-Detektor," Master's thesis, University of Bonn, 2008

[68] **ATLAS** Collaboration, "Search for neutral MSSM Higgs bosons decaying to $\tau^+\tau^-$ pairs in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS Experiment," Tech. Rep. ATLAS-CONF-2011-024, CERN, Geneva, Mar, 2011

[69] **ATLAS** Collaboration, "ATLAS Twiki." `twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults`, May, 2011

[70] C. Ruwiedel, N. Wermes, and H. Ströher, *Identification of hadronic τ decays using the τ lepton flight path and reconstruction and identification of jets with a low transverse energy at intermediate luminosities with an application to the search for the Higgs boson in vector boson fusion with the ATLAS experiment at the LHC. oai:cds.cern.ch:133219*. PhD thesis, University of Bonn, Bonn, 2010

[71] M. Schmitz, *Higgs production in Vector Boson Fusion in the $H \to \tau\tau \to \ell\ell + 4\nu$ final state with ATLAS: a sensitivity study*. PhD thesis, University of Bonn, Bonn, 2011

[72] A. Elagin, P. Murat, A. Pranko, and A. Safonov, "A New Mass Reconstruction Technique for Resonances Decaying to di-tau," `arXiv:1012.4686 [hep-ex]`.

[73] M. Schmitz, "Studie zur Bestimmung des Untergrundes aus Daten und der Higgs-Boson-Masse in Vektorbosonfusion mit $H \to \tau\tau \to \mu\mu + 4\nu$ mit dem ATLAS-Detektor am LHC," Master's thesis, University of Bonn, 2006

[74] T. Sjöstrand *et al.*, "Z Physics at LEP 1," vol. 3, p. 327. 1989. CERN Yellow Report 89-08

[75] N. Möser, M. Schmitz, J. Kroseberg, M. Schumacher, and N. Wermes, "Estimation of $Z^0 \to \tau\tau$ Background in VBF $H \to \tau\tau$ Searches from $Z^0 \to \mu\mu$ Data using an Embedding Technique," Tech. Rep. ATL-PHYS-INT-2009-109, CERN, Geneva, Dec, 2009

[76] C. Isaksson, M. Flechl, N. Möser, and M. Schmitz, "Embedding Technique for the ttbar Background Estimation in Charged Higgs Boson Searches," Tech. Rep. ATL-PHYS-INT-2010-041, CERN, Geneva, Apr, 2010

[77] S. Menke and G. L. Usai, "Proposal for Compactification of CaloCell Objects," Tech. Rep. ATL-SOFT-2004-005. ATL-COM-SOFT-2004-007. CERN-ATL-COM-SOFT-2004-007, CERN, Geneva, 2004

[78] "Data-driven estimation of the background to charged Higgs boson searches using hadronically-decaying tau final states in ATLAS," Tech. Rep. ATLAS-CONF-2011-051, CERN, Geneva, Apr, 2011

[79] **ATLAS** Collaboration, "Measurement of the $W \to \tau\nu$ Production Cross Section in pp Collisions at $\sqrt{s} = 7\,\text{TeV}$ with the ATLAS Experiment,"

[80] "A measurement of the total $W^{\pm}$ and Z/γ^* cross sections in the e and μ decay channels and of their ratios in pp collisions at $\sqrt{s} = 7\,\text{TeV}$ with the ATLAS detector," Tech. Rep. ATLAS-CONF-2011-041, CERN, Geneva, Mar, 2011

[81] J. Campbell and R. K. Ellis, 2007. User Guide available at `http://mcfm.fnal.gov/`

[82] S. S. Wilks *Ann. Math. Statist.* **9** (1938) 60–62

[83] A. Wald, "Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large," *Transactions of the American Mathematical Society* **54** no. 3, (1943) 426–482. `http://www.jstor.org/stable/1990256`

[84] E. Gross and O. Vitells, "Trial factors for the look elsewhere effect in high energy physics," *The European Physical Journal C - Particles and Fields* **70** (2010) 525–530, `arXiv:1005.1891v3 [hep-ex]`.

Danksagung (Acknowledgements)

Abschließend möchte ich mich gerne bei allen bedanken, die zum guten Gelingen dieser Arbeit beigetragen, die mit mir Frust oder Freude geteilt, die mich begleitet haben.

Mein Dank gilt zuallererst meiner Familie: Ohne euren Rückhalt wäre diese Arbeit schlichtweg nicht möglich gewesen. Ich hoffe, ihr seht mir vergessene Weihnachtsgeschenke, chaotische Geburtstagsfeiern und generelle geistige Abwesenheit nach.

Bei Professor Markus Schumacher und Professor Norbert Wermes möchte ich mich bedanken, dass sie es mir ermöglicht haben, diese Doktorarbeit – in der besonderen Konstellation Bonn-Freiburg – zu schreiben. Die große Unterstüztung, auch in persönlich schwieriger Zeit, hat mir viel bedeutet.

Ein besonderer Dank geht an Jürgen Kroseberg, zu dem ich immer mit allen Fragen kommen konnte, und der notfalls auch nach acht Stunden Zugfahrt nachts meine Folien für einen Vortrag am nächsten Tag durchgesehen hat.

Vielen Dank an Martin Schmitz für die gute Zusammenarbeit – besonders bei der Entwicklung der Embedding-Methode – die vielen, nicht nur fachbezogenen, Diskussionen und die jedes Mal exzellenten Weinbestellungen. Den Mitgliedern der BNWTGS danke ich für die wunderbaren Sonderseminare. Ich bedanke mich bei allen Kollegen, insbesondere Jan Schumacher, Marc Lehmacher, Guilherme Nunes Hanninger und Thomas Schwindt: Ich habe mich in der Arbeitsgruppe immer sehr wohl gefühlt. Besten Dank auch an die Reculet Hikers' Society für die schöne Zeit, die wir zusammen am CERN verbracht haben. Auch bei meinen "externen" Korrekturleserinnen Jana Schaarschmidt und Anna Engels-Putzka möchte ich mich auf diesem Wege bedanken.

Zuletzt, aber ganz sicher nicht am wenigsten, danke ich meinen Freunden, die mit mir zusammen Studium, Diplom- und Doktorarbeit, Höhen und Tiefen bestritten haben. Danke an die "Grillgruppe" – Christine, Claudia, Daniel, Markus, Sarah, Sascha, Sebastian und Verena – für die unzähligen gemeinsam verbrachten Stunden: beim Büffeln, im Urlaub und natürlich beim Grillen. Besonders möchte ich mich bei Claudia bedanken, die immer ein offenes Ohr und ein aufmunterndes Wort für mich hatte. Danke Anna und Jens, die Konzerte und die langen Abende mit gutem Essen werde ich immer in guter Erinnerung behalten. Vielen Dank auch an alle meine Freunde, die ich hier nicht namentlich erwähnt habe: Ohne euch alle wäre die Zeit hier nur halb so schön gewesen.

Vielen Dank, merci bien, thanks a lot!