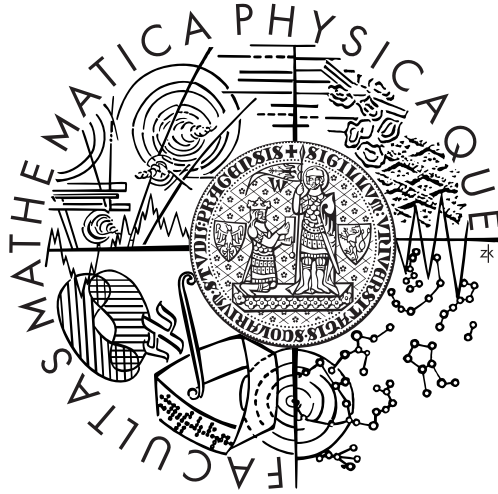


Charles University in Prague
Faculty of Mathematics and Physics

DIPLOMA THESIS



Jana Nováková

Study of the Higgs $\rightarrow \tau\tau$ Decay Reconstruction

Institute of Particle and Nuclear Physics

Supervisor: RNDr. Tomáš Davídek, Ph.D.

Study programme: Nuclear and Subnuclear Physics

I would like to express my big gratitude to my supervisor Tomáš Davídek for his help with both the physical analysis and the computing problems. My great acknowledgments belong also to Asen Christov for the data samples and for explanation of several terms about the tau leptons reconstruction.

I declare that I wrote my diploma thesis independently and exclusively with the use of cited sources. I approve of lending this thesis.

Prague, 19th April 2007

Jana Nováková

Contents

1	ATLAS detector at the LHC	1
1.1	LHC collider	1
1.2	Nomenclature	1
1.3	ATLAS experiment	1
2	Higgs boson at the LHC	5
2.1	Higgs boson searches	5
2.2	$H \rightarrow \tau\tau$ channel	8
3	Measurement of various quantities and tau reconstruction in the ATLAS experiment	11
3.1	Pattern recognition and secondary vertex reconstruction	11
3.2	Momentum and other track parameters	12
3.3	Energy measurement and calibration	13
3.4	Missing energy measurement	15
3.5	Identification and reconstruction of hadronic tau leptons	16
3.5.1	TauRec algorithm	16
3.5.2	Tau1P3P algorithm	18
4	Measurement of the Higgs boson mass with the collinear approximation	21
4.1	Kinematics of the $H \rightarrow \tau\tau \rightarrow h\nu_\tau h\nu_\tau$ channel	21
4.2	Principles of the collinear approximation	22
4.3	Results for the Higgs boson mass 120 GeV	23
4.3.1	Collinear approximation for the Monte Carlo truth	23
4.3.2	Collinear approximation for reconstructed events	25
5	Measurement of the Higgs boson mass from the tau lepton decay length	29
5.1	Description of the method	29
5.2	Full simulated sample for the Higgs boson mass 120 GeV	30
5.2.1	Resolution of the transverse decay length	31
5.2.2	Efficiency of the tau1P3P algorithm	35
5.3	Monte Carlo simulations for different Higgs boson masses	38
5.3.1	Mean value method for the mass estimation	39
5.3.2	Fractional method for the mass estimation	41
5.3.3	Accuracy of the decay lengths methods	42
6	Conclusion	45
	Bibliography	47

Název práce: Studium rekonstrukce rozpadu Higgs $\rightarrow \tau\tau$

Autor: Jana Nováková

Katedra (ústav): Ústav částicové a jaderné fyziky

Vedoucí diplomové práce: RNDr. Tomáš Davídek, Ph.D., ÚČJF

e-mail vedoucího: Tomas.Davidek@mff.cuni.cz

Abstrakt: Dvě nezávislé metody na rekonstrukci hmoty Higgsova bosonu v procesu $H \rightarrow \tau\tau$ jsou studovány a porovnány v této diplomové práci. Hmoty Higgsova bosonu je obvykle rekonstruována pomocí kolineární aproximace, která je založena na měření ztracené energie a viditelných částí tau leptonů. Neznámé impulsy neutrin vzniklých v rozpadu tau leptonů musí být odhadnuty. Hmoty je vyjádřena jako invariantní hmoty systému dvou tau leptonů. Alternativní metoda, tzv. doletová metoda, která je založena na měření rozpadové délky tau leptonů, je studována v této práci. Základem pro tuto metodu je rekonstrukce sekundárního vertexu, narozdíl od měření a kalibrace energie, které jsou velmi důležité pro kolineární aproximaci. Není očekáváno, že doletová metoda bude dávat tak přesné výsledky jako kolineární aproximace, ale může být použita jako nezávislý test výsledků kolineární aproximace.

Klíčová slova: kolineární aproximace, doletová metoda, Higgsův boson, tau lepton, experiment ATLAS

Title: Study of the Higgs $\rightarrow \tau\tau$ Decay Reconstruction

Author: Jana Nováková

Department: Institute of Particle and Nuclear Physics

Supervisor: RNDr. Tomáš Davídek, Ph.D., IPNP

Supervisor's e-mail address: Tomas.Davidek@mff.cuni.cz

Abstract: Two independent methods of the Higgs boson mass reconstruction in the channel $H \rightarrow \tau\tau$ are studied and compared in this thesis. The Higgs boson mass is usually reconstructed by means of a collinear approximation which is based on the measurement of missing energy and visible parts of tau leptons. Unknown momenta of neutrinos originating in the tau leptons decays have to be estimated. The mass peak is reconstructed as an invariant mass of the two tau leptons system. An alternative method, so-called decay length method, based on the tau lepton range measurement is studied in this paper. The secondary vertex reconstruction is the basic issue for this method as opposed to the energy measurement and calibration which are of crucial importance for the collinear approximation. The decay length method is not expected to give as precise results as the collinear approximation. However, it can be used as an independent test of the results of the collinear method.

Keywords: collinear approximation, decay length method, Higgs boson, tau lepton, ATLAS experiment

Chapter 1

ATLAS detector at the LHC

1.1 LHC collider

The Large Hadron Collider (LHC) is a proton-proton accelerator which is being built in the European Organisation for Nuclear Research (CERN). First proton-proton collision is planned in the end of year 2007. Both proton beams will be accelerated to energy of 7 TeV. The design luminosity of the collider is $10^{34} \text{ cm}^{-2}\text{s}^{-1}$. Head-on collisions of the proton bunches occur every 25 ns. At the design luminosity there are about 23 interactions per crossing.

1.2 Nomenclature

The beam direction defines the z axis, the xy plane is the plane transverse to the beam line. The positive direction of the x axis is defined from the interaction point to the centre of the LHC ring and the positive y axis is pointing upwards. The azimuthal angle ϕ is measured around the beam axis and the polar angle θ is the angle from the beam axis. The transverse quantities such as transverse momentum p_T , transverse energy $E_T = E \sin \theta$ and missing transverse energy E_T^{miss} are defined in the xy plane.

Nomenclature used in the text:

$\eta = -\ln \tan(\theta/2)$	pseudorapidity
$\Delta R = \sqrt{\Delta^2 \eta + \Delta^2 \phi}$	distance in the $\eta - \phi$ space, cone size
$\text{RMS} = \sqrt{\sum_i (r_i - \langle r \rangle)^2 / N}$	root mean squared of a variable r (statistical quantity)
m_H	Higgs boson mass
m_τ	tau lepton mass
L_{xy}	decay length (range) in the transverse plane

Natural units $c = \hbar = 1$ are used in the paper.

1.3 ATLAS experiment

The ATLAS (A Toroidal LHC ApparatuS) is one of the LHC experiments. The purposes of the ATLAS experiment are precise measurements of the Standard Model (SM) parameters and discoveries beyond the SM. One of the main goals for the ATLAS collaboration is to detect the Higgs boson, which is associated with the mass generation, if it exists. New particles which are predicted by theories beyond the SM can be discovered at the

LHC energy. Supersymmetric particles and not yet observed heavy gauge bosons can be detected with the ATLAS detector. Quarks or leptons compositeness could be observed at the LHC. Precise measurements of the W boson and top quark masses as well as gauge couplings and CP violation in B mesons decays will be done with the ATLAS detector.

The ATLAS detector is 22 m high and 44 m long with weight of 7000 tons. The overall detector layout can be seen in the figure 1.1. Basic design criteria of the detector come from the physical programme of the ATLAS experiment [1]:

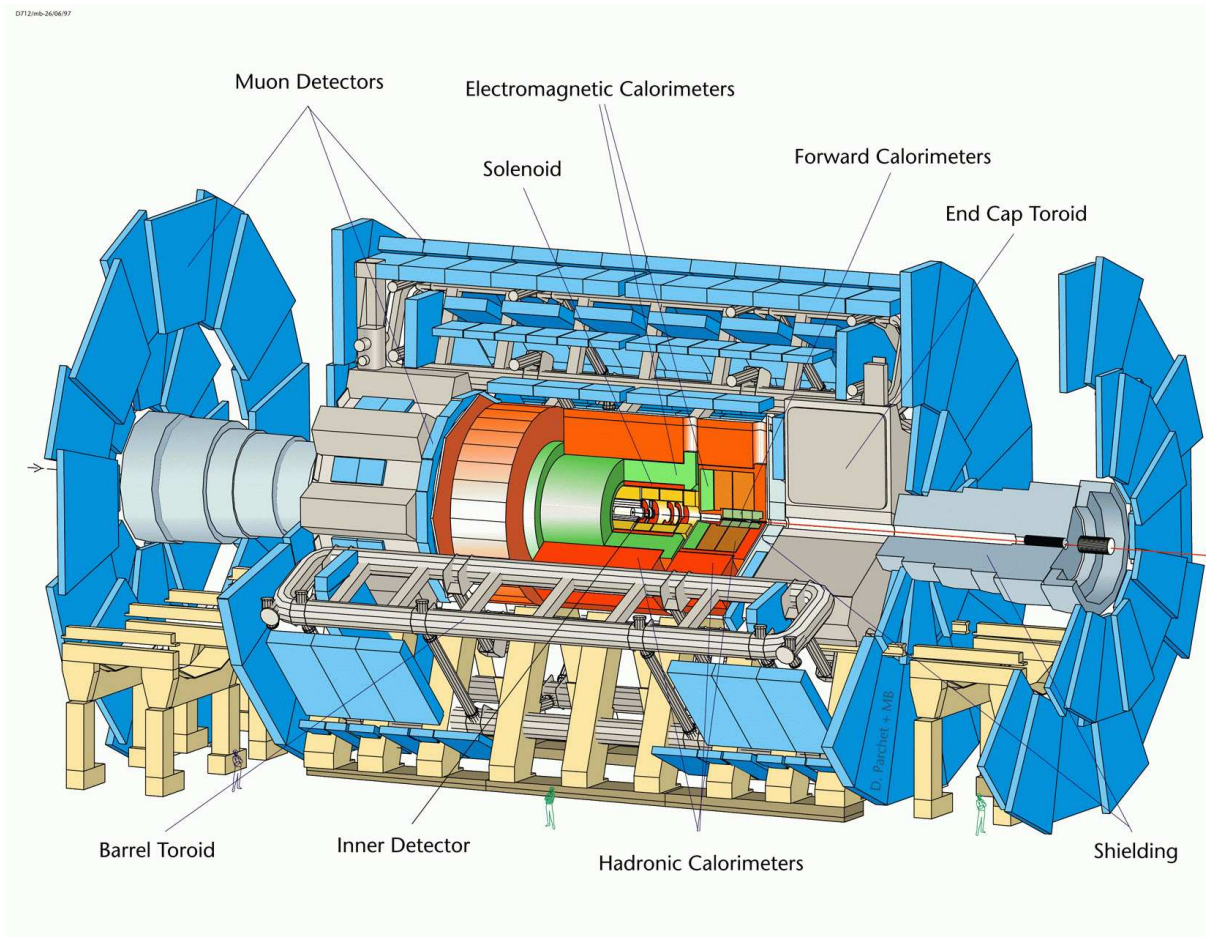


Figure 1.1: ATLAS detector overview [1]

- **Inner Detector**

The detection system located the most close to the interaction point in a solenoidal magnetic field of 2 T is the inner detector. The inner detector measures tracks of charged particles in the magnetic field. The main tasks of the inner detector are pattern recognition, momentum and vertex measurements which can be achieved only with high granularity detectors. The inner detector provides full tracking coverage over pseudorapidity range $|\eta| < 2.5$.

The highest granularity is performed around the primary vertex region using semiconductor pixel detectors. The pixel system contains a total of 140 millions detector elements, each $50 \mu\text{m}$ in $R\phi$ and $300 \mu\text{m}$ in the z direction. Typically three pixel layers are crossed by each track. The spatial resolution is $12 \mu\text{m}$ in $R\phi$ and $66 -$

$77 \mu\text{m}$ in z . The next inner detector layer is performed by a silicon semiconductor tracker (SCT). The SCT is designed to provide eight measurements per track in the intermediate radial range. The spatial resolution is $16 \mu\text{m}$ in $R\phi$ and $580 \mu\text{m}$ in z . Tracks can be distinguished if separated by more than $\sim 200 \mu\text{m}$. The last layer is a transition radiation tracker (TRT) which is based on the use of straw detectors. Each straw is 4 mm in diameter and it is filled with a xenon gas mixture to add the electron identification capability (transition-radiation photons created in a radiator between straws can be detected with the xenon gas). The TRT system allows a large number of measurements, typically 36, to be made on every track. Position accuracy of $170 \mu\text{m}$ in the $R\phi$ can be achieved by average straw, but the large number of straws per track improves the resolution to $50 \mu\text{m}$ at the LHC design luminosity.

- **Calorimeters**

Basic requirements for the electromagnetic calorimeter (EM) are a very good electron and photon identification and measurements. The EM calorimeter is a lead/liquid-argon (LAr) detector. The EM consists of a barrel part ($|\eta| < 1.5$) and two end-caps ($1.4 < |\eta| < 3.2$). The barrel as well as the end-caps are further segmented into longitudinal sections with a different granularity in η and ϕ in order to achieve very good position measurement, especially in η . The total thickness of the EM calorimeter is more than 24 radiation lengths (X_0) in the barrel and more than 26 X_0 in the end-caps.

The hadronic calorimeter is supposed to provide full-coverage for accurate jet and missing transverse energy measurements. A sampling detector with an iron absorber and plastic scintillator plates (tiles) as the active material is used over the range $|\eta| < 1.7$ for the barrel and the extended barrel creating the Tilecal. At larger pseudorapidities the liquid-argon detectors are used. The hadronic end-cap calorimeter (HEC) is used in the region $1.5 < |\eta| < 3.2$, while the range $3.1 < |\eta| < 4.9$ is covered by the high density forward calorimeter (FCAL).

- **Muon chambers**

The muon spectrometer measures the deflection of muon tracks in the large superconducting air-core toroid magnets. The precise measurement of muon tracks can be made in the $R - z$ projection, in a direction parallel to the bending direction of the magnetic field. In the barrel region, tracks are measured in chambers arranged in three cylindrical layers around the beam axis. In the transition and end-cap regions, the chambers are installed vertically. Over most of the pseudorapidity range, a precision measurement of the track coordinates is provided by Monitored Drift Tubes (MDTs). The MDTs can provide a single-wire resolution of $80 \mu\text{m}$. At larger pseudorapidities and close to the interaction point, Cathode Strip Chambers (CSCs) with higher granularity are used. The muon system involves independent fast trigger chamber system.

The ATLAS trigger and data-acquisition (DAQ) system is based on three levels of online event selection. The interaction rate of 10^9 Hz is expected at the LHC design luminosity. The rate of selected events must be reduced to about 100 Hz for permanent storage. Excellent trigger efficiency is required for observing physical processes with very low cross-sections which are the main goals for the ATLAS experiment. The level one (LVL1) trigger makes the first selection from reduced-granularity information of a subset

of detectors. The LVL1 trigger decision is based on combinations of objects required in coincidence or veto. The first trigger is supposed to reduce the rate of selected events to 75 kHz. The LVL2 trigger works with data from small fractions of the detector around objects indicated by the LVL1 trigger, so-called regions of interest. The LVL2 trigger is expected to reduce the rate to 1 kHz. After LVL2 trigger the last online selection is performed by the Event Filter (EF). The EF uses offline methods which are adapted to the online analysis. It performs the final selection of events which will be stored for the offline analysis. The output rate is supposed to be 100 Hz which corresponds to an output data rate of the order of magnitude of 100 MB/s.

Chapter 2

Higgs boson at the LHC

2.1 Higgs boson searches

Experimental observations of the Z and W bosons as well as the top quark detection strongly supported the Standard Model theory (SM). The Higgs boson is the only not yet observed elementary particle which is predicted by the SM. The Higgs boson is of crucial importance for an explanation of the mass symmetry breaking on the electroweak level.

The theoretical model of electroweak interactions is a quantum field theory based on the $SU(2) \times U(1)$ symmetry group. The Lagrangian is required to be invariant under this symmetry group. This construction leads to the existence of the intermedial bosons W^\pm , Z^0 and γ . It is not possible to introduce mass terms neither for the particles of matter (fermions) nor for the gauge bosons without breaking the symmetry at this point. The problem can be solved by the Higgs mechanism. One doublet of complex scalar fields is assumed in the SM. The Lagrangian remains invariant under the required symmetry group, but the ground state of the system is no longer invariant, so-called spontaneous symmetry breaking. This process leads to the existence of a new neutral scalar particle in the SM - the Higgs boson. The fermions and the intermedial bosons acquire their masses in interactions with the Higgs boson field. In extensions of the SM not only one doublet is considered which leads to more Higgs particles. In this paper only the SM Higgs boson is taken into the account.

The Higgs boson mass is not predicted in the SM. However, constrains on its mass can be found in both the theory and the experiment. The present-day Standard Model is supposed to be a perturbative theory. For the Higgs boson with the mass of the order of ~ 1 TeV the full decay width is almost as large as its mass [7]. The Higgs boson cannot be any longer considered as a particle and the perturbative calculations become inconsistent. Measurements on the LEP collider excluded the Higgs particle below 114.4 GeV with 95% confidence level (C.L.) [12]. Other clues are indirect measurements of the Higgs mass. If radiative corrections to the W boson propagator are assumed, the W mass is proportional to the top mass squared and the logarithm of the Higgs boson mass. It is possible to constrain the Higgs boson mass from precise measurements of the W boson and top quark masses, see the figure 2.1. The most probable value is $m_H = 76^{+33}_{-24}$ GeV with 68% C.L. [8] which has already been excluded by the measurements at the LEP collider. The Higgs particle is expected below 144 GeV with 95% C.L from the indirect measurements. The limit increases to 189 GeV if the LEP 2 direct search for the Higgs boson is included. The ATLAS detector is designed to detect the Higgs boson or bosons over the whole possible mass range up to 1 TeV.

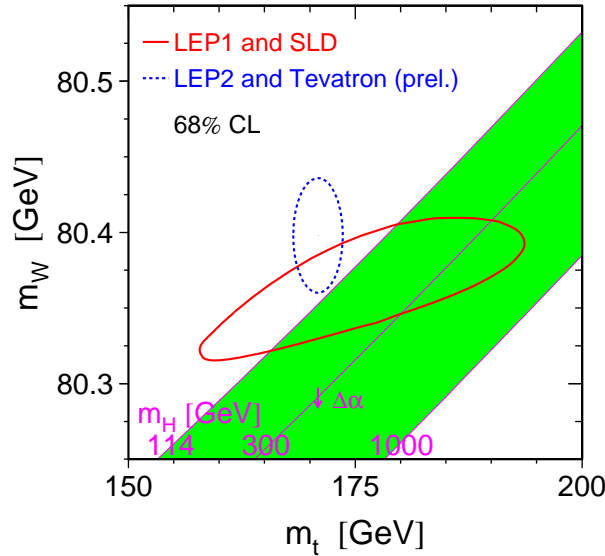


Figure 2.1: Indirect Higgs boson mass measurements - m_W vs m_{top} plot [8]

The Higgs boson production cross-sections depend on its mass. Feynman diagrams for two modes with the largest cross-sections at the LHC are shown in the figure 2.2 and the cross-sections for various production channels over the whole mass region can be seen in the figure 2.3. The Higgs boson production is dominated by the gluon-gluon fusion where the Higgs boson is produced through the top triangle. There is only the Higgs boson in the final state. This mode is called a direct production. The second largest cross-section comes from the vector boson fusion (VBF) which contributes by $\sim 10\% - 15\%$ to the total cross-section in the region where the Higgs mass is lower than double Z boson mass and becomes more important with increasing mass. The production is accompanied by two jets in the opposite detector region than the Higgs boson. There is a large pseudorapidity gap between the jets and the Higgs boson. The QCD jets tend to have higher activity in the central detector regions in comparison with the vector boson fusion events. This can be used to trigger the Higgs boson signal against the QCD background.

The Higgs particle is very unstable and decays immediately. The branching ratios of the Higgs decay modes depend on the Higgs mass. Possible discovery channels at the LHC with its signal significance for the Higgs mass below 200 GeV are shown in the figure 2.4. The signal significance [15] is calculated from the number of expected signal S and background B events as S/\sqrt{B} . The most important channels for the detection of the Higgs particle with mass lighter than double Z boson mass at the LHC are [1]:

- $H \rightarrow b\bar{b}$: This is the decay mode with the highest cross-section in the low mass region. There is a huge QCD jets background which has to be rejected. The direct production cannot be efficiently triggered. This channel can be observed only in the associated production with a W or Z boson or a $t\bar{t}$ pair. The leptonic decays of the Z and W bosons or semileptonic decays of the top quark provide an isolated high p_T lepton which can be easily triggered. Excellent b-tagging is required.
- $H \rightarrow \gamma\gamma$: The branching ratio for this channel is rather low ($< 0.5\%$), but it can

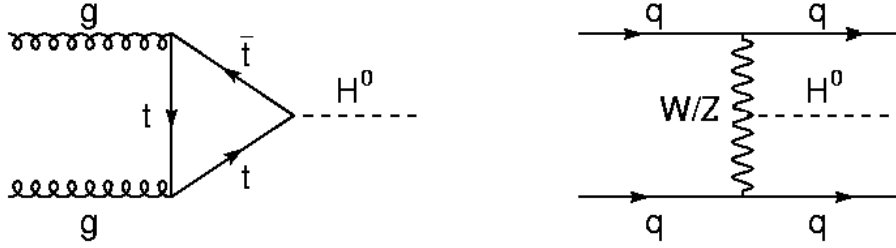


Figure 2.2: Feynman diagrams for the Higgs boson production through the gluon-gluon fusion (plot on the left) and through the vector boson fusion (plot on the right)

provide a very good mass resolution. In order to observe the narrow mass peak above the $\gamma\gamma$ continuum, very good energy and angular resolution of the electromagnetic calorimeter is required. Other important task is a good particle identification to reject the QCD jets background and background from the $Z \rightarrow ee$ decays in case the Higgs mass is close to the Z boson mass.

- $H \rightarrow ZZ^* \rightarrow 4\ell$: The symbol ℓ denotes both an electron (e) or a muon (μ), final states with $2e + 2\mu$, $4e$ and 4μ are studied. This channel provides a rather clean signature. The branching ratio is higher than in the $\gamma\gamma$ case. The main background comes from ZZ^* and $Z\gamma^*$ continuum or $t\bar{t}$ and $Zb\bar{b}$ production.
- $H \rightarrow WW^{(*)} \rightarrow \ell\nu_\ell \ell\nu_\ell$: This channel is an important mode for Higgs boson masses in the range between 150 and 190 GeV. The branching ratio is higher than for the previous channel in this mass region. Due to the presence of two neutrinos from the W bosons decays, the transverse mass distribution will be measured in order to estimate the Higgs mass.
- $H \rightarrow \tau\tau$: This channel will be described in more details in the next section.

If the Higgs boson mass is higher than double Z boson mass, the dominant channel will be the one with two real Z bosons as the decay products. The most promising channel is the mode with both Z bosons decaying into a pair of leptons ($Z \rightarrow ee$ or $Z \rightarrow \mu\mu$), so-called gold-plated channel, which gives a rather clean signature in the range $180 \text{ GeV} < m_H < 700 \text{ GeV}$. The measured leptons can be combined into two pairs with an invariant mass in a given mass window around the mean Z boson mass. The expected background (mainly the continuum production of Z boson pairs) is smaller than the signal. The detection and measurement of the final state leptons with high momenta does not put any further requirements on the detector. This channel is not observable for the Higgs masses higher than 700 GeV due to low probability. Channels with higher branchings, $H \rightarrow ZZ \rightarrow \ell\ell \nu_\ell \nu_\ell$, $H \rightarrow ZZ \rightarrow \ell\ell jj$ and $H \rightarrow WW \rightarrow \ell\nu_\ell jj$ (where the symbol j stands for a jet), can be used in the very high mass region.

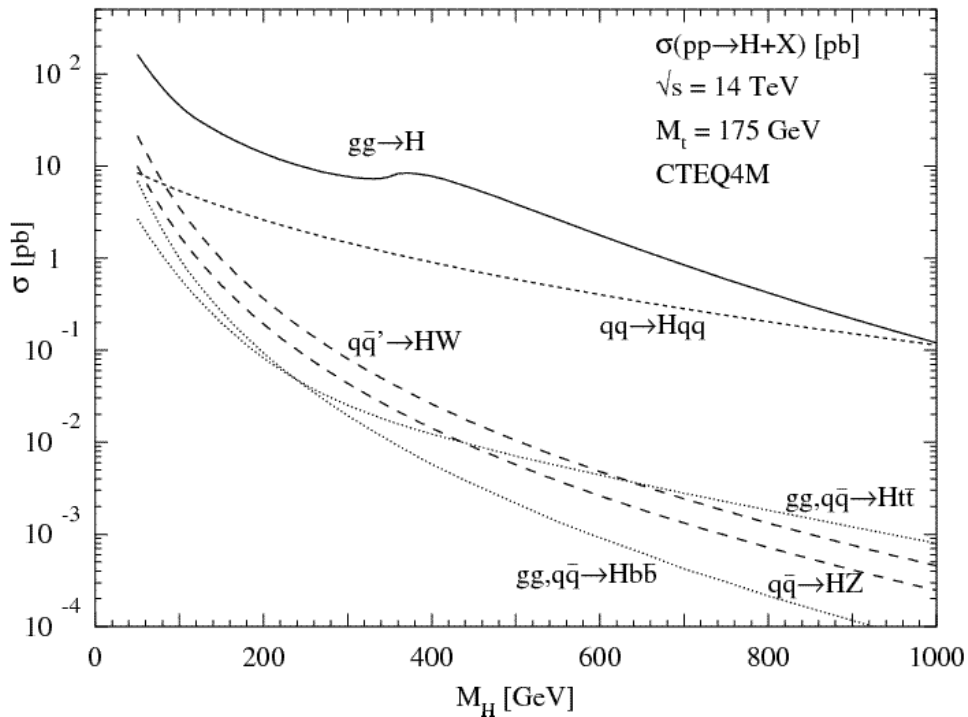


Figure 2.3: Cross-sections for various Higgs boson production modes at the LHC [15]

2.2 $H \rightarrow \tau\tau$ channel

The Feynman diagram for a Higgs boson decay into a fermion-antifermion pair at the tree level is shown in the figure 2.5.

The matrix element for this process can be written as

$$\mathcal{M} = -\frac{g}{2} \frac{m_f}{m_W} \bar{u}(p, s_1) v(k, s_2) \quad (2.1)$$

where g is the electroweak coupling, m_W is the W boson mass, m_f is the fermion mass, u, v are Dirac spinors, p, k denote four-momenta and s_1, s_2 spin states of the fermion, antifermion.

The decay width can be calculated from the knowledge of the matrix element (unpolarised final state is considered)

$$\Gamma = \frac{N_c}{8\pi} \left(\frac{g}{2} \frac{m_f}{m_W} \right)^2 m_H \left(1 - 4 \left(\frac{m_f}{m_H} \right)^2 \right)^{\frac{3}{2}} \quad (2.2)$$

where N_c stands for a number of colour states of the fermion ($N_c = 1$ for leptons, $N_c = 3$ for quarks). The term $(1 - 4(m_f/m_H)^2)^{\frac{3}{2}}$ in the equation (2.2) can be considered as one, because the Higgs boson mass is much higher than a mass of any known fermion ($m_f^2 \ll m_H^2$).

The ratio of the Higgs boson decays into different fermion types can be derived from the formula (2.2). The ratio scales as the square of the fermion mass. This means $\text{BR}(H \rightarrow \tau\tau) : \text{BR}(H \rightarrow \mu\mu) : \text{BR}(H \rightarrow ee) = m_\tau^2 : m_\mu^2 : m_e^2$ for leptonic final states. Decays $H \rightarrow \mu\mu$ are suppressed by a factor of 3.5×10^{-3} in comparison with the $\tau\tau$ channel. The decay width for the double tau channel is 2.5×10^{-4} GeV for the Higgs

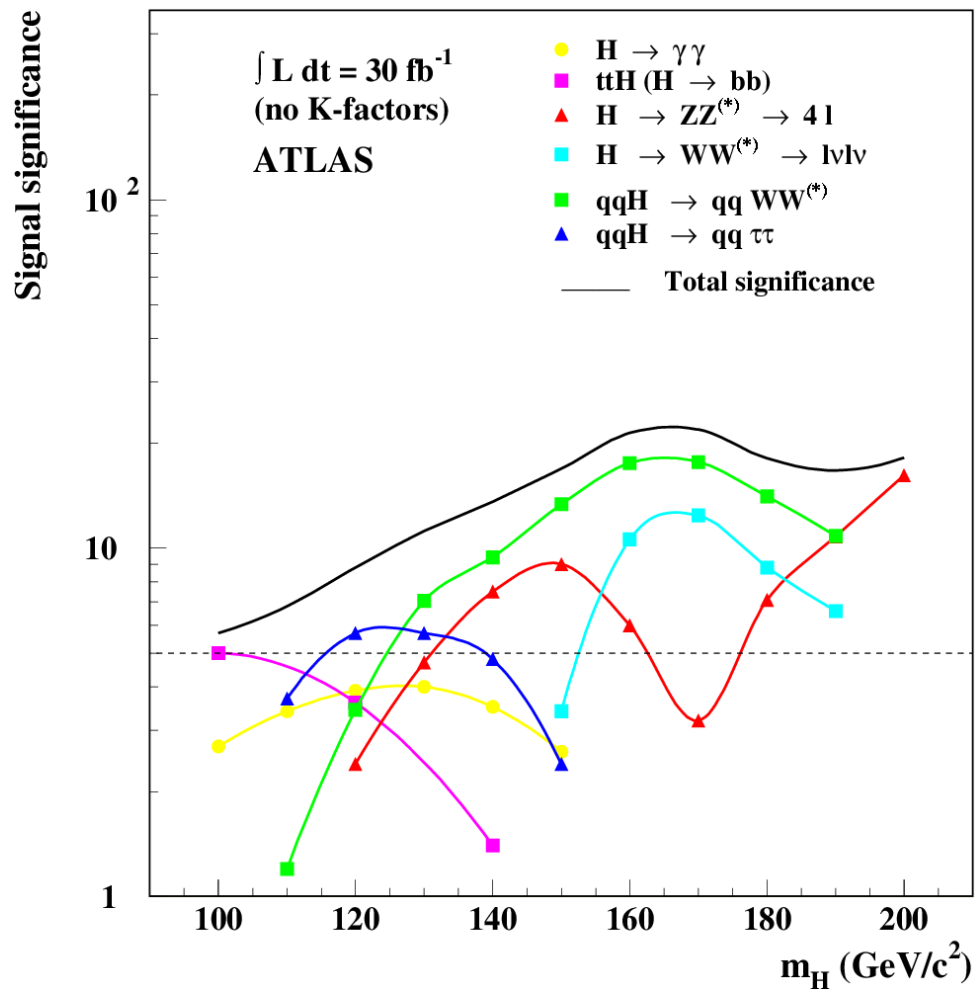


Figure 2.4: Signal significance for several Higgs boson discovery channels at the LHC [15]

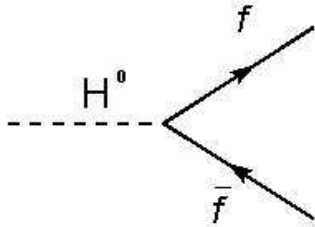


Figure 2.5: Feynman diagram for the $H \rightarrow f\bar{f}$ decay

boson mass 120 GeV. Similar considerations can be made about quark-antiquark final states. The top-antitop final state does not contribute to the Higgs boson width in the low mass region (below the double W/Z boson mass), because it is too heavy ($m_t \approx 174$ GeV). Only the decay modes with two b-quarks ($m_b \approx 4.2$ GeV) or c-quarks ($m_c \approx 1.25$ GeV) have branchings higher than 10^{-2} for the light Higgs boson. Decays into other quarks are strongly suppressed, because of their low mass (the next heaviest quark is the s-quark with the mass $m_s \approx 0.095$ GeV).

The ATLAS experiment is not sensitive to the $H \rightarrow c\bar{c}$ decays. The tau-tau channel is besides the $b\bar{b}$ mode the only direct fermionic channel measurable at the LHC. The Higgs coupling to fermions can be measured in this channel. The Higgs decay into two tau leptons is the only channel in the low mass region which has a signal significance higher than 5σ for the integrated luminosity 30 fb^{-1} , see the figure 2.4. The channel is observable for the Higgs boson mass in the range 115 – 140 GeV where it represents 4 – 7% of the Higgs decays [11].

The tau lepton mode can be efficiently triggered if the Higgs boson is produced through the vector boson fusion. Both leptonic and hadronic decays of the tau lepton (see the section 3.5 where tau leptons decay modes are described) can be used for the Higgs boson mass reconstruction. Only the mode with both taus decaying hadronically is assumed in this paper. This mode contributes by 42.3% to the $H \rightarrow \tau\tau$ events. Mainly γ^*/Z +jets and WW+jets events with two tau leptons, two jets and missing energy in the final state form the background. QCD jets wrong assigned as the tau leptons also contribute to the background events. Reliable tau identification algorithms with a very good level of tau/jet separation are required.

Chapter 3

Measurement of various quantities and tau reconstruction in the ATLAS experiment

Measurement and reconstruction of quantities which are of crucial importance for the Higgs boson mass prediction in the collinear approximation and the decay length method are described in this section. The tau lepton identification and reconstruction algorithms are also explained in this section.

3.1 Pattern recognition and secondary vertex reconstruction

The inner detector is sensitive to charged particles which create signals (hits) in the semiconductor detector elements and in the transition radiation tracker. The pattern recognition algorithms build possible tracks (track candidates) from selected hits in the inner detector. Firstly, hits which could belong to the same track are selected from a high number of hits in the inner detector. Typically, the track reconstruction is performed along paths which join the vertex region to seeds on the outer surface of the inner detector. The seeds are typically calorimeter clusters (electron/photon candidates from the EM calorimeter, jets from the hadronic calorimeter) and muons' tracks found in the muon system. Secondly, the selected hits with their various errors are fitted to a helical trajectory which is expected in an uniform magnetic field. However, the magnetic field is not perfectly homogeneous and special corrections have to be applied. Multiple scattering, energy losses in the material and bremsstrahlung radiation of electrons must be considered for the pattern recognition.

Two basic pattern recognition algorithms (iPatRec, xKalman) are used in the ATLAS experiment [2]. iPatRec finds combinations of precision hits (pixel and SCT hits) which satisfy specific criteria on the maximum curvature of the track. A track candidate is formed from the selected hits. The candidate's track is extrapolated to the TRT where appropriate hits are added. Tight cuts are made on the straw residuals and on the ratio of found to expected straw hits in order to limit high luminosity occupancy effects. The final track fitting is done on the selected hits. Conversely, the xKalman algorithm starts with TRT hits. Track candidates are found with fast histogramming of straw hits. Candidates are extrapolated to the SCT and pixels where the selected hits are fitted with a helical

curve. The accepted helix trajectories are extrapolated back into the TRT where a narrow region around the extrapolation is defined. The width of the region is determined by the errors of the helix trajectory. All straw hits within this region are included for the final track fitting. The track quality is determined for both algorithms. The track is considered as a good quality one if it passes cuts on the number of precision and TRT hits, number of layers without a signal in the precision layers, momentum of the track, χ^2 of the fit and number of maximum hits shared with another track.

The secondary vertex position is obtained from reconstructed tracks with a displacement from the beam axis. Typically, the secondary vertex is reconstructed as a 3D intersection point of selected charged tracks. The quality of individual tracks has to be taken into account. If there are two reconstructed tracks which are expected to form the secondary vertex, they do not have to intersect necessarily. In that case the point of the closest approach is considered as the secondary vertex. If there are more tracks, they usually do not intersect in one precise point, but they rather define a region where the vertex can be found. The vertex can be reconstructed with the least squares method. The vertex finding can be improved by excluding tracks with a high displacement from the vertex region or by considering weights for individual tracks.

3.2 Momentum and other track parameters

Transverse momenta of charged particles are measured in the inner detector from the curvature of tracks in the magnetic field. The momentum of the particle in the plane perpendicular to the uniform magnetic field equals

$$p_{\perp} = 0.3 qBr \quad (3.1)$$

where q denotes a charge of the particle as a multiple of an elementary charge, B is the intensity of the magnetic field and r is the radius of the projection of the helical trajectory in the plane perpendicular to the magnetic field. The numerical factor 0.3 comes from the conversion to the natural units. If B is measured in Tesla and r in meters, the momentum is expressed in GeV.

Solenoidal magnets placed around the inner detector generate magnetic field which can be approximately considered as homogeneous with the intensity $B = 2$ T in the direction of the beam axis in the ATLAS detector. The sagitta s of the track projection in the transverse plain can be calculated from the transverse position of hits measured along the track. The sagitta is proportional to the position measurements which means that it has a Gaussian error [2]. The transverse radius of the track is proportional to the reciprocal value of the sagitta. The quantity measured with a Gaussian error is $1/p_T$ ($1/p_T \propto s$). If the position and resolution of measurements, the material distribution and the magnetic field are uniform as a function of the transverse radius, the resolution in $1/p_T$ can be expressed as $\sigma(1/p_T) = A \oplus B/p_T$ [1]. The first term is associated with the spatial resolution of measured hits along the track and the second term with interactions of the charged particle with the detector material. However, the tracking system is not perfectly uniform and the magnetic field differs from the uniform field. The resolution can be written as (p_T in GeV)

$$\sigma\left(\frac{1}{p_T}\right) \approx 3.6 \times 10^{-4} \oplus \frac{1.3 \times 10^{-2}}{p_T \sqrt{\sin \theta}} \quad (3.2)$$

The relative resolution of the transverse momentum can be derived as $\sigma(p_T)/p_T = p_T \sigma(1/p_T)$. It can be seen that the resolution of the transverse momentum is getting worse with increasing momenta.

Not only the transverse momentum, but also other track parameters can be measured in the inner detector. These are the azimuthal angle ϕ , the polar angle θ and the impact parameter d_0 (the smallest distance of the track to the beam axis in the transverse plane). Similar parametrisation of the resolution as for $1/p_T$ can be used (p_T in GeV):

$$\sigma(\phi) \approx 0.075 \oplus \frac{1.8}{p_T \sqrt{\sin \theta}} \quad [\text{mrad}] \quad (3.3)$$

$$\sigma(\cot \theta) \approx 0.70 \times 10^{-3} \oplus \frac{2.0 \times 10^{-3}}{p_T \sqrt{\sin^3 \theta}} \quad (3.4)$$

$$\sigma(d_0) \approx 11 \oplus \frac{73}{p_T \sqrt{\sin \theta}} \quad [\mu\text{m}] \quad (3.5)$$

3.3 Energy measurement and calibration

Energy measurement is proceeded in calorimeters. Calorimeters are the only detecting devices sensitive to both charged and neutral particles. The energy measurement is based on the total absorption of the initial particle. In the electromagnetic or strong interaction of the high energy primary particle with the detector material new high energy particles are produced. The secondary particles with high energy loose their energy mainly in new reactions. The shower is developed by particles continuing in the multiple production until the secondary particles have low energy for new reactions. Charged secondary particles loose their energy mainly by ionisation in the lower energy region. The ionisation mechanism does not further develop the particle shower.

All ATLAS calorimeters are sampling devices with different types of absorber and active material. The absorber is made from a material where is a high cross-section for the interaction of the primary particle with the absorber material. The absorber in the EM calorimeter is made of material with a high atomic number (lead) where is very high probability of the interaction between the photon and atoms of the material. Hadrons have the highest cross-section for the interaction in the high density material. That is the reason why iron is used as the absorber in the Tilecal. The active material is placed between separate absorber layers. The ionisation charge (liquid argon calorimeters) or scintillation light (Tilecal) are produced and collected in the active material. Not only the total energy deposited in the calorimeter cells can be measured, but also the longitudinal profile of the shower can be observed by this method.

Difference between hadronic and electromagnetic showers have to be briefly described in order to understand different calibration of the EM and hadronic calorimeters. The particle production mechanism is more complicated in the hadronic showers than in the electromagnetic showers [16]. The hadronic showers consist of two components:

- *Electromagnetic component* comes from π^0 and η^0 mesons which decay to a $\gamma\gamma$ pair in the most cases.
- *Hadronic component* contains besides hadrons, which further continue in the multiple production of secondary particles, also invisible energy. The binding energy of the nucleus which is broken via the strong interaction, as well as neutrinos and muons which escape the detection contribute to the undetectable energy.

Since the hadronic showers contain invisible energy, the signal from the electron is larger than the signal from the hadron at the same energy ($e/\pi > 1$). The fraction of the initial hadron energy carried by the electromagnetic component (f_{em}) fluctuates with a large amplitude and the distribution is non-Gaussian. On the top of it, the mean value of f_{em} is energy dependent ($\langle f_{em} \rangle \propto \ln E$). These features of hadronic showers have a negative effect on the energy resolution. An important quantity which describes the calorimeter is the ratio e/h where e denotes the calorimeter response to the electromagnetic component and h to the non-electromagnetic component of the hadron showers. The ratio e/h is more than 1 for the ATLAS calorimeters. The quantity e/h is energy independent, but it is not directly measurable. It can be derived from a fit to the e/π values measured at different energies.

Very important issue is the energy calibration of the calorimeter. The collected signal is proportional to the energy of the primary particle. Calibrated energy is usually considered as a function of energy deposited in the calorimeter cell and a position of the given cell in the calorimeter. Calibration parameters are expected to minimise a square of the difference between the calibrated and true energy. The collected charge in the EM calorimeter can be considered to be a linear function of the true energy of the initial particle in the first approximation. The calibration of the EM calorimeter involves corrections for energy deposited in a dead material (e.g. cryostats) and in transition regions between different parts of the calorimeter (cracks). The calibration of the hadronic calorimeter is more complicated. Firstly, the deposited energy is calibrated to the electromagnetic scale in the ATLAS experiment. Secondly, the energy calibrated to the EM scale has to be corrected for the two-component structure of hadronic showers.

The sampling method [1] is one of the jet calibration methods used in the ATLAS experiment. The calibrated energy is calculated as a linear combination of energies measured in different parts of the calorimeter. The easiest way of the energy calibration is based just on two different calorimetric regions - electromagnetic and hadronic calorimeter:

$$E^{\text{calib}} = A(E_{\text{truth}}) E^{\text{elmg}} + B(E_{\text{truth}}) E^{\text{had}} \quad (3.6)$$

where the calibration weights A, B are energy dependent and $E^{\text{elmg}}, E^{\text{had}}$ stand for energy measured in the EM, hadronic calorimeter collected from cells in a given cone from the jet centre. The energy dependence of the calibration parameters is expected to be logarithmic ($A = A_1 + A_2 \ln E_{\text{truth}}$, same for B). The pseudorapidity dependence of the fit parameters is taken into account by dividing the jet sample into several pseudorapidity bins. The minimisation of the difference between the calibrated and true energy is done for each bin separately. This method works well for higher jet energies and lower pseudorapidity range.

The H1-style calibration [1] is a different calibration method which is based on the energy deposition in individual calorimetric cells. The individual cell energy is calibrated using weights dependent on the energy density in the given cell. The response of a cell with a small signal (cell with a high hadronic component) is corrected upwards to make its response equal to that of cells with large (typically electromagnetic) deposited energy. This method gives better resolution than the sampling method. The disadvantage of the method is a higher number of fitting parameters in comparison with the sampling method. The H1-style calibration is used also for tau leptons energy calibration in a tauRec algorithm (see the section 3.5.1).

The resolution of the energy measurement in calorimeters can be written as

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \quad (3.7)$$

where the first term is associated with statistical processes in the showers, the second one is associated with electronic noise of the detecting apparatus and the constant term c is caused by non-uniformities and non-linearities in photomultipliers, proportional counters, etc. Contrary to the momentum measurement the energy resolution improves with increasing energy.

Resolution of the ATLAS calorimeters from the test beams 2004 with single particles (energy in GeV):

- *electromagnetic calorimeter* - electron beam [3]

$$\frac{\sigma(E)}{E} = \frac{0.1139}{\sqrt{E}} \oplus \frac{b}{E} \oplus 0.0045 \quad (3.8)$$

The noise term b is of the order of 0.2 GeV.

- *Tilecal* - pion beam [4]

$$\frac{\sigma(E)}{E} = \frac{0.54}{\sqrt{E}} \oplus 0.059 \quad (3.9)$$

The noise term can be neglected in the Tilecal.

3.4 Missing energy measurement

Missing energy is an important signal for several interesting physical processes at the LHC. High missing energy appears in SUSY particles' decays or in some of the Higgs boson decay modes. Mainly neutrinos contribute to the missing energy. Neutrinos interact only via the weak interaction. They have a very low cross-section for the interaction with matter (10^{-38} cm² for neutrino with energy $E_\nu = 1$ GeV). They do not leave a direct signature in the detector (missing information). Neutrinos are not the only particles which escape the detection. For example, undetectable energy appears also in hadronic interactions. This energy should not contribute to the measured missing energy, but it is included in the hadronic energy calibration.

By measuring all detectable particles in the event and applying the momentum and energy conservation, the x and y component of the missing objects' momentum can be calculated

$$p_x^{\text{miss}} = - \sum_i p_{x_i} \quad (3.10)$$

$$p_y^{\text{miss}} = - \sum_i p_{y_i} \quad (3.11)$$

where $p_{x(y)}^{\text{miss}}$ denotes the appropriate component of the missing momentum and the sum on the right hand side of the equation runs over all detected objects in the event. The z component of the missing momentum cannot be measured at the LHC. The z components of the colliding partons' momenta are not fixed, but they are distributed according to the

parton probability density function. That is why the total z component of the momentum is not known in the hadron collisions.

One of the important condition for a precise missing transverse energy measurement is an accurate calibration of the detecting system (calorimeters and the muon system). The energy in the calorimeter cells can be calibrated in different ways. The standard approach is the H1-style calibration of cells where weights optimised for jets are used. The calibration can be improved by the so-called refined calibration [10]. The basic idea is to decompose reconstructed final state objects back into calorimetric cells and then apply different weights, which take into account the type of the reconstructed objects (electrons/photons, tau leptons, jets, rest), to appropriate cells. The calibration of the calorimeter accounts for energy losses in cryostats and cracks. The contribution of low energy cells which are outside clusters has to be considered. Only cells with energy higher than a given threshold are added to the missing energy calculation. The threshold has to be carefully set up in order to remove cells with only electronic noise and pile-up, but not to reject cells with real deposited energy. Non-linearity of the calorimeter response in the low energy region has to be considered. Another important condition for a good missing energy resolution is a sufficient detector coverage.

The resolution of each component of the missing transverse energy can be parametrised

$$\sigma(E_{x(y)}^{\text{miss}}) = k \times \sqrt{\sum E_{\text{T}}} \quad (3.12)$$

where the term $\sum E_{\text{T}}$ denotes the total transverse energy measured in the calorimeters. The term $k \approx 0.46 \text{ GeV}^{1/2}$ for $A \rightarrow \tau\tau$ events [1] (A is a neutral CP-odd Higgs boson in the Minimal Supersymmetric Standard Model), the resolution and E_{T} are expressed in GeV.

3.5 Identification and reconstruction of hadronic tau leptons

The tau lepton has a very short mean lifetime ($87 \mu\text{m}$). The signal in the detector comes from its decay products. The most important tau decay channels with their branching ratios (BR) are listed in the table 3.1. The tau decays are classified as 1 prong (one charged particle in the final state) and 3 prong (three charged particles in the final state). Final states with more than three charged particles are very rare (0.1% of the tau decays). Other classification is based on the presence of hadrons (hadronic decays) or electrons and muons (leptonic decays) in the final state. Only hadronic decays are considered in this paper. These decays contribute by 65% to the tau decays.

Two basic algorithms for the identification and reconstruction of tau leptons decaying into hadrons are used in the ATLAS experiment - tauRec [5] and tau1P3P [13]. The hadronic tau lepton looks like a narrow jet in the detector. QCD jets can be mistakenly identified as tau leptons. That is why a very good level of tau/jet separation is a basic issue of the tau identification algorithms.

3.5.1 TauRec algorithm

The tauRec algorithm is based by default on clusters. Clusters with $E_{\text{T}} > 15 \text{ GeV}$ are considered to be a possible tau signal, so-called tau candidates. In the next step tracks

decay channel	BR[%]
$\tau^- \rightarrow e^- \bar{\nu}_e \nu_\tau$	17.8
$\tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau$	17.4
$\tau^- \rightarrow \pi^- \nu_\tau$	10.9
$\tau^- \rightarrow \pi^- \pi^0 \nu_\tau$	25.5
$\tau^- \rightarrow \pi^- 2\pi^0 \nu_\tau$	9.5
$\tau^- \rightarrow \pi^- 3\pi^0 \nu_\tau$	1.0
$\tau^- \rightarrow K^- \nu_\tau$	0.7
$\tau^- \rightarrow K^- \pi^0 \nu_\tau$	0.5
$\tau^- \rightarrow \pi^- \pi^- \pi^+ \nu_\tau$	9.3
$\tau^- \rightarrow \pi^- \pi^- \pi^+ \pi^0 \nu_\tau$	4.6

Table 3.1: Tau lepton decay modes according to [12]

with $p_T > 2$ GeV in the inner detector within $\Delta R < 0.3$ from the cluster centre are associated to the tau candidate and cell energies in different cones around the initial seed are collected. The calorimeter clusters are calibrated with the H1-style.

Identification variables based on the tau decays features are constructed:

- *Electromagnetic radius*

$$R_{em} = \frac{\sum_i E_{T_i} \Delta R_i}{\sum_i E_{T_i}} \quad (3.13)$$

where the index i runs over all electromagnetic cells in the cluster within $\Delta R < 0.4$ from the cluster centre. It describes a transversal shower profile of the tau jet. This requirement is based on the fact that tau jets are well collimated in the space. This feature is used also in the following criteria.

- *Isolation fraction*

$$\Delta E_T^{12} = \frac{\sum_j E_{T_j}}{\sum_i E_{T_i}} \quad (3.14)$$

where the sum in the numerator runs over all electromagnetic cells in the cluster with $0.1 < \Delta R < 0.2$ and the sum in the denominator runs over cells with $\Delta R < 0.4$.

- *Transverse energy width in the η strip layer*

$$\Delta\eta = \sqrt{\frac{\sum_i E_{T_i} (\Delta\eta_i)^2}{\sum_i E_{T_i}}} \quad (3.15)$$

where the sum runs over cells in the cluster in the first layer of the EM calorimeter barrel (η strip layer) with $\Delta R < 0.4$.

- *Number of strips*

Cells in the η strip layer of the EM calorimeter within $\Delta R < 0.4$ around the cluster centre with deposited energy above a given threshold are counted. Jets usually give more strips than tau jets.

- *Number of associated tracks*

It stands for the number of reconstructed tracks with $p_T > 2$ GeV which are within

$\Delta R < 0.3$ from the cluster centre. Tracks have to satisfy the track quality criteria. Events with one charged track represent 76% of hadronic taus' decays and events with three tracks 24%. Events with more than three charged particles contribute only by 0.2%.

- *Tau lepton charge*
The charge is reconstructed from the associated tracks.
- *Lifetime signed pseudo impact parameter significance*

$$\sigma_{\text{IP}} = \frac{d_0}{\sigma_{d_0}} \times \text{sign}(\sin \Delta\phi) \quad (3.16)$$

The symbol d_0 denotes a pseudo impact parameter, σ_{d_0} its error and $\Delta\phi$ stands for the azimuthal angle between the calorimeter cluster and the leading track at the point of the closest approach. The parameter d_0 is defined in the transverse plane as the smallest distance of a track to the beam axis. A sign is assigned to this parameter. Due to the term $\text{sign}(\sin \Delta\phi)$ in the definition (3.16) the sign of d_0 is not necessary the same as the sign of σ_{IP} . The identification variable σ_{IP} is defined to be positive if the decay happened in the flight direction. See [5] for further details.

- *Reconstructed E_{T} over p_{T} of the first track*
Transverse energy is collected from clusters with $\Delta R < 0.4$. Tau jets have in general a high fraction of their energy in the leading track (first track). On the contrary, QCD jets are expected to have more uniform distribution of the p_{T} .

The likelihood for each tau candidate is calculated. The calculation is based on probability distribution functions of the identification variables.

The tau energy is calibrated with the H1-style method which is developed primarily for the jet calibration. Hadronic products of the tau leptons have similar features as jets and therefore the similar way of the calibration can be used. Weights used for the H1-style calibration of the tau leptons are fitted to jets. These weights are not optimal for taus, because the hadronic and electromagnetic composition and depth of the energy deposition are not the same as for jets. Corrections for the special features of the tau leptons' energy distribution in the detector are included in the additional factor (dependent on E_{T} and η) which the H1-calibrated energy in the given cell is multiplied with.

A very good level of tau/jet separation can be achieved over a broad transverse energy range (tenths to hundreds GeV). TauRec algorithm gives good results in many physical channels in the region from light Higgs boson to heavy SUSY particles.

3.5.2 Tau1P3P algorithm

The tau1P3P algorithm is optimised for the identification of hadronic tau leptons with transverse visible energy in the range 20 – 70 GeV which is an interesting region for light Higgs boson (see the section 5.2) or soft SUSY particles.

The tau lepton is not considered as an object similar to a jet, but it is formed by one or three charged tracks (π^{\pm}) in the inner detector and a signal in the electromagnetic calorimeter (π^0), all objects are well collimated in the space. The algorithm is based on finding a good quality leading hadronic track in the inner detector with $p_{\text{T}} > 9$ GeV. The track has to pass several quality cuts ($d_0 < 1\text{mm}$, number of precision hits > 8 , number of

TRT hits > 9 and $\chi^2 < 1.7$). Tau decay modes with kaons are excluded by the tau1P3P. Modes with more than three charged particles are neglected due to low probability (0.2% of hadronic decays). No nearby tracks with $p_T > 1$ GeV within $\Delta R < 0.2$ are allowed for 1 prong events and one or two nearby tracks satisfying $p_T > 1$ GeV are required for 3 prong events. The charge consistency with a total sum of ± 1 is required for 3 prong events with three reconstructed tracks.

Multivariable analysis is used for the acceptance or rejection of the tau candidate. Tau identification observables are built and combined into one discriminant value. Identification quantities are calculated in a narrow cone $\Delta R < 0.2$. A cone $\Delta R < 0.4$ is used only for the isolation criterion. The pseudorapidity-azimuthal distance is calculated from a seed which is defined by the leading track at vertex for 1 prong events and by the bary-centre of nearby hadronic tracks at vertex weighted by the tracks' momenta for 3 prong events. The identification variables are (if not stated otherwise, the meaning is the same as for the tauRec and the cone $\Delta R < 0.2$ is used):

- *Electromagnetic radius*
- *Isolation fraction*
- *Number of strips*
- *Width of the energy deposition in strips*

$$W_{\text{strip}} = \frac{\sum_i (\Delta\eta_i)^2 \cdot E_{T_i}}{\sum_i E_{T_i}} - \frac{\sum_i (\Delta\eta_i \cdot E_{T_i})^2}{(\sum_i E_{T_i})^2} \quad (3.17)$$

The width denotes the variance of η in the strip layer in the EM calorimeter weighted by the deposited transverse energy in a given strip. The sum runs over all cells in the given layer with $\Delta R < 0.2$.

- *Hadronic transverse energy over p_T of track(s)*
Ratio of the energy deposited in the hadronic calorimeter and the track transverse momentum. The sum of p_T of all charged tracks is used for 3 prong events.

- *Isolation criterion*

$$\frac{\sum_j E_{T_j}^{\text{EM}} + \sum_k E_{T_k}^{\text{Had}}}{\sum_i E_{T_i}} \quad (3.18)$$

where the sums in the numerator run over electromagnetic and hadronic cells with $0.2 < \Delta R < 0.4$ and the sum in the denominator involves all calorimeter cells within $\Delta R < 0.4$. The hadronic energy is calibrated to the EM scale.

The energy scale is defined using the energy-flow technique [13] based on information from the inner detector and the electromagnetic calorimeter. The energy is calculated using different components

$$E_T^{\text{flow}} = E_T^{\text{emcl}} + E_T^{\text{neuEM}} + \sum p_T^{\text{track}} + \sum res E_T^{\text{chrgEMtrk}} + res E_T^{\text{neuEM}} \quad (3.19)$$

The momentum resolution of tracks is better than the energy resolution for low energies. That is why the momenta of charged tracks p_T^{track} in the distance $\Delta R < 0.2$ from the bary-centre are used for the energy calculation. Neutral pions decay into a pair of photons before reaching the EM calorimeter. The photons from the pions decays contribute to

two terms - E_T^{emcl} and E_T^{neuEM} . The symbol E_T^{emcl} denotes pure electromagnetic energy with no significant hadronic leakage. The seeded clusters are required to be isolated from the qualified tracks. Energy is collected from a very narrow window. The neutral electromagnetic energy E_T^{neuEM} is seeded by the track at vertex. Energy is collected from the remaining cells in a cone $\Delta R < 0.2$. Two remaining terms, $\sum res E_T^{\text{chrgEMtrk}}$ and $res E_T^{\text{neuEM}}$, stand for a parametrisation of corrections for residual effects (e.g. energy leakage outside the narrow cone).

The package tau1P3P is still under development. The identification criteria will be further optimised.

Chapter 4

Measurement of the Higgs boson mass with the collinear approximation

The Higgs boson mass can be reconstructed as an invariant mass of the tau leptons pair in the $H \rightarrow \tau\tau$ channel. Four-momenta of both tau leptons are needed for the invariant mass calculation. Both tau leptons decay very quickly and the signal in the detector comes from their decay products. The products contain at least two neutrinos (one from each tau lepton) which escape the detection. Energy and directions of the visible parts of the tau leptons (neutrinos excluded) and the missing transverse energy in the event can be measured in the detector.

Only the decay mode with both taus decaying hadronically $H \rightarrow \tau\tau \rightarrow h\nu_\tau h\nu_\tau$ is considered in this paper. It is the channel with the lowest possible number of neutrinos. The momenta of both neutrinos represent six unknown variables which have to be estimated in order to reconstruct the Higgs boson mass.

4.1 Kinematics of the $H \rightarrow \tau\tau \rightarrow h\nu_\tau h\nu_\tau$ channel

Our problem is described with a set of five kinematic equations:

$$E_x^{\text{miss}} = p_x^{\nu_1} + p_x^{\nu_2} \quad (4.1)$$

$$E_y^{\text{miss}} = p_y^{\nu_1} + p_y^{\nu_2} \quad (4.2)$$

$$m_\tau^2 = (P_{h_1} + P_{\nu_1})_\mu (P_{h_1} + P_{\nu_1})^\mu \quad (4.3)$$

$$m_\tau^2 = (P_{h_2} + P_{\nu_2})_\mu (P_{h_2} + P_{\nu_2})^\mu \quad (4.4)$$

$$m_H^2 = (P_{h_1} + P_{\nu_1} + P_{h_2} + P_{\nu_2})_\mu (P_{h_1} + P_{\nu_1} + P_{h_2} + P_{\nu_2})^\mu \quad (4.5)$$

The first two equations shows that the missing energy is represented by two neutrinos. The equations stand for the components of the missing energy in the transverse plane. The other equations represent the conservation of the four-momenta. The symbol $P_\mu P^\mu$ denotes the scalar product of appropriate four-momenta. Subscripts h_1, h_2 stand for the visible parts of the tau leptons and ν_1, ν_2 for the neutrinos. The given set of the equations can be easily simplified into a set of three non-linear equations. If the Higgs boson mass is fixed (e.g. known from other channel), we have three equations with four unknown components of neutrinos momenta. The set of the equations can be further analytically simplified which leads to four branches of one non-linear equation with two unknown

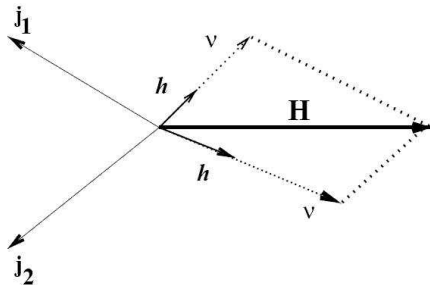


Figure 4.1: Geometry of the collinear approximation for the Higgs boson produced through the vector boson fusion. The symbols j_1, j_2 stand for trigger jets.

quantities. The problem can be considered as an equation with one unknown and one arbitrary parameter. The arbitrary parameter has to be fixed. The whole possible range of the appropriate component of the neutrino momentum can be divided into several bins and the parameter is selected as a random number in each bin. The remaining unknown can be obtained as a numerical solution of our equation. No solution is found for most of the bins, but still several solutions can be found over the whole momenta range. The problem is how to select the correct solution.

The collinear approximation is a simple way how to estimate the neutrinos momenta without solving the given set of the non-linear equations. It is the basic method for the Higgs boson mass reconstruction in the double tau channel. It is used as the reference method to the alternative decay length method in this paper.

4.2 Principles of the collinear approximation

Tau leptons produced in the Higgs boson decay have high momenta due to the high difference between the Higgs and tau lepton masses. The Higgs boson momentum also contributes to the tau leptons momenta. Decay products of taus are therefore well collimated in space in the most cases. The collinear approximation is based on the assumptions that the tau leptons are massless and neutrinos from the tau decays have the same directions as the visible parts of the tau leptons, see the figure 4.1.

The unknown neutrinos energies $E_{\nu_{1(2)}}$ can be expressed as [14]

$$E_{x(y)}^{\text{miss}} = E_{\nu_1} u_{1x(y)} + E_{\nu_2} u_{2x(y)} \quad (4.6)$$

where $u_{1x(y)}$ and $u_{2x(y)}$ are components of unitary vectors in the directions of the detected parts of the tau leptons and $E_{x(y)}^{\text{miss}}$ denotes the appropriate component of the transverse missing energy. The system of the equations (4.6) can be solved if the determinant is non-zero. The determinant equals $\sin \Delta\phi^{\text{vis}}$, where $\Delta\phi^{\text{vis}}$ stands for the azimuthal angle between the detected parts of the tau leptons. The visible parts of the taus cannot be emitted back to back in the transverse plane in order to find a solution for the given set of the equations.

The Higgs boson mass is evaluated as

$$m_H = \sqrt{2(E_{h_1} + E_{\nu_1})(E_{h_2} + E_{\nu_2})(1 - \cos\theta)} \quad (4.7)$$

where $E_{h_{1(2)}}$ denotes energy of the visible tau lepton products and θ stands for the angle between the directions of the detected parts of the tau leptons.

4.3 Results for the Higgs boson mass 120 GeV

The collinear approximation was applied to the full simulated data $H \rightarrow \tau\tau \rightarrow h\nu_\tau h\nu_\tau$ for $m_H = 120$ GeV. The sample was simulated in Athena 11.0.5 and reconstructed in Athena nightly version 12.0.x. The Higgs boson was produced only through the vector boson fusion (VBF). The Pythia version 6.323 generator with the CTEQ6L parton distribution function (PDF) was used for the Higgs boson production and decays. The tau leptons decays were simulated in the TAUOLA library where spin correlations between the taus from the Higgs boson decays were correctly considered.

Basic selection criteria were applied on the hadronic parts of the tau leptons:

$$p_T^{\text{vis}} > 12 \text{ GeV} \quad (4.8)$$

$$|\eta^{\text{vis}}| < 2.7 \quad (4.9)$$

$$|\Delta\phi^{\text{vis}}| < 2.9 \quad (4.10)$$

where the superscript “vis” stands for the visible part of the tau leptons. The last selection criterion does not allow the detected parts of the tau leptons to be emitted back to back in the transverse plane which coincides with the condition for finding a solution for the collinear set of the equations (4.6).

4.3.1 Collinear approximation for the Monte Carlo truth

The basic assumption of the collinear approximation [14] is that the directions of the neutrino and the rest of the tau decay products are very close. This feature was checked with Monte Carlo (MC) true information. The distance in the pseudorapidity-azimuthal plane $\Delta R_{\nu\tau}$ between the neutrino and the visible part of the tau lepton is shown in the figure 4.2. The distance $\Delta R_{\nu\tau}$ is smaller than 0.1 for both taus from the Higgs decay in 68% of events. The directions of the tau products are really very close in the space as expected.

First, the collinear approximation was done for Monte Carlo true information to check the sensitivity of the method. Solving the set of equations (4.6) leads to an unphysical result with negative neutrino energy in 4% of events. The events with the bad solution have to be excluded from the analysis. The unphysical solutions can be caused only by the assumption on the neutrinos directions for the MC truth. The number of events without any physical solution and the RMS of the reconstructed mass distribution was studied in different $\Delta R_{\nu\tau}$ bins. See the table 4.1 where results for events with one tau lepton satisfying $\Delta R_{\nu\tau} < 0.1$ and the second one in various bins are displayed. The number of events without any solution increases significantly for taus with $\Delta R_{\nu\tau} > 0.1$. Similarly, the RMS of the mass distribution increases with the $\Delta R_{\nu\tau}$ distance as expected. The assumption on the neutrinos directions broaden significantly the natural width of the Higgs boson mass peak, the RMS of the generated Higgs mass distribution is only 0.01 GeV.

The reconstructed mass distribution was fitted with a Gaussian curve to estimate the Higgs mass. The result of the collinear approximation for the MC true data is $m_H = (119.8 \pm 0.1)$ GeV for the number of events corresponding to the integrated luminosity 30 fb^{-1} . The noted deviation is the error of the Gaussian fit. The collinear approximation works with a very good accuracy on the MC truth.

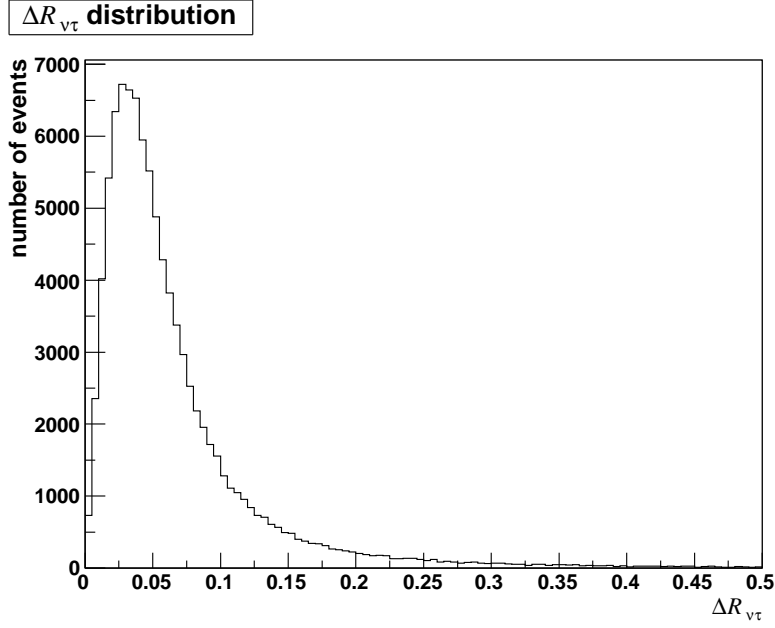


Figure 4.2: Pseudorapidity-azimuthal distance between the neutrino and the detected part of the tau lepton in the Higgs boson decays

$\Delta R_{\nu\tau}$	fraction of events [%]	RMS [GeV]	bad solution [%]
0 – 0.025	15.0	5.0	0.9
0.025 – 0.05	26.5	5.4	0.6
0.05 – 0.1	26.7	6.4	1.1
0.1 – 0.3	13.1	7.9	6.5
> 0.3	1.8	11.0	34.5

Table 4.1: Collinear approximation for events with one tau lepton with $\Delta R_{\nu\tau} < 0.1$ and the second one in various $\Delta R_{\nu\tau}$ bins. The fraction of events means the fraction of events in different bins and the total number of events (no restrictions on $\Delta R_{\nu\tau}$).

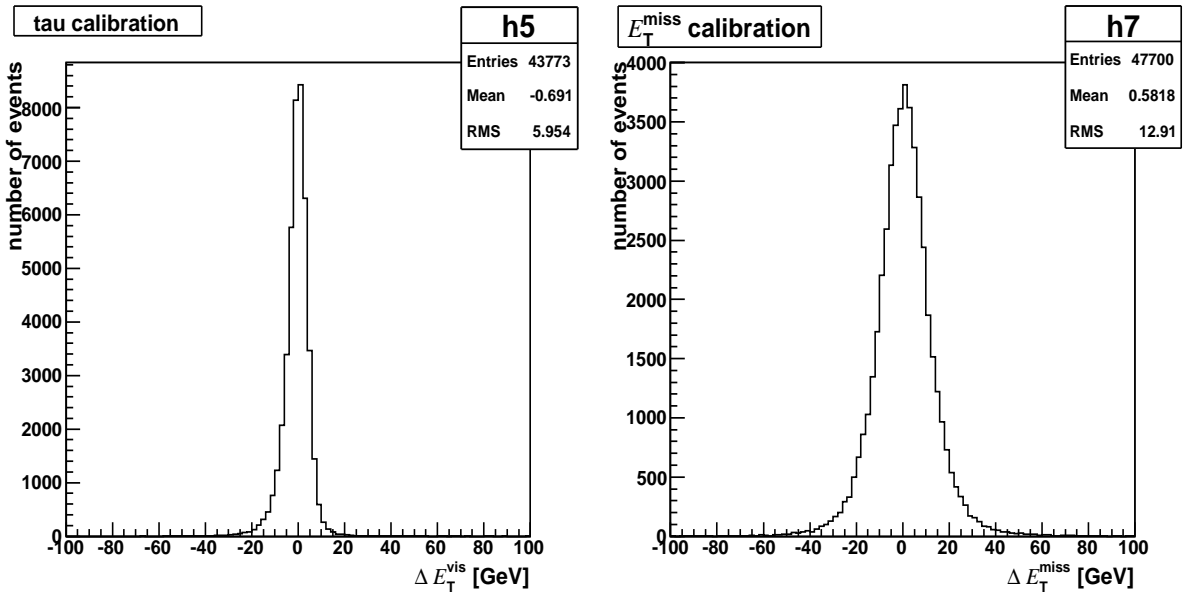


Figure 4.3: The accuracy of the calibration of the tau visible transverse energy and missing transverse energy

4.3.2 Collinear approximation for reconstructed events

The full reconstructed events were investigated in the next step. The energy measurement and calibration are of crucial importance for the collinear approximation. That is why the transverse energy resolution, the difference between the reconstructed and true value $E_T^{\text{reco}} - E_T^{\text{truth}}$, is studied. The resolution for the visible parts of the tau leptons and for the missing energy can be seen in the figure 4.3. The missing energy was reconstructed in every event. Only neutrinos from tau decays contribute to the missing energy in our process. The hadronic undetectable energy, which comes from the trigger jets in the VBF process, is accounted for in the jet calibration. Contrary to the missing energy, tau leptons were not reconstructed in every event. Only tau leptons reconstructed with the tauRec algorithm with likelihood higher than 4.0 (a default value) were considered. The reconstructed tau leptons were accepted if their directions were close to the true taus ($\Delta R < 0.3$) and they had the same charge. The resolution for the tau leptons is narrower than for the missing energy as expected. The means of the resolution plots are -0.7 GeV for the visible part of the taus and 0.6 GeV for the missing energy. The shifts in the calibration of the taus and the missing energy are not significant and they balance.

The influence of the energy resolution on the precision of the collinear approximation was studied. The collinear approximation was done for several cases - MC true information, reconstructed information about tau leptons + true E_T^{miss} , true taus + reconstructed E_T^{miss} , reconstructed information for both (see the figure 4.4). Only events with two reconstructed and accepted tau leptons were considered (21% of all events). The mass peak for the full reconstructed case (reconstructed tau leptons + reconstructed E_T^{miss}) has a large tail (right bottom plot in the figure 4.4) which was not considered. That was why the Gaussian fit was done only in the central part of the mass peak. The means and sigmas of the fit and numbers of events with the unphysical solution can be seen in the table 4.2. If the reconstructed taus and true missing energy are considered, the mass peak is wider than for the MC truth and there is a shift in the mean value from the true Higgs boson mass. The number of events with a bad solution is almost the same as for the

information	mean [GeV]	sigma [GeV]	bad solution [%]
tau truth + E_T^{miss} truth	119.9	2.1	4
tau reco + E_T^{miss} truth	119.3	6.5	5
tau truth + E_T^{miss} reco	118.8	12.5	34
tau reco + E_T^{miss} reco	117.6	12.0	34

Table 4.2: Results of the fit of the reconstructed mass spectra and the number of events with the unphysical solution for several combinations of true and reconstructed information. The number of events corresponds to $\sim 590 \text{ fb}^{-1}$.

$L [\text{fb}^{-1}]$	$m_H [\text{GeV}]$	$\sigma_{m_H} [\text{GeV}]$
30	118.7	0.9
100	118.5	0.5
300	118.1	0.3

Table 4.3: Results of the collinear approximation for various integrated luminosities L . The symbol σ_{m_H} denotes the statistical deviation.

MC truth. On the contrary, if the reconstructed missing energy and true taus are taken into account, the number of events with no physical solution increases rapidly (34% of the events). The missing energy measurement uncertainties contribute also to the higher width (two times more than in the previous case) and tails of the reconstructed mass distribution (left bottom plot in the figure 4.4). The shift in the mean value is higher than in the previous case. If reconstructed information is taken for both the taus and the missing energy, the results are nearly the same as in the case with reconstructed missing energy and true taus. The sigma of the Gaussian fit is slightly better than in the previous case. This is caused by the balance of the mean values of the resolution plots (figure 4.3). The high number of events without any physical solution, the shift in the mean value of m_H and the broad mass spectrum is caused mainly by the missing transverse energy measurement. The accurate missing energy measurement is of crucial importance for the collinear approximation.

The accuracy of the collinear approximation was tested with the full simulated sample for a number of events corresponding to various integrated luminosities. The cross-section for the signal process (Higgs boson produced through the vector boson fusion decays into a tau lepton pair, both tau leptons decay hadronically and pass the selection criteria) is 81 fb. Reconstructed tau leptons and reconstructed missing energy were considered which should correspond to the real detector data. Results are shown in the table 4.3. The shift between the reconstructed and generated mass represents 1.1-1.6% of the generated mass. The shift can be smaller if especially the missing energy measurement improves. The statistical errors, which are displayed in the table, scales as one over a square root of the number of events as expected. The collinear approximation predicts the Higgs boson mass with sufficient accuracy even for the integrated luminosity 30 fb^{-1} where the statistical uncertainty is only 0.8%. If the integrated luminosity 300 fb^{-1} is considered, the reconstructed Higgs boson mass is $m_H = (118.1 \pm 0.3) \text{ GeV}$. This luminosity is expected to be achieved in three years of running of the LHC accelerator at the design luminosity. The same luminosity is used for the following reconstruction method.

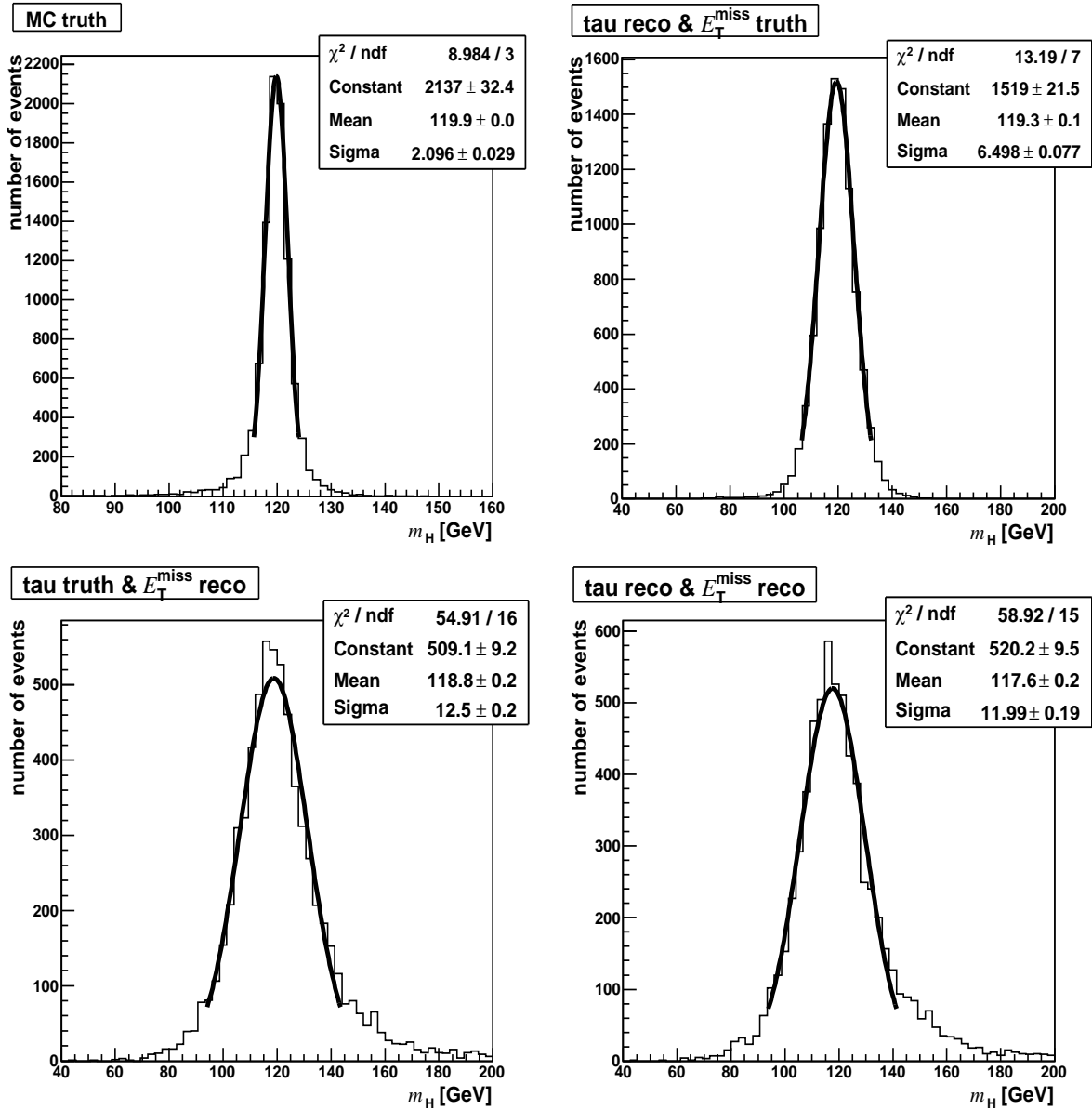


Figure 4.4: The Higgs boson mass reconstruction for the MC truth and for reconstructed events. The number of events corresponds to $\sim 590 \text{ fb}^{-1}$. For further details see the text, section 4.3.2.

Chapter 5

Measurement of the Higgs boson mass from the tau lepton decay length

The energy measurement of the decay products is a direct way how to calculate the parent's particle mass. However, there are also other quantities which are correlated with the mass of the primary particle. For example, the decay lengths distribution of the secondary particles. The secondary vertices measurement of the B mesons from the $t \rightarrow bW$ decay is expected to estimate the top quark mass with accuracy comparable to methods based on the energy measurement at the LHC, see [6] for further details. In principle, it is possible to use the so-called decay length method for the Higgs boson mass reconstruction in the $H \rightarrow \tau\tau$ channel. The question is whether the secondary vertices of tau leptons can be measured with adequate accuracy even for taus with a very short range in order to estimate the Higgs mass with a reasonable precision.

5.1 Description of the method

In two body decays the products have opposite momenta with equal magnitude in the centre of mass system (CMS). In the case of $H \rightarrow \tau\tau$ decay the tau leptons momenta equal

$$p_{\text{CMS}} = \frac{m_H}{2} \sqrt{1 - 4 \left(\frac{m_\tau}{m_H} \right)^2} \quad (5.1)$$

Due to the fact that the Higgs boson mass is much higher than the tau lepton mass, the momentum in the CMS frame can be considered as $p_{\text{CMS}} \approx m_H/2$.

The Higgs boson produced at the LHC will have a non-zero relativistic boost in the most cases, the average value of the gamma factor is about 2.5 (depending on the Higgs boson mass). The Lorentz transformation has to be considered to estimate the tau leptons momenta in the laboratory frame.

The decay length of the each tau lepton is proportional to its momentum and therefore correlated with the Higgs boson mass. To avoid the dependence on the tau lifetime let us consider the mean decay length:

$$\langle L \rangle = t_0 \langle \beta\gamma \rangle = t_0 \frac{\langle p \rangle}{m_\tau} \quad (5.2)$$

where t_0 denotes the mean lifetime of the tau lepton in its CMS system.

The transverse decay length L_{xy} can be measured with better accuracy than the whole decay length because of higher uncertainty in the z axis direction [6]. This is caused by the fact that the longitudinal momenta of colliding partons are not known. However, the sum over components of momentum in the transverse plane of all objects in the event must be zero.

It is possible to estimate the Higgs boson mass by measuring the mean transverse decay length of tau leptons if the shape of the $\langle L_{xy} \rangle(m_H)$ dependence is known. The advantage of such approach is that the measurement of the decay length is independent on the energy calibration. The basic issue is the secondary vertex reconstruction as opposed to the energy measurement and calibration. The decay length method is highly uncorrelated with the collinear approximation.

Samples with different Higgs masses were generated in order to estimate the correlation between the Higgs boson mass and the transverse decay length. The only full simulated sample was the $H \rightarrow \tau\tau \rightarrow h\nu_\tau h\nu_\tau$ for the Higgs mass $m_H = 120$ GeV, so-called H(120) sample, which was used for the collinear approximation method. No other full simulated samples were available. Simulations for various Higgs boson masses, which were used for the construction of the $\langle L_{xy} \rangle(m_H)$ dependence, were only Monte Carlo truth with the same Pythia setup as in the full simulated sample. The H(120) full simulated sample was studied to estimate the detector and reconstruction effects to emulate the detector output on true information. Namely, the resolution of the tau transverse decay length measurement and efficiency of the tau reconstruction algorithm were studied (section 5.2) and afterwards applied on the MC true samples (section 5.3).

Independence of decays of the first and the second tau in each event was assumed. Only events with both taus decaying hadronically were considered. Both tau leptons had to satisfy the same basic selection criteria as for the collinear approximation, equations (4.8) - (4.10). The tau secondary vertices were reconstructed only for taus with more than one associated track. This means that the decay length cannot be measured for tau leptons decaying to one charged particle. Only 3 prong taus (tau leptons decaying into three charged particles), which represent 24% of the hadronic taus, were studied.

5.2 Full simulated sample for the Higgs boson mass 120 GeV

Two tau identification algorithms were used in the H(120) sample - tauRec and tau1P3P, see the section 3.5 for description of both methods. The tauRec performs almost constant efficiency of the tau lepton reconstruction in the transverse energy range $E_T^{\text{vis}} > 20$ GeV [9]. The tau1P3P algorithm is optimised for low tau leptons energy $20 < E_T^{\text{vis}} < 70$ GeV [13]. If the mass of the hadronic part of the tau is neglected (this is a good approximation in this energy range), the transverse momentum equals the transverse energy. The spectrum of the transverse momenta of the visible part of tau leptons p_T^{vis} is shown in the figure 5.1. Tau leptons with $p_T^{\text{vis}} > 70$ GeV represent 20% of the taus satisfying the basic selection criteria (4.8) - (4.10). The tau1P3P algorithm can be used for the identification of tau leptons from the Higgs boson decays in the low Higgs boson mass region.

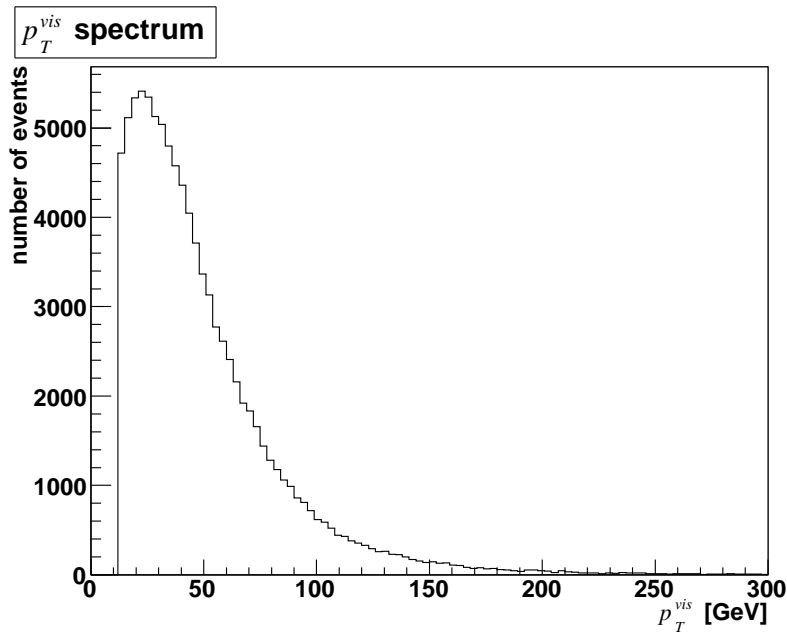


Figure 5.1: Transverse momentum of the visible part of the tau leptons

5.2.1 Resolution of the transverse decay length

The basic criterion for deciding which tau reconstruction algorithm (tauRec or tau1P3P) was better for the decay length measurement was the resolution of this quantity. The resolution, which was defined as the difference between the reconstructed and true transverse decay length $L_{xy}^{\text{reco}} - L_{xy}^{\text{truth}}$, was considered to be dependent on the pseudorapidity and on the true transverse decay length. The ATLAS detector was fully symmetric in the azimuthal angle and therefore no ϕ dependence was assumed. The sample was divided into five decay length bins for three different pseudorapidity regions. The resolution for both tau identification algorithms in different pseudorapidity and decay length bins is shown in the figure 5.2. The resolution for the tau1P3P algorithm (bottom plots) has a mean value closer to zero, lower RMS and almost no overflow. It is obvious that the tau1P3P algorithm provides better resolution. This is not surprising because this algorithm is based on tracks as opposed to the tauRec, which is based on calorimeter clusters. From this point on only the tau1P3P algorithm was used.

The next step was to look at the correlation between reconstructed and true decay lengths. Profile plots in different L_{xy}^{truth} bins over all allowed pseudorapidity region were drawn, see the figure 5.3. The straight line drawn in each picture corresponds to $L_{xy}^{\text{truth}} = L_{xy}^{\text{reco}}$ which is expected in the ideal case. As can be seen, the reconstructed lengths in the first bin $L_{xy}^{\text{truth}} < 0.5$ mm are almost independent on the truth. However, with higher lengths we get much better agreement. The transverse decay length is counted as the transverse distance between the secondary vertex coordinates and the origin of the coordinate system. The primary vertex measurement is in a 0.01 mm scale in the transverse plane which cannot influence the value of the decay lengths even in the first bin where the average shift from the truth is nearly one millimetre. The difference between the primary vertex coordinates and the origin could be therefore neglected. With respect to the results of the tau decay lengths reconstruction in the first bin, only tau leptons which satisfy $L_{xy}^{\text{truth}} \geq 0.5$ mm were used in the analysis.

In order to estimate the reconstruction of the tau transverse decay lengths, a resolution

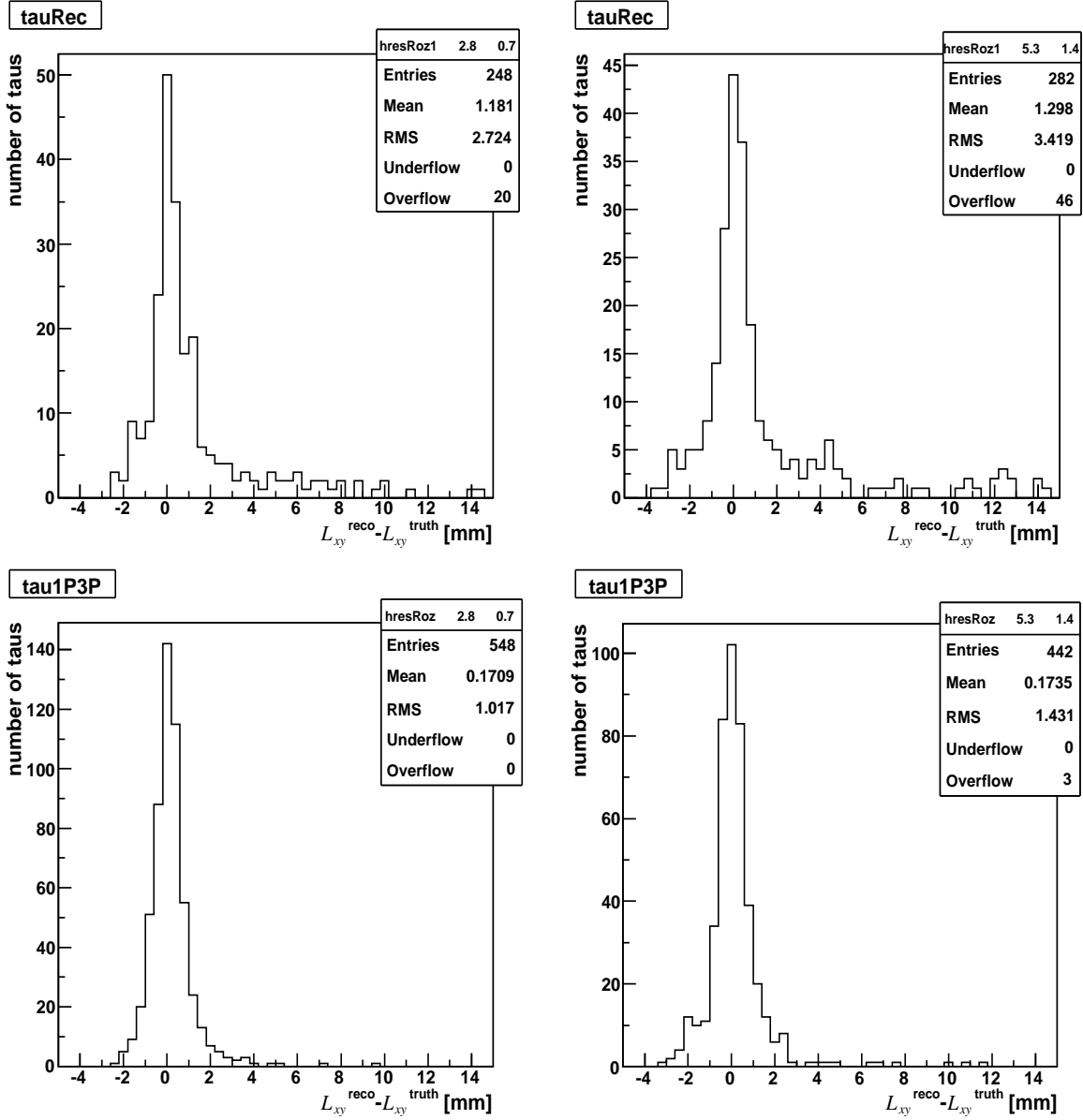


Figure 5.2: The transverse decay length resolution for the tauRec (top) and tau1P3P (bottom) algorithms. The plots on the left show events with $L_{xy} \in (1.4, 2.8) \text{ mm} \wedge |\eta| \in (0, 0.7)$, the plots on the right $L_{xy} \in (2.8, 5.3) \text{ mm} \wedge |\eta| \in (0.7, 1.4)$.

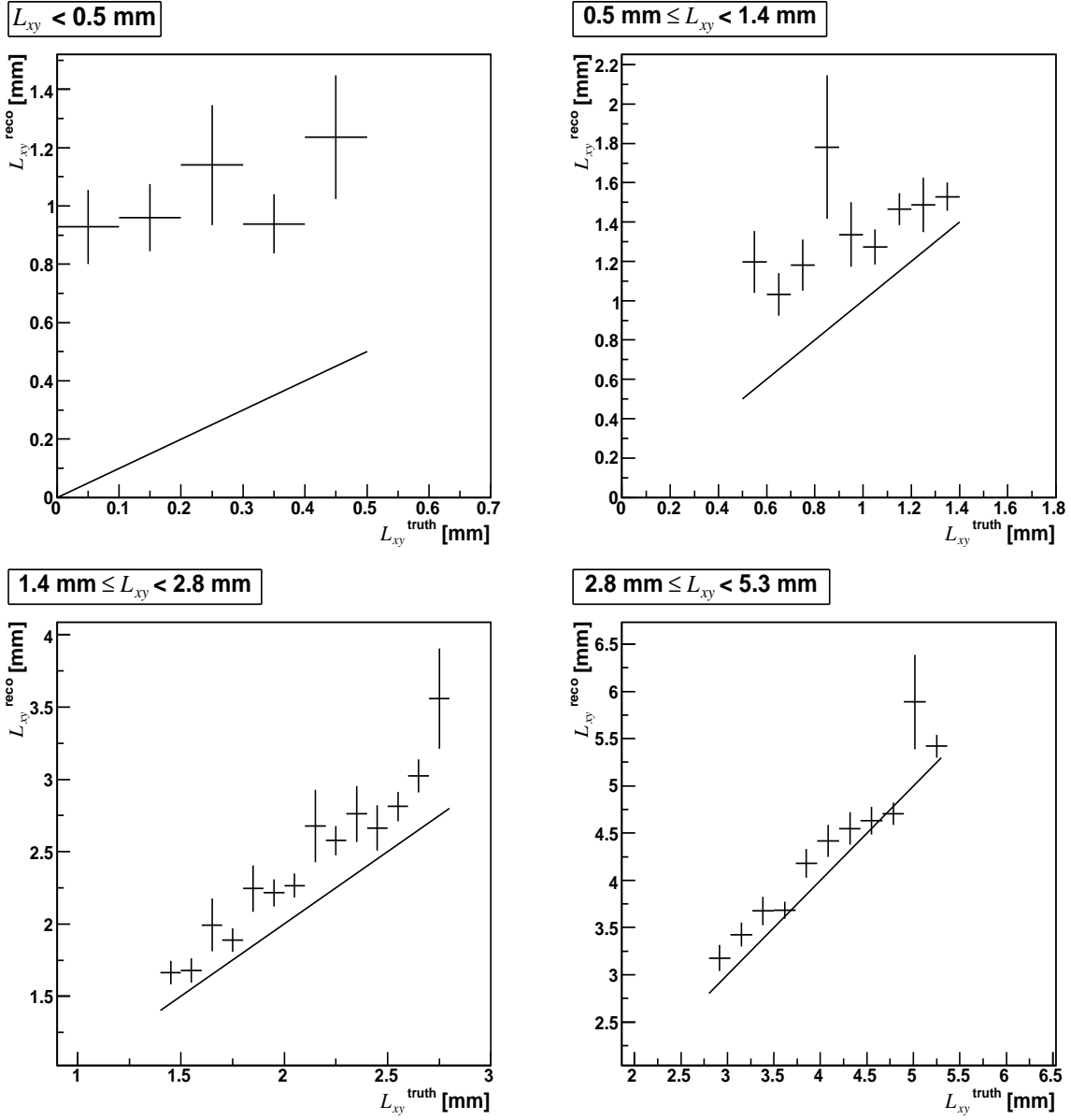


Figure 5.3: Profile plots for the dependence of reconstructed tau transverse decay lengths on the true values in different L_{xy}^{truth} bins

of this variable was studied in the full simulated events. Besides the $H \rightarrow \tau\tau \rightarrow h\nu_\tau h\nu_\tau$ sample also $H \rightarrow \tau\tau \rightarrow \ell\nu_\tau\nu_\ell h\nu_\tau$ and $Z \rightarrow \tau\tau$ full simulated samples were used in the resolution plots to get higher statistics. The resolution of the secondary vertex was supposed to be independent on the process which the tau lepton was produced in. Since only tau leptons with $L_{xy}^{\text{truth}} \geq 0.5$ mm were considered, the resolution was plotted in 12 separate bins (4 in $L_{xy}^{\text{truth}} \times 3$ in η).

Initially, only tau leptons with the secondary vertex reconstructed with the tau1P3P algorithm were considered. The resolution plots of the transverse decay lengths were not symmetrically distributed around zero, events with $L_{xy}^{\text{reco}} > L_{xy}^{\text{truth}}$ slightly dominated (see the figure 5.2 bottom). That was the reason why the resolution plots were not fitted to a Gaussian to get parameters for Gaussian smearing of L_{xy}^{truth} . Random numbers distributed according to the resolution in different bins from appropriate histograms (figure 5.2 bottom) were used to smear the true values of the transverse decay length L_{xy}^{truth} for 3 prong tau leptons reconstructed with the tau1P3P algorithm. The reconstructed and the emulated transverse lengths distributions were compared in separate L_{xy}^{truth} bins over all pseudorapidity region in order to check the precision of the emulation method. We got quite a good agreement except for the events with reconstructed lengths close to zero in the first bin where $0.5 \text{ mm} \leq L_{xy}^{\text{truth}} < 1.4 \text{ mm}$. When a relative resolution $(L_{xy}^{\text{reco}} - L_{xy}^{\text{truth}})/L_{xy}^{\text{truth}}$ was considered instead of the absolute resolution, the emulated spectrum corresponded to the reconstructed one even in this L_{xy}^{truth} bin. Kolmogorov test was used to decide which decay length smearing (according to histograms with relative or absolute resolution) corresponded better to reality. According to results of the test, the relative resolution was used for the emulation in the first L_{xy}^{truth} bin and the absolute resolution in the rest of the bins. The reconstructed and emulated distributions in different L_{xy}^{truth} bins are shown in the figure 5.4 and the whole spectra are plotted in the figure 5.5.

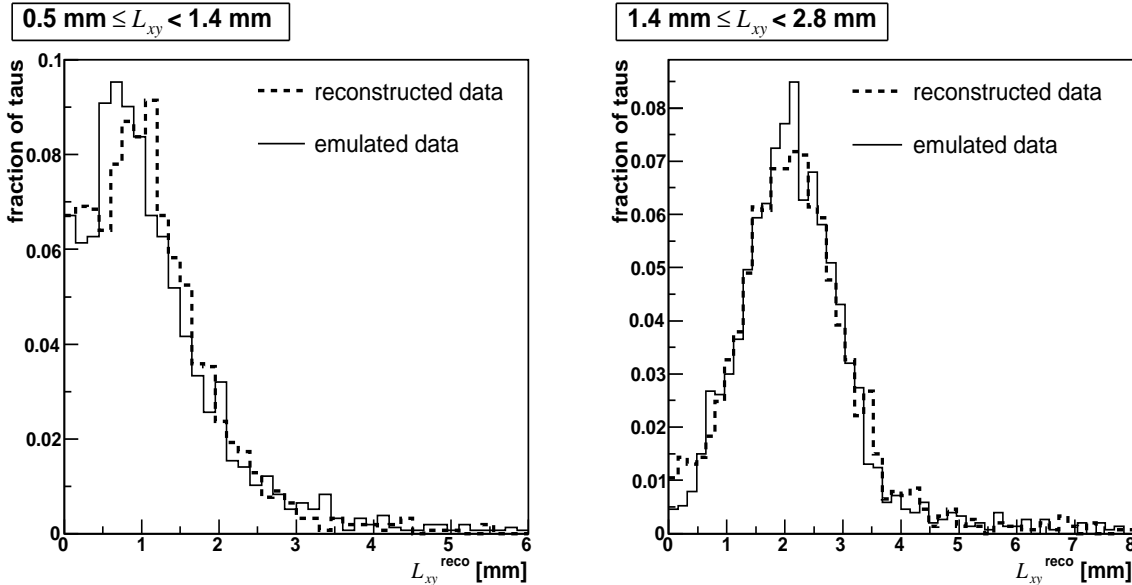


Figure 5.4: Reconstructed and emulated transverse decay lengths in the first two L_{xy}^{truth} bins. For details on the emulation method see the text, section 5.2.1.

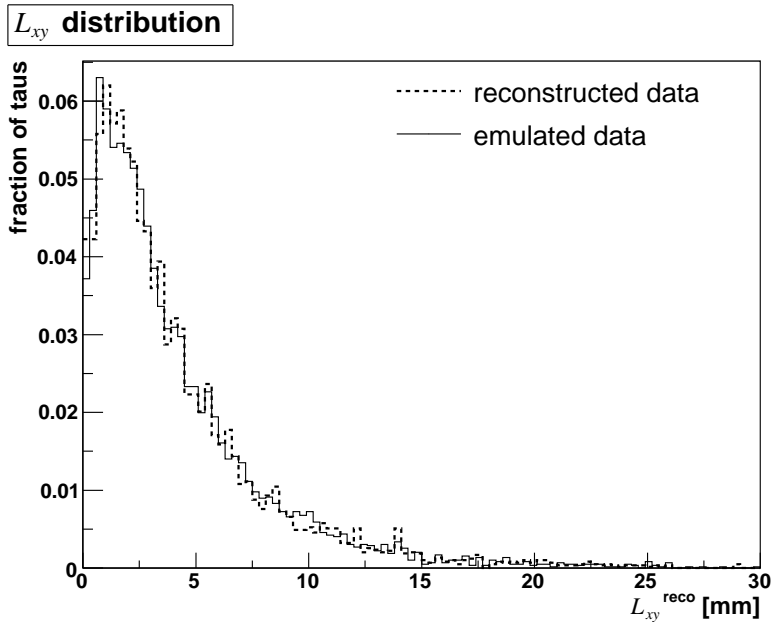


Figure 5.5: Reconstructed and emulated transverse decay length spectra. For details on the emulation method see the text, section 5.2.1.

5.2.2 Efficiency of the tau1P3P algorithm

Only events with 3 prong tau leptons reconstructed with the tau1P3P algorithm were investigated in the previous section. When working with a Monte Carlo true sample we are supposed to create a subset of the sample with similar attributes as tau1P3P reconstructed events have. To be able to select such a subset efficiency of the tau1P3P algorithm has to be studied.

Only true tau leptons decaying into three charged pions were selected. Other 3 prong decays, which contributed only by 4% to all 3 prong tau decays and were not considered by the tau1P3P algorithm, were not taken into account. The transverse momenta of the charged pions had to satisfy $p_T > 9$ GeV for the pion with the highest transverse momentum (corresponding to the leading track) and $p_T > 1$ GeV for the other pions as required by the tau1P3P reconstruction method. The visible part of the tau was required to be in the $|\eta^{\text{vis}}| < 2.5$ pseudorapidity range which represented the coverage of the inner detector.

Efficiency was considered to be dependent on the true transverse decay length and pseudorapidity. First, efficiency as a function of pseudorapidity was studied. Efficiency was evaluated in three η^{vis} bins ($0 \leq |\eta^{\text{vis}}| < 0.7$, $0.7 \leq |\eta^{\text{vis}}| < 1.4$, $1.4 \leq |\eta^{\text{vis}}| < 2.5$). The division in pseudorapidity corresponds to geometric attributes of the inner detector. Efficiency in the first two bins has the same value (38%). Efficiency is slightly lower in the last bin (35%) which involves end-caps. However, the difference is not significant. On the other hand, the dependence of efficiency on the true transverse decay length is more remarkable, see the figure 5.6. Surprisingly, efficiency is falling with increasing decay lengths for $L_{xy}^{\text{truth}} > 5$ mm. The possible explanation is that the tau1P3P algorithm is optimised for the reconstruction of tau leptons with transverse visible energy in the range 20 – 70 GeV [13]. Higher momenta are correlated with higher decay lengths, see the figure 5.7. The tauRec algorithm has a better performance than the tau1P3P in the high E_T range. That is the reason why the tauRec is expected to have higher efficiency

efficiency as a function of L_{xy}

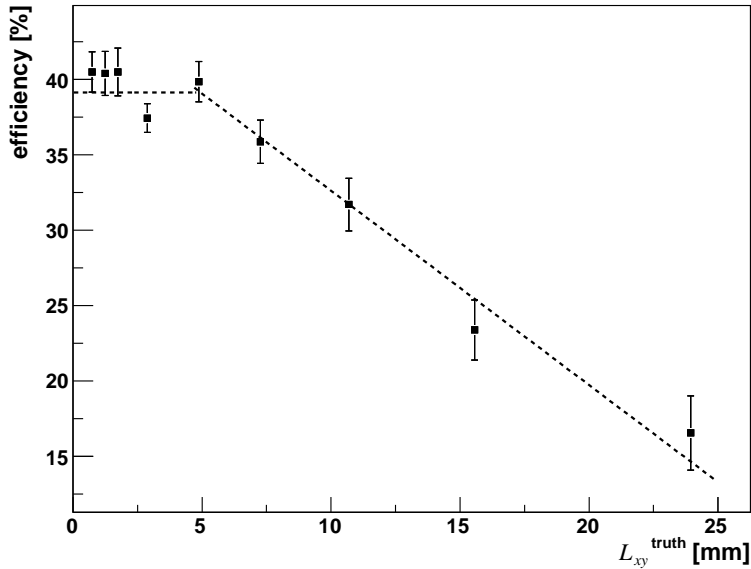


Figure 5.6: Efficiency as a function of the true transverse decay length

for higher decay lengths than the tau1P3P. Indeed, efficiency of the tauRec is higher for transverse decay lengths larger than 10 mm. In order to increase the statistics, it is possible to consider the tau1P3P algorithm for lower decay lengths ($L_{xy}^{\text{truth}} < 10$ mm) and the tauRec for higher lengths region ($L_{xy}^{\text{truth}} \geq 10$ mm). However, the difference between the number of taus reconstructed with the mixture of methods as described and only by the tau1P3P algorithm in the whole L_{xy}^{truth} range represents only 3% of all taus reconstructed with the tau1P3P. The effect on statistics is therefore negligible. Only the tau1P3P algorithm with its efficiency was considered.

Efficiency of the tau1P3P algorithm as a function of L_{xy}^{truth} can be fitted with a constant function up to ~ 5 mm and with a linear fit for higher lengths. The efficiency function reaches zero for $L_{xy}^{\text{truth}} \sim 35$ mm. There is only one 3 prong tau lepton reconstructed with the tau1P3P algorithm which has $L_{xy}^{\text{truth}} > 35$ mm. The efficiency function is considered to have a zero value for $L_{xy}^{\text{truth}} > 35$ mm. The efficiency function is represented with a dash line in the figure 5.6. Only the dependence of efficiency on the transverse decay length is considered. When the same selection criteria are applied, the efficiency function is assumed to be independent on the Higgs boson mass.

Several quantities were compared in order to check if a subset selected by the use of the estimated efficiency function on the MC true data has similar features as the subset reconstructed with the tau1P3P algorithm, see figures 5.8 and 5.9. Pseudorapidities for tau leptons reconstructed with the tau1P3P algorithm and for taus selected from the whole Monte Carlo true sample by means of the efficiency function were plotted. The distributions are in a very good agreement. This confirms our assumption that the pseudorapidity dependence of efficiency is negligible. The transverse momenta distribution is wider for our selection than for the tau1P3P set. Our rather simple model which emulates the selection of reconstructed taus cannot take into account all features of the complex event reconstruction process with correlations between different quantities. The quantity of the highest priority for our study is the true transverse decay length which is emulated with a very good accuracy, see the figure 5.9.

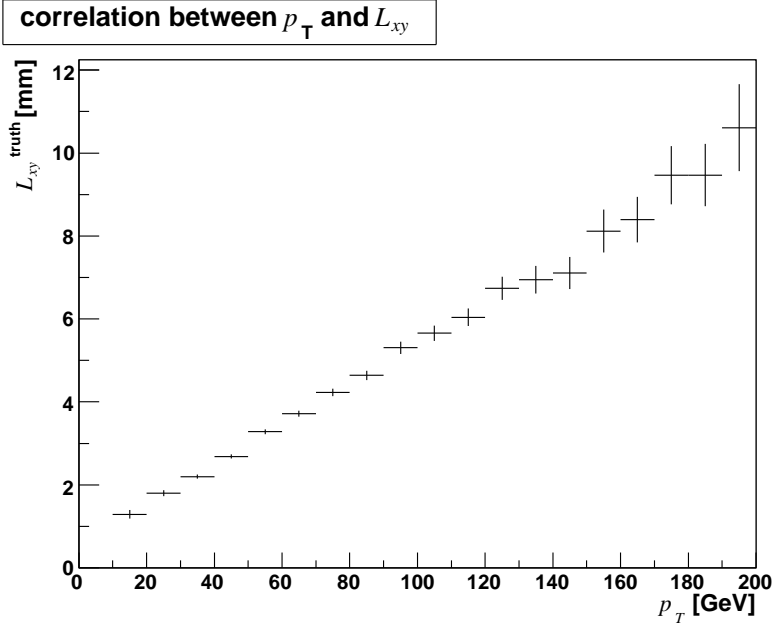


Figure 5.7: Correlation between the transverse momenta of tau leptons and their transverse decay lengths (profile plot)

algorithm	L_{xy}^{reco} range [mm]	$\langle L_{xy}^{\text{reco}} \rangle$ [mm]	$\sigma_{\langle L_{xy}^{\text{reco}} \rangle}$ [mm]
tau1P3P	0 – ∞	4.20	0.06
emulation	0 – ∞	4.13	0.05
tau1P3P	0 – 35	4.14	0.05
emulation	0 – 35	4.13	0.05

Table 5.1: Comparison of the transverse decay lengths reconstruction and emulation performed at the same MC sample. The symbol $\sigma_{\langle L_{xy}^{\text{reco}} \rangle}$ denotes a statistical error of the mean value of the reconstructed/emulated transverse decay length.

The emulation of the transverse decay lengths reconstruction and the reconstructed decay lengths from the full simulation are quantitatively compared in the table 5.1. The comparison is done in two intervals of decay lengths $L_{xy}^{\text{reco}} \in (0, \infty)$ mm and $L_{xy}^{\text{reco}} \in (0, 35)$ mm where L_{xy}^{reco} denotes the emulated or tau1P3P reconstructed transverse decay length. The emulation of the reconstructed decay lengths distribution was done in two steps. Firstly, efficiency function was applied on the whole Monte Carlo true sample in order to select a subset of tau leptons. Secondly, L_{xy}^{truth} smearing according to the resolution plots (figure 5.2 bottom) was done on the selected subset. The errors displayed in the table 5.1 represent statistical deviations. The mean emulated value in the restricted interval $L_{xy}^{\text{reco}} < 35$ mm agrees with the tau1P3P reconstructed average transverse decay length within its statistical deviation. However, this is not true about the mean value estimated in the whole L_{xy}^{reco} interval. Tau leptons with $L_{xy}^{\text{reco}} > 35$ mm contribute only by 0.1% to taus reconstructed with the tau1P3P algorithm in the whole L_{xy}^{reco} range. The difference in the number of tau leptons in the whole and the restricted L_{xy}^{reco} interval is negligible. Moreover, tau leptons with $L_{xy}^{\text{reco}} > 35$ mm have a relative error $|L_{xy}^{\text{reco}} - L_{xy}^{\text{truth}}| / L_{xy}^{\text{truth}}$ greater than 100% in almost all cases. Other point is that there is no tau lepton with $L_{xy}^{\text{reco}} > 35$ mm in the emulation. Only the restricted interval $L_{xy}^{\text{reco}} \in (0, 35)$ mm is considered further on.

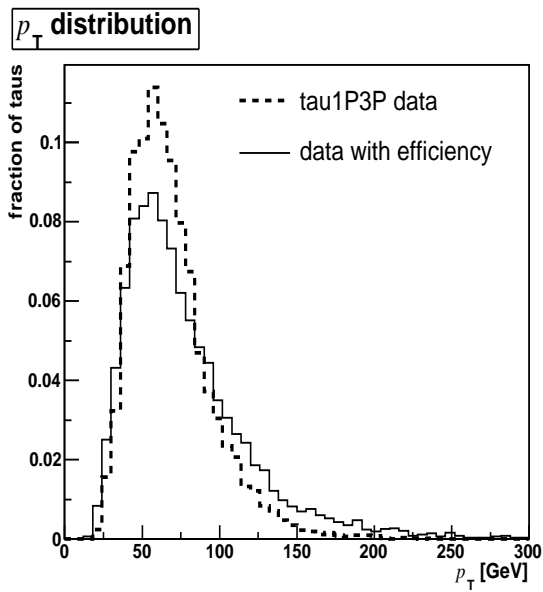
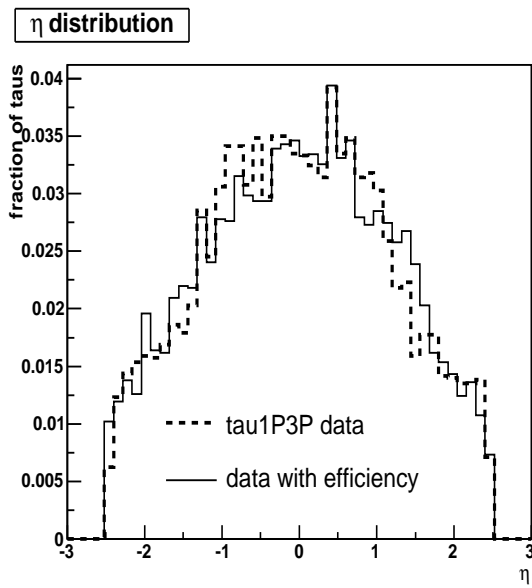


Figure 5.8: Pseudorapidity and transverse momenta distributions for taus reconstructed with the tau1P3P algorithm (dash line) and for taus selected according to the efficiency function from the whole MC true sample (solid line)

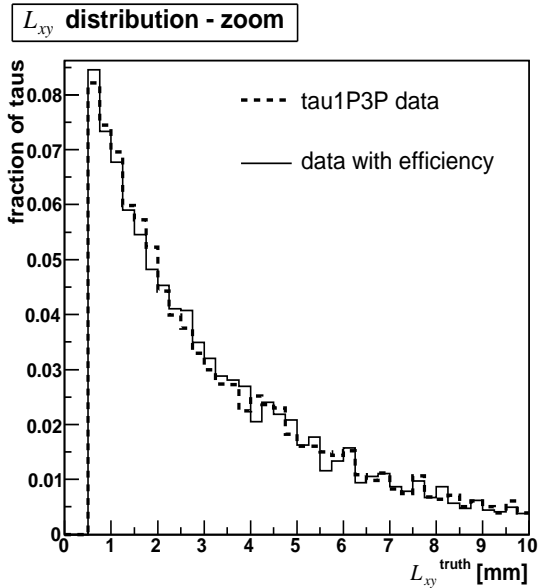
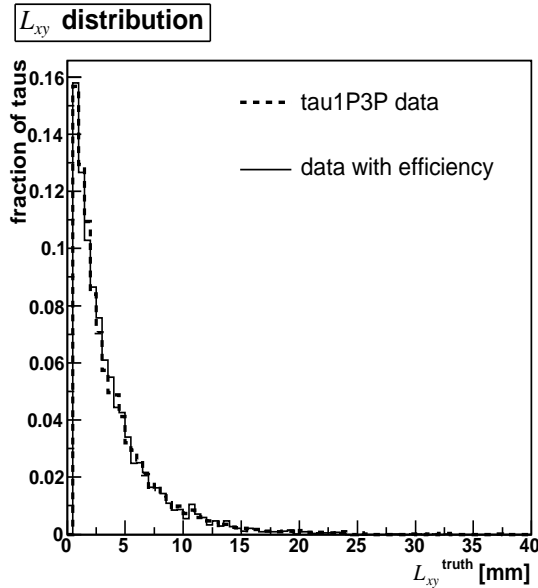


Figure 5.9: True transverse decay length distributions for taus reconstructed with the tau1P3P algorithm (dash line) and for taus selected according to the efficiency function from the whole MC true sample (solid line). The first plot shows the whole L_{xy}^{truth} spectra, the second one only the beginning of the L_{xy}^{truth} distributions (smaller bins are used).

5.3 Monte Carlo simulations for different Higgs boson masses

Monte Carlo truth simulation was done for six values of the Higgs boson mass in the range 115 – 140 GeV. The same version of Athena and Pythia as for the full simulated

sample was used. Only the Higgs boson production via the vector boson fusion as in the full simulated data was considered. 250000 $H \rightarrow \tau\tau$ events were generated for each Higgs mass. The samples included tau leptons decaying into both hadrons and leptons. Both taus decaying hadronically represented 42% of all events. After applying the basic selection criteria (4.8) - (4.10) 24% of $H \rightarrow \tau\tau$ events remained. Only 3 prong tau decays with three charged pions were considered (23% of hadronic tau decays). The cut on the inner detector coverage ($|\eta^{\text{vis}}| < 2.5$) and the minimum transverse decay length ($L_{xy}^{\text{truth}} \geq 0.5$ mm) was applied. The efficiency function derived in the section 5.2.2 was considered. The detector smearing of the L_{xy}^{truth} value was emulated and the cut $L_{xy}^{\text{reco}} < 35$ mm was applied. After this procedure, the number of taus decreased to 7300 – 7900 depending on the Higgs boson mass. The dependence $\langle L_{xy} \rangle(m_H)$, so-called mass estimator, was calculated from this sample.

The mean value is not the only possibility how to evaluate an unknown distribution. The mean value does not describe the shape of the dependence. Therefore the average value may not be the best estimator for the decay lengths distribution. The best way would be to fit the L_{xy} distribution with a theoretically predicted curve. However, there are two parts contributing to the decay lengths spectrum, the tau lepton lifetime and momentum. The lifetime distribution is an exponential, but the shape of the transverse momenta distribution is known only from simulations and it can not be fitted with a simple function. The L_{xy} distribution can be fitted with an exponential, but the fit is not optimal, especially in the low decay lengths region with the highest statistics. The fit parameter of the exponential curve is therefore not expected to be a good number evaluating the decay lengths distribution. Besides the mean value $\langle L_{xy} \rangle$ (section 5.3.1), a fraction f of events with L_{xy} smaller and higher than an intersection point of decay lengths distributions for various Higgs boson masses (section 5.3.2) was used to create another mass estimator.

Due to the fact that the cross-section of the Higgs boson production varied with its mass, the number of accepted tau leptons for different Higgs masses did not correspond to the same luminosity. Small statistical errors of $\langle L_{xy} \rangle$ and f for all generated Higgs masses are of crucial importance for construction of a reliable mass estimator. That was the reason why all generated events, not only number of events corresponding to the same luminosity, were used. In principle, it was possible to reduce the statistical errors to a negligible level if bigger statistics generated. Due to this fact confidence level intervals around the predicted curve were not considered.

5.3.1 Mean value method for the mass estimation

Mass estimators for the mean value of the true transverse decay lengths (lower line) and for the reconstructed lengths (upper line) obtained with the same data subsets are shown in the figure 5.10. The dependence of the mean transverse decay length on the Higgs boson mass can be fitted with a linear curve $\langle L_{xy} \rangle = a m_H + b$. The results of the fit can be seen in the table 5.2. Two data sets display a shift in the average transverse decay length, but their slopes differ negligibly. Similarity of the slopes shows that the systematic uncertainties are the same in both cases. The difference in the absolute term points out that the reconstructed lengths are overestimated. This is in agreement with the resolution plots (figure 5.2 bottom) where events with $L_{xy}^{\text{reco}} > L_{xy}^{\text{truth}}$ dominate.

To verify the parameters of the fit, let us consider the most simple case when the Higgs boson is produced in the rest in the laboratory frame. Momenta of both tau leptons from

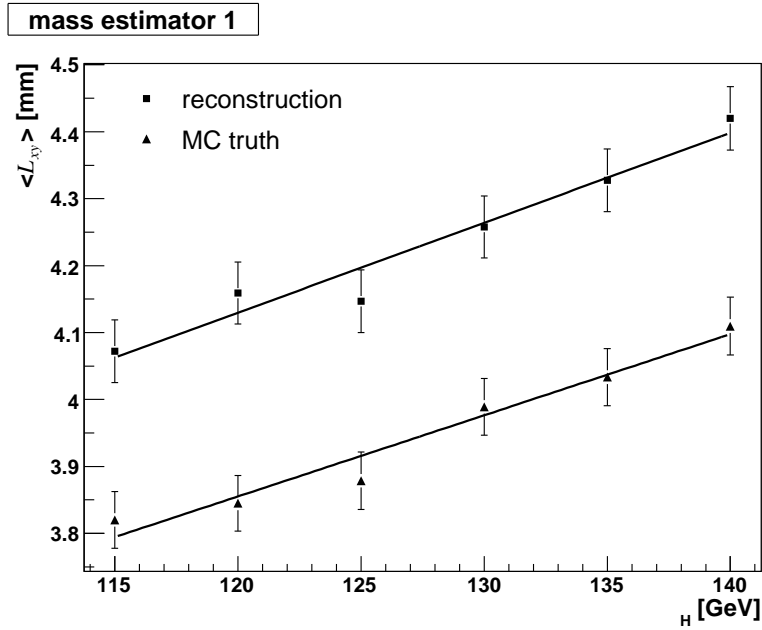


Figure 5.10: The mass estimator for the mean value of true and reconstructed transverse decay lengths obtained with the same data subsets. Parameters of the fit are further discussed in the text, see the section 5.3.1.

MC truth	a [mm · GeV ⁻¹]	$(1.2 \pm 0.2) \times 10^{-2}$
	b [mm]	(2.4 ± 0.3)
reconstruction	a [mm · GeV ⁻¹]	$(1.3 \pm 0.2) \times 10^{-2}$
	b [mm]	(2.5 ± 0.3)

Table 5.2: Parameters of the $\langle L_{xy} \rangle = a m_H + b$ dependence

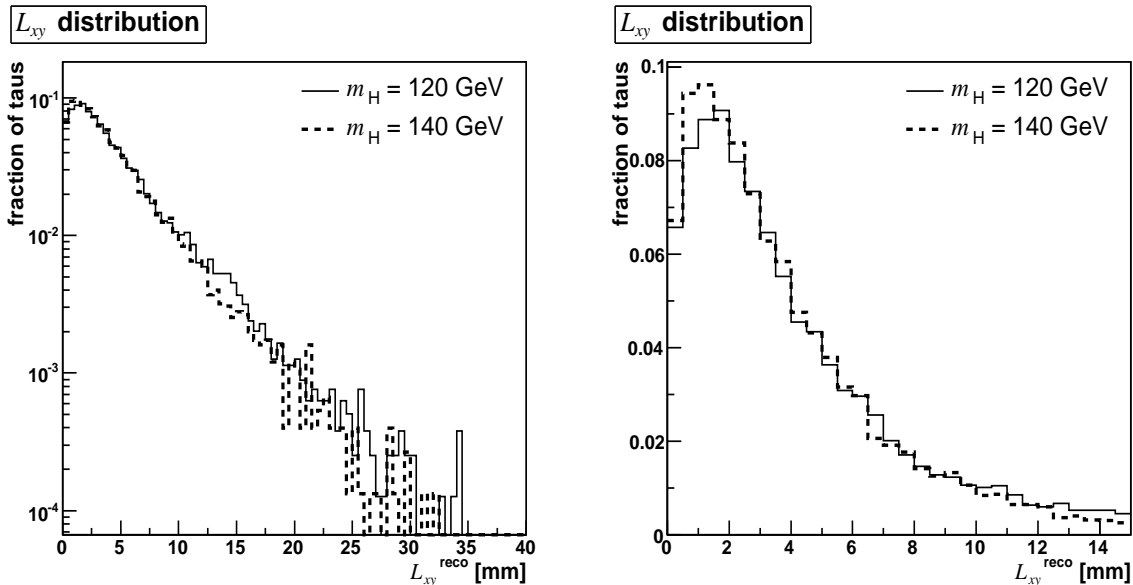


Figure 5.11: Normalised distributions of reconstructed decay lengths for two Higgs boson masses. The left plot shows the whole spectra, the right one only the beginning.

the Higgs decay have the same magnitude in this case, see the equation (5.1), which can be approximated as $p_\tau \approx m_H/2$. Using the equation (5.2) the mean transverse decay length then obeys

$$\langle L_{xy} \rangle \approx t_0 \frac{m_H}{2} \frac{\langle \sin \theta \rangle}{m_\tau} \quad (5.3)$$

The mean value of sinus of the polar angle θ is calculated over the area covered by the inner detector. The function is linear in m_H , no absolute term is present. The coefficient of the proportionality equals $1.9 \times 10^{-2} \text{ mm} \cdot \text{GeV}^{-1}$ which is in a qualitative agreement with the parameter a of the fit. The absolute term and the difference in the value of the slope is associated with effects which are not included in our calculation. The absolute term appears with the relativistic boost. The difference in the coefficient of the proportionality is associated mainly with the efficiency function. The efficiency function influences also the value of the absolute term. The effects of cuts on L_{xy}^{truth} and p_T are not so significant.

5.3.2 Fractional method for the mass estimation

To compare the decay lengths distributions for different Higgs masses, the intersection point of the normalised distributions was found. The distributions for all masses were very similar and they intersected in the same region. The transverse decay length value 5.0 mm, where the distributions were almost the same, was considered as a working point. The L_{xy}^{reco} spectra for two Higgs boson masses are shown in the figure 5.11. The fraction $f = N(L_{xy} < 5.0 \text{ mm})/N(L_{xy} > 5.0 \text{ mm})$, where N denotes the number of tau leptons with L_{xy} in the given region, was derived for each Higgs boson mass. The linear regression was used to fit the $f(m_H)$ dependence as in the previous case, see the figure 5.12. The fraction f decreases with the increasing Higgs mass. This means that the decay lengths distribution contains higher lengths for higher Higgs boson masses which is in agreement with previous results and our expectations.

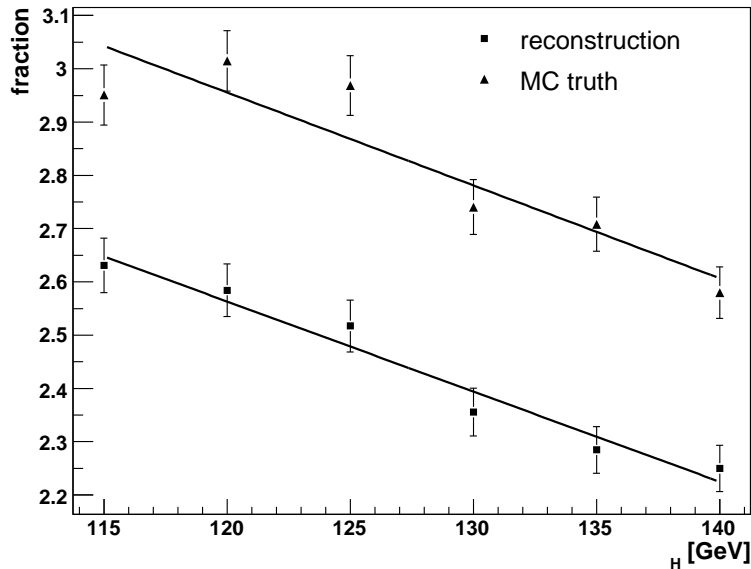


Figure 5.12: The mass estimator for the fraction f of events with $L_{xy} < 5$ mm and $L_{xy} \geq 5$ mm for true and reconstructed transverse decay lengths obtained with the same data subsets.

5.3.3 Accuracy of the decay lengths methods

The full simulated sample H(120) was used to estimate the accuracy of the mean value and fractional methods. The accuracy of both methods was estimated for the integrated luminosities 300 fb^{-1} and 500 fb^{-1} . The mean value of the reconstructed transverse decay lengths $\langle L_{xy}^{\text{reco}} \rangle$ and the fraction f was calculated for the 3 prong tau leptons reconstructed with tau1P3P algorithm in the H(120) full simulated sample. The number of events corresponding to the required luminosity was considered for the calculation. The derived values of $\langle L_{xy}^{\text{reco}} \rangle$ and f were considered to represent the values from the real experiment. The Higgs boson mass with its statistical error was predicted by the use of the mass estimators $m_H = (\langle L_{xy}^{\text{reco}} \rangle - b)/a$, similarly for f . See the figure 5.13 for the mean value method where the solid horizontal line represents the average value of the transverse decay length and the dash lines its statistical errors derived from the H(120) sample for the integrated luminosity 300 fb^{-1} . The mass estimator is drawn with the bold solid line in the plot. The results for both decay length methods are listed in the table 5.3. Both methods give similar results with a statistical error in the range $5 - 6 \text{ GeV}$ for the integrated luminosity 300 fb^{-1} . The statistical deviation scales as one over a square root of a number of events which can be derived from the errors for different luminosities. All predicted Higgs boson masses correspond to the true value $m_H = 120 \text{ GeV}$ within their statistical errors. Since the mean value method gives slightly better resolution, this method is considered as a decay length method further on.

It is possible to estimate the error of the Higgs boson mass measurement for different masses from the knowledge of the error for $m_H = 120 \text{ GeV}$ and expected number of events. The tau products have higher mean momenta for higher Higgs boson masses. That is why there is a higher number of tau leptons which satisfy the basic selection criterion on p_T^{vis} and the cuts on the transverse momenta of the tracks for higher Higgs masses. The fraction of generated 3 prong tau leptons satisfying the basic selection criteria (4.8) - (4.10) and

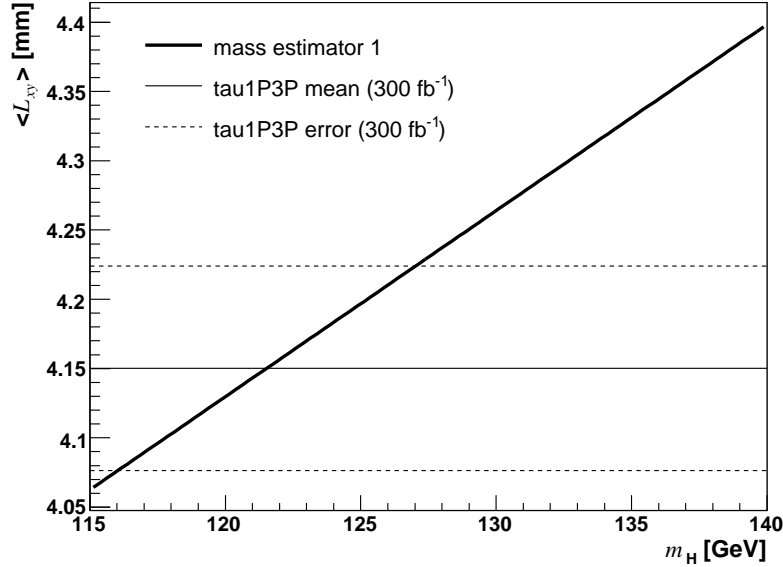


Figure 5.13: Accuracy of the mean decay length method for the integrated luminosity 300fb^{-1} . For further details see the text, section 5.3.3.

method	m_{H}^{300} [GeV]	$\sigma_{\text{stat}}^{300}$ [GeV]	m_{H}^{500} [GeV]	$\sigma_{\text{stat}}^{500}$ [GeV]
mean value	121.5	5.5	117.5	4.2
fraction	122.2	6.0	119.8	4.8

Table 5.3: Results of the decay length methods for the Higgs mass $m_{\text{H}} = 120$ GeV. The symbol σ_{stat} stands for the statistical error. The upper index denotes the integrated luminosity which the predicted Higgs boson mass and its error correspond to. For further details see the text, section 5.3.3.

m_H [GeV]	σ [fb]	$N_{3\text{prong}}$	σ_{m_H} [GeV]
120	81	11660	5.5
130	60	8640	6.4
140	35	5040	8.4

Table 5.4: Measurement of the Higgs boson mass for the integrated luminosity 300 fb^{-1} . The symbol σ denotes the cross-section for the VBF $H \rightarrow \tau\tau \rightarrow h\nu_\tau h\nu_\tau$ process with both taus satisfying the basic selection criteria (4.8) - (4.10), $N_{3\text{prong}}$ stands for the expected number of tau leptons decaying into three charged particles and σ_{m_H} the statistical error of the decay length method.

the expected number of reconstructed tau leptons satisfying additional cuts ($|\eta^{\text{vis}}| < 2.5$, $L_{xy}^{\text{truth}} > 0.5 \text{ mm}$, $L_{xy}^{\text{reco}} < 35 \text{ mm}$) equals $\sim 26.2\%$ for the Higgs mass 120 GeV and $\sim 27.3\%$ for the mass 140 GeV. The difference in efficiency of the reconstruction is only 1% and can therefore be neglected. The efficiency of the tau leptons reconstruction is considered to be constant over the considered Higgs mass region. The precision of the decay length method is shown in the table 5.4 for three Higgs boson masses. The displayed errors are the statistical deviations. As the cross-section decreases with the increasing Higgs boson mass, the measurement accuracy is worse for higher masses. The statistical uncertainties are between 5% and 6%.

Chapter 6

Conclusion

Two independent methods for the Higgs boson mass reconstruction were investigated - the collinear approximation and the decay length method. Accuracy of both methods was compared for the Higgs boson mass 120 GeV and the number of events corresponding to the integrated luminosity 300 fb^{-1} . Similar selection criteria were used. Different cuts on pseudorapidity were used (only region covered by the inner detector was assumed for the decay length method). Only 3 prong events were studied in the decay length method. Tau leptons were reconstructed with different tau identification algorithms in the studied mass reconstruction methods. The results of both methods are listed in the table 6.1.

method	m_H [GeV]	σ_{stat} [GeV]
collinear app.	118.1	0.3
decay length	121.5	5.5

Table 6.1: Comparison between the collinear approximation and the decay length method for the integrated luminosity 300 fb^{-1}

The statistical deviation is 0.3% for the collinear approximation, whereas 4.5% for the decay length method. The mass resolution is significantly worse for the decay length method than for the collinear approximation. The collinear approximation is a better method for the Higgs boson mass reconstruction. On the other hand, the advantage of the decay length method is that it is independent on the tau energy calibration and transverse missing energy measurement. The energy measurement can be used only for the signal events selection in order to reject the background events. This method can be used as an additional method to the collinear approximation as a check of consistence.

True information from the Monte Carlo simulation was used in the analysis. This will not be available in the real experiment. Especially the cut $L_{xy}^{\text{truth}} > 0.5 \text{ mm}$ has to be replaced. Similar cut has to be applied on the reconstructed decay lengths if real events are considered.

Statistical errors of the measured mean value of the decay length are crucial for the Higgs boson mass uncertainties. The statistics will grow four times if also 1 prong events can be added in the analysis. The secondary vertex is not reconstructed for tau leptons with one associate track, but the impact parameter is reconstructed for every tau candidate. It is possible that by measuring the impact parameter we would have new information correlated with the Higgs boson mass which can lead to a more precise result.

Systematic errors were not investigated. The systematic deviations are associated

mainly with the Monte Carlo simulations. Processes depending on the MC generator which can affect the uncertainties of the Higgs boson mass reconstruction methods are supposed to be mainly the parton density functions of the colliding protons, the initial and final state QCD radiation. The systematic deviation of the MC modelling could be estimated by using at least two independent Monte Carlo simulation programs. Only Pythia generator for the Higgs boson production + decays and Tauola for the tau leptons decays were used in the analysed MC samples.

No background studies have been done. Only samples with signal events were used. Additional uncertainties on the Higgs mass will arise once the background processes (e.g. $Z \rightarrow \tau\tau$) are considered.

Bibliography

- [1] ATLAS Collaboration: *ATLAS Detector and Physics Performance Technical Design Report*, CERN/LHCC/99-14, 1999.
- [2] ATLAS Collaboration: *ATLAS Inner Detector Technical Design Report*, CERN/LHCC/97-16, 1997.
- [3] M. Delmastro: *Status of LAr EM performance and measurements for CTB*, talk at ATLAS Overview Week, February 2006.
- [4] V. Giangiobbe: *2004 Combined Test Beam*, talk at ATLAS Physics Workshop, June 2005.
- [5] M. Heldmann, D. Cavalli: *An Improved τ -identification for the ATLAS Experiment*, ATL-PHYS-PUB-2006-008, 2005.
- [6] C.S. Hill, J. R. Incandela, J. M. Lamb: *A method for measurement of the top quark mass using the mean decay length of b hadrons in $t\bar{t}$ events*, Phys.Rev.D71:054029, 2005.
- [7] J. Hořejší: *Fundamentals of Electroweak Theory*, The Karolinum Press, 2002.
- [8] <http://lepewwg.web.cern.ch/LEPEWWG/>.
- [9] A.Kaczmarska: *tauRec: integration, algorithms, performance*, talk at Tau mini-workshop, February 2007.
- [10] P. Loch: *Jet/etmiss/tau combined performance*, talk at ATLAS Overview Week, February 2007.
- [11] B. Mellado, G. Unal, S. L. Wu: *Higgs Production Cross-Sections and Branching Ratios for the ATLAS Higgs Working Group*, <http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/HIGGS/higgs-xsec/xsec-br.html>, 2004.
- [12] Particle Data Group: *Particle Physics Booklet*, 2006.
- [13] E. Richter-Was, L. Janyst, T. Szymocha: *The tau1P3P algorithm: implementation in Athena and performance with CSC data samples*, ATL-COM-PHYS-2006-029, 2006.
- [14] E. Richter-Was, T. Szymocha: *The light Higgs decay into tau-lepton pair: reconstruction in different production processes*, ATL-COM-PHYS-2004-010, 2004.
- [15] J.Tanaka *Discovery potential of the Standard Model Higgs at the LHC*, ATL-COM-PHYS-2004-075, 2004.

- [16] R. Wigmans: *Advances in Hadron Calorimetry*, Annu. Rev. Nucl. Part. Sci. 1991.41, 1991.