UNIVERSITY OF COPENHAGEN
NIELS BOHR INSTITUTE

Almut Maria Pingel

Tau lepton identification and studies of associated Higgs boson production with the ATLAS detector

THIS THESIS HAS BEEN SUBMITTED TO THE PHD SCHOOL OF
THE FACULTY OF SCIENCE, UNIVERSITY OF COPENHAGEN

Dissertation for the degree of Doctor of Philosophy in physics

Almut Maria Pingel
Copenhagen 2015

Academic advisors:

Associate Professor Mogens Dam, Niels Bohr Institute
Associate Professor Stefania Xella, Niels Bohr Institute

Assessment committee:

Professor Ian Bearden, Niels Bohr Institute (Chair)
Professor Ulrik Egede, Imperial College London, UK
Professor Paula Eerola, University of Helsinki, Finland

Laß die Moleküle rasen,
was sie auch zusammenknobeln!
Laß das Tüfteln, laß das Hobeln,
heilig halte die Ekstasen!
(Christian Morgenstern, Galgenlieder)

Abstract

In this thesis two main topics are presented: the identification of hadronic tau decays and a search for Higgs bosons that decay to tau leptons and are produced in association with a Z boson.

In 2012, the ATLAS detector at the LHC recorded $20\,\text{fb}^{-1}$ of data at a centre-of-mass energy of 8 TeV. Many studies on Standard Model (SM) processes, as well as searches for physics beyond the SM were carried out, which all rely on a good particle identification. This thesis focuses on the identification of hadronically decaying tau leptons. Multivariate techniques are used to reject backgrounds that mimic the tau signature in the detector and are falsely reconstructed as tau leptons. The focus is on the optimisation of the algorithm using boosted decision trees (BDT). The method is prepared for the 2012 data taking period and made robust against run conditions with many simultaneous proton-proton interactions per bunch crossing (pile-up). Furthermore, the dependence of the identification method on the tau momentum is studied, and the usage of additional algorithms that explore the pion content of the tau decay.

The second part concerns the Higgs boson which was discovered in 2012. Higgs boson production via a top quark-loop (gg fusion) as well as via vector boson fusion (VBF) has been observed, but no evidence for the production in association with a vector boson (VH) has been claimed yet. The thesis presents a study of this production mode, with a leptonically decaying Z boson in the final state, and the Higgs boson decaying to a pair of tau leptons. Both the hadronic decay of the tau lepton as well as the leptonic decays (τ_e, τ_μ) are considered. The background is estimated from simulation, and renormalised in side bands to match the data. The contribution is then transferred to the signal region using the fake factor method. Combining the result from the four final states $\tau_{\text{had}}\,\tau_{\text{had}}$, $\tau_e\tau_{\text{had}}$, $\tau_\mu\tau_{\text{had}}$ and $\tau_e\tau_\mu$, a limit is set on the cross section of the process $Z(\rightarrow ee, \mu\mu)H(\rightarrow \tau_{\text{had}}\tau_{\text{had}})$. No deviation from the Standard Model is found. An upper observed (expected) limit of 8.96 ($6.18^{+8.94}_{-4.45}$) times the SM cross section is set at 95% confidence level for a Higgs boson mass of 125 GeV.

Resume

I denne afhandling er to emner præsenteret: Identifikation af hadroniske henfaldende tau-leptoner med ATLAS-detektoren ved CERN samt en søgen efter en Standard Model (SM) Higgs-boson ved associeret produktion med en Z-boson.

I 2012, optog ATLAS detektoren ved LHC $20\,\mathrm{fb}^{-1}$ data med en Kollisionsenergi på 8 TeV. Mange studier af såvel Standardmodellen og mere avancerede teoretiske modeller, blev udført med stor succes, på dette data. Alle disse studier byggede på en god identifikation af partikler.

Den første del af afhandlingen fokuserer på identifikation af hadroniske henfald af tau-leptonen. Multivariate statistiske metoder er brugt til at afvise baggrundshypotesen som er en signatur, der ligner tau-leptonens i detektoren, og derfor fejlagtigt rekonstruktueres som en tau-partikel. Fokus i denne opgave er på optimering af tau identifikationen ved brug af et Boosted Decision Trees (BDT). Metoden er anvendt på data optaget i 2012, og er lavet robust for at håndtere situationer med mange overlappende proton-proton interaktioner per bundt passage (pile-up). Yderligere, er afhængigheden af identifikationsmetoden af tauimpulsen også studeret, samt muligheden for at benytte nye algoritmer, der udnytter Pioner dannet ved tauhenfald.

Den anden del af afhandlingen omhandler Higgs-bosonen, som blev opdaget i 2012. Higgs-bosonen er observeret via produktion af et top kvark-loop (gg fusion) eller Vektor-Boson-Fusion (VBF). Der er dog ingen dokumentation for produktion af Higgs-bosonen i forbindelse med en vektor-boson (VH). I afhandlingen præsenteres et studie af denne produkton med en Z-boson, der henfalder leptonisk og en Higgs-boson, der henfalder til et par af tau-leptoner. Både den hadronisk henfaldende tau-lepton såvel som den leptoniske henfaldende tau (taue, taumu) er behandlet. Baggrunden er estimeret ud fra monte-carlo simulering og normaliseret til såkaldte side-bands ved at tilpasse data. Bidraget overføres derefter til signal området ved hjælp af "fake-faktor" metoden. Kombineres alle henfaldsprodukterne for Higgs-bosonen: tauhadtauhad, tauetauhad, taumutauhad og tauetaumu, opnås en observeret (forventet) øvre grænse for produktionstværsnittet på 8.96 ($6.18^{+8.94}_{-4.45}$) gange SM tværsnit på 95% konfidensniveau for en Higgs-boson med hvilemasse på 125 GeV.

Contents

1. Introduction

Preface

Particle physics stands for the studies of the very small, and is concerned with the most fundamental matter constituents and the forces that bind them. Even though the processes at this scale are so different from what we experience in our every day life, mankind has reached an astonishing level of understanding. The understanding is based on mathematical models that describe the processes and predict interactions. The processes themselves sometimes appear alien to us, because there is no analogue in human experience. Particles are created out of pure energy and vice versa, interactions and whereabouts are based on probabilities, and the boundaries of 'particle' and 'wave' blur. We are constantly penetrated by neutrinos without being harmed by them and without even perceiving them. Yet, we can build detectors sensitive and large enough to measure neutrino properties. The continued exploration of particles and of the fundamental forces requires nowadays large experiments, powerful accelerators and endurance to deduct an experiment over many years until enough data is collected to claim evidence for the observation of a rare process. The LHC at CERN currently is the world's largest (man-built) accelerator. Proton-proton collisions have just been started again and are recorded at 13 TeV centre-of-mass energy - an energy never reached before in an accelerator. That being said, naturally produced particles (cosmic rays) reache such and much higher energies, and from that angle the proton-proton collisions at the LHC are not a unique happening. However, the LHC provides the possibility to built a detector around the collision point, so that the collision and the out-flying particles can be studied in an experimentally well-defined setup. Already after the first years of data taking at 7 and 8 TeV, the experiments have expanded our knowledge of particle physics. Pre-eminently in this regard is the discovery of the long-searched Higgs boson, the missing piece in the current formulation of the Standard Model.

My personal way into particle physics started with a small project within the CMS collaboration on jet reconstruction and gained speed with my diploma thesis within the ATLAS collaboration. When I first searched for a topic for my diploma thesis, I talked to the Peter Schleper, Johannes Haller and Philip Bechtle, who led CMS and ATLAS research groups at the University of Hamburg and at DESY. Independently of each other, and even though the groups had different research profiles, they all proposed projects with a focus on tau leptons. This settled my path, and the tau lepton has since been the barycentre of my work. The tau lepton also led my way to the Niels Bohr Institute, which was very enganged in the ATLAS Tau Combined Performance Working Group. The work of Mogens Dam and Stefania Xella caught my interest and they supported me to work on tau identification. And this is why this thesis focuses on the tau lepton.

Tau lepton identification

Any analysis with tau leptons in the final state relies on a high tau identification efficiency and a high rejection of objects that mimic the tau detector signature. In contrast to the other leptons, tau leptons are very short-lived and typically decay before reaching the detector. Therefore, they can only be identified by their decay products. Because the leptonic decay mode can hardly be distinguished from prompt light leptons, only the hadronic mode is considered. The largest background to tau identification are QCD-jets, which are produced at high rate at the LHC. Multivariate techniques are used

to reject this background. The focus of the studies presented in this thesis is on the algorithm using Boosted Decision Trees. Several aspects of tau identification are studied, with the goal of optimising the performance in terms of pile-up[1] robustness, momentum dependence, and background rejection. The studies on the pile-up dependence were carried out in 2011, resulting in the development of a pile-up robust tau identification for the 2012 data taking period. The other studies were carried out during the LHC shut-down phase. The results have influenced the algorithms prepared for the new data taking period.

Associated Higgs boson production

The discovery of the Higgs boson in 2012 brought about the challenge to explore all aspects of the new boson. To begin with, the clear focus was on the observation of the frequent or experimentally clean production and decay mode. Towards the end of the Run I data-taking period, the focus was driven towards the rarer modes, even though the available amount of data was not expected to be enough to observe these modes. The production of the Higgs boson in association with a Z boson, in which the Higgs boson decays into a pair of tau leptons, is part of this category and presented in this thesis. The idea was to use the shut-down phase of the LHC to explore the signature and current sensitivity towards the signal. Furthermore, the task is to identify potential difficulties of the analysis and to test control and validation regions. In order to obtain the broadest possible overview, all six final states (generated by the leptonic or hadronic decay of the tau leptons) are analysed in a common approach. While a growing number of analyses bases the background estimation on fully data-driven techniques, a mixed approach is used here. Simulated samples are used throughout all analysis steps to model the background. Corrections for the normalisation and critical variables are then obtained from data. This approach is chosen in order to provide a cross check for the fully data-driven approaches, and with that gain confidence in the background modelling. The mixed approach turned out considerably more complex than expected. To a large extend this is driven by the statistical limitation of the signal regions, the considered side bands and the simulated samples. It becomes clear that a larger amount of simulated events is needed in order to fully understand the composition of all regions. Especially the validation regions close to the signal regions show deviations between the data and the prediction, and an understanding of the deviations will strengthen the confidence in the background modelling of the signal region considerably. The four final states $e\mu$, $e\tau_h$, $\mu\tau_h$ and $\tau_h\tau_h$ are eventually used to set a limit on the cross section of the process $Z(\rightarrow ee, \mu\mu)H(\rightarrow \tau\tau)$.

Outline

The thesis is structured as follows. Chapter 2 gives a brief overview of the particles and forces described in the Standard Model of particle physics as well as the underlying mathematic description and limitations. Furthermore, it includes a section on the characteristics of proton-proton collisions. The experimental setup is described in Chapter 3, introducing both the LHC accelerator and the ATLAS detector. Chapter 4 focuses on the collected data and its preparation. The reconstruction of the detector signal and the identification of particles is discussed briefly, as well as the ATLAS trigger. Furthermore, a short description of efficiency measurements is provided, which are used to correct mis-modelling in simulated samples. Chapter 5 is dedicated to the identification of hadronically decaying tau leptons with the ATLAS detector. Besides containing a description of my own studies concerning the identification with boosted decision trees, an overview is given of the other aspects of tau reconstruction, namely the lepton vetos, the energy calibration and efficiency measurements. The study of associated Higgs boson production is described in Chapter 6. After a short introduction

[1] The term 'pile-up' describes the occurrence of several proton-proton interactions within the same bunch crossing.

of the Higgs boson and the searches and measurements carried out previously, my own analysis is presented. The results and the conclusions are presented in Chapter 7.

Acknowledgements

I would like to greatly thank my supervisor Mogens Dam, for supporting me throughout my PhD study, for critical examination of numerous histograms and plots, for motivating discussions and well thought-through and helpful comments.

I also would like to thank Stefania Xella, for working with me on tau identification and quickly getting me engaged in the tau working group. Being active within the tau working group has been very motivating for me, and I want to thank all the (current and future) tau conveners, namely Soshi Tsuno, Stan Lai, Stefania Xella, Martin Flechl, Attilio Andreazza, Will Davey, and Pier-Olivier DeViveiros, for giving me a great time in the tau working group. I learned a lot from presenting my work in the regular meetings, being challenged with paper editing, having fun on the tau workshops and experiencing conferences in Japan and Spain. I regret every day that I have not spend on boosting tau performance.

Furthermore, I would like to say thanks to all my colleagues at NBI. Especially Pavel, Peter and Ingrid who helped me get started at NBI; Kristian, Ask, Lars, Morten M. and Valentina for relaxing coffee breaks; Alex and Sasha for first introducing me to bouldering; Lisa and Michael for never being tired of solving bouldering problems; Craig for dragging me away from my desk from time to time; James for not being surprised that the MC modelling is not always perfect; And Axel for transforming many hours in the train into a journey to far away galaxies;

Special thanks go to Lotte for helping me with the Danish abstract, Michael for fighting with plotting macros, and Morten D. for designing the title page. Furthermore, my parents, who spend many hours reading and correcting this thesis.

2. Theory

This chapter gives an introduction to particle physics. The characteristics of the fundamental particles and forces are presented in Sec. 2.1, follwed by a brief description of the mathematical framework in Sec. 2.2. The shortcomings of the Standard Model are addressed in Sec. 2.3. The last section, Sec. 2.4, focusses on the particularities of particles physics at a hadron collider.

2.1. The Standard Model of Particle Physics

Mankind has since long been fascinated by the question of the fundamental basis of our world. Two competing concepts existed, without scientific basis for one or the other: on one hand the idea that everything was made of a few undividable building blocks, and on the other hand the concept of a continuum. Manifold studies of elements and their chemical properties convinced many scientists in the 19th century of the building block theory, with atoms (from the Greek ἄτομος, indivisible) as the fundamental bricks [1]. However, in 1897 J. J. Thomsen discovered the electron and it became clear that the atoms were not the end of the road.

More than 100 years later, atoms have been studied in detail. They can be ionised, excited, split, fused and even the first antiatoms have been created [2]. They are clearly not the smallest building block, but they remained the dominating object to describe the world as we see it in everyday life. Looking inside the atom and understanding the constituents and what binds them is the entrance to the world of particle physics.

The atom consists of a nucleus orbited by electrons. While the nucleus is not elementary, the electron is. Together with the muon and tau lepton, it forms the group of charged *leptons* (Tab. 2.1). The lepton-family is completed by three associated neutral and extremely light neutrinos. Neutrinos do not form part of the atomic bound state, but are emitted in radioactive decays. The atomic nucleus is composed of protons and neutrons (except for the hydrogen atom, which consists of a single proton). These belong to the huge group of *hadrons*, composite particles made of *quarks*. The group splits into *mesons*, with integer spin and composed of a quark-antiquark pair, and *baryons*, with half-integer spin and composed of three quarks[1]. The constituent quarks are elementary particles, six in total. The *up-type* quarks u, c, t have an electric charge of $q_e = +2/3$, while the *down-type* quarks d, s, b have an electric charge of $q_e = -1/3$. Protons, consisting of uud, thus carry a charge of $q_e = 1$, whereas neutrons, made of udd, are electrically neutral.

The atom, or rather its matter constituents, is held together by fundamental forces. These are mediated by the force carriers, allowing the particles to interact with each other if they carry the same type of charge. The protons and electrons are both electrically charged and interact through the *electromagnetic force*. The mediator is the photon, which itself is neutral and massless. As a result, the electromagnetic field has infinite reach. The quarks are bound into hadrons by the *strong force*. The associated particle is the gluon, which couples to colour charge. It is massless and electrically neutral, but carries colour charge itself. The gluon can therefore self-interact which limits its reach to ∼1 fm. The strong force is much stronger than the electromagnetic force and can bind quarks or

[1] Theoretically possible are also integer and half-integer spin combinations of more quarks: tetra-quarks (two quarks and two antiquarks) and penta-quarks (four quarks and one antiquark), respectively. The experimental evidence is discussed controversially, see for example Refs. [3–5].

hadrons of the same electromagnetic charge, which would otherwise feel a repelling force. The strong force vanishes for very small distances or very high energies (*asymptotic freedom*) but is constant otherwise. Two diverging quarks have a linearly growing potential between them. At a certain point, enough energy is stored in the field to create quark-antiquark pairs - which form new hadrons with the primary quarks. This *confinement* leads to the fact that free quarks can not be observed and all hadrons are, at distance, colour neutral. Within a nucleus, protons and neutrons are close enough to feel each others colour field and to be attracted by it. The last missing force is the *weak force*, mediated by the neutral Z boson and the charged $W^{\pm}$ bosons. The associated charge is the weak charge (or weak *isospin*), carried by all matter constituent types[2]. The weak force is responsible for radioactive decays: down-type quarks can transform into up-type quarks (and vice versa) by the exchange of a W boson and charged leptons can transform into the associated neutrinos (and vice versa). The W and Z bosons are both massive, which limits their interaction range to $< 1\,\text{fm}$. A summary of the fundamental particles is given in Tab. 2.1.

The formalisms to describe the elementary particles and their interactions are collectively called the Standard Model of particle physics. The term was first introduced by Treiman [6, 7] in the 1970's, when the mathematic formulation of the electromagnetic, the weak and the strong interaction was established.

2.2. Mathematic description

The Standard Model of Particle Physics (SM) is the mathematic description of the fundamental particles and their interactions. It has been extremely successful so far. In several cases, particles were predicted before their direct experimental observation. Calculations of process rates agree very well with the experimental observation. One of the greatest successes is the calculation of the magnetic dipole moment of the electron. When the Dirac equation was introduced (see Sec. 2.2.1), it predicted a factor 2 compared to the dipole moment of a rotating mass, which was in agreement with experimental observations at that time. When higher precision data became available deviations were found, at the order of 1‰. But also the theoretical predictions improved once higher order corrections were added. The factor is today the most precisely measured quantity in physics, known to 12 digits [8]. So far, no deviation from the theoretical calculation has been found.

The mathematical framework to describe elementary particles and their interactions is the quantum field theory, in which both fermions and bosons are expressed as quanta of fields [9]. A major concept is to use symmetries in order to describe the system's dynamics. There are different types of symmetries in the SM, but all have in common that a symmetry is associated with a conserved quantity (Noether-Theorem [10]). Examples known from classical dynamics are energy, momentum and angular momentum conservation, which are connected with time, translation and rotation invariance, respectively. These are *global* symmetries, i.e. the symmetry holds at all points in space-time. The mathematic descriptions of the electromagnetic, weak and strong interactions are based on *local* symmetries, meaning that the transformation is dependent on the space-time coordinates (*gauge theory*).

The following sections give an overview over the mathematical description of the SM, without claim of completeness. Detailed introductions can be found in Refs. [9, 11, 12], which are also the basis of the following sections.

[2] Only the left-handed component of quarks and leptons (and right-handed component of the corresponding antiparticle) is weakly charged. See also Sec. 2.2.4.

Table 2.1: Summary of elementary particle properties [8]. To every fermion, there is also an antifermion, with the same mass, but opposite charge. Bosons are in general all integer spin particles, while fermions have half-integer spin.

		name	symbol	mass	em charge	weak charge$^{\alpha}$	colour charge	spin
Fermions	Quarks up-type	up	u	2.3 MeV	$2/3$	$1/2$	✓	$1/2$
		charm	c	1275 MeV	$2/3$	$1/2$	✓	$1/2$
		top	t	173 GeV$^{\beta}$	$2/3$	$1/2$	✓	$1/2$
	Quarks down-type	down	d	4.8 MeV	$-1/3$	$-1/2$	✓	$1/2$
		strange	s	95 MeV	$-1/3$	$-1/2$	✓	$1/2$
		bottom	b	4 GeV	$-1/3$	$-1/2$	✓	$1/2$
	Leptons neutral	e neutrino	ν_e	< 2 eV	0	$1/2$	✗	$1/2$
		μ neutrino	ν_μ	< 0.19 MeV	0	$1/2$	✗	$1/2$
		τ neutrino	ν_τ	< 18.2 MeV	0	$1/2$	✗	$1/2$
	Leptons charged	electron	e	0.511 MeV	-1	$-1/2$	✗	$1/2$
		muon	μ	106 MeV	-1	$-1/2$	✗	$1/2$
		tau lepton	τ	1777 MeV	-1	$-1/2$	✗	$1/2$
Bosons		photon	γ	0^{γ}	0^{γ}	0	✗	1
		Z boson	Z	91.2 GeV	0	0	✗	1
		W boson	W	80.4 GeV	± 1	$\pm 1/2$	✗	1
		gluon	g	0^{δ}	0	0	✓	1
		Higgs boson	H	≈ 125 GeV$^{\epsilon}$	0	$-1/2$	✗	0

$^{\alpha}$Third component T^3 of the weak isospin. The value applies to left-handed fermions only, all right-handed fermions have $T^3 = 0$.

$^{\beta}$Direct measurement

$^{\gamma}$Upper measurement bounds are $m < 10^{-18}$ eV, $q < 10^{-35}$ e.

$^{\delta}$Theoretical value. A mass of a few MeV can not be excluded.

$^{\epsilon}$See Sec. 6.1.3 for the latest measurement

2.2.1. Dynamics of free massive spin-1/2 particles

Historically, the first equation to fully describe fermions was the Dirac equation. It is based on both quantum mechanics and special relativity and describes a free spin-1/2 particle with mass m, momentum $\vec{p}$ and energy $E^2 = \vec{p}^2 + m^2$:[3]

$$(i\gamma^\mu \partial_\mu - m)\psi = 0 \ , \tag{2.1}$$

where ψ is the 4-component Dirac-spinor (the particle's wave function) and γ^μ the Dirac matrices. The equation has four solutions, each describing a possible particle state. Special implications are:

- The particle has *spin*, a quantum number.
- If the particle carries electromagnetic charge, the spin is connected with a magnetic dipole-moment, which is twice the strength of a magnetic moment induced by a rotating mass.

[3]Throughout the thesis, natural units are used, $\hbar = c = 1$. Double appearing indices imply summation $\gamma^\mu \partial_\mu \equiv \sum_{\mu=0}^{3} \gamma^\mu \frac{\partial}{\partial x^\mu}$.

- Each particle has an antiparticle, with the same mass and spin, but opposite charge and magnetic dipole-moment [4].

The exceptionality of the equation was and is the provision of a theory and common basis for several unexplained observations at that time: the high magnetic moment of electrons, details in the hydrogen spectrum and Pauli's theory of spin. In addition, it also predicted the antielectron (positron).

2.2.2. Quantum electrodynamics

The description of the electromagnetic interaction is based on the U(1) symmetry group. The Lagrangian density (from now on simply called *Lagrangian*) is constructed to be locally gauge invariant:

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.2}$$

$$= \underbrace{\bar{\psi}i\gamma^\mu\partial_\mu\psi}_{\substack{\text{fermion}\\\text{kinematics}}} - \underbrace{\bar{\psi}m\psi}_{\substack{\text{fermion}\\\text{mass term}}} + \underbrace{\bar{\psi}e\gamma^\mu A_\mu\psi}_{\substack{\text{electromagnetic}\\\text{interaction}}} - \underbrace{\frac{1}{4}F_{\mu\nu}F^{\mu\nu}}_{\substack{\text{photon}\\\text{kinematics}}} . \tag{2.3}$$

The covariant derivative D_μ is defined as $D_\mu = \partial_\mu - ieA_\mu$, where e is the electric charge. The vector field A_μ represents the photon and transforms in coordination with the Dirac spinor ψ:

$$\begin{aligned} \psi &\rightarrow e^{i\theta(x)}\psi \ , \\ A_\mu &\rightarrow A_\mu + \frac{1}{e}\partial_\mu\theta(x) \ . \end{aligned}$$

The electromagnetic field tensor is $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. As labelled in Eq. 2.3, the different terms describe different properties of the system: the kinematics of both the charged fermion and the photon, the fermion mass and the electromagnetic interaction.

2.2.3. Non-Abelian gauge theories

The weak and strong interactions are described by higher dimension symmetry groups SU(n) with $n = 2$ and $n = 3$ for the weak and strong interactions, respectively. The Dirac fields are then n-plets (douplets or triplets) and transform as:[5]

$$\psi \rightarrow e^{i\theta_a(x)T_a}\psi \ . \tag{2.4}$$

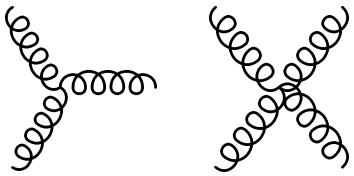
The T_a matrices are the n^2-1 linearly independent generators of the symmetry group. The conventional choice in case of SU(2) are the three Pauli matrices, $T_a = 1/2\sigma_a$, and the eight Gell-Mann matrices in the case of SU(3), $T_a = 1/2\lambda_a$. Per generator, a gauge field G^a_μ is introduced, transforming at first order as:

$$G^a_\mu \rightarrow G^a_\mu - \frac{1}{g}\partial_\mu\theta_a - f_{abc}\theta_b G^c_\mu \ . \tag{2.5}$$

The additional last term is introduced to achive local gauge invariance despite non-communiting generators, $[T_a, T_b] = if_{abc}T_c$. Symmetry groups with this behaviour are called *non-Abelian*, in contrast to *Abelian* groups, where all f_{abc} vanish. The U(1) symmetry group is an example of the latter.

[4] Based on his equation, Dirac predicted the existence of the antielectron state, even though his explanation was not well accepted. Feynman and Stueckelberg finally interpreted the negative energy solutions as antiparticles with positive energy [11].

[5] Summation over the double appearing index a is again implied, $a = 1, ..., (n^2 - 1)$.



(a) (b)

Figure 2.1: Gluon self-interaction vertices

The real-valued constants f_{abc} are called the structure constants of the group[6]. With the coupling constant g and the covariant derivative $D_\mu = \partial_\mu + igT_aG^a_\mu$, the Lagrangian takes the form

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}G^a_{\mu\nu}G^{a,\mu\nu} \tag{2.6}$$

$$= \underbrace{\bar{\psi}i\gamma^\mu\partial_\mu\psi}_{\substack{\text{fermion}\\\text{kinematics}}} - \underbrace{\bar{\psi}m\psi}_{\substack{\text{fermion}\\\text{mass term}}} \underbrace{-\bar{\psi}g\gamma^\mu T_aG^a_\mu\psi}_{\text{interaction}} \underbrace{-\frac{1}{4}G^a_{\mu\nu}G^{a,\mu\nu}}_{\substack{\text{boson}\\\text{kinematics}}} . \tag{2.7}$$

The field strength tensor is defined as $G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - gf_{abc}G^b_\mu G^c_\nu$. The last term in Eq. 2.7 therefore contains additional terms describing self-interactions:

$$-\frac{1}{4}G^a_{\mu\nu}G^{a,\mu\nu} = -\frac{1}{4}(\partial_\mu G^a_\nu - \partial_\nu G^a_\mu)(\partial^\mu G^{a,\nu} - \partial^\nu G^{a,\mu}) \quad \text{kinematic term} \tag{2.8}$$

$$-\frac{1}{2}gf_{abc}(\partial_\mu G^a_\nu - \partial_\nu G^a_\mu)G^{b,\mu}G^{c,\nu} \quad \text{3-point self interaction} \tag{2.9}$$

$$-\frac{1}{4}g^2f_{abe}f_{cde}G^a_\mu G^b_\nu G^{c,\mu}G^{d,\nu} \quad \text{4-point self interaction} \tag{2.10}$$

The strong interaction is fully described by the SU(3) Lagrangian. The gauge fields represent eight massless gluons, each carrying colour and anticolour. The gluons self-interact, as shown in Fig. 2.1.

The weak interaction is a more complicated case, which is discussed in the followings section.

2.2.4. Glashow-Weinberg-Salam theory

The weak interaction cannot be described by a simple SU(2) Lagrangian, due to the following peculiarities:

- The observed vector bosons of the weak interaction, W and Z bosons, are massive. It is however not possible to add a mass term to the Lagrangian in Eq. 2.7 without violating local gauge invariance.
- The charged weak interaction mediates only between left-handed[7] fermions and right-handed antifermions.
- The coupling strength of the neutral weak interaction is dependent on the electromagnetic charge and chirality of the fermion.

[6]For SU(2), f_{abc} is the Levi-Civita-Symbol ϵ_{abc}; For the structure constants of SU(3), see section 44, "SU(3) isoscalar factors and representation matrices" in Ref. [8].

[7]Handedness in terms of chirality. For massless particles, chirality is identical to helicity, the spin-projection onto the direction of motion. A spin parallel to the momentum gives positive helicity (right-handed), while anti-parallel spin gives negative helicity (left-handed).

- Besides charged current interactions of quarks within the same generation ($ud/cs/tb$) inter-generation interactions occur ($us/ub/cd/cb/td/ts$).

In order to generate masses for the W and Z bosons without breaking the gauge symmetry, a new field, the scalar Higgs field ϕ [13, 14], is introduced into the Lagrangian. The symmetry is then *spontaneously broken* and the weak bosons acquire mass through the BEH-mechanism[8]. The GWS theory, named after Glashow [16], Weinberg [17] and Salam [18], describes the experimental observations.

The electromagnetic and weak interaction are unified in a common $\mathrm{SU(2)_L}\times\ \mathrm{U(1)_Y}$ symmetry group, with the *weak hypercharge* Y as the generator of the $\mathrm{U(1)_Y}$ symmetry group and the coupling constants g and g'. Only left-handed components transform under the $\mathrm{SU(2)_L}$, which is denoted by the subscript 'L'. The fermion fields are split into their right-handed and left-handed components by applying projection operators to the wavefunction:

$$\psi_L = \frac{1}{2}(1-\gamma^5)\psi \ , \qquad \psi_R = \frac{1}{2}(1+\gamma^5)\psi \ . \tag{2.11}$$

The left-handed fields are arranged as weak iso-spin doublets χ, while the right-handed fields remain singlets under the SU(2) symmetry[9]:

$$\chi_L = \begin{pmatrix}\nu_e\\ e^-\end{pmatrix}_L, \begin{pmatrix}u\\ d'\end{pmatrix}_L, \dots \ , \qquad \psi_R = e_R, u_R, d_R, \dots \ . \tag{2.12}$$

The transformation for the fermion fields is then as follows:

$$\chi_L \to e^{i\theta_a T_a + i\eta Y}\chi_L \ , \qquad \psi_R \to e^{i\eta Y}\psi_R \ . \tag{2.13}$$

In Eq. 2.12, the down-type mass eigenstates are rotated into flavour eigenstates of the weak interaction $d'_i = V_{ij}d_j$, with V_{ij} being the Cabibbo-Kobayashi-Maskawa matrix. This allows for charged weak current interactions between quarks of different generations. An example is the decay of the lambda baryon, $\Lambda^0(uds) \to p(uud) + \pi^-(ud)$, where the strange quark transforms into an up quark under the transmission of a (virtual) W boson.
The unbroken Lagrangian is gauge invariant, with massless bosons and fermions:[10]

$$\begin{aligned}
\mathcal{L} = {} & \bar{\chi}_L\gamma^\mu(i\partial_\mu - gT_a\mathcal{W}^a_\mu - g'\frac{1}{2}Y\mathcal{B}_\mu)\chi_L + \bar{\psi}_R\gamma^\mu(i\partial_\mu - g'\frac{1}{2}Y\mathcal{B}_\mu)\psi_R && \Big\} \ \mathcal{L}_f \\
& - \frac{1}{4}\mathcal{B}_{\mu\nu}\mathcal{B}^{\mu\nu} - \frac{1}{4}\mathcal{W}^a_{\mu\nu}\mathcal{W}^{a,\mu\nu} && \Big\} \ \mathcal{L}_g \\
& + \mid (i\partial_\mu - gT_a\mathcal{W}^a_\mu - g'\frac{1}{2}Y\mathcal{B}_\mu)\phi \mid^2 - V(\phi) && \Big\} \ \mathcal{L}_h \\
& - (G_1\bar{\chi}_L\phi\psi_R + G_2\bar{\chi}_L\phi_c\psi_R + \text{hermitian conjugate}) && \Big\} \ \mathcal{L}_y
\end{aligned} \tag{2.14}$$

The gauge fields $\mathcal{B}$ and $\mathcal{W}^a$, $a = 1, 2, 3$, are a massless isospin singlet and a massless isospin triplet, respectively. The last two lines $\mathcal{L}_h$ and $\mathcal{L}_y$ are connected to the new scalar Higgs field ϕ. It is a SU(2) doublet with two complex scalar fields, hypercharge $Y = 1$ and a potential $V(\phi)$:

$$\phi = \begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}\phi_1 + i\phi_2\\ \phi_3 + i\phi_4\end{pmatrix}, \quad V(\phi) = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2 \ . \tag{2.15}$$

[8] The mechanism was formulated in parallel by several physicists, specially Brout and Englert [13], Higgs [14] and Guralnik, Hagen, and Kibble [15]. The name relates to Brout, Englert and Higgs.

[9] The right-handed neutrinos ν_R are left out. Because neutrinos are considered massless in the SM, the ν_R cannot interact and are hence not accessible.

[10] $\mid \cdot \mid^2$ is used as a shortcut for $(\cdot)^\dagger(\cdot)$.

The superscripts '+' and '0' indicate the electromagnetic charge. With $\mu^2 < 0$ and $\lambda > 0$, the potential has a local maximum at $\phi = 0$. The minima lie on a sphere, satisfying

$$\phi^\dagger\phi = \frac{1}{2}(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2) = -\frac{\mu^2}{2\lambda} \quad . \tag{2.16}$$

The potential as such is invariant under an SU(2) symmetry. However, choosing one specific minimum breaks the symmetry. In order to rebuild the observed standard model, a convenient choice is $\phi_1 = \phi_2 = \phi_4 = 0$ and $\phi_3 = \sqrt{-\mu^2/\lambda} = v$ (see also Fig. 2.2). Due to gauge invariance [11], expanding ϕ around this particular ground state (*vacuum*) simplifies to a substitution of

$$\phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \tag{2.17}$$

into the Lagrangian Eq. 2.14. From the originally four fields, only one field is left, which represents the physical Higgs boson, $h(x)$. The other three now generate masses for three vector fields, the experimentally observed vector bosons of the weak interaction:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(\mathcal{W}_\mu^1 \mp \mathcal{W}_\mu^2) \qquad \text{with mass } m_\mathrm{W} = \frac{1}{2}v\,g \quad , \tag{2.18}$$

$$Z_\mu^0 = \frac{1}{\sqrt{g^2 + g'^2}}(g'\mathcal{W}_\mu^3 + g\mathcal{B}_\mu) \qquad \text{with mass } m_\mathrm{Z} = \frac{1}{2}v\sqrt{g^2 + g'^2} \quad . \tag{2.19}$$

One vector fields remains massless and can be identified with the photon field A_μ. Defining the weak mixing angle $\tan\theta_w = g'/g$, the electrically neutral vector bosons appear as rotations of the fields $\mathcal{W}_\mu^3$ and $\mathcal{B}_\mu$:

$$\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_w & -\sin\theta_w \\ \sin\theta_w & \cos\theta_w \end{pmatrix}\begin{pmatrix} \mathcal{W}_\mu^3 \\ \mathcal{B}_\mu \end{pmatrix} \quad . \tag{2.20}$$

The electric charge e and the charge operator Q can be identified as:

$$e = \frac{gg'}{\sqrt{g^2 + g'^2}} \quad , \quad Q = T^3 + \frac{Y}{2} \quad . \tag{2.21}$$

The new scalar boson appearing in the Lagrangian, the Higgs boson h, is itself massive, even though the value as such is not predicted. The Higgs boson couples to the weak vector bosons and to itself. The relevant 3-point and 4-point interaction terms within $\mathcal{L}_h$ in Eq. 2.14 are

$$\mathcal{L}_{h,int} = \delta_V\frac{2m_V^2}{v}V_\mu V^\mu h + \delta_V\frac{m_V^2}{v^2}V_\mu V^\mu hh + \frac{m_H^2}{2v}hhh + \frac{m_H^2}{8v^2}hhhh \quad , \tag{2.22}$$

with $V = W^\pm$ or Z. It is apparent that the coupling strength to bosons is proportional to the squared boson mass. The factor δ_V is due to the different electric charges of the vector bosons, $\delta_W = 1$, $\delta_Z = 1/2$. It appears when the boson masses are identified within the Lagrangian, expecting a term of the form $m^2W^+W^-$ for charged bosons and $(1/2)m^2ZZ$ for neutral bosons.

The fermion masses are generated via the same scalar Higgs field. In the unbroken state, the fermions are massless and interact with the Higgs field through so-called Yukawa-couplings, $\mathcal{L}_y$ in Eq. 2.14. The couplings G_i are arbitrary and independent for the different fermions. After the spontaneous

[11] The expansion is performed in terms of four real fields. Due to local gauge invariance, three fields can be gauged away and do not appear in the final Lagrangian; Only $h(x)$ remains and thereby Eq. 2.17. In more pituresque words: by choosing a specific gauge, the three massless Goldstone bosons are 'eaten' by the gauge fields and give them mass.

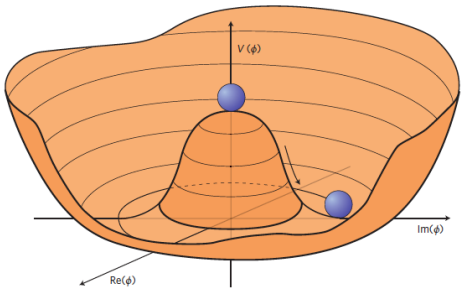


Figure 2.2: The potential $V(\phi)$ of two complex fields cannot easily be visualised. The illustration shows the potential for a field with one complex component only. This corresponds to choosing $\phi_1 = 0$ and $\phi_2 = 0$, so that $\mathrm{Re}(\phi) \equiv \phi_3/\sqrt{2}$ and $\mathrm{Im}(\phi) \equiv \phi_4/\sqrt{2}$. The local maximum is at $\phi_3 = 0$ and $\phi_4 = 0$, and the minima lie on a circle with radius $\phi_3^2 + \phi_4^2 = -\mu^2/\lambda$. Choosing any of the minima breaks the symmetry [19].

symmetry breaking, mass terms for the fermions appear. The first term in $\mathcal{L}_y$ then generates masses for the charged leptons and down-type quarks. The field ϕ_c in the second term is needed for the up-type quark masses[12]. It transform in the same manner as ϕ and is related to it as follows:

$$\phi_c = -i\sigma_2\phi^* = \begin{pmatrix} -\bar{\phi}^0 \\ \phi^- \end{pmatrix} \xrightarrow[breaking]{after} \frac{1}{\sqrt{2}} \begin{pmatrix} v + h(x) \\ 0 \end{pmatrix} . \tag{2.23}$$

Two terms appear for each massive fermion: a mass term and a coupling term to the physical Higgs boson,

$$\mathcal{L}_y = -m\bar{\psi}\psi - \frac{m}{v}\bar{\psi}\psi h . \tag{2.24}$$

In contrast to the bosonic interactions, the interaction with the Higgs boson is linearly dependent on the fermion mass.
Summarising, after spontaneous symmetry breaking the four parts of the Lagrangian describe

$\mathcal{L}_f$: the fermion kinematics and their interactions with the vector bosons γ, Z, $W^\pm$,

$\mathcal{L}_g$: the kinematics and self-interactions of the vector bosons,

$\mathcal{L}_h$: the Higgs boson kinematics and self-interactions, the masses of the heavy vector bosons and their interactions with the Higgs boson,

$\mathcal{L}_y$: the masses of the fermions and their interaction with the Higgs boson.

2.3. Shortcomings of the Standard Model

The Standard Model has been very successful in describing observed phenomena and measurements, but also in the prediction of unobserved particles. The heavy quarks, the W and Z bosons as well as the Higgs boson were predicted before experiments existed to directly measure the particles. Nevertheless, the Standard Model in its current state can not answer all questions. The most important open points are:

[12] The neutrinos are assumed to be massless and are therefore not considered here.

Dark energy and dark matter: Cosmological observations reveal that the largest fraction of the universe's energy and matter content is not described by the SM. The extra matter content, the *dark matter* [20], could be explained by additional particles. Such candidates appear in (hypothetic) extensions of the SM, i.e. supersymmetry [21], but there has been no experimental evidence so far. There are no experimentally verifiable theories to describe *dark energy* [22].

Matter-Antimatter asymmetry: The universe is matter dominated and no cluster of antimatter has been observed so far. Assuming that matter and antimatter existed in equal parts after the Big Bang, a mechanism is needed to favour matter production later on. While these mechanisms exist in the SM (e.g. CP-violation in kaon decays), the currently known processes are not large enough to explain the measured matter-antimatter asymmetry [23].

Neutrino masses: Neutrinos are assumed massless in the SM Lagrangian, even though measurements of neutrino oscillations have proven neutrinos to be massive. Experimental observations indicate neutrino mass eigenstates with masses not exceeding $\approx$1 eV. There are various models on how to extend the SM to include neutrino masses, but experiments do not provide evidence of which model to favour [24].

Fine-tuning or hierarchy problem: The Higgs mass calculation is driven by higher order loop corrections, which diverge quadratically. It is mathematically possible to fine-tune the calculation to cancel the large corrections to the relatively small Higgs boson mass of 125 GeV, but it remains an unsatisfactory feature. In the past, divergences turned out to be an indication of undiscovered particles, i.e. in the calculation of the electron mass without the positron, or the WW boson scattering amplitude without the Higgs boson. Therefore, the hierarchy problem is used as an argument for an additional, undetected group of particles at higher masses, for instance supersymmetric partners to all SM particles [25, 26].

Gravity: Gravity is not part of the SM and, so far, attempts failed to construct a quantum field theory of gravity. However, on subatomic scale gravity is much weaker than the other forces and has no measurable effect [27].

2.4. Particle physics at hadron colliders

The fundamental particles, their interactions and bound states are studied at the LHC in proton-proton collisions. In order to interpret the data, the recorded events are compared to simulations. Going from the Lagrangians to an actual prediction of differential cross sections is however complicated. In fact, exact analytic solutions do not exist. The scattering amplitude is obtained as a perturbation series in the coupling constant. As long as the latter is small, higher order terms can be neglected. In the SM[13], the coupling constants are *running*, meaning that they depend on the momentum scale they are studied at. In QED, the coupling grows with the energy scale, but is still weak at the Z boson mass scale and perturbation theory can be used. In QCD, the behaviour is reversed, with a decreasing coupling for growing energy scale. This is the earlier mentioned asymptotic freedom. For the collisions at the LHC, this is practical in the sense that hard collision processes can be calculated with perturbation theory even to very high momentum transfers. Problematic in terms of calculations is the soft regime, in which the higher order terms become exceedingly important and cannot be neglected anymore

[13] To be more precise, the effect is connected to renormalisation group theory.

(*non-perturbative regime*). The description and simulation of these processes is therefore based on measurements or phenomenological models.

As mentioned in Sec. 2.1, protons are composite particles made of quarks. In scattering processes with very low momentum transfer, the constituents are not visible and the proton behaves like a solid entity. With increased momentum transfer, the substructure becomes visible. The three constituent quarks, *uud* (*valence quarks*), are bound inside the proton by a colour field. The mediating gluons can split into more gluons and into quark-antiquark pairs, so that the valence quarks are embedded in a 'sea' of gluons, quarks and antiquarks (*sea quarks*). In a collision with high momentum transfer, any of the particles inside the proton can take part in the interaction. The *parton distribution functions* (PDFs) $f_i(x, Q^2)$ describe the probability[14] of finding a parton i, carrying the momentum fraction x in the proton. The PDFs are depending on the momentum transfer in the interaction, Q^2, as depicted in Fig. 2.3. The more energetic the interaction is, the more partons become visible. The PDFs fall into the non-perturbative regime and are obtained from measurements [28]. The fact that the partons carry only a fraction of the proton momentum has both advantages and disadvantages. In particular, it means that the initial state of the collision is not fully known. The two colliding partons carry (usually) not the same momentum fraction, so that the collision products will be boosted along the beam direction. Concerning the reconstruction of the event, it follows that the event is per se unbalanced in the longitudinal direction. The presence of particles that escape detection cannot be inferred. The situation is different in the transverse plane. The colliding partons carry no (or vanishing) transverse momentum, so that the event is per se balanced in the transverse plane. For the reconstruction of an event, it follows that the transverse momenta of all final state particles have to add up to zero. If this is not the case, it means that some final state particles escaped detection. Even though it is not possible to determine the number of escaped particles, the measurement of the direction and magnitude of this *missing transverse energy* is crucial for the understanding of the interaction. The upside of the varying collision energies is that a large energy spectrum can be scanned without changing the beam energy. This makes the LHC a perfect machine for the search (and discovery) of particles of unknown mass.

Pure colour processes dominate the interactions at the LHC. The quarks or gluons fly apart after the interaction, but they remain connected via the colour field. As described, this leads to the production of additional quarks and antiquarks, which eventually reconnect to form a jet of hadrons. The momentum direction of the primary quarks or gluons is preserved, so that two scattered coloured particles appear as two jets in the detector. If high-energy gluons are emitted, additional jets can be formed. These processes are collectively called *multi-jet* events. The dominant parton in the jet formation process influences the jet shape. The probability of a gluon to emit another gluon is higher than the probability of a quark emitting a gluon. Therefore, gluon initiated jets have a higher hadron multiplicity and are in general wider than quark initiated jets.

For the event simulation, the collision is factorised into several aspects (Fig. 2.4) [30]. The *hard process* is typically the starting point around which the remaining activity is build. The PDFs have to be taken into account in order to get the correct balance of quarks and gluons over many initial states. The PDFs also determine the momentum distribution of the initial state partons. Both the initial and final state particles radiate off photons or gluons. Since this *initial* and *final state radiation* (ISR, FSR) is independent of the hard process, it is simulated in an extra step, rather than treating it as a higher order contribution to the hard process. Due to their self-interacting nature, radiated gluons will successively radiate additional gluons, thus building up a *parton shower*. All coloured objects then have to recombine into hadrons. This *hadronisation* process can not be calculated within

[14] The PDFs are no probability density functions (pdfs) in the common sense. The normalisation is chosen to reflect the quantum numbers of the proton: $\int_0^1 f_u(x) - f_{\bar{u}}(x)\,\mathrm{d}x = 2$, $\int_0^1 f_d(x) - f_{\bar{d}}(x)\,\mathrm{d}x = 1$, $\int_0^1 f_s(x) - f_{\bar{s}}(x)\,\mathrm{d}x = 0$. Summation over all fractional momenta xp gives the total proton momentum p: $\int_0^1 xp\left(\sum_q f_{q_i}(x) + \sum_{\bar{q}} f_{\bar{q}_i}(x) + f_g\right)\mathrm{d}x = p$.

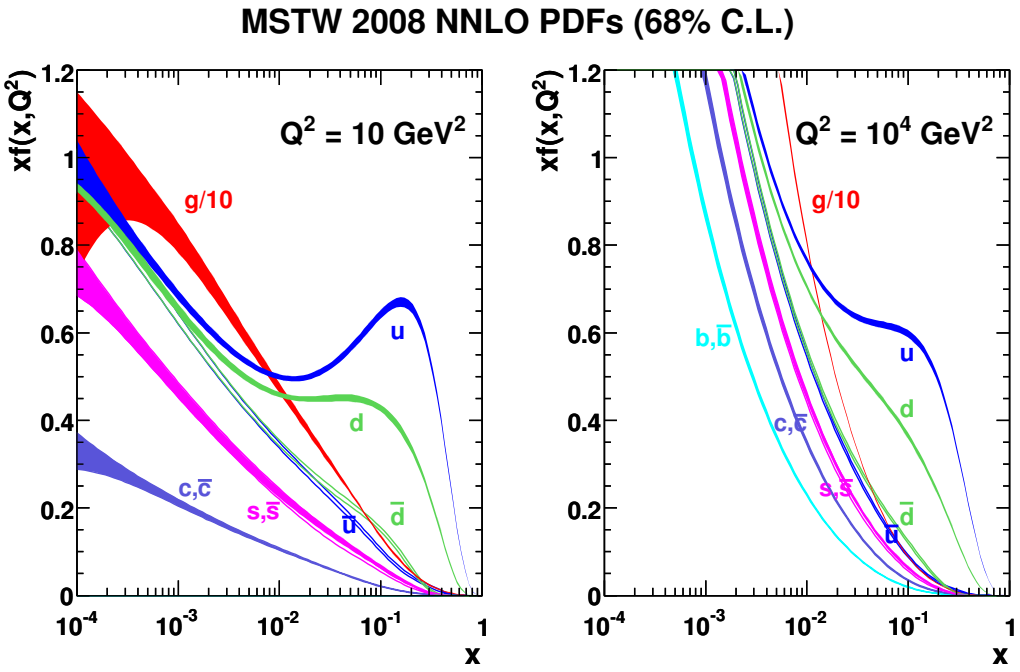


Figure 2.3: Parton distribution functions xf_i for quarks, antiquarks and gluons in a proton for two different energy scales Q^2. The gluon function is scaled by a factor 0.1 [28, 29].

perturbation theory and phenomenological models are used. Finally, the decay of all unstable particles is simulated. Besides the hard process, there is additional activity from the proton-remnants, the *underlying event*.

The different aspects of the simulation have been implemented into generators with various models and methods. For instance, the multi-purpose generator HERWIG++ [31] uses angular ordered parton showering and the cluster hadronisation model, while PYTHIA [32] orders the shower in transverse momentum and uses the Lund string model for hadronisation. Other generators focus on the simulation of one specific aspect and can be combined with other programs, i.e. PHOTOS [33], which is specifically for the modelling of QED radiative corrections. All implementations come with their pros and cons, and the best choice depends on the analysis. Because some of the event generation steps rely on random sampling techniques, simulated samples are commonly called *Monte Carlo* (MC) samples.

After the collision process, the particles traverse the detector, which is simulated by the GEANT4 framework [34]. The entire detector is implemented in the software package, for instance the geometry, the different materials, all cables and fields. GEANT4 simulates the particles passage through the detector, including electromagnetic and hadronic interactions leading to secondary particle production. The program's output are *hits*, which are energy deposits in the sensitive regions of the detector. ATLAS software then digitalises the hits, i.e. translates them into a detector signal of times and voltages, also adding electronic noise. From this point on, simulation and data can be treated alike.

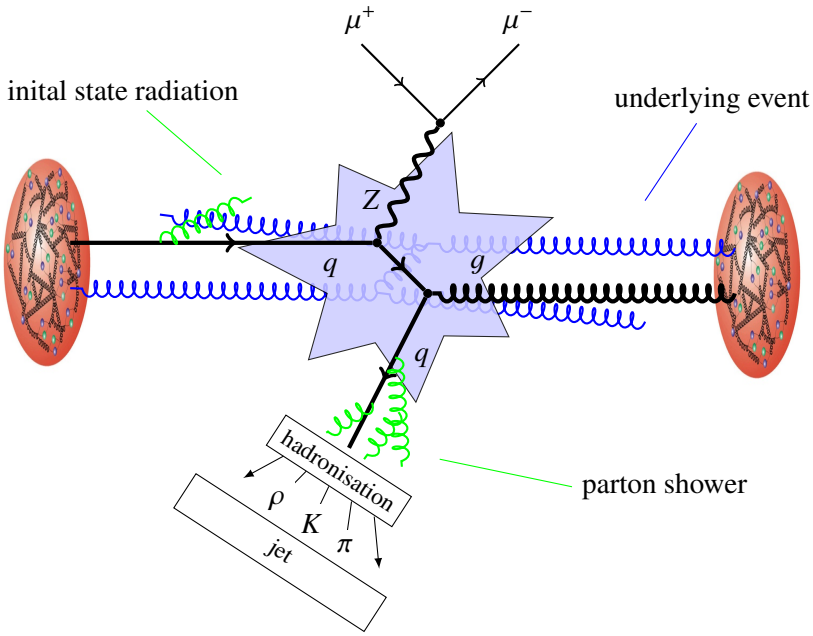


Figure 2.4: Sketch of a Z + $textrm jet$ event in a pp collision. The hard collision $qg \rightarrow Zq$, with $Z \rightarrow \mu\mu$, is visible as two opposite sign muons and one jet. The ISR, the FSR and the underlying event will contribute to the soft activity in the event. The incoming protons appear as flat disc due to the relativistic Lorentz-contraction, indicated by the slightly elliptic form of the protons.

3. LHC and the ATLAS experiment

This chapter introduces the experimental setup. The first to sections focus on the acceleration of particles, starting with a short excursus on the history of particle acceleration (Sec. 3.1), and ending with an overview of the LHC accelerator complex (Sec.3.2). The ATLAS detector with all its subsystems is described in Sec. 3.3.

3.1. Accelerating particles

CERN[1] was founded 60 years ago as a European laboratory for nuclear physics, and is located near Geneva at the Swiss-French border. Nowadays, the main focus is on particle physics, whilst maintaining a broad research program ranging from climate physics and accelerator development to antimatter production [35].

Fundamental particles and hadrons can be studied in various ways. In the early days of particle physics, natural sources such as radioactive decays and cosmic rays were used. Cosmic rays, mainly protons, passing through the earth's atmosphere, interact with it and initiate a particle shower. Intense studies of particle tracks lead to the discovery of the positron, the muon and mesons [36–40]. Measurements of the energy spectrum of electrons emitted in radioactive β-decays led to the discovery of the neutrino [41–43].

A more controlled way to produce particles is by artificially colliding protons or electrons[2]. The first to use a particle accelerator for a physics experiment were Cockcroft and Walton [44]. In 1932, they generated a high DC voltage field to accelerate protons to an energy of $\approx 0.5\,\text{MeV}$ and observed the disintegration of a lithium atom [45–47]. Before long, higher energies were reached by using alternating fields and recurring acceleration. The linear Wideroe-accelerator [48] was the first one to use alternating fields. By exposing the particle many times to the field, high acceleration is gained with a small potential differences. Modern linear accelerators are still based on this principle, using standing or travelling waves and radio-frequency cavities. By bending the particle path into a circular orbit, it was possible to traverse the same acceleration unit many times. This was first achieved with the Cyclotron [49], which used a constant magnetic field to bend the particle path and a alternating electrical field to accelerate the particle. In order to reach higher energies and to leave the non-relativistic regime, it was necessary to split acceleration, bending and focusing units. In a synchrotron, the magnetic fields to bend the particles are time-dependent and synchronised with the acceleration units, so that the particles travel on the same path every circulation [50]. The Cosmotron [51], one of the first synchrotrons, reached the multi-GeV energy range, and in 1953 protons were boosted to 3.3 GeV for the first time.

Over the years, larger and larger accelerators were build and higher and higher energies were reached. The two competing designs for high energy particle physics are synchrotrons and linear accelerators (linacs). In a synchrotron, the acceleration units are traversed many times and high energies can be reached in a cost-efficient way. However, charged particles which are bent onto a curved

[1] CERN is the acronym of the French 'Conseil EuropeÌĄen pour la Recherche NucleÌĄaair'. The council founded the European Organisation for Nuclear Research, which then used CERN as a proper name.

[2] In principle, any particles can be collided; Due to practical and technical reasons, mainly (anti)protons, electrons or heavy ions have been used in the past.

path emit light (synchrotron radiation). The energy loss per circulation is dependent on the energy and mass of the particle and the bending radius, $\Delta E \propto 1/R(E/m)^4$. This limits the maximum feasible energy for a given radius, and is a practical problem for the acceleration of electrons. The largest ever build electron synchrotron was the Large Electron Positron Collider (LEP) at CERN, with a circumference of 27 km and a top energy of 104.5 GeV per beam. The same tunnel now hosts the Large Hadron Collider (LHC), in which protons are accelerated to 4 TeV in the year 2012. Due to the much smaller synchrotron radiation, even higher energies per beam are possible. The upper limit is here given by the magnetic field strength. The problem of high energy loss through synchrotron radiation is avoided in a linear accelerator. However, every acceleration unit is only traversed once, so that high energies can only be reached with very long devices. The longest linear accelerator, the Stanford Linear Accelerator, started operation in 1966 at SLAC[3] with a total length of 3 km [52]. After an upgrade of the acceleration units, the then named SLC [4] is also the most powerful linear device ever built. From 1989 to 1998, electrons and positrons were accelerated up to 50 GeV and then brought to collision [53, 54].

The trend to build larger accelerators in order to reach higher energies continues. Design reports are being written for a linear collider with a length of 31 km (ILC [55]) and a circular collider with a circumference of 80-100 km [56]. In parallel, new concepts are investigated to reach higher acceleration gradients. The Compact Linear Collider (CLIC) [57] is a design with two beam pipes. A high-current low-energy electron beam generates accelerating radio-frequency (RF) waves for the low-current high-energy main beam. The foreseen energy gain is 100 MV/m, which is three times higher than the gradient of the ILC, using current superconducting RF technology. The potential to gain even higher energy per distance has the acceleration with plasma wakefields. A laser or a highly relativistic bunch of charged particles traverses a plasma, separating the charged ions and electron of the plasma. This creates high electric fields, which can be used to accelerate electrons in a second, following bunch. Acceleration gradients of more than 10 GV/m have already been achieved [58].

3.2. The LHC

The Large Hadron Collider (LHC) [62] at CERN is a synchrotron, with a circumference of 27 km, located 50 m to 175 m below surface in the Geneva area. Both protons and heavy ions can be accelerated and collided. Before these are injected into the LHC, they traverse a series of pre-accelerators (Fig. 3.1). Older parts of the CERN accelerator complex and existing tunnels are used as much as possible in order to build the LHC and its pre-acceleration parts in a cost-efficient way. The oldest section is the Proton Synchrotron (PS), which is CERN's first synchrotron and started operation already in 1959. Since then, it has been upgraded many times and has functioned as a pre-accelerator to all kinds of experiments.

The proton acceleration chain starts with an ordinary gas bottle, containing hydrogen gas. Protons are obtained by stripping off the electrons in an electric field and are boosted to an energy of 50 MeV in the linear accelerator Linac 2. Firstly, they are passed on to the Proton Synchrotron Booster (PSB) and then, at an energy of 1.4 GeV, to the Proton Synchrotron (PS). In the PS the protons are accelerated to 25 GeV. The PS collects several bunches from the PSB and also splits them, in order to organise the protons in the desired bunch-train structure. The train is then passed to the Super Proton Synchrotron (SPS) and the PS if filled again. Three to four trains are collected in the SPS and then boosted to an energy of 450 GeV per proton. 13 SPS fillings are injected into the LHC, both in the clockwise and anti-clockwise direction. The protons are then accelerated to their final energy of 4 TeV in 2012. The

[3] Stanford Linear Acceleration Center, nowadays SLAC National Accelerator Laboratory
[4] SLAC Linear Collider

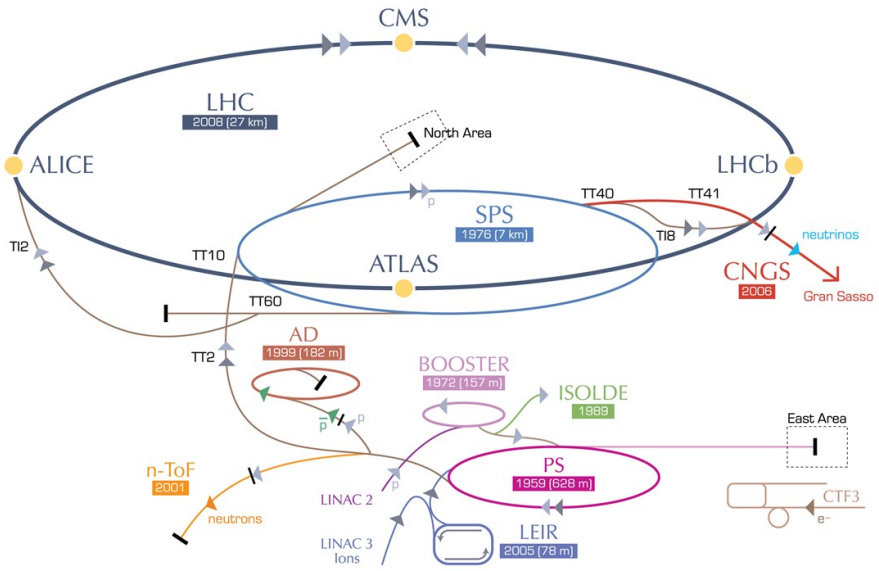


Figure 3.1: The CERN accelerator complex [59]. Protons (p) for the LHC are pre-accelerated in Linac 2, Booster, PS and SPS. Heavy ions for the LHC start in Linac 3, are passed on to Low Energy Ion Ring (LEIR) and follow from there on the same route as the protons. The neutron time-of-flight facility (n-TOF) studies nuclear neutron interactions. The neutrons are produced by directing protons from the PS onto a lead target. At the Isotope mass Separator On-Line facility (ISOLDE) low-energy beams of radioactive nuclei are studied, which are created by directing the proton beam for the BOOSTER onto special targets. The CERN Neutrinos to Gran Sasso (CNGS) project provides neutrinos for the Gran Sasso National Laboratory. Protons from the SPS are guided onto a graphite target, where hadrons are created whic, among other particles, subsequently decay to neutrinos. The Antiproton Decelerator (AD) supplies several experiments with a low-energy antiproton beam. The North and East Areas host a large variaty of experiements, which take protons or heavy ions from the SPS or PS. Finally, the Compact Linear Collider (CLIC) collaboration developing a new technology for the acceleration of (anti)electrons at the CLIC test facility (CTF3) [60, 61].

Table 3.1: Overview of accelerators, which are part of the LHC accelerator chain, with year of initial start-up, and final energy per proton beam during normal data taking periods in 2012 [61, 63].

accelerator	start-up year	length	final energy	speed in relation to speed of light
Linac 2	1978	30 m	50 MeV	31.4 %
PSB	1972	157 m	1.4 GeV	91.6 %
PS	1959	628 m	25 GeV	99.93 %
SPS	1976	6911 m	250 GeV	99.9998 %
LHC	2008	26657 m	4000 GeV	99.999997 %

bunches are circulated in the LHC over many hours and brought to collision at the different interaction points. Per bunch-crossing, only a few protons interact, so that the bunches can be re-collided several times. Due to collisions and beam dynamics, protons are lost with time and eventually the beam is dumped and the accelerator is refilled.

The LHC tunnel hosts six experiments which record the beam-beam collisions at four interaction points:

ATLAS is a multi-purpose detector[5], designed to be sensitive to all kind of physics programs, ranging from Standard Model measurements and precision test over flavour and heavy ion physics to searches for physics beyond the SM. A primary goal is the search for and study of the Higgs Boson [64].

CMS (Compact Muon Solenoid) is a multi-purpose detector, with equal goals as the ATLAS experiment, but using a different detector design and technology [65]. CMS and ATLAS are hence independent experiments, validating each others findings.

ALICE (A Large Ion Collider Experiment) is optimised for heavy ion collisions. The main goal is to study QCD at high temperature and energy densities, in order to explore the quark-gluon-plasma, a state where quarks are no longer defined into hadrons [66].

LHCb (Large Hadron Collider beauty) is designed to study heavy flavour physics, mainly CP-violation in rare decays of hadrons containing b- and c-quarks [67].

LHCf (Large Hadron Collider forward) is placed close to the ATLAS experiment to measure neutral hadrons in the very forward region of the collision. The goal is to better understand hadron interaction models for high-energy cosmic ray experiments [68].

TOTEM (TOTal cross section, Elastic scattering and diffraction dissociation Measurement at the LHC) is placed close to the CMS experiment. It's purpose is to study elastic and diffractive scattering and to measure the total proton-proton cross-section [69].

3.3. The ATLAS detector

The ATLAS detector was designed as a multi-purpose detector, able to record proton-proton collisions with final state particles in a wide range of momenta and heavy ion collisions with very high numbers of tracks. The detector has a cylindrical shape, is 25 m in height and 44 m in length, and weights

[5] ATLAS: formerly for A Toroidal LHC ApparatuS, now proper name.

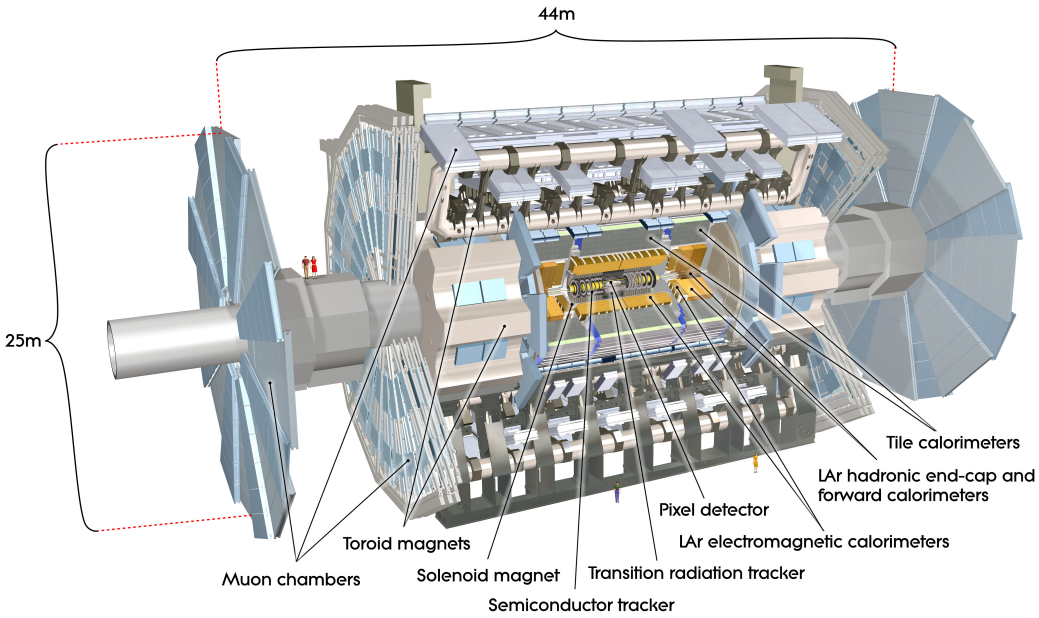


Figure 3.2: The ATLAS detector [64].

≈ 7000 tonnes. It is aligned along the beam line with the nominal interaction point (IP) in its centre (Fig. 3.2). The different detector components enclose the IP like different layers of an onion:

- Innermost is the inner detector (Sec. 3.3.3) for the reconstruction of charged particle trajectories.
- For the measurement of particle momenta, the inner detector is embedded in a magnetic field provided by a solenoid (Sec. 3.3.2).
- The solenoid is closely followed by the electromagnetic and hadronic calorimeters (Sec. 3.3.4), responsible to stop and measure the energy of most charged and neutral particles created in the collision.
- Outermost is the muon system (Sec. 3.3.5), for a precise measurement of muon tracks.
- The muon system is interleaved in a toroid magnet (Sec. 3.3.2).

The detector components have two types of arrangement: the *barrel* part, in which the detector modules are arranged as concentrical layers along the beam line, and the *end-cap* part, where the modules are installed on discs perpendicular to the beam. The detector is symmetric around the beam pipe. For an effective filtering of specific events, ATLAS is equipped with a hardware and software based trigger system (Sec 3.3.6). The experiment is completed by three smaller detector systems in the very forward region, LUCID, ZDC and ALFA. These are placed along the beam line at ± 17 m, ± 140 m and ± 240 m, respectively. LUCID and ALFA have the main purpose of measuring luminosity[6], while ZDC has mainly been used for detection of neutrons in heavy ion collisions.

In the following sections, the detector components are described in more detail. The descriptions are based on [64] if not quoted otherwise.

[6] See Sec. 4.1 for an explanation of luminosity.

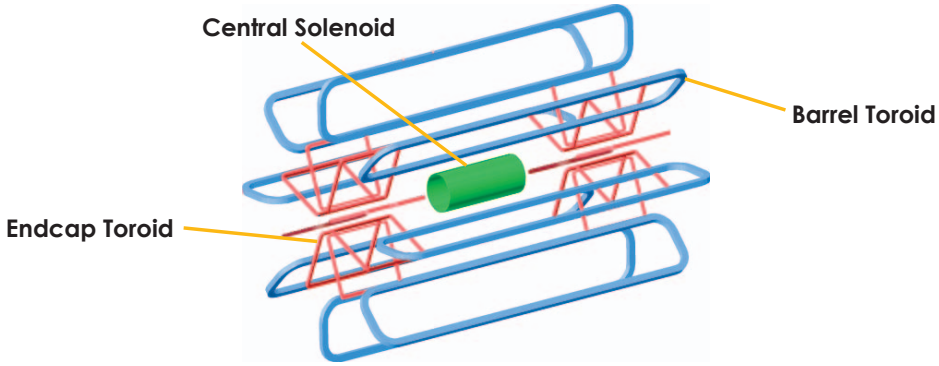


Figure 3.3: The geometry of the magnet system [70].

Table 3.2: Abstract of Magnet System properties [64].

	central solenoid	barrel toroid	end-cap toroids
axial length	5.8 m	25.3 m	5.0 m
diameter inner	2.46 m	9.40 m	1.65 m
outer	2.56 m	20.10 m	10.70 m
number of coils	1	8	2x8

3.3.1. Coordinate system

The origin of the ATLAS coordinate system is defined as the nominal interaction point. It is a right-handed Cartesian system, with the x-axis pointing towards the centre of the LHC ring, the y-axis pointing upwards and the z-axis along the beam line. The angle ϕ in the x-y plane is measured with respect to the x-axis and is positive for positive y-values, $\phi \in [-\pi, \pi]$. The polar angle θ is the angle with respect to the z-axis, $\theta \in [0, \pi]$. The pseudorapidity η is then defined as

$$\eta = -\ln \tan\left(\frac{\theta}{2}\right) \quad . \tag{3.1}$$

Transverse components, such as transverse momentum p_{T}, transverse energy E_{T} and missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ are defined in the x-y-plane. Distances between physics objects in the detector are calculated as

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \quad . \tag{3.2}$$

3.3.2. Magnet system

The ATLAS magnet system provides two types of field geometries: An axial and a toroidal field. Charged particles traversing the magnetic field $\vec{B}$ are influenced by the Lorentz force $\vec{F} = q_e \times \vec{v} \times \vec{B}$: The flight path is bent depending on the momentum $\vec{p} = \gamma m \vec{v}$ and the electromagnetic charge q_e of the particle[7]. Magnetic fields are hence an important component in detector design for a precise measurement of momentum and charge. The ATLAS magnet system is composed of four superconducting sub-systems, operated at 4.5 K (Fig. 3.3). The central solenoid provides a 2 T axial magnetic field for the inner detector and bends the particles trajectories in the R-ϕ plane. It is placed right in front of the calorimeter. In order to prevent large energy losses, the amount of material needs to be small. By

[7] Here, m is the mass of the particle and γ is the relativistic Lorentz factor, $\gamma = 1/\sqrt{1-v^2/c^2}$.

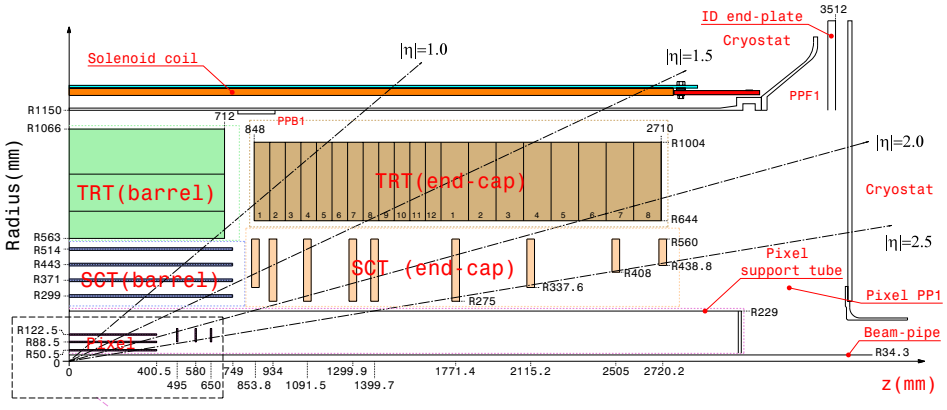
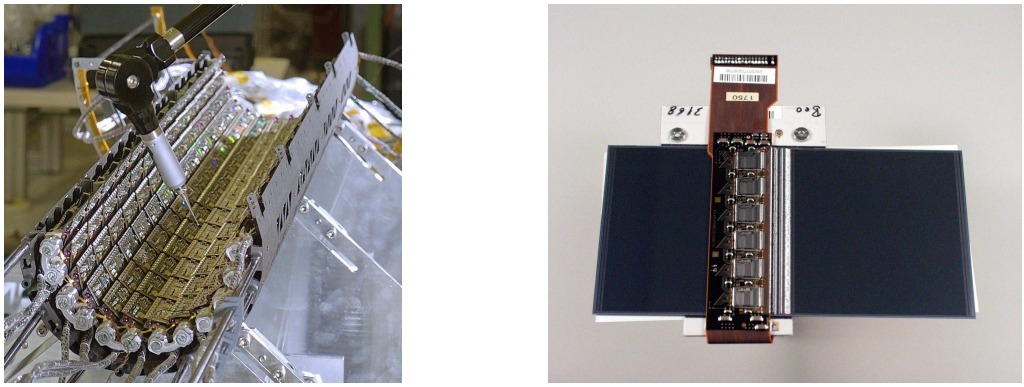


Figure 3.4: Cut-section view of the inner detector, modified [64].



(a) (b)

Figure 3.5: a) Photograph of a barrel half-shell of the pixel detector. b) Photograph of a barrel SCT module. The stereo angle of 40 mrad between the upper (black) and lower (white) SCT plates is clearly visible [64].

placing the solenoid in the same vacuum vessel as the electromagnetic LAr calorimeter, the material due to the magnet amounts to ≈0.66 radiation length[8] only. The steel of the calorimeter and the girder structures function as the return yoke. The toroidal field is provided by a huge toroid in the barrel region of the detector and a smaller one at each side. They provide a field of approximately 0.5 T and 1.0 T, respectively. The coils encompass the muon system and bend the muon trajectories in the R-z plane. The toroidal field allows for a very good momentum resolution for muons measured solely in the muon system. In order to achieve a good momentum resolution, the uncertainties from the magnet system itself need to be small. The B-field is therefore mapped constantly, using 4 NMR probes close to the interaction point and 1730 3D-Hall cards in the muon system. The latter use the Hall effect to measure the local magnetic field. Each card hosts three orthogonally mounted Hall probes to determine both value and direction of the field [71, 72].

3.3.3. Inner detector

The inner detector (ID) measures the path of charged particles. The momentum and charge sign can then be determined. A good track reconstruction is also important to identify primary and secondary

[8] See Sec. 3.3.4 for an explanation of radiation length.

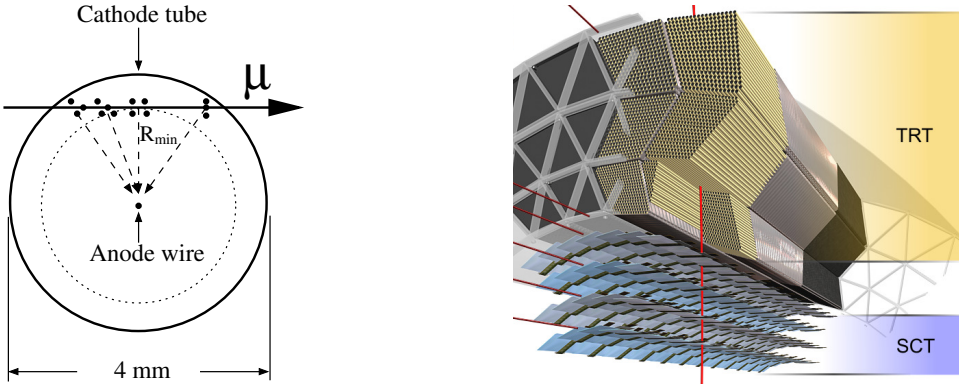


Figure 3.6: a) Schematic of a TRT drift tube. b) Schematic view of the barrel TRT module and the barrel SCT layers. Modified from [64].

vertices. The system is based on three concepts: semi-conductor detectors, drift tubes and transition radiation. Semi-conductor detectors based on silicon are used in the innermost layers of the ID. Charged particles traversing the semi-conductor produce electron-hole pairs. Pushed by an electrical field, these travel to either sides of the detector and are collected. Silicium trackers provide very good resolution and are small sized. Three layers of pixel detectors (PIX) occupy the first ≈12 cm in radius of the inner detector (Fig. 3.5(a)). Together with three discs on each side, the pixel detector provides very high resolution track points within $|\eta| < 2.5$. The semiconductor tracker (SCT) is based on the same technology, using a simpler design. Instead of pixels, strips are used to collect the signal. The strips are 6.2 cm long, with a pitch of 80 μm. A typical module is build of 4 sensors: Two sensors are connected to a 12.8 cm long unit. The two pairs are glued together back-to-back, with an angle of 40 mrad (Fig. 3.5(b)). This stereo angle is important for providing the measurement of the z-component, i.e. the coordinate along the strip direction. Both pixel detector and SCT are exposed to high radiation doses. In order to keep noise currents low, they are operated at -5°C to -10°C. Additionally, the electrical field is raised over time to ensure a good charge collection efficiency.

The transition radiation tracker (TRT) is composed of drift tubes interleaved with material causing transition radiation. While the first part functions as a tracking detector, the second part helps to identify highly relativistic electrons. The drift tubes, called straws, are only 4 mm in diameter and up to 144 cm long. They have a gold-coated tungsten anode wire in the centre, are filled with a Xenon-based gas mixture, and are operated at room temperature. Charged particles traversing the straw ionise the gas. The freed electrons drift towards the anode, as shown in Fig. 3.6(a). The measurement of the drift time allows to determine the track coordinate. The straws are interleaved with fibres or foils which function as transition radiation material, causing highly relativistic electrons to emit low-energy photons. These are detected in the straw tubes as high-threshold hits, because they yield a higher signal amplitude than the ionisation signal from charged particles (low-threshold hit). The straws are positioned parallel to the beam axis in the barrel region (Fig. 3.6(b)) and therefore only provide the measurement of the R-ϕ component. In the end-cap, the straws run perpendicular to the beam axis. The relatively low resolution (compared to PIX and SCT) is compensated by the high number of hits per track (~36). Also the long track length within the TRT ensures a good momentum resolution. The barrel module of the TRT is divided into two halves at $\eta = 0$.

Table 3.3: Abstract of inner detector subsystem properties [64]. The accuracy is given per module (PIX, SCT) or straw (TRT).

	PIX	SCT	TRT
coverage	$\|\eta\| < 2.5$	$\|\eta\| < 2.5$	$\|\eta\| < 2.0$
accuracy: barrel end-cap	 10 µm (R-ϕ), 115 µm (z) 10 µm (R-ϕ), 115 µm (R)	 17 µm (R-ϕ), 580 µm (z) 17 µm (R-ϕ), 580 µm (R)	 130 µm[α]
# components: barrel end-cap	 3 cylindrical layers 2×3 discs	 4 cylindrical layers 2×9 discs	 73 straw layers 2×160 straw layers
# hits	3 pixel hits	4 space points	22–40 low-threshold[β], 7–10 high-threshold[γ]
# read-out channels	80.4×10^6	6.3×10^6	351×10^3

[α] drift time accuracy

[β] for charged particles with $p_T > 5$ GeV

[γ] for electrons with energies > 2 GeV

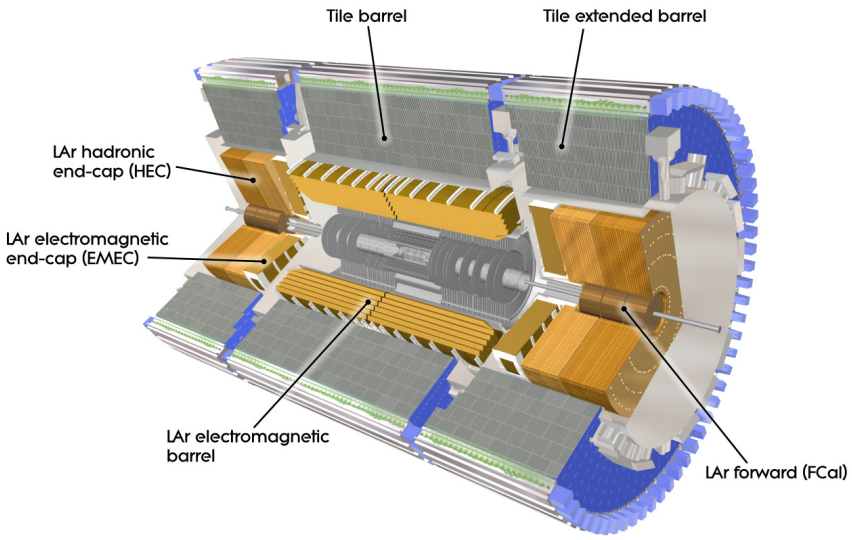
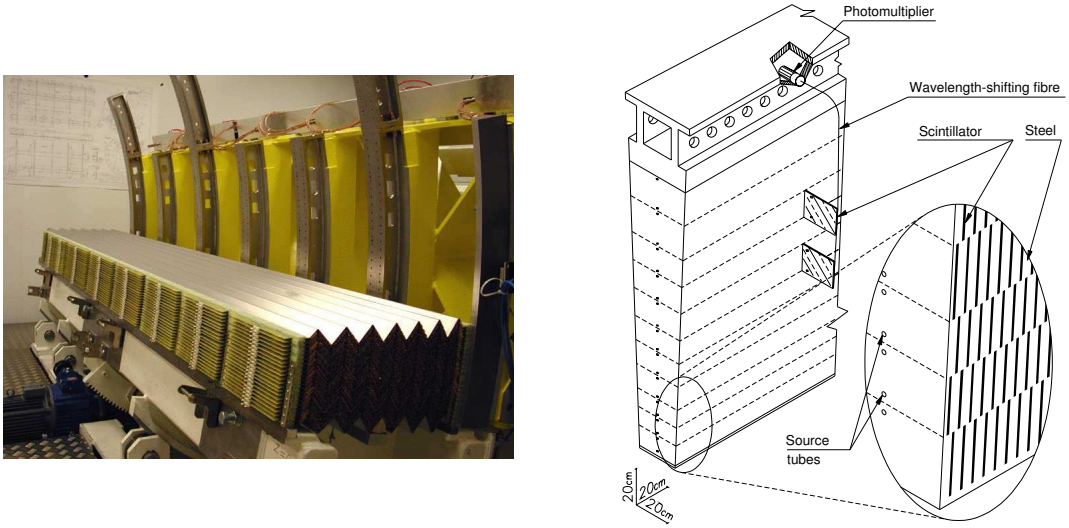


Figure 3.7: The electromagnetic and hadronic calorimeters, surrounding the inner detector. The EM barrel, EMEC, HEC and FCal calorimeters use liquid argon as active material, the hadronic Tile calorimeter uses scintillating tiles [64].



(a) (b)

Figure 3.8: a) Photograph of a partially stacked electromagnetic LAr calorimeter barrel module. Clearly visible are the accordion shaped absorbers and the metal support rings. b) Sketch of a Tile calorimeter module, with alternating layers of steel and scintillator tiles. Fibres collect the signal over the full radial length and transmit it to photomultipliers [64].

3.3.4. Calorimeters

The energy of both charged and neutral particles is measured in the calorimeter system. There are two types of calorimeters: the electromagnetic calorimeter (ECAL) and the hadronic calorimeter (HCAL), optimised for different types of particles. The concept is to stop the particle within the calorimeter volume by making it interact with the detector material. All ATLAS calorimeters are sampling calorimeters, which means that dense passive absorbers alternate with active measurement layers.

The electromagnetic calorimeter is optimised to stop electrons and photons. Electrons and photons interact electromagnetically with the atoms of the absorber material, causing the electrons to emit photons (bremsstrahlung) and the photons to generate e^+e^- pairs. The chance of such a process to happen is described by the radiation length, X_0. It is the mean distance over which the energy of a highly energetic electron is reduced to $1/e$ of its original energy due to bremsstrahlung. Furthermore, it is $7/9$ of the mean free path of a highly energetic photon for pair production [73]. By successive pair production and bremsstrahlung, a shower of particles is created until no energy is left for its continuation.

The hadronic calorimeter measures the energy of hadrons. These interact mainly via the strong force with the nuclei. The probability is described by the nuclear interaction length, λ, the mean path length before a nuclear interaction occurs. For the same material, X_0 is much smaller than λ [74], so that the ECAL appears to be 'short' in terms of nuclear interaction lengths. The hadronic shower generates a significant amount of neutral pions, which quickly decay to a pair of photons. The hadronic shower therefore has an electromagnetic and a hadronic component, which necessiates a calibration of the energy measurement.

The calorimeter system consists of various sub-detectors (Fig. 3.7), which all differ slightly in their design. The barrel ECAL has lead plates as passive layers, which are folded to an accordion shape and stabled, as shown in Fig. 3.8(a). The gap between the plates is filled with liquid argon (LAr),

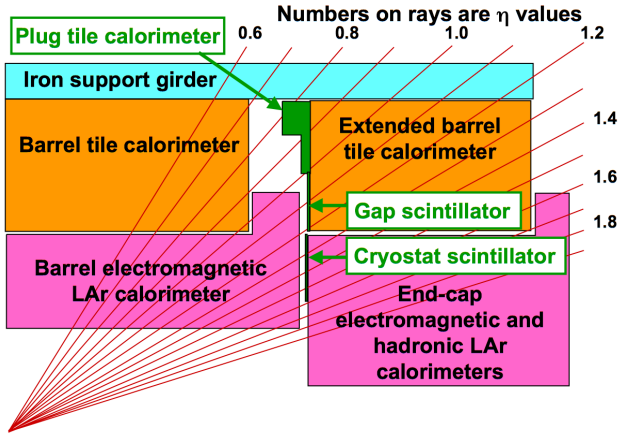


Figure 3.9: Sketch of the gap between barrel and end-cap calorimeters, with the additional scintillator detectors to measure the energy lost in the passive material within the gap (cables, services, etc. - not shown) [64].

interleaved with copper electrodes. Each time the shower passes through the active layers, LAr atoms are ionised, and the freed electrons are collected at the copper electrodes. The total signal is a direct measure of the energy of the primary electron or photon. The barrel calorimeter is $22X_0$ deep, and thus mostly deep enough to contain the full electromagnetic shower. It's nuclear radiation length amounts to less than 2λ. Three layers of different granularity ensure a precise measurement of photons and electrons.

The barrel HCAL is a Tile calorimeter, where plates of steel alternate with scintillating tiles, as shown in Fig. 3.8(b). Particles traversing the scintillator cause the emission of light which is collected by wavelength shifting fibres and measured with photomultipliers. The Tile calorimeter has a depth of 7.4λ.

The end-cap region is equipped with an accordion shaped LAr ECAL (EMEC), a copper-LAr HCAL (HEC), and the forward calorimeter (FCal). The latter extends the reach of the ATLAS detector to $\eta = 4.9$. As very high radiation doses are expected in this region, LAr is again used as the active material, due to its radiation-hardness properties. The FCal is composed of a module optimised for electromagnetic showers, with copper as the passive material, and two modules with tungsten for the hadronic shower.

The calorimeters are designed seamless in ϕ, but have a few gaps in z. The barrel ECAL is divided into two halves, leading to a small gap of 4 mm at $z = 0$. The larger and more influential gap is around $z = 3.5$ m, between barrel ECAL or Tile barrel and EMEC or Tile extended barrel (Fig. 3.9). The gap houses the ECAL cryostat walls, cables, power supply and other services for all inner detectors. Extra scintillators and steel-scintillator detectors are installed to partly recover the energy measurement in this region. Nevertheless, the energy resolution of electromagnetic showers in the region $1.37 < |\eta| < 1.52$ ('Crack region') is reduced significantly. Finally, energy is also lost in the material in front of the calorimeter system, especially in the solenoid. An active layer of liquid argon, the presampler, is installed prior to the first ECAL passive layer to measure this energy loss.

3.3.5. Muon system

The muon system forms the outermost part of the ATLAS detector. Given enough energy, muons penetrate the inner detector and calorimeters loosing only a limited amount of energy. Muons with

Table 3.4: Abstract of calorimeter properties. The granularity varies with η and R [64].

	coverage	granularity	layers	# readout channels
presampler	$\lvert\eta\rvert < 1.8$	$\Delta\eta = 0.025$ $\Delta\phi = 0.1$	1	9344
EM barrel	$\lvert\eta\rvert < 1.475$	$0.025/8 < \Delta\eta < 0.075$ $0.025 < \Delta\phi < 0.1$	3	101760
Tile	$\lvert\eta\rvert < 1.7$	$0.1 < \Delta\eta < 0.2$ $\Delta\phi = 0.1$	3	9852
EMEC	$1.375 < \lvert\eta\rvert < 1.475$	$0.025/8 < \Delta\eta < 0.1$ $0.025 < \Delta\phi < 0.1$	3	62208
HEC	$1.5 < \lvert\eta\rvert < 3.2$	$0.1 < \Delta\eta < 0.2$ $0.1 < \Delta\phi < 0.2$	4	5632
FCal	$1.3 < \lvert\eta\rvert < 4.9$	$3.0 < \Delta x[\text{cm}] < 5.0$ $2.6 < \Delta y[\text{cm}] < 4.7$	3	1762

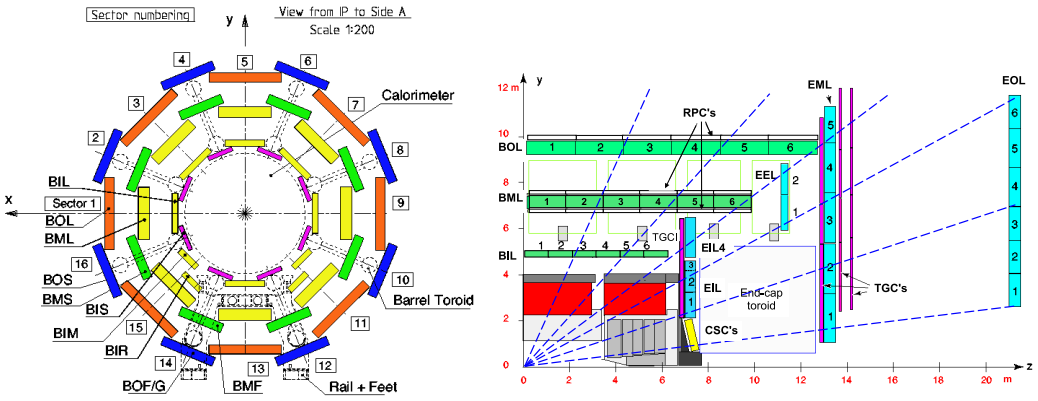


Figure 3.10: Cross-section view of the muon system. a) View of the three barrel layers of MDTs in the x-y-plane. Each layer consists of alternating small and large chambers, providing hermetic coverage in ϕ. b) View of both barrel and end-cap detectors in the y-z-plane (bending plane). The trigger RPCs envelope the second and third layer of the barrel MDTs. Only one layer of TGCs is placed within the magnetic field in front of the first MDT end-cap wheel, the other three layers surround the second MDT wheel. The blue dashed lines indicate the path of muons with infinite momentum [64]. The abbreviations are explained in the list of acronyms (see table of contents).

Table 3.5: Abstract of muon system properties [64].

	MDTs	CSCs	RPCs	TGCs
Function	tracking	tracking	triggering, second coordinate	triggering, second coordinate
coverage	$\|\eta\| < 2.7$	$2 < \|\eta\| < 2.7$	$\|\eta\| < 1.05$	$1.05 < \|\eta\| < 2.4^*$
resolution	35 µm(z/R), no ϕ	40 µm(R), 5 mm (ϕ)	10 mm (z), 10 mm (ϕ)	(2–6) mm (R) (3–7) mm (ϕ)
layers	3	1	3	4
# of hits	20	4	6	9
# readout channels	339×10^3	31×10^3	359×10^3	318×10^3

$^*|\eta| < 2.7$ for second coordinate

a minimum momentum of ≈3 GeV can be measured in the muon system. The design goal was to provide a precise measurement of the muon momentum and charge by only using the muon system. It is therefore embedded into a toroidal magnetic field (Sec. 3.3.2). In order to gain sufficient bending even for very high momentum (super-TeV) muons, the muon system has a large depth (≈ 5 m). The muon system consists of several sub-detectors, which are optimised for either tracking or triggering. Monitored Drift Tubes (MDTs) are the main component of the muon system. There are three layers of MDTs, both in the barrel and the end-cap regions. A module consist of 2×3 or 2×4 layers of $(1 - 6)$ m long and 30 mm in diameter drift tubes. When a muon passes through, it ionises the gas and the freed electrons drift towards the anode wire in the centre of the tube. The drift time is measured to determine the minimal distance between track and wire and thereby the track coordinate. The tubes are oriented along ϕ in both barrel and end-cap, meaning that the ϕ component of the track cannot be measured. In the innermost end-cap and at high eta, the rates are too high for the operation of MDTs and Cathode Strip Chambers (CSCs) are used. These are multi-wire proportional chambers, in which the cathode is segmented into strips. The CSC provides a measurement of both the R and ϕ coordinates. As the MDTs are too slow to provide triggering, faster Resistive Plate Chambers (RPCs) and Thin Gap Chambers (TGCs) are used in the barrel and end-cap, respectively. The trigger chambers are placed close to the MDTs. They measure the coordinate ϕ as well as z (barrel) or R (end-cap). The trigger chambers provide the missing component for the MDTs. The granularity is highest in the end-cap region in order to achieve good performance despite the fact that the outer TGCs are outside of the magnetic field and placed relatively close to each other. The most central detector region (up to $|\eta| < 0.1$) is not covered by trigger nor tracking chambers, leaving a gap to provide services for all inner detector parts, i.e. the solenoid, the calorimeters and the inner detector. Due to the supporting structure of the detector (feet and rails) the muon system has further areas of reduced detector coverage and reduced momentum resolution. The installation of additional chambers and the combination of information from several detector systems nevertheless reduces the acceptance loss.

3.3.6. Trigger system

The rate of collision exceeds the capability of processing and saving of events by many orders of magnitude. Therefore, ATLAS has a trigger system that filters out events for further analysis. More specifically, events with muons, electrons, tau leptons, high-momentum jets, large missing transverse energy, or large total transverse energy are collected. The system consists of three levels, successively reducing the event rate while taking into account more and more measurement details. The first level trigger (L1) is hardware-based. Signals from the calorimeters, using a reduced granularity, and from

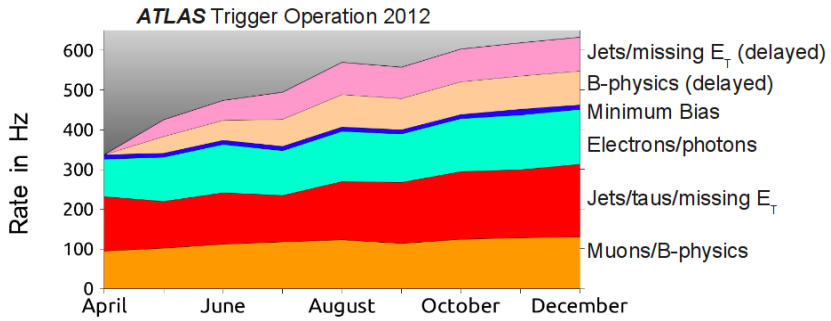


Figure 3.11: Output rates of the EF trigger in 2012, per month. Periods with unstable beam conditions or special runs are not considered. The '(delayed)' streams were recorded in 2012, but processed only during the LHC shut-down phase in 2013 and 2014 [75].

the RPCs and TGCs are taken into account. The output rate is 75 Hz, with a processing time per event of less than 2.5 µs. The L1 trigger defines Regions of Interest (RoIs), which are passed on to the next trigger level. Both the level 2 trigger (L2) and the event filter trigger (EF) are software based. At L2, the full detector information within a RoI is used for a refined object selection. In addition to the clustering algorithms, a fast tracking algorithm for tracks in the inner detector is available. At EF, nearly the full event information is available. Fine granularity, advanced tracking and multi-variate techniques are used to identify physics objects. The design output rate of the EF trigger is 200 Hz. However, computing resources developed better than expected, so that average output rates of more than 600 Hz were achieved in 2012, as shown in Fig. 3.11 [75].

Depending on the desired process and its expected rate, the trigger requires a minimum of one characteristic object (i.e. one muon) or a combination of several objects (i.e. a muon and an electron, or a jet and missing transverse energy). Furthermore, the objects have to satisfy selection criteria, typically a minimum amount of energy or momentum. Also requirements on event topologies are used, for example putting an upper threshold on the separation of two objects. The recorded data is organised in different data streams, depending on the physics objects that triggered the event. The main streams for physics analyses are

- 'Egamma' for events triggered by electrons and photons,
- 'Muons' for events triggered by muons,
- 'JetTauEtmiss' for events triggered by jets, tau leptons, missing transverse energy or total transverse energy.
- 'MinBias' is a collection of inelastic collisions, without a (specific) hard process.

The data streams are inclusive, meaning that the same event can be recorded in several streams [76, 77].

4. Data Taking and Preparation

This chapter is about the preparation of the recorded data, i.e. the steps necessary to go from the actual collision to reconstructed objects that can be used in an analysis. Section 4.1 describes important quantities related to the proton-proton collision, such as the luminosity determination and pile-up. An introduction to the reconstruction of detector objects and particle identification is given in Sec. 4.2. It is follwed by a description of trigger objects in Sec. 4.3, concentrating on triggers that are used in the Higgs boson search. Efficiency measurements and the determination of scale factors, which are needed to correct for mis-modelling in simulated samples, is described in Sec. 4.4.

4.1. Data Taking

The produced number of events of a specific process, e.g. Higgs boson events, depends on the physics of the process and on the parameters of the incoming proton beams. The latter is parametrised via the *instantaneous luminosity*, L, a measure of the number of incoming particles per time and area [8]:

$$L = \frac{n_b f_r n_1 n_2}{2\pi \Sigma_x \Sigma_y} \quad . \tag{4.1}$$

Here, n_b is the number of colliding bunch pairs, f_r is the revolution frequency, n_i is the number of protons per bunch, and $\Sigma_{x,y}$ is the beam width in horizontal or vertical direction. Bunch length, collision angle and bunch overlap will further influence the rate. Integrated over time, the luminosity is a measure for the amount of data collected by the experiment, because it is directly proportional to the number N of events collected:

$$N = \sigma L_{\text{int}} \ , \qquad L_{\text{int}} = \int L \, \mathrm{d}t \ . \tag{4.2}$$

The cross section σ depends on the process and on the collision energy, as shown for a number of processes in Fig. 4.1. The production cross section of W and Z bosons is six orders of magnitude lower than the total cross section. The Higgs bosons production rate is lower by another three to four orders of magnitude. The figure underlines the importance of high-performance triggering to filter out events of specific interest.

The luminosity is measured using multiple independent detectors and methods, in order to obtain a robust measurement. The main devices are the beam conditions monitor (BCM) and the LUCID detector, which is especially designed for this measurement. The detectors each consist of sub-detectors on both sides of the nominal ATLAS interaction point, at $z = \pm 17$ m and $z = \pm 184$ cm, respectively. Both LUCID and BCM are fast detectors, able to measure the luminosity per colliding bunch pair. This is done by either counting inelastic scattering events or detector hits. Since each detector and method has a different counting efficiency, the measurements have to be calibrated. This is done in beam-separation scans, the so called van der Meer scans [79]. In special runs, one beam is displaced with respect to the other, and the interaction rate $R(\delta)$ is measured as a function of the distance δ between the two beams. Independent of the beam shape and the specific monitor used, the beam width Σ_i in Eq. 4.1 is then given as:

$$\Sigma_i = \frac{1}{\sqrt{2\pi}} \frac{\int R_i(\delta) \, \mathrm{d}\delta}{R_i(0)} \quad . \tag{4.3}$$

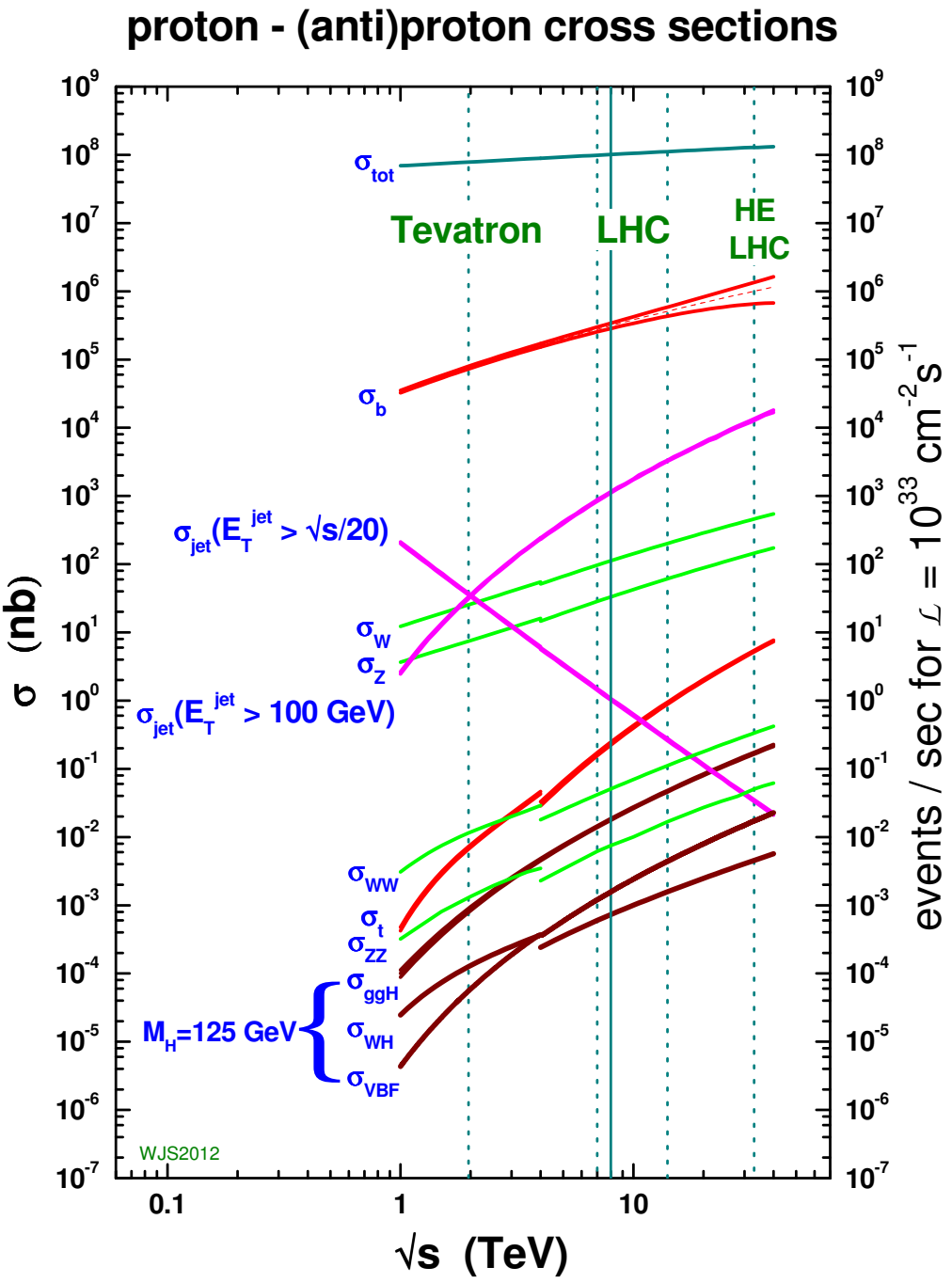
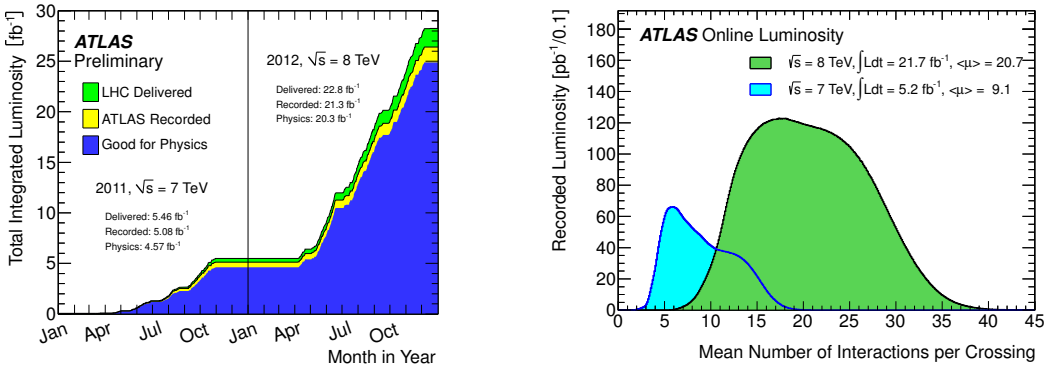


Figure 4.1: Cross sections in proton-antiproton collisions ($\sqrt{s} < 4\,\text{TeV}$) and proton-proton collisions ($\sqrt{s} > 4\,\text{TeV}$) as a function of centre-of-mass energy $\sqrt{s}$. The vertical lines indicate the centre-of-mass energies of three colliders: the Tevatron with a final energy of $\sqrt{s} = 1.96\,\text{TeV}$, the LHC with $\sqrt{s} = 7\,\text{TeV}$ (2011), $\sqrt{s} = 8\,\text{TeV}$ (2012) and $\sqrt{s} = 14\,\text{TeV}$ (design) and the HE-LHC (High Energy LHC), a proposed LHC upgrade, with $\sqrt{s} = 33\,\text{TeV}$ [78].



(a) (b)

Figure 4.2: a) Cumulative luminosity versus time delivered by the LHC (green), recorded with the ATLAS detector (yellow) and satisfying certain quality criteria (blue) [80, 83]. b) Mean number of interactions within one bunch crossing for the data recorded in 2011 and 2012 [80, 83].

With knowledge of the product $n_1 n_2$ from LHC beam current measurements, the luminosity can be determined using Eq. 4.1 [80].

The luminosity is monitored in time units of typically one minute. Within such a *luminosity block* (LB), conditions are considered to be stable. Other information, such as data quality or trigger settings, are provided for the same time units. A collection of specific LBs then forms a data sample, and the integrated luminosity is given by the sum of the integrated luminosities per LB [80].

In 2012 (2011), the LHC delivered the ATLAS detector with a total amount of 22.8 fb^{-1} (5.46 fb^{-1}) of stable beam at a centre-of-mass energy of 8 TeV (7 TeV). Figure 4.2(a) shows the accumulation of data over the two years. Not all collisions delivered are actually recorded due to the start up of the detector and inefficiencies in the acquisition of data. Furthermore, data used in a physics analysis need to be of good quality. That means in particular that all detector components are in operation and working well. In 2012 (2011), this was the case for about 95% (90%) of the recorded data. The large difference in the total amount of data in 2012 compared to 2011 is caused mainly by a change of beam parameters. In 2011, the LHC started with a lower number of bunches, which was then gradually increased over the first half of the data taking period. When the maximum number of 1380 bunches with a spacing of 50 ns was reached, the bunches were filled successively with more protons. While the first change mainly increases the number of bunch crossings, the second change increases the chance of an interaction. Eventually, this leads to multiple interactions occurring in the same bunch crossing, commonly referred to as *pile-up*. The average number of interactions per crossing is shown in Fig. 4.2(b) for the years 2011 and 2012. A smaller transverse beam size at the collision point also shifts the distribution to higher values. Clearly visible is the double peak structure in the 2011 distribution, which is due to an improved focusing of the beam ($\beta^* = 1.5\,\text{m} \rightarrow \beta^* = 1.0\,\text{m}$) in the last quarter of the 2011 data taking period. In 2012, the LHC was operated with $\beta^* = 0.6\,\text{m}$, resulting in an even higher number of simultaneous interactions. A reduced emittance delivered by the pre-accelerator chain (injectors) contributed further to a steady increase of the peak luminosity over both years [81, 82].

4.2. Reconstruction

A collision event leaves signals in millions of readout channels in the detector. In order to interpret the event, these signals have to be translated into particle traces. First of all, basic structures like tracks and clusters are reconstructed. They are the basis for building particle candidates. Rather loose criteria are applied in the *reconstruction* step, so that very little signal is lost. The candidates can overlap, meaning that the same track or cluster can be associated to several candidates. A higher level of background rejection and candidate separation is reached by applying particle *identification* algorithms, which are optimised to reduce specific backgrounds. Selected reconstruction and identification algorithms are outlined in the following. The focus is on the methods most relevant for the Higgs search described in chapter 6, for instance the reconstruction within the central part of the detector ($|\eta| < 2.47$), and approaches used during the 2012 data taking period.

Tracks and vertices: The basic track reconstruction algorithm [64, 84] is an 'inside-out' approach. It is optimised for the reconstruction of tracks with a minimum momentum of 400 MeV. The algorithm starts with track seeds from the first few inner detector layers. Moving outwards, additional SCT-hits are added. Finally, the track candidates are extended into the TRT and refitted. Secondary particles from the subsequent decay of (long lived) particles might not be reconstructed by the inside-out approach. For their reconstruction, a 'back-tracking' approach is used. It starts in the TRT and moves inwards. For the track fitting, methods like Kalman filtering [85] or Gaussian-sum filtering [86] are used. The reconstructed tracks are then the input for an iterative algorithm that identifies all collision vertices. The vertex with the highest $\sum(\text{track } p_{\text{T}}^2)$ is called the *first primary vertex* in the event. It defines the origin of the event's coordinate system. The other collision vertices are called *pile-up vertices* [1]. The variable *number of vertices*, N_{vtx}, is often used to measure the pile-up activity of an event. The variable counts the primary vertex and the pile-up vertices. The primary vertex is here required to have at least four associated tracks, while the pile-up vertices are required to have at least two associated tracks. The variable N_{vtx} is related, but not identical, to the mean number of interaction per crossing shown in Fig. 4.2(b). The latter is deduced from the measured luminosity.

Clusters: Two cluster reconstruction algorithms are used: the sliding-window approach and the topocluster approach [87, 88]. The sliding-window algorithm is used for the reconstruction of photons and electrons, and therefore considers the electromagnetic calorimeter only. Calorimeter towers of size $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ are created by summing over the longitudinal ECAL layers. A fixed window of 3×5 towers ($\eta \times \phi$) is used to find and define clusters, by searching for local maxima of contained energy across the tower grid.

Topoclusters are formed from cells that have an energy exceeding the expected noise level by a specific amount. Thereby, the expected noise level is given by the width of the noise distribution, which includes both electronic noise and noise from pile-up [89]. The algorithm is seeded by cells with a signal-to-noise ratio above some high threshold t_{seed}. Iteratively, neighbouring cells are added, if they are above a second, lower threshold $t_{\text{neighbour}}$. If a neighbouring cell is below that threshold but above a third threshold t_{cell}, it is also added to the cluster but not used to continue the iteration. Neighbouring cells are defined as the (usually eight) surrounding cells within the same layer, and typically overlapping cells in adjacent layers and calorimeter subsystems. In this way, a three-dimensional cluster with arbitrary shape is created. An additional algorithm can be applied to split clusters that merged during the process.

[1] For completeness, another group of vertices are the *secondary vertices*. They arise from the decay of particles produced in the collision.

The topocluster energy is estimated at the *EM scale*, meaning that it reproduces the energy deposited by an electromagnetic shower. Since ATLAS is a non-compensating detector, hadronic showers have a lower response and the cluster needs to be recalibrated in order to correctly describe the energy response of the hadronic part of the shower. This is done using the local hadron calibration (LC) scheme [90]. Besides the non-compensation, the energy at LC scale is also corrected for signal losses due to noise threshold effects and energy losses due to non-instrumented detector regions [91].

Electrons and photons: The electron and photon reconstruction is seeded from a sliding-window electromagnetic cluster with at least 2.5 GeV. A track-cluster match is performed, taking into account the bending direction of the track. If only a track segment (failing the default track reconstruction) can be matched to a cluster, the track fitting is redone with an electron-hypothesis to account for higher bremsstrahlung losses. Clusters with a matched track are then classified as electrons. The cluster is rebuild using 3×7 (barrel) or 5×5 (end-cap) cells in $\eta\times\phi$, taking into account the energy distribution in each layer. Clusters without a matched track are classified as unconverted photons. Converted photons, i.e. photons undergoing a conversion into an electron-positron pair, have to be recovered from the group of electrons. A dedicated conversion vertex finding is performed, considering the photon hypothesis ($m_\gamma = 0$). The conversion vertices are, if possible, associated to the clusters, and all electron candidates are re-evaluated [87, 92].

The identification of electrons is done using a multivariat approach. Distributions of the longitudinal and transversal shower shapes, track properties and track-cluster matching are used to build likelihood functions for signal and background. Discriminators are constructed by combining the signal and background functions for multiple sets of variables. Based on the discriminators, three working points (WP), *loose*, *medium* and *tight*, are defined for increasing background rejection. While the loosest WP focuses on the rejection of light-flavour jets, the tightest WP takes into account additional variables to reject heavy-flavour jets and conversions [87].

In addition to the electron ID, requirements on the isolation can be used to reject backgrounds. For the isolation in the inner detector, the *track isolation* $p_{\mathrm{T}}^{\mathrm{iso}(r)}$, the transverse momenta of the tracks surrounding the electron in a cone of $\Delta R < r$ are summed (the electron track is not considered). The momentum sum is commonly put in relation to the electron momentum, which is then denotes as $p_{\mathrm{T}}^{\mathrm{iso}(r)}/p_{\mathrm{T}}$. Likewise, the *calorimeter isolation* $E_{\mathrm{T}}^{\mathrm{iso}(r)}$ is defined. The transverse energies of the topoclusters in a cone of $\Delta R < r$ around the electron axis are summed. The energy from the elecron is removed by subtracting the energy from the central 5×7 cells. As for the track isolation, the calorimeter isolation is commonly set in relation to the electron momentum, denoted here as $E_{\mathrm{T}}^{\mathrm{iso}(r)}/p_{\mathrm{T}}$.

Muons: Muons are reconstructed mainly using tracks in the inner detector and muon system. Several reconstruction types are provided, using a statistical combination of tracks in the detector sub-systems (STACO muons). Standalone (SA) muons are reconstructed using the muon system only. To obtain vertex information, the track is extrapolated towards the interaction region. Combined (CB) muons are built from tracks that are independently reconstructed in the inner detector and muon system. Segment-tagged (ST) muons are based on inner detector tracks, which are extrapolated to the muon system and can be matched to track segments there. The CB muon reconstruction achieves the highest purity, but fails in low instrumented areas or for very low momentum muons. The SA and ST algorithms help to recover these muons [93].

Similar as for the electrons, requirements on the isolation in the inner detector and the calorimeter can be used to enhance the purity of a sample of muons.

Jets: The jet objects considered in this thesis are built from 4-2-0 topoclusters, which are calibrated either at EM or LC scale. The signal-to-noise thresholds are indicated as $t_{\mathrm{seed}} - t_{\mathrm{neighbour}} - t_{\mathrm{cell}}$. The jet

is reconstructed using the anti-k_t jet clustering algorithm [94]. The algorithm iteratively collects all topoclusters into jets. It starts with the high-energy (*hard*) clusters, successively adding the clusters in the area around. Two measures control the jet formation process:

$$d_{ij} = \min(k_{\mathrm{T}i}^{-2}, k_{\mathrm{T}j}^{-2}) \frac{\Delta_{ij}^2}{R^2} , \tag{4.4}$$

$$d_{iB} = k_{\mathrm{T}i}^{-2} . \tag{4.5}$$

Here, $k_{\mathrm{T}i}$ is the transverse momentum of an object i, and $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)$, with y and ϕ being the rapidity and azimuth. The objects i and j are topoclusters or, later in the process, pseudojets (the sum of several topoclusters). The radius parameter R determines the maximum size of the final jet. In this thesis, only jets built with $R = 0.4$ are used. The algorithms searches for the minimum value among all possible d_{ij} and d_{iB}. If the minimum is of type d_{ij}, the objects i and j are combined, forming a new object. If the minimum is of type d_{iB}, the object i is considered a final jet, and removed from the list of objects. The procedure is repeated until no objects are left. Due to the negative exponent of $k_{\mathrm{T}i}$, the jet-finding process and the appearance of the final collection of jets is driven by the hard clusters[2]. A resulting jet is perfectly conical with radius R, if there are only soft clusters within the distance $2R$ around a hard cluster seed. If there is a second hard cluster within distance $2R$, but further than distance R, two jet are formed. The association of the soft clusters between them then depends on the energy balance of the two jets. If the second hard cluster is closer than R, a single jet is created. The anti-k_t algorithm is both infrared and collinear safe, meaning that the addition of clusters from soft emission or the splitting of a hard cluster into two will not change the set of reconstructed hard jets.

The jet momentum is given by the sum of four-momenta of the constituent topoclusters, which are assumed massless. Jet specific corrections are applied to account for pile-up contributions, the first primary vertex in the event[3], and a MC and data based energy calibration. Four working points are defined to reject jets from non-collision backgrounds, like beam-gas events (proton hitting residual gas inside the beam-pipe), beam-halo events (beam-halo interacting with collimators), cosmic-ray muons overlapping with collision events, and calorimeter noise. Discriminating variables are mainly based on energy fractions of calorimeter sub-systems or layers, signal pulse shapes in cells, and timing [89, 96].

Tau leptons: The reconstruction of tau leptons is seeded by anti-k_t jets, with parameter $R = 0.4$ and topoclusters calibrated at LC scale as input. Only jets fulfilling $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ and $\eta < 2.5$ are considered. Clusters and tracks within a cone of $\Delta R < 0.2$ around the barycentre define the tau candidate. The energy calibration is optimised for hadronically decaying tau leptons, taking into account the specific hadron mixture as well as reducing contributions from pile-up. A multivariate approach is used to reject the dominant multi-jet background. Information of shower shapes, track distances and jet sub-structure is combined in a boosted decision tree discriminator and three working points are defined. Additional algorithms provide rejection of electrons and muons. Even though reconstructed and identified tau objects are provided for any number of associated tracks, the algorithms are optimised for tau candidates with one or three tracks. A detailed description of tau lepton reconstruction and identification is given in chapter. 5.

[2] This is in contrast to other jet finding algorithms, for instance the k_t algorithm [95]. It follows essentially the same procedure, but uses a positive exponent, k_T^2, and therefore starts from soft clusters.

[3] The jet finding is performed using the geometrical centre of the detector and only later the positions of all topoclusters are recalculated with respect to the first primary vertex.

Missing transverse energy: Neutrinos traverse the detector without leaving a signal[4]. The same holds for some additional particles appearing in models that describing physics beyond the SM, such as Dark Matter candidates. As mentioned in Sec. 2.4, information on undetected particles are inferred via the missing transverse energy in the event. It is determined by taking the vector-sum of the transverse momenta of all objects in the event. The magnitude of the resulting vector determines the absolute value of the missing transverse energy, $E_{\mathrm{T}}^{\mathrm{miss}}$. The direction of the resulting vector is reversed to give the ϕ-component of $E_{\mathrm{T}}^{\mathrm{miss}}$. The objects in the event are counted as follows. Calorimeter deposits are associated with reconstructed and calibrated objects, in the order electrons, photons, hadronically decaying tau leptons and jets. The matched objects contribute to the calculation of $E_{\mathrm{T}}^{\mathrm{miss}}$. Muons are added if they are at least partly reconstructed in the muon system. To avoid double counting, muon energy deposits in the calorimeter are subtracted. Finally, topoclusters that are not associated to any reconstructed object are added as well (*soft term*). In order to increase the pile-up robustness, an additional pile-up suppression is applied using tracking information. Jets are only considered if they most likely come from the primary vertex in the event. The soft term is scaled by the soft term vertex fraction (STVF). For its determination, the momenta of tracks not associated to any reconstructed object are summed up. The STVF is defined as the fraction of momentum that is also associated to the primary vertex [98, 99].

4.3. Triggering

The decision of the trigger system (Sec. 3.3.6) is based on dedicated trigger objects. The reconstruction of the trigger objects resembles the reconstruction described before, but is not identical. The complexity of the trigger objects increases with the trigger level, because more time is available and less events have to be evaluated. The description here focusses on the electron and muon triggers needed in the Higgs analysis.

Muon trigger: The L1 muon trigger is seeded by spatial and temporal coinciding hits in the RPCs and TGCs, which are pointing to the interaction region. The momentum is estimated by evaluating the degree of deviation from an infinite momentum track. If a given p_{T} threshold is passed, a RoI of typically 0.1×0.1 (RPCs) and 0.03×0.03 (TGCs) in $\Delta\eta \times \Delta\phi$ is passed to the next trigger level. At L2, tracks within the RoI are reconstructed using MDT hits, and are combined with the closest track in the inner detector. At EF trigger level, the full detector information within the RoIs is used for a more advanced track reconstruction and combination with inner detector tracks. Requirements on the inner detector isolation can be used to reduce backgrounds by summing up the momentum of additional tracks with at least 1 GeV within a cone around the muon track [100].

The muon trigger used in the Higgs boson search described in chapter 6 is a combination ('or') of two single muon triggers, called `mu24i_tight` and `mu36_tight`. At least one muon is required to pass the momentum threshold of 24 GeV and 36 GeV at EF level, respectively. The lower p_{T} trigger has an additional isolation requirement. The extra track momentum in a cone of $\Delta R < 0.2$ is not allowed to exceed 12% of the muon momentum. Both triggers were operated unprescaled[5] throughout 2012.

[4] Neutrino detection is not possible with the ATLAS detector. Nevertheless, with specifically designed detectors neutrino detection is possible - even though only a small fraction of the total neutrino flux is detected, due to the tiny interaction cross section. An example of such a neutrino detector is IceCube at the south pole [97].

[5] The prescale of a trigger indicates the fraction of triggered events allowed to pass. A prescale of one (unprescaled) means that all events that satisfy the trigger requirements are accepted.

Electron trigger: The L1 electron trigger searches for local maxima in the calorimeter using a sliding-window of 4×4 towers. The trigger towers span over the electromagnetic and hadronic calorimeter (excluding the last HCAL layer), using a reduced granularity of 0.1×0.1 in $\Delta\eta \times \Delta\phi$. If the combined energy of two neighbouring cells within the 2×2 core of the window is above a certain threshold, the window is passed as a RoI to the next trigger level, L2. Within the RoI, fast clustering and tracking is done using the full detector granularity. Simple shower shape related quantities are used to enhance the purity. At EF level, an advanced electron reconstruction is performed. Requirements on the longitudinal and transversal shower shape and track quality reduce the non-electron background [101, 102].

The electron trigger used in the Higgs boson search described in chapter 6 is a combination ('or') of two single electron triggers, called `e24vhi_medium1` and `e60_medium1`. The cluster energy thresholds at EF level are 24 GeV and 60 GeV, respectively. The first trigger puts additional requirements on the absolute leakage into the HCAL at L1 and on the longitudinal shower shape in the ECAL at EF. Additionally, the extra track momentum in a cone of $\Delta R < 0.2$ is not allowed to exceed 10% of the cluster energy [103]. Both triggers were operated unprescaled throughout 2012.

4.4. Efficiency measurements and scale factors

Due to imperfect modelling of shower shapes and track properties, reconstruction and identification efficiencies of electrons, muons, and tau leptons can differ in data and simulation. The MC efficiency is therefore corrected using a multiplicative factor on each reconstructed or identified object to restore the efficiency observed in the data. These *scale factors* (SF) are obtained by measuring the efficiency ε in data and simulation and forming the ratio $\varepsilon_{\text{data}}/\varepsilon_{\text{MC}}$. The efficiency measurement in data requires a clean sample of the object under investigation. This is typically obtained by performing a *tag-and-probe* analysis. A sample enhanced in specific events (e.g. $Z \to ee$ events) is selected by requiring a *tag* and a *probe* object. Tight selection criteria are applied on the tag (e.g. one of the electrons), in order to reject events with similar signatures. The efficiency measurement is then performed on the probe object (e.g. the other electron), which has to satisfy minimal requirements only. The efficiency scale factors for each object type are provided within the ATLAS collaboration by the respective working groups.

Electrons: The electron scale factors are measured on samples enhanced in $Z \to ee(\gamma)$ and $J/\psi \to ee$ events. The different components from triggering, reconstruction and identification are obtained individually. In the end, they are combined (multiplied) to a global scale factor. Since the efficiencies are dependent on the transverse energy and pseudorapidity of the electron, the scale factors are obtained in bins of E_{T} and η. The J/ψ enhanced sample is predominantly used in the low-E_{T} bins, starting at 7 GeV, while the Z enhanced sample becomes exceedingly important above 10–15 GeV. The identification efficiency is very momentum dependent. For the tightest working point, it increases from ≈65% to 90% within the measurement's momentum range. Differences in the identification efficiency between data and simulation are mostly small, but can reach up to 10% in the low E_{T} or high η bins. The scale factor uncertainties range from 0.5% to 10%. The higher uncertainties are connected to the transition region of the calorimeters. The reconstruction efficiency, i.e. the efficiency to match a track to a cluster, is better than 95%. The data-MC differences are mostly at the permille level and not exceeding 1–2%. Uncertainties on the scale factors are 0.5–2% above 15 GeV, which is the measurement threshold for the reconstruction efficiency. Below 15 GeV, a scale factor of 1 is assumed and uncertainties are assigned of 2% and 5%in the barrel and end-cap regions, respectively [87].

The trigger efficiency of the combined `e24vhi_medium1` and `e60_medium1` triggers is measured in a $Z \to ee$ analysis. The trigger efficiency increases notably up to around 40 GeV and then starts to level off. At ≈60 GeV, the non-isolated trigger joins in, resulting in an abrupt increase of the efficiency [103].

A sample enhanced in $Z \to ee$ events is also used for an in-situ energy calibration. While the final calibration is applied on data events, an energy smearing is applied on simulated events [87].

Muons: The muon scale factors are measured on samples enhanced in $Z \to \mu\mu$ events and are binned in η and ϕ. No dependence on momentum is observed. Different maps for different data taking periods are produced, in order to account for varying detector conditions. An average map is created for the full 2012 data taking period, taking into account each period's contribution to the full luminosity. A momentum scale correction and resolution smearing is obtained from $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ events. The corrections are measured separately for inner detector and muon system tracks and then propagated to the combined muon momentum reconstruction. The scale corrections are at maximum 0.2% (1%) for inner detector (muon system) tracks, and resolution smearing corrections are below 10% (15%). The reconstruction efficiency for CB+ST muons is uniformly around 99% for muons with momenta above 4 GeV, except for the most central detector region $\eta < 0.1$ [93].

The trigger efficiency for the `mu24i_tight` and `mu36_tight` trigger combination is measured in a sample enhanced in $Z \to \mu\mu$ events. The efficiency plateau for the trigger combination has a smooth behaviour and is reached at ≈25 GeV. Trigger efficiency scale factors are binned in η and ϕ, with typical uncertainties of 1% (statistical) and 0.6% (systematic) [100].

Jets: The jet selection efficiency of the loosest working point is above 99.8% over the full momentum and pseudorapidity range. The measurements performed on data and on simulated samples are in excellent agreement.

The jet energy calibration and uncertainty measurements are obtained in data events enhanced in $Z \to ee$ + 1 jet events, γ + 1 jet events and multi-jet events, depending on the momentum range studied. The momentum of the Z boson, the photon or the multi-jet system balances the (leading) jet momentum and can be measured precisely. Uncertainties on the jet energy scale of about 3% are determined [89].

Tau leptons: The tau identification efficiency scale factors are obtained in events enhanced in $Z \to \tau\tau$ events, with one leptonically decaying tau lepton (tag) and one hadronically decaying tau lepton (probe). For $p_{\mathrm{T}} > 20$ GeV, no p_{T} dependence is observed, and the scale factors are provided in two η bins and seperately for candidates with one or three tracks. The uncertainties are of the order of 2-5%. For $p_{\mathrm{T}} < 20$ GeV, uncertainties increase to a maximum of 15%. Efficiency scale factors are also provided for the electron veto. They are obtained in a $Z \to ee$ analysis, where one of the electrons is mis-identified as a tau lepton. The scale factors are binned in η, with uncertainties of about 10% for the loose WP [104].

5. Tau identification

A high reconstruction and identification (ID) efficiency of tau leptons is important for any analysis with tau leptons in the final state. Equally important is a high rejection of objects with a similar detector signature, so that it is possible to select a sample that is dominated by actual tau leptons. The reconstruction and identification algorithms are required to perform well over the LHC-accessible momentum range as well as in 'busy' events, e.g. caused by numerous simultaneous proton-proton collisions. Therefore, the tau algorithms are updated and improved continuously.

This chapter reviews the author's studies concerning tau identification. Boosted decision trees (BDTs) are used to discriminate the signal (tau leptons) from the main background of jets from multi-jet events. Three distinct cases are investigated:

- the dependence of the BDT tau ID on varying pile-up conditions,
- the dependence of the BDT tau ID on the momentum of the tau lepton,
- the inclusion of variables related to the substructure of the tau lepton decay.

The three sections are embedded in a description of the tau reconstruction and identification chain. The chapter starts with an introduction of the tau lepton with a focus on its hadronic decay modes and methods of identification (Sec 5.1.1). Concrete requirements and ID algorithms as used in the 2011 data taking period are described in Sec. 5.2. The 2011 tau ID is the starting point of the investigations concerning the pile-up dependence of the BDT tau ID (Sec. 5.3). Section 5.4 gives the full overview over the steps taken from the reconstruction of tau objects to the usage in an analysis. It includes the description of the final 2012 tau ID, the lepton vetos, the tau energy scale and the tag-and-probe measurements to determine correction factors. The 2012 tau ID is the starting point for the investigations concerning the momentum dependence of the BDT (Sec. 5.5) and the inclusion of substructure information (Sec. 5.6).

5.1. Introduction

5.1.1. The tau lepton

The tau lepton was discovered in electron-positron collisions at SLAC in 1975 [105]. It is the heaviest lepton with a mass of 1777 MeV, which makes it heavier than the lightest mesons. The tau lepton therefore decays either leptonically into an electron or muon and neutrinos, or hadronically into mainly charged and neutral pions. The dominating decay modes are shown in Fig. 5.1. The mean lifetime of the tau lepton is 290×10^{-15} s [8]. As a consequence, a typical 50 GeV tau lepton travels $\approx$2 mm and decays before it reaches the first layer of the ATLAS detector. The tau lepton can hence not be detected directly and is identified by its decay products. In an analysis, the whole event structure is considered. A final state tau lepton is then regarded with all decay components: a lepton or hadronic jet and missing transverse energy from the neutrino(s). In the object reconstruction based on detector hits, only the detectable decay components are of interest. Since neutrinos leave the detector unseen, the visible decay products are the lepton or the hadronic jet only. The electron or muon from

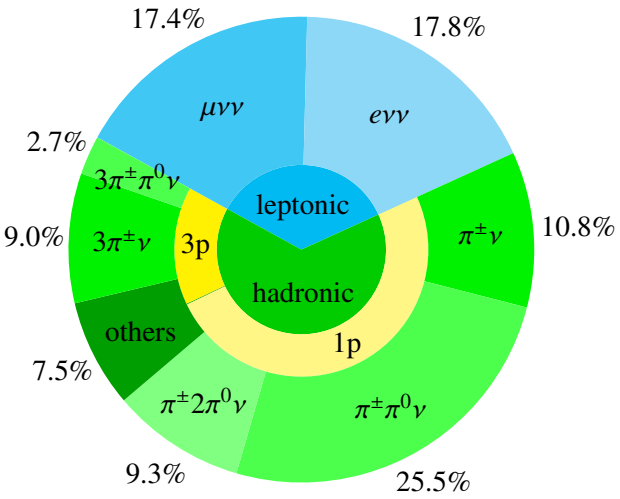


Figure 5.1: Overview of the dominating tau lepton decay modes. The category 'others' contains decays with other charged or neutral mesons (mainly kaons), and higher number of neutral or charged pions.

the leptonic decay is nearly indistinguishable from primary (*prompt*) electrons or muons. In the object reconstruction these are hence treated as electron or muon candidates. The tau reconstruction and identification algorithms focus exclusively on the hadronic decay products[1]. The symbolic notation τ_h is used here for the tau detector object.

5.1.2. Detector signature and discriminating variables

In the majority of hadronic tau decays, the hadrons are one or three charged pions and up to two neutral pions. The neutral pions immediately decay into two photons and are seen in the electromagnetic calorimeter. The charged pions leave a track in the inner detector and are stopped mainly in the hadronic calorimeter. The detector signature resembles jets in multi-jet events (here referred to as *QCD-jets*). As described in Sec. 4.1, multi-jet events occur at very high rates at the LHC and consequently QCD-jets represent the highest background to tau identification. The fact that hadronic tau decays consist of one or three charged hadrons is a starting point for the rejection of such jets. Figure 5.2 shows the number of reconstructed tracks associated to a hadronic tau decay compared to a QCD-jet. In the case of tau leptons, the distributions peak at 1 and 3, while it is more spread out in the case of QCD-jets. Furthermore, the average number of tracks increases with energy for QCD-jets, while it is roughly momentum independent in the case of tau leptons[2]. Requiring one or three tracks therefore reduces the number of QCD-jets significantly. In the tau identification algorithms, decays with one or three charged hadrons are treated differently, in order to allow for an optimal QCD-jet rejection in each case. The decays are then referred to as *1-prong* and *3-prong*. The tau leptons emerging from a collision event are mostly boosted and the decay products appear in a narrow cone

[1] In this thesis, terms like 'reconstructed tau decays' therefore always refer to the visible part of an hadronic decay, even if not explicitly mentioned.

[2] The decay mode probability is independent of the tau momentum. However, the efficiency to reconstruct a track and the association to the tau candidate is not. Therefore, there is some migration between the bins, which nevertheless confines itself to the low multiplicity bins.

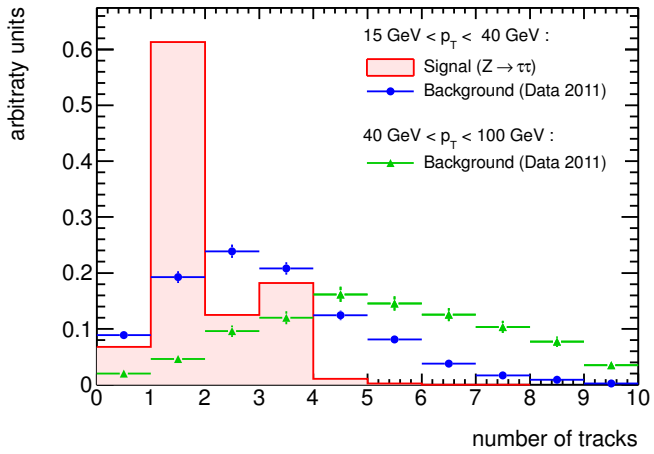


Figure 5.2: Number of reconstructed tracks associated to a tau candidate for simulated decays in $Z \rightarrow \tau\tau$ events (signal) and QCD-jets from multi-jet events in 2011 data (background) at $\sqrt{s} = 7$ TeV centre-of-mass energy. The QCD-jet multiplicity spectrum is shown for two momentum regions. All three distributions are normalised to unity.

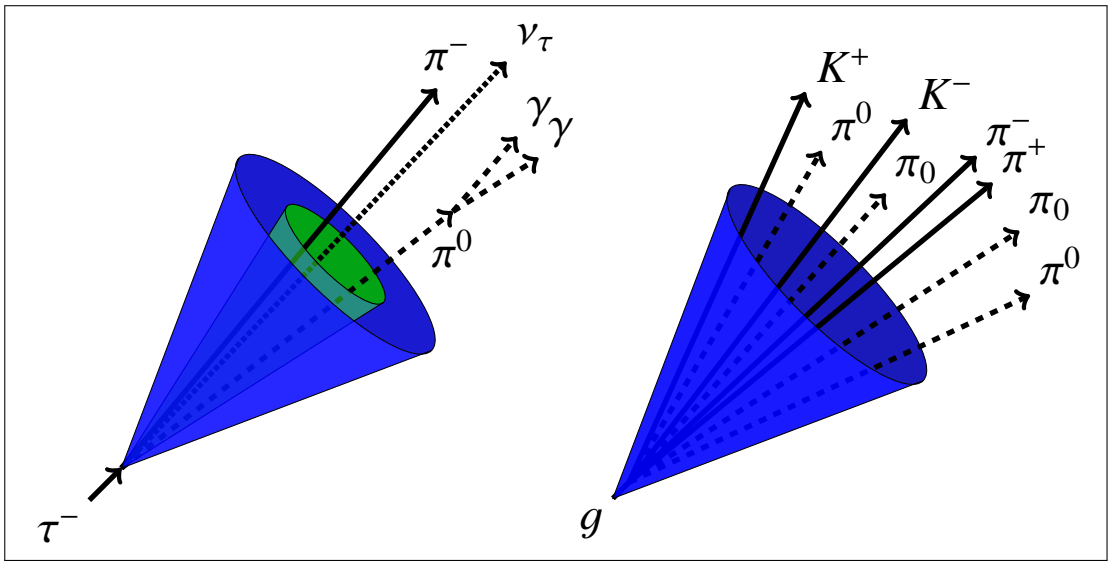


Figure 5.3: Sketch of a hadronic 1-prong tau decay (right) and a QCD-jet (left).

in the direction of flight. The decay products are reconstructed as a jet and gather closely around the jet barycenter, as illustrated in Fig. 5.3. This *core cone*, defined by the area $\Delta R < 0.2$, is surrounded by a depleted *isolation annulus*, defined as the area with $0.2 \leq \Delta R < 0.4$ around the barycenter. The decay cone of a tau lepton is on average narrower than the cone of a QCD-jet of the same momentum (see Fig. 5.3), meaning that the particles of a QCD-jet spread more evenly over the full reconstruction cone of $\Delta R < 0.4$. This leads to an intense usage of the cone size in discriminating variables, such as the number of isolation tracks, the isolation energy, the average track or cell distance and the ratio of energies deposited in different cones. Also the fact that the tau lepton travels a small distance before it decays can be used for tau identification. This leads to a larger transverse impact parameter of the tracks and the possibility to reconstruct a secondary vertex in the case of a 3-prong decay. These following variables are therefore calculated during tau reconstruction and were used in 2011 for the purpose of jet discrimination. Distances are measured with respect to the tau axis, which is given by the calorimetric barycenter.

R_{track} The *track radius* describes the p_{T}-weighted mean distance between the tracks and the tau axis. Both core and isolation tracks are considered.

R_{cal} The *calorimeter radius* describes the E_{T}-weighted mean distance between the calorimenter cells and the tau axis. All cells that are associated to tau topoclusters and that are within the full cone $\Delta R < 0.4$ are considered.

f_{track} The *leading track momentum fraction* is defined as the ratio of the momentum of the leading track and the calorimeter energy deposit (in the cone $\Delta R < 0.4$).

f_{cent} The *central energy fraction* is defined as the calorimeter energy fraction in the innermost cone of $\Delta R < 0.1$.

m_{tracks} The *track mass* is the invariant mass of both core and isolation tracks, assuming zero mass for each track.

m_{clus} The *cluster mass* is the invariant mass of all topoclusters associated to the tau candidate. Each cluster is assumed massless.

m_{effClus} The *effective cluster mass* is the invariant mass of the clusters that dominate the total calorimeter energy.

$f_{\text{NleadClus}}$ The *energy fraction of the leading N clusters* is the fraction of calorimeter energy carried by the most energetic N topoclusters.

$N_{\text{track}}^{\text{iso}}$ The *number of isolation tracks* counts the tracks in the isolation annulus $0.2 \leq \Delta R < 0.4$.

S_{leadTrk} The *leading track impact parameter significance* is given by the leading track's impact parameter divided by the its uncertainty.

$S_{\text{T}}^{\text{flight}}$ The *secondary vertex decay length significance* is given by the decay length of the secondary vertex divided by its uncertainty, for candidates with more than one core track. Only core tracks are used in the secondary vertex fit.

ΔR_{max} The *maximum track distance* is given by the maximal distance between a core track and the tau axis.

$E_{\text{T}}^{\text{iso}}$ The *isolation energy* is defined as the sum of topocluster energies in the isolation annulus.

Distributions of a few powerful variables for signal and background are shown in Fig. 5.4. Detailed definitions are provided in Appendix A, along with figures for all variables.

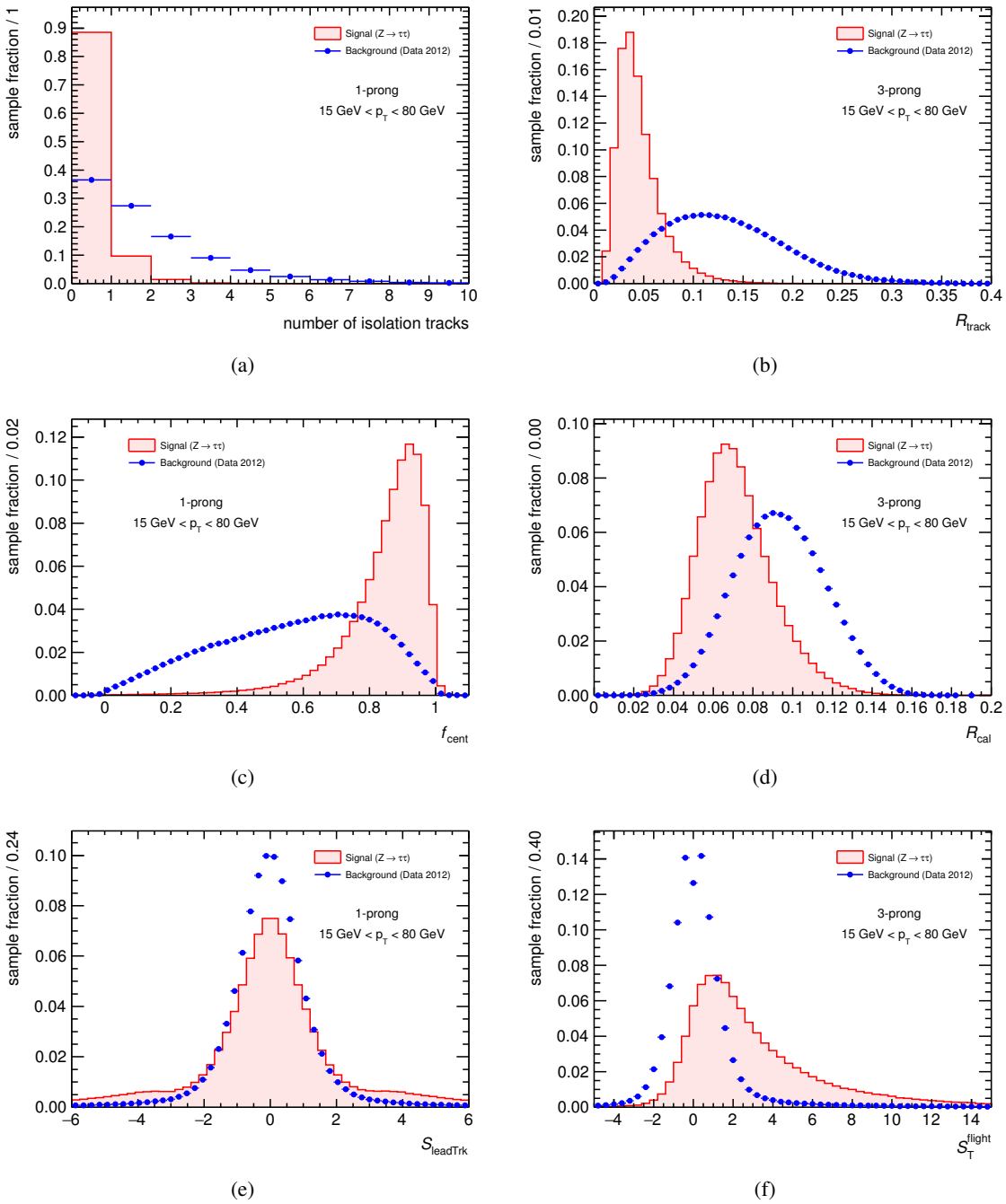


Figure 5.4: Selection of variables for tau identification highlighting the different sources of discrimination: tracks (top), energy deposits in the calorimeter (center) and Variables used for tau identification use are based on tracks (a,b), the calorimeter energy deposits (c,d) and decay length (c,d) Variables from used for the identification of hadronic tau decays: (a) number of isolation tracks, (b) track radius, (c) central energy fraction, (d) calorimeter radius, (e) impact parameter significance of the leading track and (f) secondary vertex decay length significance. The signal candidates are from simulated $Z \to \tau\tau$ events, with one track (1-prong) or three tracks (3-prong) within the cone $\Delta R < 0.2$, matched to a true hadronic tau decay with one or three charged hadrons, respectively. The background candidates are from multi-jet events in 2012 data, with one or three reconstructed tracks within the cone $\Delta R < 0.2$. All distributions are normalised to unity.

5.1.3. Identification methods

Three different algorithms are available to identify hadronically decaying tau leptons, which all make use of these discriminating variables: a cut-based identification and multivariate approaches using projective Likelihoods (LLHs) and boosted decision trees (BDTs).

Cut-based identification

The simplest method is the cut-based identification. The tau candidates have to pass a set of cuts on a set of discriminating variables. A relatively small number of 4-5 variables is selected. The thresholds can be defined momentum dependent[3], in order to obtain a p_{T} independent identification efficiency.

LLH-based identification

The LLH-based algorithm interprets the shape of the discriminating variable distributions as probability density functions (pdfs). For both signal and background, likelihood functions, $L_{\mathrm{S,B}}$, are defined as products of the single variable distributions. The final LLH score is then given by the log-likelihood ratio:

$$\text{LLH score} = \ln\left(\frac{L_{\mathrm{S}}}{L_{\mathrm{B}}}\right) = \sum_i \ln\left(\frac{p_i^{\mathrm{S}}(x_i)}{p_i^{\mathrm{B}}(x_i)}\right) \ , \tag{5.1}$$

where i runs over the input variables and $p_i^{\mathrm{S,B}}(x_i)$ is the signal (S) and background (B) probability density function of the discriminating variable x_i. The LLH score is a continous score. Identification working points are provided by defining score thresholds.

BDT-based identification

A decision tree is based on the repeated application of cuts. In contrast to a simple cut-based algorithm, objects are not discarded if they fail a cut, but are probed further. There are different ways to build a decision tree regarding the selection of variables, cut thresholds and the repeated growing of the tree (boosting). The focus of this description is on the settings used in the 2011 and 2012 tau ID.

As shown in Fig. 5.5, a decision tree is started at the root node, which contains all signal and background objects. All discriminating variables are tested in order to find the best cut to split the sample into objects passing or failing the cut. Two child nodes are created containing these objects. The best cut is the one yielding the highest separation gain. The separation is quantified by the *gini-index* $g = p(1-p)$, which is based on the node's signal fraction or *purity* p, i.e. the ratio of signal objects to all objects in the node. The separation gain is determined by comparing the separation of the root node (or parent node) and the joined separation of the child nodes. The latter is given by the sum of the children's gini-indizes weighted by the relative fraction of objects in each child node. One child node is characterised by a higher signal purity, while the other one is characterised by a lower signal purity. The process is then recursively called on the child nodes until at least one exit condition is fulfilled. Such exit conditions are for example the maximum depth of a tree, the total number of tree nodes or the minimum number of objects in a node. The purity of a final node determines the *score* of its contained objects. The nodes are classified as *signal nodes* or *background nodes*, for purities larger than or smaller than 0.5, respectively. This classification is used in the boosting step, in which the tree is regrown. Before the new tree is started, a weight is assigned to each object. Previously misclassified objects, i.e. signal objects in background nodes and vice versa, receive high weights, while correctly classified objects receive low weights. The tree is then trained on the reweighted sample,

[3] See for example Ref. [106].

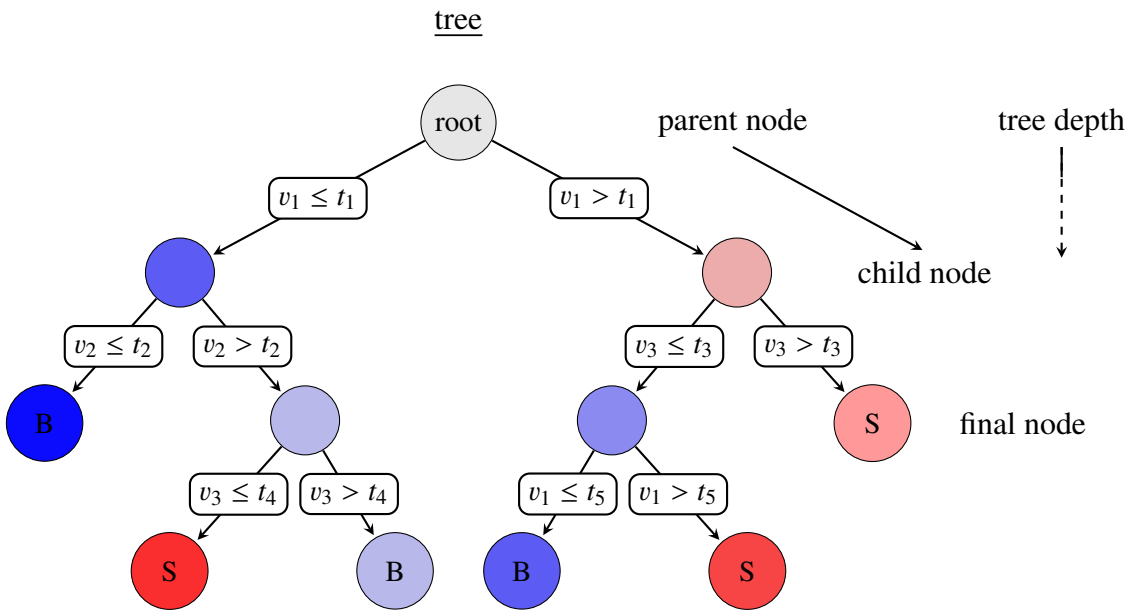


Figure 5.5: Sketch of a decision tree with three discriminating variables v_1, v_2, v_3, five cut thresholds t_1, ... , t_5 and a depth of three. The final nodes are classified as signal (S) or background (B) depending on the node's purity.

and hence focuses on the previously mis-classified objects. The boosting is typically performed a few ten to a few hundred times, until the error fraction or boosting weights stabilise. For each objects, the final BDT score is a weighted average of the scores obtained in each boosting step. By using a boosted tree, the score becomes less spiky and, more important, the overall performance increases. The building or growing of the trees is called the BDT *training*, while the performance evaluation is the *testing*. Independent samples are used for the two steps, in order to allow for overtraining checks. Overtraining describes the situation in which the BDT learns to discriminate background and signal based on statistical fluctuations in the training sample rather than on real trends. The performance seen in the training and testing samples would then differ significantly, and the overtraining eventually leads to a reduced overall discrimination strength. The BDT based tau identification uses variable sets of typically five to ten variables. Either the BDT score is used directly within an analysis, or predefined tau ID working points are used. These WP are created by defining BDT score thresholds that often depend on kinematic, event or decay quantities, such as p_{T}, N_{vtx} and N_{track} [107].

5.1.4. Defintions and terminology

The identification algorithms are prepared separately for 1-prong and 3-prong decays. The following definitions are used to classify reconstructed tau candidates:

- 1-prong candidates have one reconstructed track. Signal candidates are matched to true 1-prong decays.
- 3-prong candidates have three reconstructed tracks. Signal candidates are matched to true 3-prong decays.
- Multi-prong candidates have two or three reconstructed tracks. Signal candidates are matched to true 3-prong decays.

- Multi*-prong candidates have two or more reconstructed tracks. Signal candidates are matched to true 3-prong decays.

Only 1-prong and 3-prong candidates are used for the preparation (training) of the algorithms. The 3-prong algorithms are then used for the classification of all candidates with more than one track.

In order to validate the performance of the identification algorithms, the signal and background efficiencies are plotted as a function of several variables, mainly p_{T}, η, and N_{vtx}. The efficiency consists of two parts: the reconstruction efficiency, $\varepsilon_{\mathrm{reco}}$, and the identification efficiency, $\varepsilon_{\mathrm{ID}}$. The first one describes how many true simulated hadronic tau decays are reconstructed as tau candidates, and is therefore defined for signal candidates only. It is driven to a large extend by the track reconstruction and association efficiency. The identification efficiency describes how many reconstructed tau candidate pass a specific identification algorithm. It is therefore entirely driven by the performance of the ID algorithms. The definitions are as follows:

$$\varepsilon_{\mathrm{reco+ID}} = \varepsilon_{\mathrm{reco}} \times \varepsilon_{\mathrm{ID}} \ ,$$

$$\varepsilon_{\mathrm{reco}} = \frac{\text{\# reconstructed n-prong tau candidates}}{\text{\# true n-prong decays}} \ ,$$

$$\varepsilon_{\mathrm{ID}} = \frac{\text{\# reconstructed n-prong tau candidates passing ID}}{\text{\# reconstructed n-prong tau candidates}} \ .$$

For background candidates $\varepsilon_{\mathrm{reco}}$ is set to one and therefore $\varepsilon_{\mathrm{reco+ID}} = \varepsilon_{\mathrm{ID}}$. Instead of the background efficiency often the background rejection is plotted, which is defined as the inverse of the efficiency.

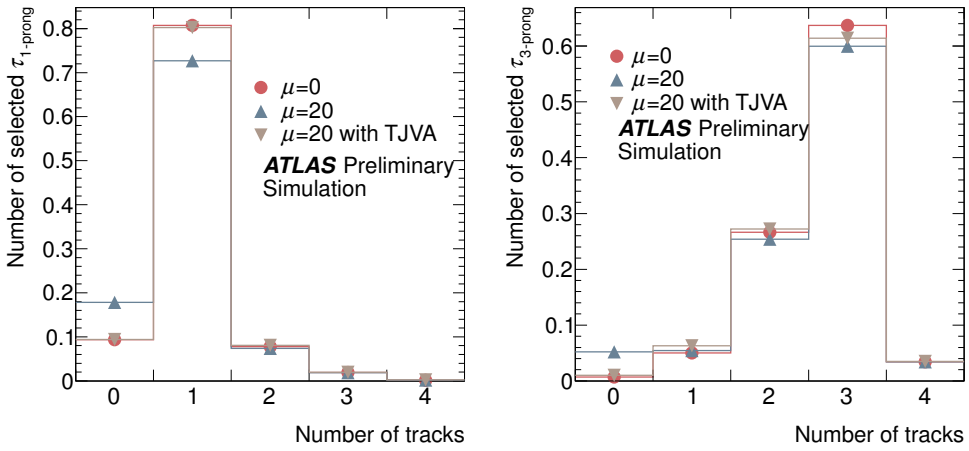
5.2. Reconstruction and identification of hadronic tau decays in 2011 data

The tau reconstruction is initialised by anti-k_t jets (parameter $R = 0.4$) which are constructed from topoclusters calibrated at LC scale. The jet has to fulfill $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ and $|\eta| < 2.5$. The tau direction and energy is determined from the four-vector sum of the clusters within $\Delta R < 0.2$ around the jet barycenter (jet axis)[4]. Tracks are associated to the tau candidate, if they are within $\Delta R < 0.2$ (*core tracks*) or within $0.2 \leq \Delta R < 0.4$ (*isolation tracks*) around the jet axis. The tracks have to fulfill the following criteria:

- $p_{\mathrm{T}} > 1\,\mathrm{GeV}$,
- number of PIX hits ≥ 2,
- number of PIX hits + number of SCT hits ≥ 7,
- $|d_0| < 1.0\,\mathrm{mm}$,
- $|z_0 \sin\theta| < 1.5\,\mathrm{mm}$.

The distances of closest approach in the transverse plane, d_0, and in the longitudinal plane, z_0, are measured between the track and the vertex with the highest jet vertex fraction. The *jet vertex fraction*,

[4] An energy calibration is applied that is optimised for hadronically decaying tau leptons. In Sec. 5.4.3, the 2012 version is described. For details concerning the 2011 version, see Ref. [108].



(a) (b)

Figure 5.6: Number of tracks associated to a tau candidate, that is matched to a true hadronic tau decay with (a) one charged hadron and (b) three charged hadrons, in events without (μ=0) and with (μ=20) pile-up. The tracks quality is either measured with respect to the first primary vertex in the event (circles, triangles) or with respect to the vertex with the highest jet vertex fraction (up-side-down triangles, denoted 'with TJVA'). The tau candidates have a momentum of at least 15 GeV and are taken from simulated $Z \rightarrow \tau\tau$ events [109].

f_{JVF}, is defined as the fraction of track momentum of a jet that is associated to a given vertex:

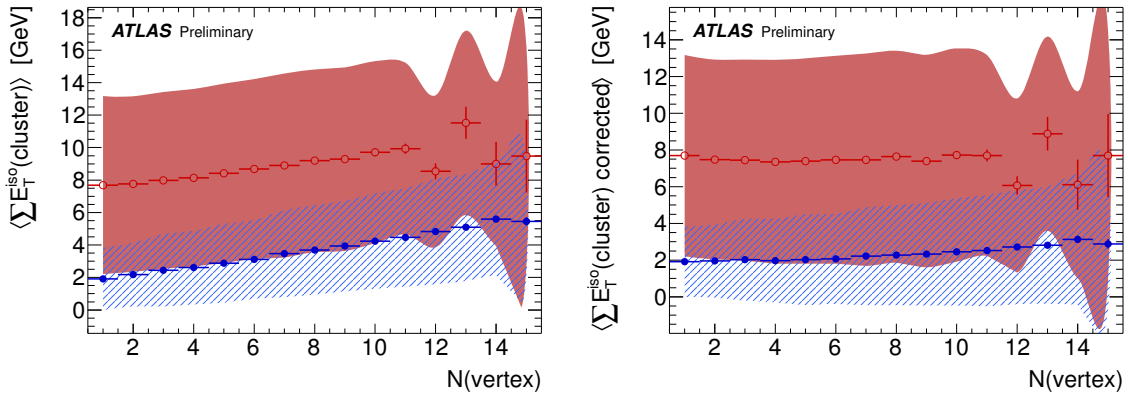
$$f_{\text{JVF}}(\text{jet}|\text{vtx}) = \frac{\sum p_{\text{T}}^{\text{trk}|\text{vtx}}}{\sum p_{\text{T}}^{\text{trk}}} \quad . \tag{5.2}$$

The index 'trk' refers to the tracks in the core region of the jet seed that satisfy the first three requirements in the list above, and 'trk|vtx' is the subset of tracks that is matched to the given vertex. Using the vertex with the highest jet vertex fraction instead of the first primary vertex in the event increases the pile-up robustness of the track association. This is illustrated in Fig 5.6 for simulated true hadronic tau decays. For high pile-up events, more often tracks are disregarded if the first primary vertex is used (default in 2010), and hence fewer tracks are associated to the tau candidates. The f_{JVF} approach reduces the effect significantly, especially in the 1-prong case [109].

The discriminating variables are calculated on the reconstructed tau candidates. The associated core and isolation tracks are used, as well as the associated topoclusters. Instead of using the clusters directly, some variables are based on the cells associated to the topoclusters. The decay cone size is directly dependent on the momentum of the decaying tau lepton, the higher the boost, the smaller is the cone. Hence, variables relying on fixed cone sizes will often show a momentum dependence, and highly boosted tau candidates appear more 'tau like'. If these variables are used in an identification algorithm, they can cause a momentum dependent efficiency. The ID algorithms use different p_{T}-flattening procedures to circumvent this. Furthermore, the discrimination variables are influenced by the pile-up conditions of the event. The additional activity in the event leads to more reconstructed topoclusters, but also adds energy to existing clusters. On average, pile-up leads to a more 'background like' appearance of the tau decay.

Table 5.2: Variables used by the 2011 and 2012 identification algorithms. The star (⋆) indicates variables that are redefined for the 2012 BDT ID: the calorimeter cone size is reduced to $\Delta R < 0.2$ and an additional pile-up correction term is added. The open circles (○) indicate variables that were added later to be part of the final 2012 ID. The definition of all variables is given in Appendix A.

	Cuts 2011		BDT 2011		LLH 2011		BDT 2012		LLH 2012		
	1p	3p	1p	3p	1p	3p	1p	3p	1p	3p	
R_{track}	•	•	•	•	•	•	•	•	•	•	track variables
$N^{\text{iso}}_{\text{track}}$	•	•	•	•	•		•		•		track variables
S_{leadTrk}			•	•			•		•		track variables
$S_{\text{T}}^{\text{flight}}$		•		•		•		•		•	track variables
ΔR_{max}				•		•		•		•	track variables
m_{tracks}				•		•		•		•	track variables
f_{track}	•	•	•	•			⋆	⋆	⋆	⋆	mixed
f_{cent}			•	•	•	•	⋆	⋆	⋆	⋆	calorimeter variables
R_{cal}			•	•	•						calorimeter variables
$f_{\text{2leadClus}}$					•						calorimeter variables
$f_{\text{3leadClus}}$			•	•							calorimeter variables
m_{effClus}			•	•							calorimeter variables
$E^{\text{iso}}_{\text{T,corr}}$	•	•									calorimeter variables
N_{π^0}							○	○			substructure variables
$p_{\text{T}}^{\pi^0\text{+tracks}}/p_{\text{T}}$							○	○			substructure variables
$m_{\pi^0\text{+tracks}}$							○	○			substructure variables



(a) (b)

Figure 5.7: The dependence of the isolation cone energy on the number of vertices for (a) the uncorrect version $E_{\mathrm{T}}^{\mathrm{iso}}$ and (b) the pile-up corrected version $E_{\mathrm{T,corr}}^{\mathrm{iso}}$ for 1-prong tau candidates. The points indicate the mean in each bin and the bands indicate the standard deviation. The filled blue points are signal candidates from simulated $Z \to \tau\tau$ and $W \to \tau\nu$ events, while the open red points are background candidates from a dijet sample in 2011 data [110].

5.2.1. Cut-based identification

The cut-based approach uses a small set of variables, as listed in Table 5.2. The focus is on track variables to identify narrow and isolated showers. The track variables are relatively pile-up robust, due to the track impact parameter requirements. This is however not the case for the main calorimeter variable, $E_{\mathrm{T}}^{\mathrm{iso}}$, which quantifies the cluster energy deposited in the calorimeter isolation cone. A correction procedure is applied to reduce the effect of pile-up contributions, by making use of the jet vertex fraction of the tau lepton candidate in terms of its vertex[5]. An estimation of the local pile-up contribution from charged particles to the tau candidate is obtained by

$$p_{\mathrm{T}}^{\text{pile-up}} = (1 - f_{\mathrm{JVF}}) \sum(\text{track } p_{\mathrm{T}}) \ . \tag{5.3}$$

In order to determine the contribution to the isolation cone alone and to account for neutral contributions, a calibration parameter α is introduced to scale $p_{\mathrm{T}}^{\text{pile-up}}$. After tests with different values for α, a value of 1 was chosen. With an upper limit of the correction of 4 GeV, the pile-up corrected isolation cone variable $E_{\mathrm{T,corr}}^{\mathrm{iso}}$ is defined as

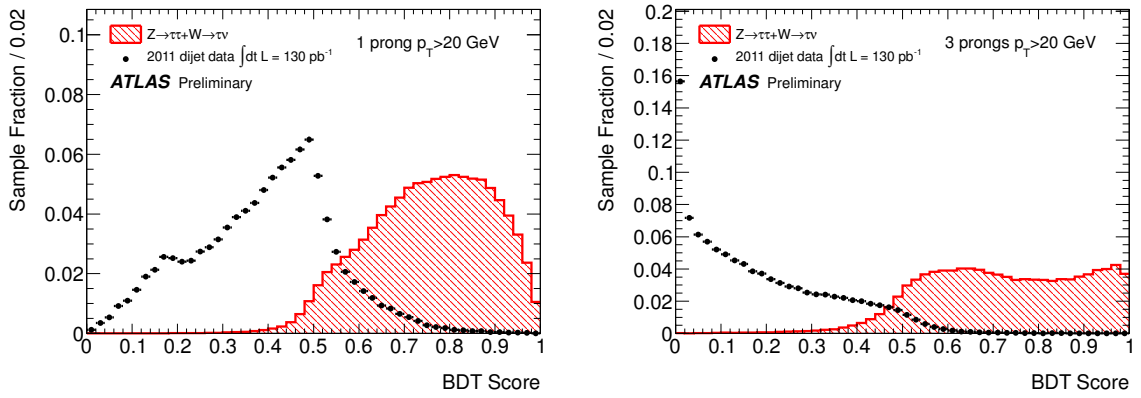
$$E_{\mathrm{T,corr}}^{\mathrm{iso}} = E_{\mathrm{T}}^{\mathrm{iso}} - \min\Big(\alpha(1 - f_{\mathrm{JVF}}) \sum(\text{track } p_{\mathrm{T}}),\ 4\,\mathrm{GeV} \Big) \ . \tag{5.4}$$

Three signal efficiency working points are defined and the optimal cuts are obtained separately for candidates with one track and more than one track. The cuts on the track radius are parameterised in p_{T} in order to yield approximately flat efficiencies [106, 110].

5.2.2. Boosted Decision Tree identification

The 2011 BDTs are trained separately for candidates with one track or three tracks, and for events with 1–5 or >5 primary vertices. The variable set is a broad mixture of variables, describing the

[5] At the time when the pile-up correction was developed, the track quality was determined with respect to the first primary vertex in the event. Consequently, also the jet vertex fraction was determined with respect to the first primary vertex. However, the differences are small and of no importance here.



(a) (b)

Figure 5.8: BDT score for (a) 1-prong and (b) 3-prong tau candidates for the 2011 BDT tau ID. The signal candidates are from simulated $Z \to \tau\tau$ and $W \to \tau\nu$ samples, and the background candidates are from a dijet sample selected from 2011 data.

shower shape and energy distribution as well as the track impact parameters. Furthermore, invariant masses of clusters and tracks are included, as listed in Table 5.2. The BDT score is shown in Fig. 5.8. Cut thresholds on the BDT score are defined targeting signal efficiencies of 60%, 50% and 30% for 1-prong candidates, and 65%, 55% and 35% for multi*-prong candidates. In order to compensate for the p_{T} dependence of the BDT score, the thresholds are defined as a function of momentum. Furthermore, the score is influenced by the pile-up conditions. The p_{T} dependent thresholds are therefore obtained in bins of the number of primary vertices: 1–3, 4–5, 6–7 and ≥ 8. The signal efficiency as a function of the number of vertices is shown in Fig. 5.9. Within the finely tuned region ($N_{\mathrm{vtx}} < 8$), the efficiency shows only minor bin-by-bin fluctuations. The high pile-up region ($N_{\mathrm{vtx}} \geq 8$) gives a notion of the underlying pile-up dependence of the BDT score [109, 110].

5.2.3. Likelihood identification

The LLH functions are created separately for candidates with one track or three tracks and with a momentum of 0–45 GeV, 45–100 GeV or >100 GeV, and for events with 1–3, 4–7 or ≥8 primary vertices. All variables used in the cut-based and BDT approach are studied during the optimisation of the LLH identification. Any variables contributing less than a few percent to the overall performance are taken out of the final set, in order to reduce potential sources of systematic uncertainties. Five variables per prong category remain, as listed in Table 5.2. In order to obtain a smooth likelihood score in momentum despite the three momentum bins, a linear extrapolation is applied close to the bin borders. Three working points are defined aiming at the same signal efficiencies as the BDT identification. The LLH score thresholds are obtained as a function of the tau momentum to yield a flat signal efficiency. Furthermore, the thresholds are adjusted for different values of N_{vtx} [106, 109, 110]. The signal efficiency for the medium working point is shown in Fig. 5.9.

5.3. BDT optimisation studies: pile-up dependence

As mentioned in Sec. 5.2.2, the identification with BDTs shows a pile-up dependence. Two steps are taken to reduce the effect on the signal efficiency: Firstly, the BDT is trained in two regions of the number of vertices (N_{vtx}), and secondly, the working points are defined in four bins of N_{vtx}.

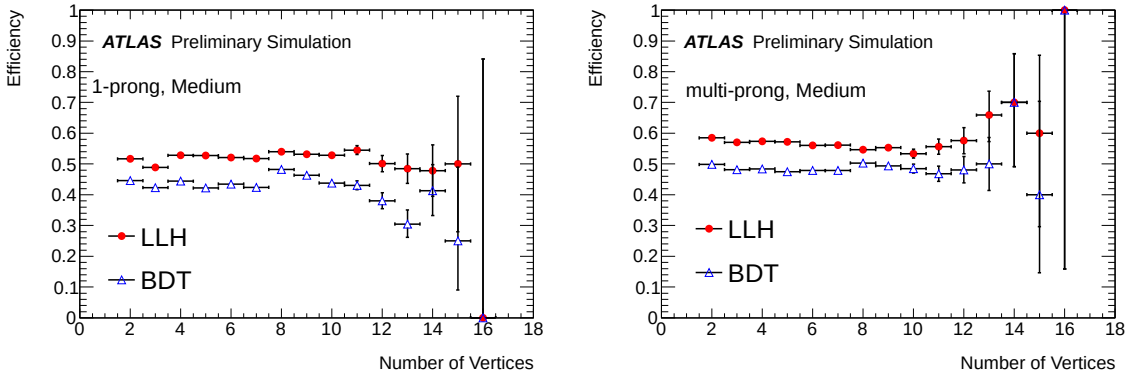


Figure 5.9: Signal efficiency $\varepsilon_{\text{reco+ID}}$ for the medium identification working point as a function of the number of vertices for (a) 1-prong candidates and (b) multi*-prong candidates. The behaviour of the BDT is shown as unfilled triangles, while the behaviour of the LLH approach is depicted with circles [109].

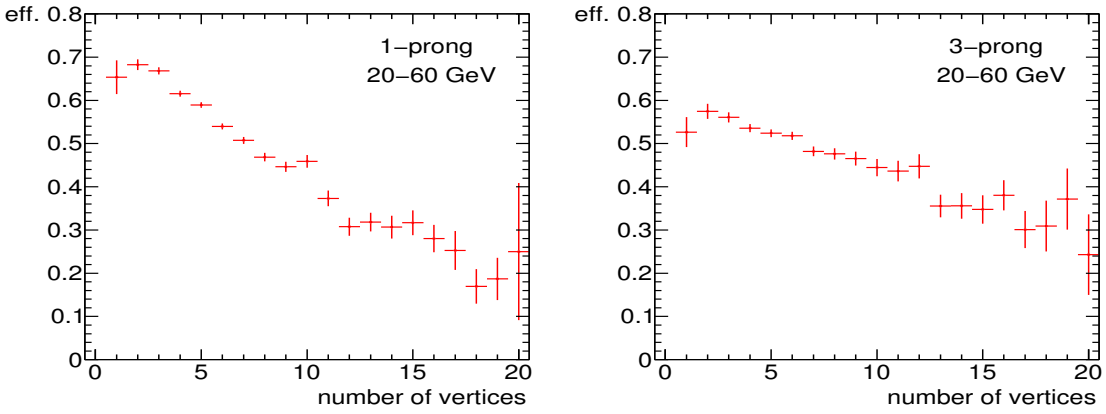
Figure 5.9 shows the signal efficiency for the medium WPs. The remaining dependence is small and only apparent in the high N_{vtx} region. Without the steps taken, the BDT score is however significantly influenced by the pile-up conditions, as shown in Fig. 5.10. The goal of this study is a reduction of the dependence of the BDT score on the pile-up conditions. The study focuses on the performance in the momentum region 20–60 GeV where the dependence is most critical and most apparent. Only the identification efficiency ε_{ID} is under investigation.

5.3.1. Samples and setting

Studies on tau identification are performed per object, meaning that not entire events are studied but single tau candidates. The signal candidates are predominantly taken from simulated $Z \to \tau\tau$ events, sometimes mixed with $W \to \tau\nu$ events. Simulated $Z' \to \tau\tau$ events can be added to increase the p_{T} spectrum on the high momentum tail[6]. Simulated *true* hadronic tau decays have to fulfil $p_{\text{T,vis}}^{\text{true}} > 10\,\text{GeV}$, $|\eta_{\text{vis}}^{\text{true}}| < 2.5$ and have one or three charged hadrons. The subscript 'vis' indicates the visible parts of the decay, ignoring the neutrino(s). Reconstructed hadronic tau decays are considered as signal candidates if they are within $\Delta R < 0.2$ of the visible position of the true decay (*truth-matched*). They have to be well-reconstructed (satisfying certain quality criteria), be within $|\eta| < 2.5$, satisfy $p_{\text{T}} > 15\,\text{GeV}$ and have at least one core track. Background candidates are predominantly taken from data. A set of jet triggers selects a sample enhanced in multi-jets events with prescales tuned to give a smooth and falling jet p_{T} spectrum. In the early studies, a simulated sample of dijet events is used. The results are evaluated using data events, and no significant differences are found. The same quality criteria and kinematic requirements of $|\eta| < 2.5$, $p_{\text{T}} > 15\,\text{GeV}$ are applied to the background candidates. Furthermore, the candidates are required to have at least one core track. Both the reconstructed signal and background tau objects will be denoted 'tau candidates' here, or the symbolic notation τ_h will be used.

If Z' events are included, the p_{T} spectra of signal and background candidates differ significantly. The signal objects are then reweighted so that the p_{T} spectra match. Due to the falling jet momentum

[6] A Z' boson is a hypothetical particle with the same properties as the SM Z boson, except for a higher mass. Samples with Z' masses of 250 GeV, 500 GeV, 750 GeV, 1000 GeV and 1250 GeV are used here and are thus a source of tau candidates with much higher momentum.



(a) (b)

Figure 5.10: Signal efficiency $\varepsilon_{\mathrm{ID}}$ as a function of the number of vertices for(a) 1-prong τ_h and (b) 3-prong τ_h. The cut on the BDT score is neither tuned in p_{T} nor N_{vtx}. It targets an overall 50% signal identification efficiency. The BDT is trained inclusive for all pile-up conditions using the same variables and parameters as the 2011 BDT tau ID.

spectrum, the high p_{T} candidates receive low weights and the sample as a whole is dominated by low p_{T} tau leptons.

The BDT training options are taken over from the 2011 training. The toolkit TMVA [107] is used for the BDT training. As described before, the BDT score is defined by the final node's purity and the separation is quantified using the gini-index. Variables are divided into 500 bins for the cut threshold scan. The minimum number of objects in a node defines the exit condition and is dependent on the total number of training objects N_{train} and number of variables N_{var} as $\max(40, N_{\mathrm{train}}/(N_{\mathrm{var}}^2)/10)$. Arbitrary large trees are allowed in terms of the number of nodes and the tree depth. AdaBoost [111, 112] is used for the decision tree boosting, with the learning rate parameter $\beta = 0.2$.[7] The purity threshold to classify a node as signal or background is $p = 0.5$, and the total number of trees is set to 50.

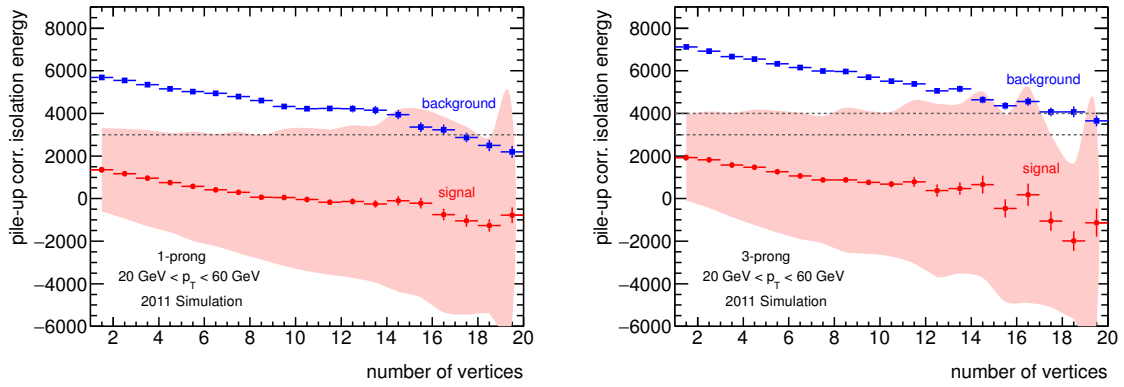
5.3.2. Factorisation of calorimeter core and isolation region

The first approach is to follow the cut-based algorithm and to make use of the pile-up corrected isolation cone energy. The idea behind this is to factorise the calorimeter variables into the core region and the isolation region. By being redefined in a smaller cone, the calorimeter variables gain pile-up robustness. The discrimination power of the calorimeter isolation cone is concentrated in one variable, which is corrected for pile-up. Because the isolation cone appears in several variables in the BDT ID, the approach subdivides into two tasks:

- All calorimeter variables have to be refined in the core cone ($\Delta R < 0.2$).
- The variable $E_{\mathrm{T,corr}}^{\mathrm{iso}}$ needs to be incorporated into the identification process.

Results of preparatory studies indicate to using the number of vertices N_{vtx} or the mean number of interactions per crossing $\langle\mu\rangle$ as a measure of the global pile-up activity instead of the local measure

[7] A small parameter β reduces the strength of the reweighting and thereby slows down the boosting procedure.



(a) (b)

Figure 5.11: Pile-up corrected isolation energy $E^{\text{iso}}_{\text{T,corr}}$ as a function of N_{vtx} for (a) 1-prong and (b) 3-prong tau candidates. The points represent the mean value in each bin, while the filled area shows the standard deviation (signal only). For a cut on the variable targeting a pile-up independent signal efficiency of about 85%, the parameter α is set to 500 MeV. The cut thresholds are indicated by the dashed lines, at 3 GeV for 1-prong candidates and 4 GeV for 3-prong candidates. A cut on 3 GeV on 3-prong candidates leads to an approximately flat efficiency of about 80%.

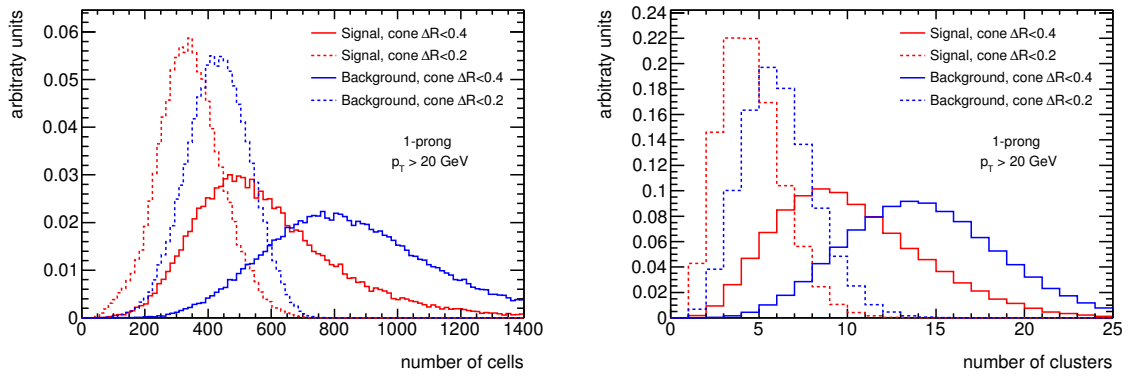
f_{JVF} [113]. Both N_{vtx} and $\langle\mu\rangle$ performed better especially in cases of very low ($\approx$0) and very high ($\approx$1) values of f_{JVF}. Therefore, the following definition of $E^{\text{iso}}_{\text{T,corr}}$ is used in this study:

$$E^{\text{iso}}_{\text{T,corr}} = E^{\text{iso}}_{\text{T}} - \alpha * N_{\text{vtx}} \quad , \tag{5.5}$$

with the correction strength parameter α. The variable $E^{\text{iso}}_{\text{T,corr}}$ is shown in Fig. 5.11 as a function of $N_{\text{vtx}}$for a parameter $\alpha = 500$ MeV. This value gives an approximately flat efficiency for a cut on $E^{\text{iso}}_{\text{T,corr}}$ targeting 85% signal efficiency.

Concerning the redefinition of the calorimeter variables, it is important to monitor the change in discrimination power, in order to avoid a large performance loss. Each calorimeter variable is therefore evaluated before and after the factorisation: a cut on $E^{\text{iso}}_{\text{T,corr}}$ is applied and the signal efficiency versus background rejection of the original and redefined variable is compared. Five BDT variables are dependent on the cone size: R_{cal}, f_{track}, f_{cent}, m_{effClus} and $f_{\text{3leadClus}}$. The first three rely on the sum of cell energies in the cone $\Delta R < 0.4$. Changing the cone size to $\Delta R < 0.2$ results in a smaller number of cells available for the calculation, as shown in Fig. 5.12(a). The mean number of cells reduces from $\langle n_{\text{cells}}\rangle \approx 580$ to $\langle n_{\text{cells}}\rangle \approx 340$ for 1-prong signal candidates. Also in the smaller cone, the number of cells is sufficiently large not to disturb the calculability of the variables as such. The variables m_{effClus} and $f_{\text{3leadClus}}$ are based on clusters. The mean number of clusters decreases from $\approx$10 in $\Delta R < 0.4$ to $\approx$4 in $\Delta R < 0.2$ for 1-prong signal candidates (Fig. 5.12(b)). Both for signal and background, the leading cluster energy ratio $f_{\text{3leadClus}}$ is more often at its limit of one, which reduces the discrimination power significantly. The ratio can be redefined to either taking into account the leading one or two clusters only, or any number of leading cells. Different versions of the variable are tested. Also the effective cluster mass m_{effClus} is pushed more often to its limit, even though the effect is less dramatic. For cases with one cluster only, the mass reaches its lower limit of zero. The problem persists when the cluster mass m_{clus} is used instead, which includes all clusters in the calculation[8]. Both versions of

[8]The effective cluster mass focuses on the high energetic clusters. In case of two clusters only, both are included, so that there is no difference in m_{effClus} and m_{clus} in the case of very few clusters.



(a) (b)

Figure 5.12: Number of (a) cells and (b) clusters for the two cone sizes $\Delta R < 0.4$ (continuous line) and $\Delta R < 0.2$ (dashed line) for 1-prong signal and background candidates. Only cells associated to clusters are considered. The signal candidates are from simulated $Z \rightarrow \tau\tau$ decays, and the background candidates are from a simulated dijet sample. The distributions are normalised to unity.

the mass are included in the study.

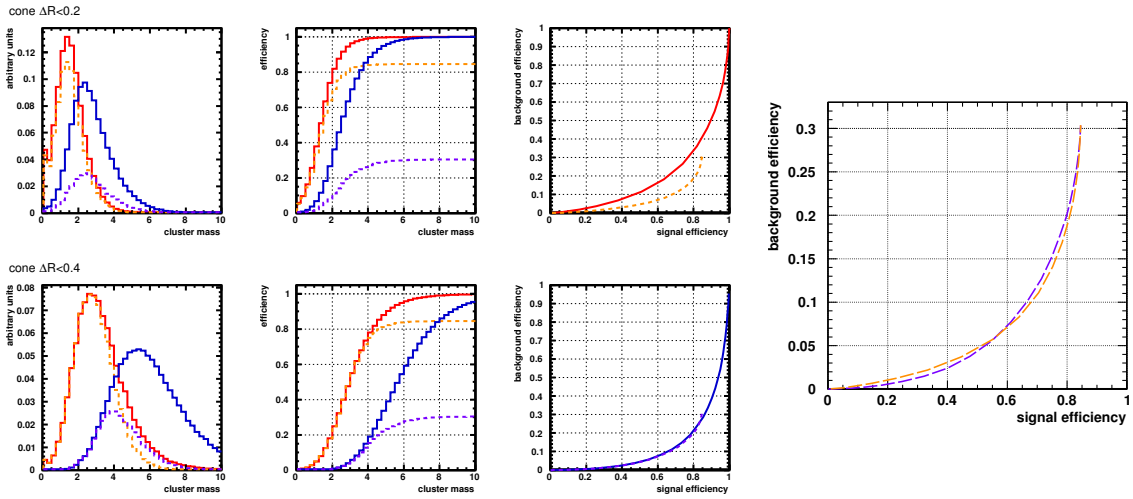
An example of the evaluation plots is shown in Fig. 5.13 for the variables m_{clus} and f_{cent}. The veto on $E^{\text{iso}}_{\text{T,corr}}$ is chosen to yield ≈85% signal efficiency independent of the pile-up conditions, which is achieved by setting $\alpha = 500\,\text{MeV}$. The cluster mass defined in the 0.2 cone alone has a lower discrimination power compared to the original variable. However, in combination with a cut on $E^{\text{iso}}_{\text{T,corr}}$, the two variables have an equivalent performance. The performance of the central energy fraction is much less affected by the redefinition. The small loss is nearly fully compensated by the cut on $E^{\text{iso}}_{\text{T,corr}}$. The other variables show a similar behaviour. The discrimination power of the single, redefined variable decreases, but the loss is very small in combination with a cut on $E^{\text{iso}}_{\text{T,corr}}$. Figure 5.14 shows that the calorimeter variables indeed gain pile-up independence by being redefined in the core cone. For the sum of cell energies, the pile-up induced energy surplus is reduced by about a factor ¼ in the smaller cone. The factor corresponds to the reduction in the surface area of the two cones, and behaves therefore as expected for uniformly distributed energy deposition from pile-up[9]. As mentioned, the sum of cell energies is part of several calorimeter variables, which gain pile-up dependence accordingly.

I think you could express this a little stronger by saying that the factor 1/4 roughly corresponds to the reduction in calorimeter surface area of the two cones. And that this is what you would expect for uniformly distributed PU.

Two ways are tested to incooperate $E^{\text{iso}}_{\text{T,corr}}$ into the BDT ID: a veto on the corrected isolation energy prior to the BDT, and using the variable as a discriminator within the BDT. The first approach is a continuation of the variable evaluation procedure. The BDT is trained and tested on the signal and background events passing the cut on $E^{\text{iso}}_{\text{T,corr}}$. This way offers maximal control and allows to tune the pile-up correction parameter in order to give a pile-up independent veto efficiency. The second approach exploits the BDT strength by leaving it to the training procedure to choose the optimal cut. The BDTs are trained and evaluated separately for 1-prong and 3-prong candidates, but inclusive

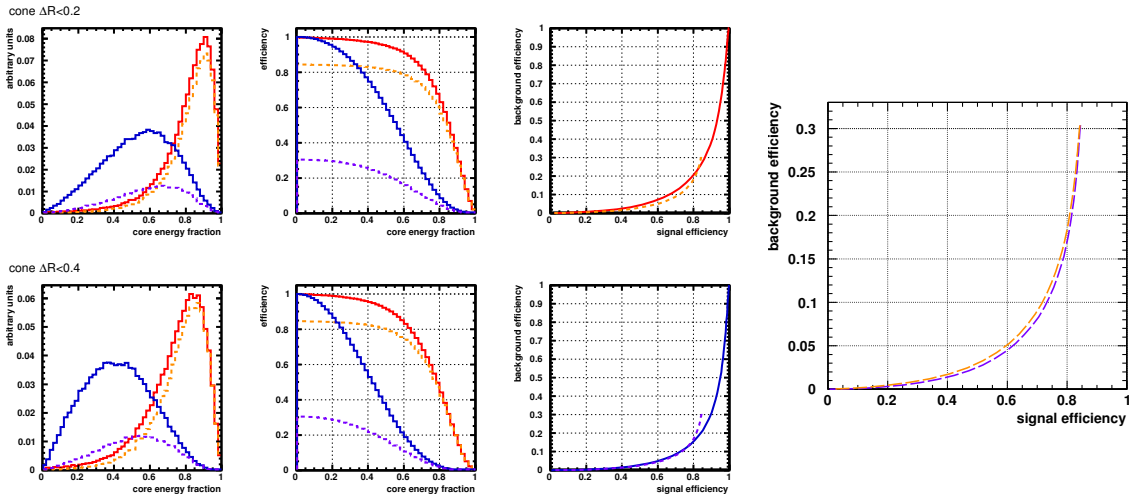
[9] It is observed that this expectation does not hold for arbitrary small cone sizes. First of all, the tau candidate is built from topoclusters, meaning that not all cells within the cone contribute to the discriminating variables. Secondly, contributions from pile-up are already surpressed during the topocluster reconstruction.

I: cluster mass



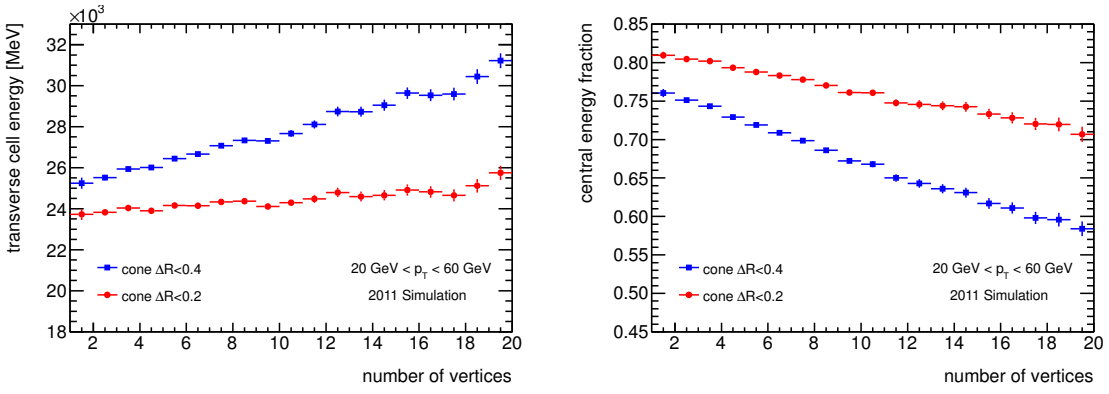
(a) (b) (c) (d)

II: central energy fraction



(a) (b) (c) (d)

Figure 5.13: Evaluation of the redefined variables cluster mass m_{clus} (I) and central energy fraction f_{cent} (II). The top row shows in each case the behaviour of the variable defined in $\Delta R < 0.2$, while the bottom row shows the original variable. The continues lines illustrate the behaviour without a cut on $E_{\text{T,corr}}^{\text{iso}}$, while the dashed lines show the behaviour with the cut applied. (a) Normalised distribution for signal candidates (red) and background candidates (blue). (b) Efficiency as a function of a cut value for signal (red) and background (blue). The efficiency is calculated from the signal side, i.e. in the range (0, cut value) for m_{clus} and (cut value, 1) for f_{cent}. (c) Signal vs background efficiency. (d) Signal vs background efficiency for the case with cut on $E_{\text{T,corr}}^{\text{iso}}$ only, comparing the redefined (orange) with the original (purple) variable (zoom of (c)). The signal candidates are from simulated $Z \rightarrow \tau\tau$ decays, and the background candidates are from a simulated dijet sample. The candidates have at least one track.



(a) (b)

Figure 5.14: Comparison of the pile-up dependence of variables defined in the cone $\Delta R < 0.4$ versus the cone $\Delta R < 0.2$ for signal tau candidates. (a) Sum of transverse cell energies belonging to the topoclusters of a tau candidate. The variable represents the denominator of the variables f_{cent}, $R_{\text{cal}}$and f_{track}. (b) Central energy fraction f_{cent}, defined as the energy fraction within cone $\Delta R < 0.1$.

in N_{vtx}. Three BDTs are compared: a BDT using the same variables as in 2011, a BDT using the redefined calorimeter variables and applying a veto on $E^{\text{iso}}_{\text{T,corr}}$, and a BDT using the variable $E^{\text{iso}}_{\text{T,corr}}$ in addition to the redefined calorimeter variables. The track variables remain unchanged in all cases. To begin with, an identical parameter $\alpha = 500\,\text{MeV}$ is chosen for both the veto and the implementation of $E^{\text{iso}}_{\text{T,corr}}$ into the BDT. As shown in Fig. 5.11, this leads to a flat veto efficiency, but overcompensates the average pile-up contribution of the variable. The result of the BDT training is shown in Fig. 5.15. The background rejection is maintained, especially when $E^{\text{iso}}_{\text{T,corr}}$ is part of BDT variable set. The veto leads to reduced performance for high signal efficiencies, which is especially apparent in the 3-prong case for the tighter veto on 3 GeV. The pile-up dependence is significantly reduced in either approach. The highest improvement shows the 1-prong case where $E^{\text{iso}}_{\text{T,corr}}$ is treated as an input variable. Studies with different parameters α show that the flattness is highly dependent on the parameter chosen. This is illustrated in Fig. 5.16, where α is chosen so that the mean values of the variable are independent of N_{vtx}. In terms of a flat BDT signal efficiency the performance is then compatible with the veto approach. Additional studies are performed with a varying veto strength and a varying definition of $E^{\text{iso}}_{\text{T,corr}}$ (using for instance the original definition with f_{JVF}), but do not lead to significantly better results.

Summarising, the factorisation of isolation and core cone in addition to the pile-up treatment of the isolation energy lead to a reduced pile-up dependence of the BDT tau ID. The implementation of $E^{\text{iso}}_{\text{T,corr}}$ gives better results in terms of background rejection compared to the veto, but reacts on the other hand more sensitive to changes of the correction strength parameter α. While the improvement in terms of pile-up dependence is potentially large enough for an inclusive training of the BDT in N_{vtx}, the procedure of flattening the tau ID working points in bins of N_{vtx} remains necessary. The question remains whether it is possible to further reduce the pile-up dependence of the BDT score. This could be done by sorting out the pile-up dependent variables or by introducing pile-up terms to additional variables. The question is investigated in the next section.

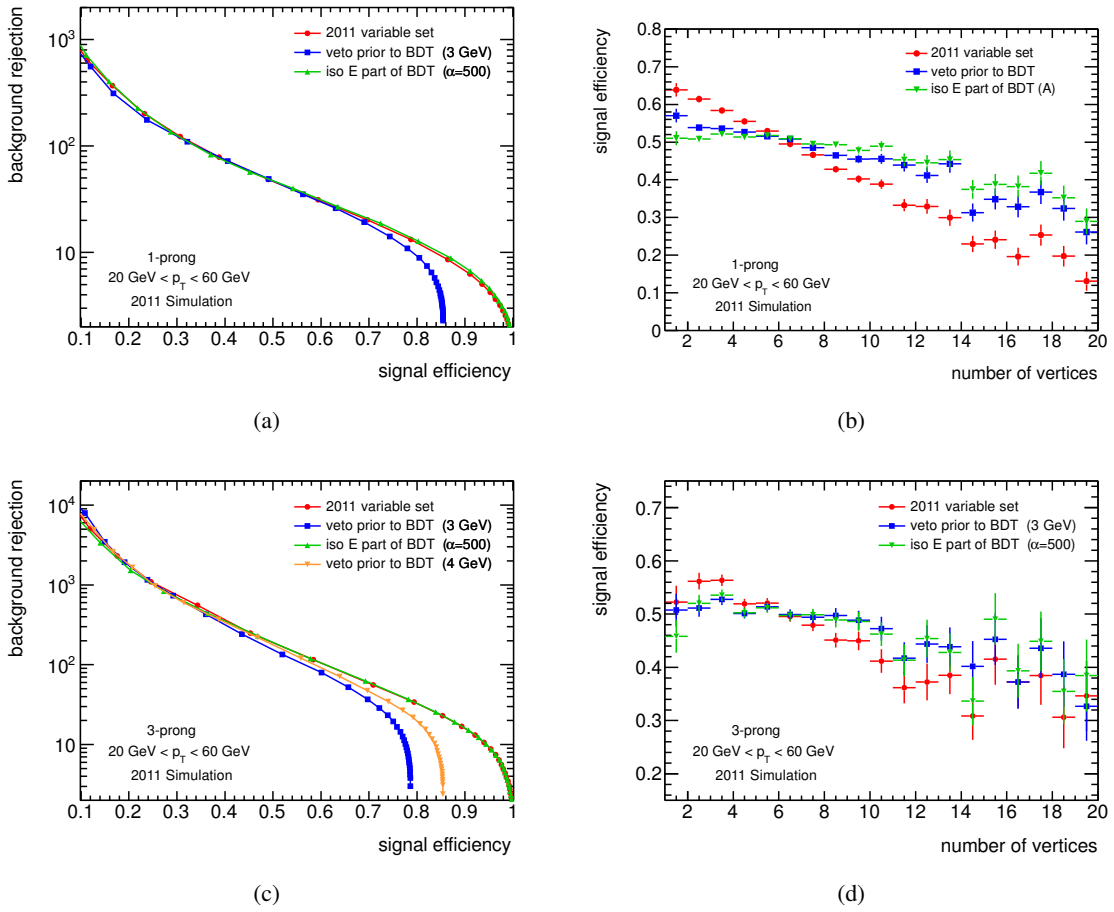
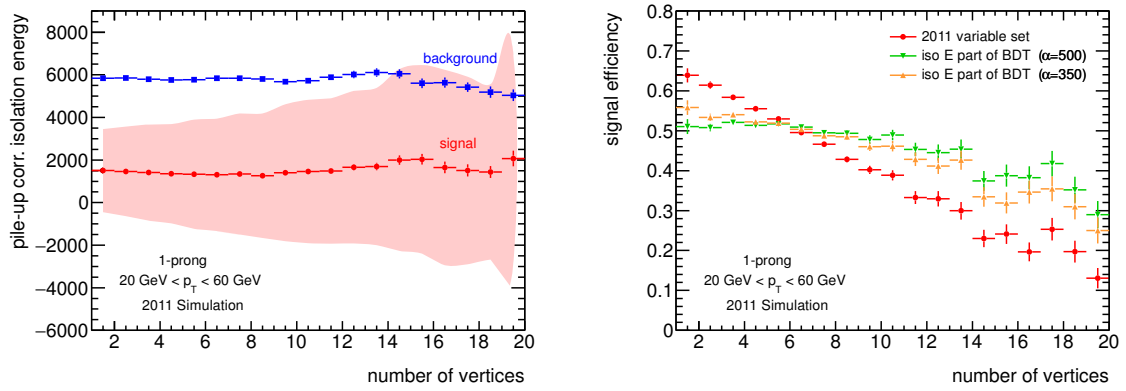


Figure 5.15: Performance comparison for the three BDTs for 1-prong (top) and 3-prong (bottom) candidates. (a,c) Background vs signal efficiency $\varepsilon_{\mathrm{ID}}$. The end-point of the orange curve represents the signal and background efficiency of the veto applied prior to the BDT. (b,d) Signal and background efficiency $\varepsilon_{\mathrm{ID}}$ as a function of N_{vtx} for a flat cut on the BDT score targeting an overall signal efficiency of 50%. The signal candidates are from simulated $Z \to \tau\tau$ and $W \to \tau\nu$ events, and the background candidates are from a simulated dijet sample.



(a) (b)

Figure 5.16: Performance comparison for the three 1-prong BDTs. (a) Background vs signal efficiency $\varepsilon_{\mathrm{ID}}$. The end-point of the orange curve represents the signal and background efficiency of the veto applied prior to the BDT. (b) Signal and background efficiency $\varepsilon_{\mathrm{ID}}$ as a function of N_{vtx} for a flat cut on the BDT score targeting an overall signal efficiency of 50%. The signal candidates are taken from simulated $Z \to \tau\tau$ and $W \to \tau\nu$ events and background candidates are taken from simulated dijet events.

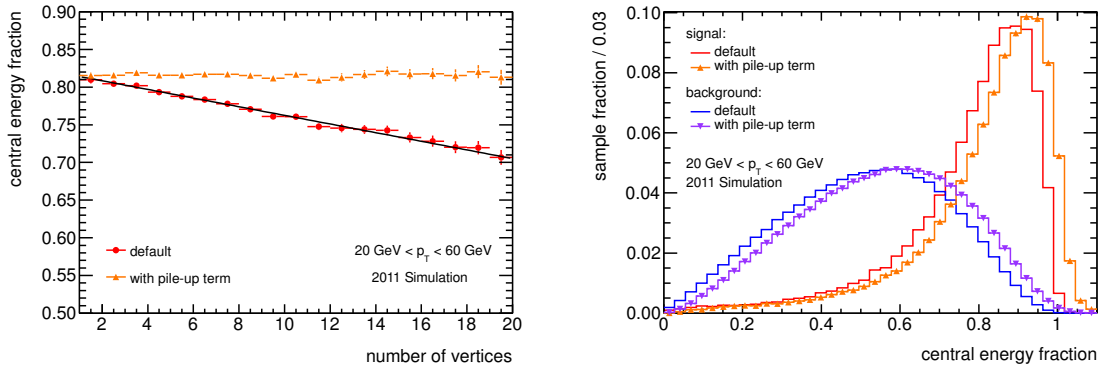
5.3.3. Minimisation of the pile-up dependence

The goal of this part of the study is to further reduce the pile-up dependence of the BDT score. An overall small loss in rejection power is accepted if pile-up independence is gained instead.

Even though nearly all variables show a certain degree of pile-up dependence, not all of them have an impact on the N_{vtx} dependence of the BDT score. The focus is therefore on the BDT score, rather than on the single variables. The starting point is a training with the track variables only, which results in a pile-up independent BDT score. The performance loss from removing all calorimeter variables is not acceptable, but the result indicates that the calorimeter variables are the main source of pile-up dependence. From the full variable set, variables are then removed one at a time to study their influence on the pile-up dependence and their impact on the discrimination power. Both the redefined variables developed in the last section as well as the original version are tested. Variables not contributing significantly to the background rejection are dropped, especially if they increase the pile-up dependence. Removing variables from the set has an influence on the importance of the remaining variables, so that the search for the critical variables (in terms of background rejection) involves many iterations. It is found that the variable $E_{\mathrm{T}}^{\mathrm{iso}}$ is not crucial for a good background rejection as long as the isolation region is covered by other variables, for instance by the much less pile-up dependent track variables. A similar observation is made for the other calorimeter variables, so that the more pile-up robust versions (redefined in the smaller cone) are favoured over the original versions.

Either of the variables f_{cent} or R_{cal} is found crucial for a good background rejection, but both result in a N_{vtx} dependent BDT score. As a new strategy, a pile-up correction term for such variables is tested. The correction follows a similar logic as before, adding a correction term $\alpha \times N_{\mathrm{vtx}}$ to the variable. The factor α is obtained by plotting the variable mean as a function of N_{vtx} and fitting a linear function, as shown in Fig. 5.17 for f_{cent}. The slope is directly defining α. Compared to the isolation energy $E_{\mathrm{T}}^{\mathrm{iso}}$ correction, the pile-up term is small relative to the variable values, so that the shape and variable range is not altered drastically.

The variable set is revised again under consideration of pile-up correction terms for calorimeter



(a) (b)

Figure 5.17: Redefined central energy fraction f_{cent} with and without pile-up correction term for tau candidates with any number of tracks. (a) f_{cent} as a function of N_{vtx} for signal candidates. The line indicates the linear fit to determine the pile-up correction strength parameter. (b) Normalised distribution for signal and background candidates.

variables. Furthermore, high momentum tau leptons from Z' samples and high momentum jets from data are included in order to test the high p_{T} behaviour. The final set is shown in Table 5.2 in comparison with the 2011 variables sets. The number of calorimeter variables is considerably reduced. The variable f_{cent} is the only variable remaining that is based on calorimeter information only. The redefined version using the smaller calorimeter cone is used, as it is the case for the mixed track and calorimeter variable f_{track}. For both cases, the additional pile-up term is used to remove the remaining N_{vtx}-gradient. The pile-up sensitivity of f_{cent} is most apparent in the low momentum region and decreases with increasing tau momentum. An unfavourable overcorrection is observed at high p_{T}. An upper threshold for the correction is defined at $p_{\text{T}} = 80\,\text{GeV}$, and the uncorrected variable is used above[10] Figure 5.18 shows the result of the training with the new variable set. The BDT is trained inclusively in momentum and N_{vtx}, which means that only one training is done for the entire momentum range and any pile-up condition. As before, the cut thresholds of the predefined working points are dependent on the tau momentum. In contrast to the previous BDT working points, no binning in N_{vtx} is applied. Nevertheless, the new working points have a stable signal efficiency (as well as background efficiency) for varying pile-up conditions, and the background rejection is at a compatible level.

5.3.4. Implementation

The proposed variable changes and the smaller set of variables is implemented in the BDT tau identification for the 2012 data taking period. The changes are also deployed in the likelihood-based tau ID [114]. The updated variable set leads to pile-up independent LLH ID working points, even though the previously used binning in N_{vtx} is dropped. Furthermore, the smaller calorimeter cone for the calculation of calorimeter variables and the new (smaller) variable set is taken over by the tau identification at EF trigger level. However, the variable N_{vtx} or similar information is not available

[10] The alternative to a rather abrupt cut-off is a smooth momentum dependent pile-up correction. Due to a time constrains, this alternative could not be tested and evaluated properly and the simpler approach was given preference. Another alternative is to correct the cone energies for pile-up and to derive f_{cent} and f_{track} from the corrected energies, which makes the cut-off redundant. However, the results were not as good as using the subsequent correction in terms of both background rejection at very low momenta and the BDT score's pile-up dependence.

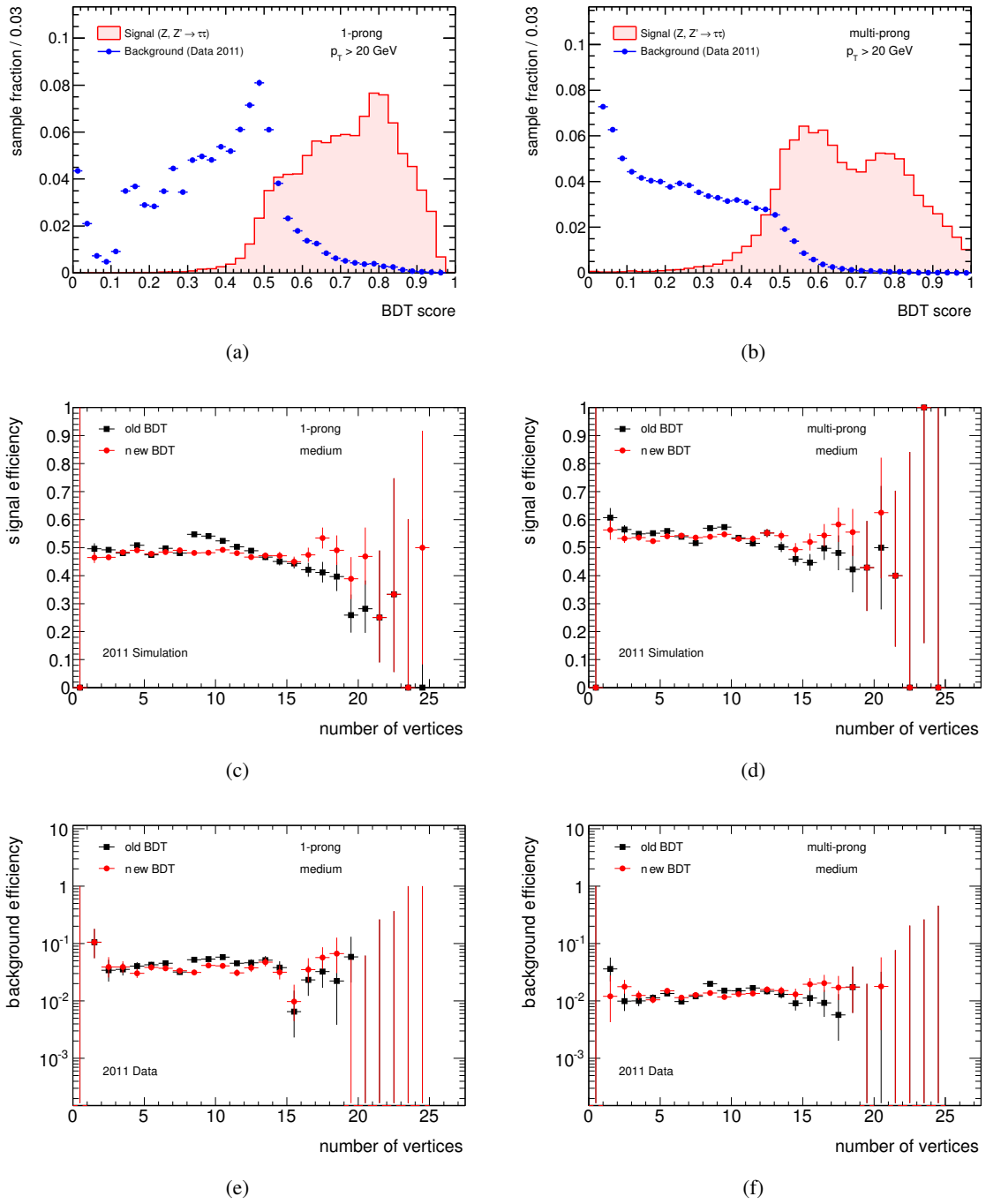


Figure 5.18: Results of the BDT training with the revised variable set for 1-prong (left) and multi-prong (right) tau candidates: (a,b) BDT score, and (c–f) total efficiency $\varepsilon_{\text{reco+ID}}$ as a function of N_{vtx} for the medium working point for signal (c,d) and background (e,f) candidates. The old BDT (squares) is trained separately for events with 1–5 and >5 number of vertices. The WP cut thresholds are dependent on the momentum and obtained in 4 bins of N_{vtx}: 1–3, 4–5, 6–7 and ≥ 8. The new BDT is trained inclusively for all pile-up conditions, and the cut thresholds are dependent on the momentum only. The pile-up robustness is obtained by using the revised discriminating variables. The candidates are obtained from simulated $Z, Z' \rightarrow \tau\tau$ samples (signal) and multi-jet events in 2011 data (background). The candidates are required to have a minimum momentum of 20 GeV.

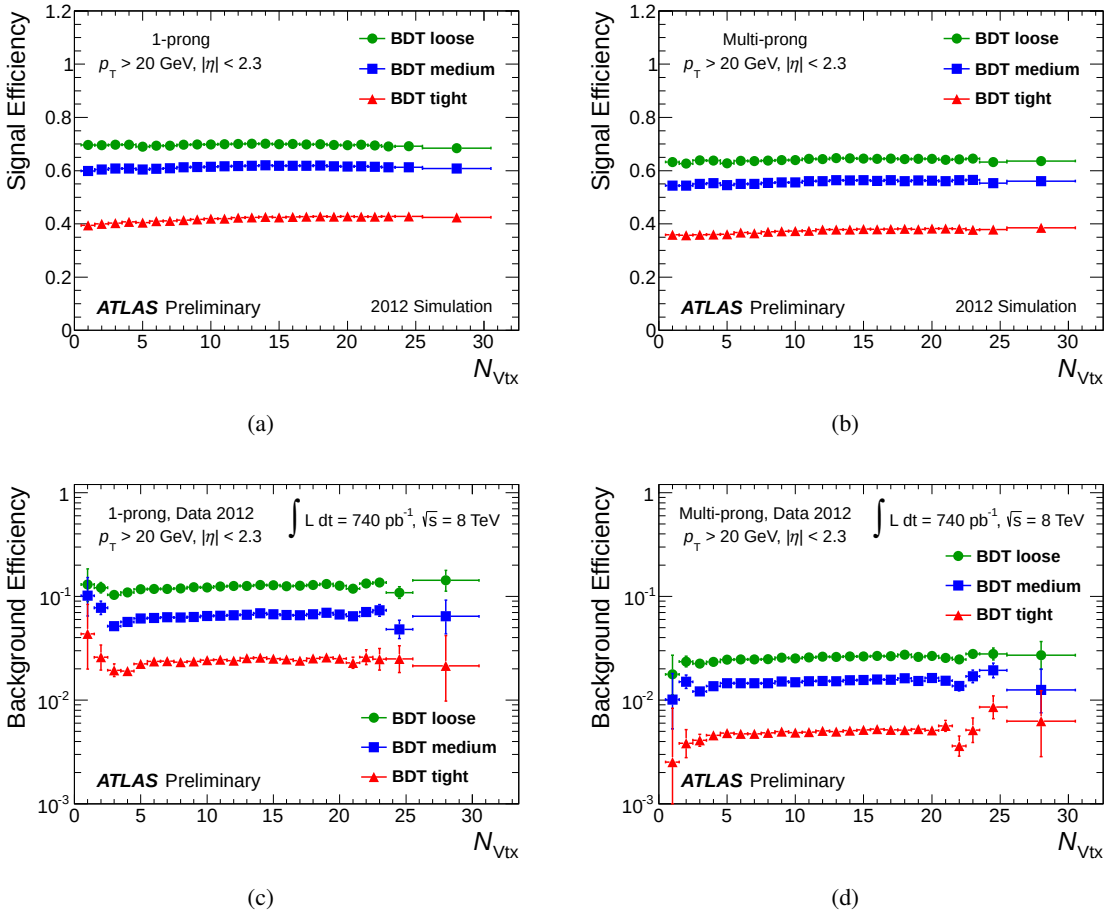


Figure 5.19: Signal (a,b) and background (c,d) efficiency $\varepsilon_{\text{reco+ID}}$ as a function of N_{vtx} for 1-prong (left) and multi-prong (right) τ_h candidates for all working points of the BDT tau ID. The signal candidates are from $Z, Z' \rightarrow \tau\tau$ and $W \rightarrow \tau\nu$ simulated samples. The background candidates are obtained from multi-jet events in 2012 data [114].

at trigger level so that the pile-up correction term cannot be used. The main technique to reduce the pile-up dependence of the tau trigger is rejecting tracks which are far away from the leading track in the RoI in terms of Δz_0 [104].

In the transition to 2012, many reconstruction algorithms were updated and prepared to cope with the expected higher pile-up conditions. Among other things, the topocluster thresholds were adjusted to be more robust against energy deposits from pile-up. The changes are passed on to the tau candidate. Since the pile-up studies were performed on a sample from the 2011 data taking period, the correction term had to be re-obtained and the BDT was re-trained using samples processed with a preliminary 2012 software release. As a result of the updated noise thresholds, the pile-up correction term reduced by approximately a factor 2.

The final evaluation on 2012 samples is shown in Fig. 5.19 for all working points.

5.4. Hadronic tau decays in 2012

5.4.1. Reconstruction and identification

As in 2011, anti-*kt* jets (parameter $R = 0.4$) are the seeds of tau reconstruction. The jet is build from topoclusters calibrated at LC scale and has to fulfill $p_T > 10\,\text{GeV}$ and $|\eta| < 2.5$. The tau vertex treatment has been updated. The vertex with the highest jet vertex fraction is now used as the coordinate system's origin (*tau vertex*). To begin with, the tau detector axis is defined as the barycenter of the clusters within $\Delta R \leq 0.2$ around the jet seed axis, assuming massless clusters. The position of these clusters is recalculated with respect to the tau vertex and their four-vector sum is recalculated. The result defines the tau momentum at LC scale and the tau intermediate axis[11]. Tracks are associated to the tau candidates, if they are within $\Delta R \leq 0.2$ (core tracks) or within $0.2 < \Delta R \leq 0.4$ (isolation tracks) around the intermediate axis. While the core tracks define the multiplicity (variable N_{track}), the isolation tracks (variable $N_{\text{track}}^{\text{iso}}$) are used only for the calculation of discriminating variables. The track quality criteria are unchanged with reference to the 2011 reconstruction [104, 114].

The discriminating variables are calculated with respect to the intermediate axis, as specified in Appendix A. As a result of the studies described in the previous section, a reduced variable set is adapted for the BDT and LLH tau ID, in order to enhance the pile-up robustness of the algorithms. For the same purpose, the variables f_{cent} and f_{track} are redefined in the smaller cone $\Delta R < 0.2$ and receive a linear pile-up correction term. The final BDT tau ID variable set is enhanced by three variables, due to a new algorithm describing the neutral pion content of the decay cone, which became available during the year 2012.

The new pion reconstruction algorithms [115, 116] splits into two parts: a π^0 counter and a π^0 finder. It is optimised on signal tau candidates which pass a medium tau identification working point.

The counting algorithm is based on the sequential application of two BDTs. The first BDT separates candidates with and without neutral pions. The second BDT then distinguishes decays with one or two neutral pions. The number of neutral pions found by the counting algorithm is thus limited to $N_{\pi^0} = 0, 1, 2$. The two BDTs are based on the same list of input variables. The variables are based on the general shower structure and not on the identification of specific clusters. They make use of the fact that the neutral pions decay to photons and thus appear as an electromagnetic shower in the first layers of the EM calorimeter, while the charged pions produce a hadronic shower reaching out into the hadronic calorimeter. The variables used are:

- $1/f_{\text{track}} = E_T/p_T^{\text{leadtrk}}$: The track fraction (or its inverse) peaks at 1 in decays with only one charged hadron, i.e. π^0-less 1-prong decays.
- $E_{\text{Ecal}}/p_{\text{tracks}}$: The ratio of the energy in the electromagnetic calorimeter and the total track momentum is smaller than 1 for decays without neutral pions, because the energy deposition for the charged hadrons is expected to reach out to the hadronic calorimeter.
- $f_{\text{Ecal}}^{\pi^\pm} = (p_{\text{trks}} - E_{\text{Hcal}})/E_{\text{Ecal}}$: The ratio of the estimated charged hadron energy in the electromagnetic calorimeter and the measured energy in the electromagnetic calorimeter peaks at 1 if no neutral pions are present. Otherwise, values lower than one are expected.
- N_{strip}: The number of cells in the strip layer is higher if neutral pions are present.
- $f_{\text{PSS}} = E_{\text{PSS}}/(E_{\text{Ecal}} + E_{\text{Hcal}})$: The energy fraction is the presampler and strip layer is, similar to N_{strip}, higher for decays with neutral pions.

[11] The final tau axis is defined after the tau energy calibration (Sec. 5.4.3).

The distributions for the different decay modes are shown in Fig. 5.20. The BDTs are trained on signal candidates from $Z \rightarrow \tau\tau$ and $Z' \rightarrow \tau\tau$ samples. The first BDT is trained on true hadronic decays without neutral pions versus decays with (any number of) neutral pions. The second BDT is trained on 1-prong tau candidates with a single π^0 versus candidates with more than one π^0. Finally, sequential cuts on the two BDT scores are defined to give optimal separation. The final distribution is shown in Fig. 5.20(f).

Another algorithms is responsible for finding the cluster or pair of clusters which meets the appearance of a neutral pion best. The so-called π^0 Finder works independent of the π^0 counting algorithm, meaning that the variable N_{π^0} is independent on the number of neutral clusters identified. In fact, the π^0 Finder identifies at least one neutral cluster in all cases, and a maximum of two clusters. The search for two clusters is done in order to catch cases in which the two photons of the pion decay are sufficiently separated to create two clusters. The algorithms considers all clusters within the cone $\Delta R < 0.3$. Assuming that an actual tau decay is confined within the core cone, the energy in the outermost ring ($0.3 < \Delta R < 0.4$) is used to estimate the contribution from pile-up and underlying event activity to each inner cluster. For the clusters with positive energy (after the correction) the effective number of clusters is calculated, so that only the most energetic clusters are used in the calculation of the π^0 likeness score. Both, single clusters and any pair (sum) of two clusters are considered. The score is mainly based on the cluster's (or cluster-pair's) energy in the presampler and strip layer, and the energy fraction in the hadronic calorimeter. The π^0 likeness score is high if the former is large and the latter is small. The cluster or cluster-pair with the highest score gets selected as the π^0 cluster. Finally, an energy correction is applied on the chosen π^0 cluster to reduce effects of hadronic shower contamination. The π^0 Finder picks the cluster closest to the generator-level π^0 in 76% of the time.

The summation of the reconstructed π^0 clusters and the core tracks lead to a new tau four-vector, denoted π^0-and-track-system here. For the summation, the charged and neutral pion masses of 140 MeV and 135 MeV are assigned to the tracks and cluster, respectively. If a cluster-pair is found, the two clusters are combined before the mass is assigned, assuming that the two clusters represent two massless photons rather than two neutral pions. From the π^0 reconstruction information, three new variables are constructed and included into the BDT variable set:

- N_{π^0}: The output of the π^0 counting algorithm.
- $p_{\mathrm{T}}^{\pi^0+\mathrm{tracks}}/p_{\mathrm{T}}$: The momentum of the π^0-and-track-system compared to the calorimeter-only measurement. The latter is the 2012 default tau momentum, fully calibrated.
- $m_{\pi^0+\mathrm{tracks}}$: Invariant mass of the π^0-and-track-system.

The distributions are shown in Fig. 5.21. The additional variables improve the background rejection by 15-20% for 1-prong tau candidates and 25-30% for 3-prong candidates, depending on the working point[12]. Since the π^0 counting and reconstruction is constructed in a pile-up robust way, the BDT pile-up strategy remains unchanged.

5.4.2. Lepton vetos

Electron veto

After the rejection of QCD-jets, the main background comes from leptons, in particular electrons. A BDT based veto is used to reject electrons. In addition to the variables used for the rejection

[12]The numbers are heavily dependent on the background considered, especially the quark-gluon mixture of the jet sample, the momentum range, the working point etc. The numbers here relate to the samples used for the tau ID studies in this thesis, i.e. a $Z \rightarrow \tau\tau$ signal sample and multi-jets from 2012 data, for a momentum range of 15 GeV to 80 GeV.

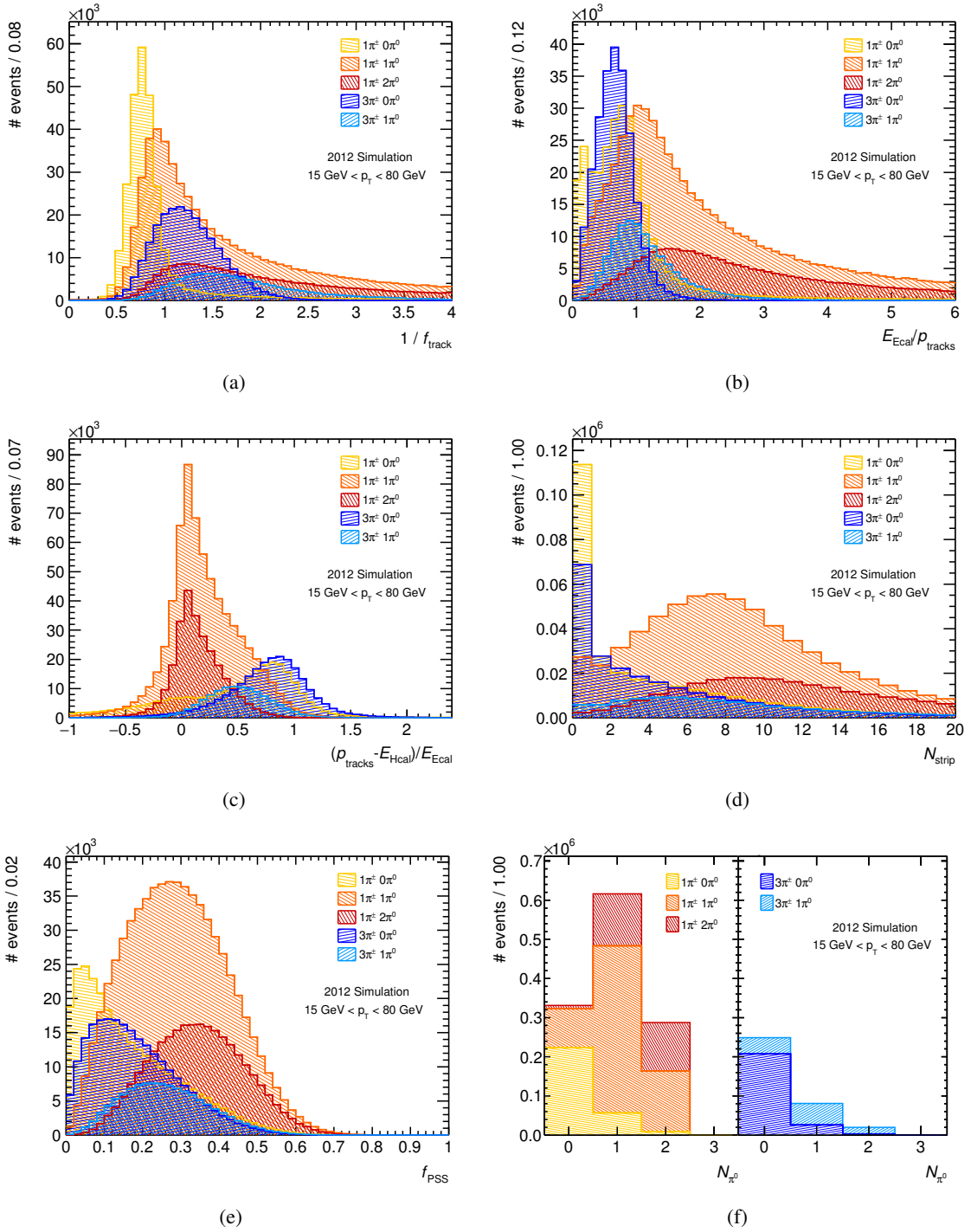


Figure 5.20: (a)–(e) Input variables for the BDT to determine the number of neutral pions, and (f) the resulting distribution. The distributions are shown for 1-prong and 3-prong signal candidates. They subdivide into the main decay modes, given by the number of neutral hadrons of the associated true tau decay (truth-matched). The distributions are not normalised, so that they reflect the composition of the signal candidates. The histograms are overlapping in (a)–(e), and stacked in (f). The signal candidates are taken from simulated $Z \to \tau\tau$ events.

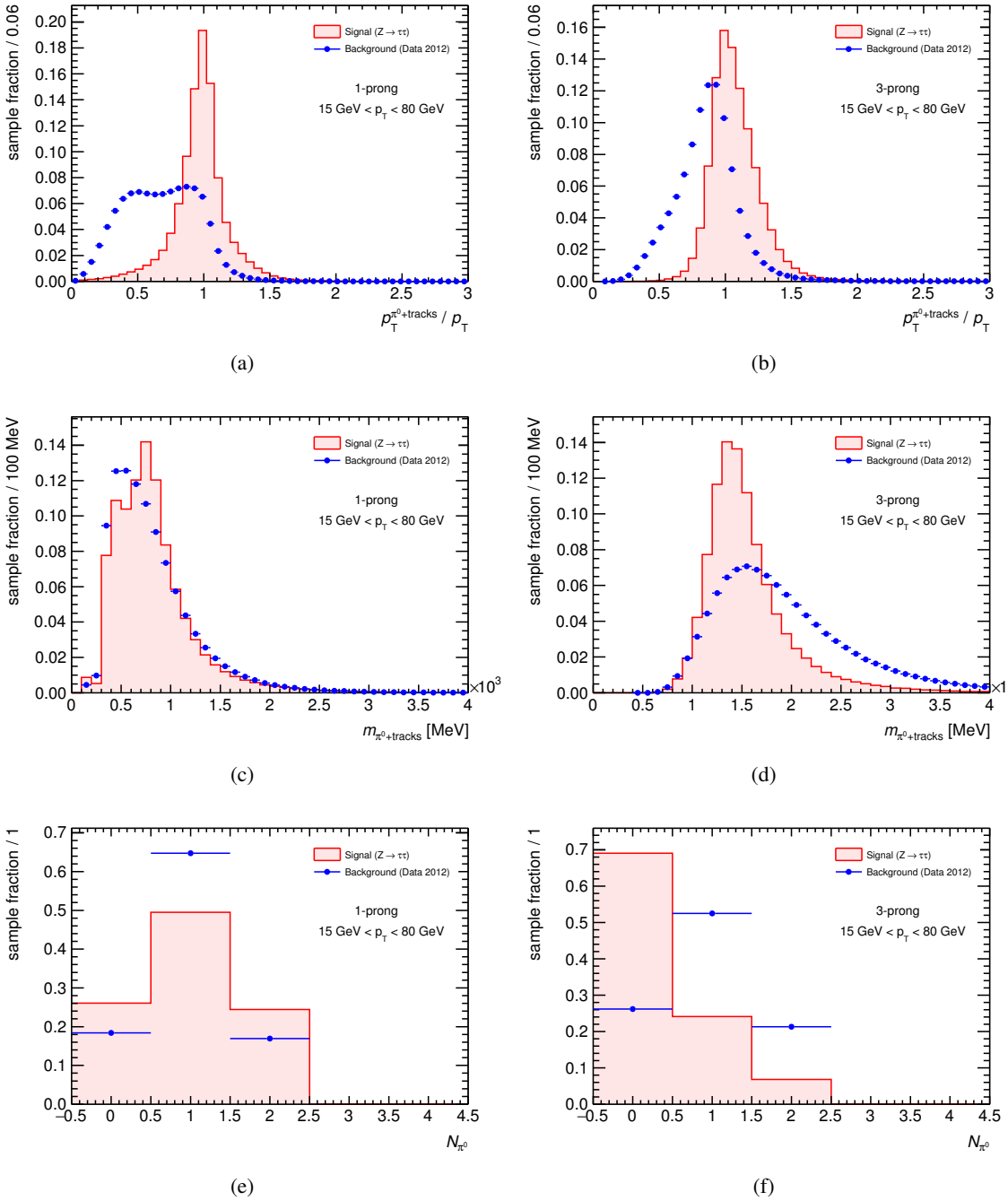


Figure 5.21: Additional variables for the BDT tau ID which are based on the decay substructure for 1-prong (left) and 3-prong (right) candidates. The distributions are normalised to unity. The signal candidates are taken from simulated $Z \rightarrow \tau\tau$ events and the background candidates are obtained from multi-jet events in 2012 data.

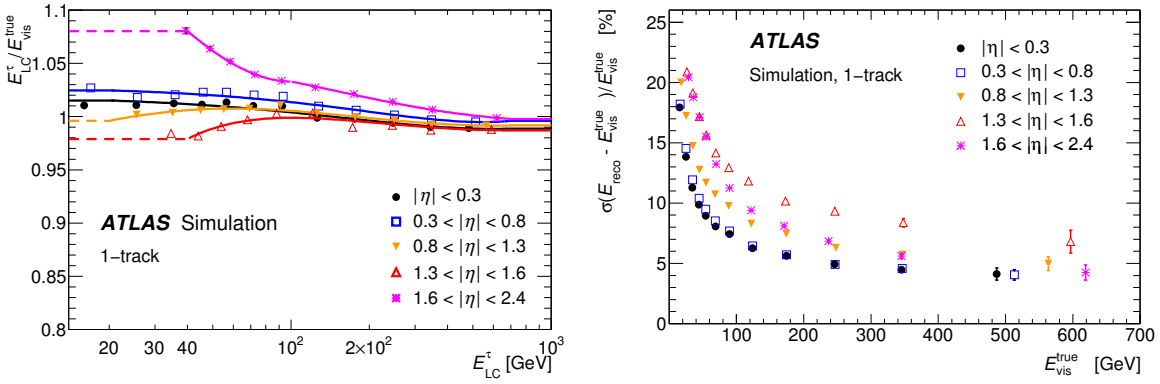
of QCD-jets, information from specific detector subsystems are of great importance. The electron creates an electromagnetic shower which does not reach into the hadronic calorimeter. The shower fraction within the ECAL is therefore a good discriminator. Electrons passing through the transition radiation tracker are more likely to emit transition radiation than pions. This fact provides the other most important variable for the rejection of electrons, namely the ratio of high-threshold hits over low-thresholds hits. This variable peaks higher for electrons than for pions. BDTs are trained separately for different pseudorapidity regions. A main reason for this is that the TRT information is not available for $\eta > 2.0$. Three working points for the electron veto are defined targeting signal efficiencies of 75%, 85% and 95% [104, 114].

Muon veto

Muons are not easily mistaken with hadronic tau decays, because they are unlikely to deposit enough energy in the calorimeter. Furthermore, they can generally be avoided by rejecting any tau candidates that overlap geometrically with reconstructed muons. A very loose muon reconstruction quality enhances the performance of the overlap removal thereby considerably. A cut-based muon veto can be applied to reject muons that are not rejected by the overlap removal, because they fall into inefficient detector regions or are stopped before reaching the muon detector system. The muon veto is mainly based on two variables, the fraction of calorimeter energy deposited in the ECAL and the ratio of track momentum and calorimeter energy. The calorimeter deposit of muons reconstructed as tau leptons are found to concentrate in either the HCAL or ECAL, but seldomly spread over both, as it is the case for real hadronic tau decays. The EM energy fraction therefore shows extreme values for mis-identified muons. Furthermore, the track momentum does not match the calorimeter energy deposition, so that the ratio of the two tends to be unsually high or low. The muon veto has a signal efficiency better than 96% and rejects about 40% of the muons. The numbers are however highly dependent on the tau lepton and muon identification working points [104, 114].

5.4.3. Tau energy scale

The tau energy is determined through the calorimeter energy deposit. A calibration step optimised for hadronically decaying tau leptons brings the energy from the LC scale to the *tau energy scale* (TES). As described in Sec. 4.2, the LC scale is applied to all jet objects and corrects for the non-compensating detector, energy deposits outside the reconstructed clusters, and insensitive detector regions. The TES takes into account the specific tau characteristics, for instance, the composition of the decay (mainly charged and neutral pions) and the smaller reconstruction cone ($\Delta R < 0.2$). In the main calibration step the energy is brought in agreement with the true energy within a few percent on the base of simulated decays. It also removes significant dependencies on the energy, the pseudorapidity, the track multiplicity and the pile-up conditions. Reconstructed true hadronic tau decays satisfying the medium tau identification with $p_{\mathrm{T}} \geq 15\,\mathrm{GeV}$ are used to obtain the calibration. The ratio of the reconstructed energy and the true energy is obtained in intervals of true energy, and separately for various η regions and 1-prong or multi*-prong candidates. For each interval, the distribution of the ratio is plotted and fitted with a Gaussian function in order to determine the mean value. The mean values as a function of the average reconstructed energy in each given interval are then fitted with an empirically derived function and represent the response curve (see Fig. 5.22(a)). Additional small corrections are applied to correct for insensitive detector regions and contributions from pile-up. The energy resolution for the fully calibrated 1-prong tau candidates is shown in Fig. 5.22(b)). For low momentum tau leptons, it is about 20% and saturates for tau leptons of a few hundred GeV at about 5%.



(a) (b)

Figure 5.22: (a) Response curves as a function of the reconstructed energy at LC scale, E_{LC}, for 1-prong tau candidates and the different pseudorapidity regions. The markers cover approximately the range with transverse energy $E_{\mathrm{T,LC}} > 15$ GeV. Only statistical uncertainties are shown. (b) Energy resolution of 1-prong tau candidates for different pseudorapidity regions. The resolution is given by the standard deviation of a Gaussian fit to the distribution $(E_{\mathrm{reco}} - E_{\mathrm{true,vis}})/E_{\mathrm{true,vis}}$ in a given range of $E_{\mathrm{true,vis}}$ and $|\eta_{\mathrm{true,vis}}|$ [104].

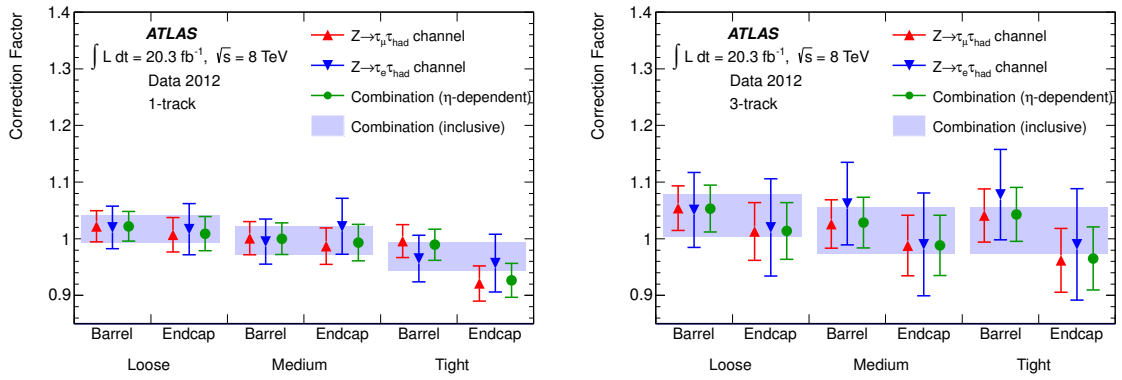
In a second step, the tau energy scale is validated and small corrections are obtained to correct for modelling differences in simulation and data. Furthermore, the TES uncertainties are determined. Two complementary approaches are used: a deconvolution method and a tag-and-probe measurement.

The first method deconvolutes the hadronic tau decay into its single components. Each single particle response is studied and then propagated to the final calorimeter response under consideration of the decay branching ratios. The response of low momentum charged hadrons is determined from in-situ E/p measurements in low pile-up data. The response of high-momentum charged hadrons is obtained from test beam measurements for the central detector region ($|\eta| < 0.8$), and comes from simulation otherwise. Electromagnetic showers are studied in $Z \to ee$ decays and used for the response of neutral pions. For the determination of the total TES uncertainty, contributions due to the detector modelling, underlying event and pile-up activity, the non-closure of the calibration method and the hadronic-shower model are added. The total uncertainty is 2%–3% for 1-prong tau candidates, and 2%–4% for multi*-prong candidates, depending on E_{T} and η. Furthermore, a TES shift of about 1% is observed (and corrected for).

The alternative approach is an in-situ measurement using $Z \to \tau\tau$ events, in which one of the tau leptons decays leptonically to a muon and the other one decays hadronically. The visible mass peak of muon and the hadronic tau decay allows to obtain the TES shift and its uncertainty. The shift is determined by changing the measured tau energy in small steps until the peak position in data and simulation matches. The observed shift confirms the findings of the deconvolution approach [104, 117].

5.4.4. Performance measurement: tau ID

The efficiency of the tau identification algorithms in data and simulation is compared in a tag-and-probe analysis. A sample enriched in $Z \to \tau\tau$ events is selected, in which one tau lepton decays leptonically to either a muon or electron (tag) and the other tau lepton decays hadronically (probe). Both channels are first measured individually and are then combined to improve the measurement precision. The main background in this analysis are other Z decays, W events and multi-jet events.



(a) (b)

Figure 5.23: Correction factors needed to bring simulated tau candidates to the level of data for the three pre-defined BDT tau ID working points, for (a) 1-prong and (b) 3-prong candidates with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$. These are the data/MC scale factors described in the text. The correction factors are shown for the individual lepton channels as well as the combination [104].

The Z and W events are suppressed by requirements on the missing transverse energy, for instance on the transverse mass of the lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$ and the relative direction to the lepton and the tau candidate. The purity is further enhanced by requirements on the visible transverse mass of the lepton and the tau candidate. The multi-jet background is estimated in a template fit of the extended track multiplicity of the tau candidate. This variable counts tracks in a wider radius around the candidate, but suppresses tracks from pile-up and underlying event activity with relative momentum and separation requirements. It is given by the sum of core tracks and *outer tracks*. Outer tracks have to be within $0.2 < \Delta R < 0.6$ and satisfy $p_{\mathrm{T}} > 500\,\mathrm{MeV}$ and $p_{\mathrm{T}}(\text{core track}) / p_{\mathrm{T}}(\text{track}) \times \Delta R(\text{core track}, \text{track}) < 4.0$ with at least one of the core tracks. The extended track multiplicity is fitted using separate signal templates for simulated 1-prong and 3-prong tau candidates and a background template estimated from data. The fit is performed for tau candidates before and after applying different levels of tau identification, and the identification efficiencies are obtained. Scale factors are obtained by comparing the efficiencies measured in data and simulation [104].

The data/MC scale factors are shown in Fig. 5.23. They are consistent with unity for the medium working point of the BDT tau ID, and are slightly higher (lower) for the loose (tight) working point. For tau candidates with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$, uncertainties amount to 2.5% and 4% for 1-prong and 3-prong candidates, respectively. For tau leptons with lower momentum, the uncertainties increase to a maximum of 15% [104]. Cross check analyses performed in a sample enhanced in $W \to \tau\nu$ and $t\bar{t}$ events show consistent results, but come with higher uncertainties [114][13].

For the future it is planned to provide *continuous* scale factors, which are not confined to the predefined working points, but provided for the BDT score directly. This opens the possibility to explore the entire BDT score in an analysis and simplifies the creation of working points individually for an analysis.

Scale factors are also measured for the electron veto. A tag-and-probe analysis is performed in a sample enhanced in $Z \to ee$ events, where one of the electrons is identified as such (tag) and the other one is mis-identified as an hadronically decaying tau lepton. The latter may not overlap

[13] The cross check analysis have been performed on the earlier version of the BDT tau ID, i.e. the version without the substructure variables.

geometrically with an identified electron. Requirements on the momentum of the tag electron and the transverse visible mass suppress contamination from $Z \rightarrow \tau\tau$ events. Scale factors are obtained for different combinations of identification levels of the overlap removal, the tau ID working point, the electron veto working point, and for different pseudorapidity regions. Uncertainties are about 10% for the loose electron veto, but can be significantly larger for the tighter working points or specific η regions [104].

5.5. Momentum dependence of the BDT

The discriminating variables show a clear dependence on the tau momentum, due to the decreasing decay cone size for a boosted tau lepton. The major SM processes with tau leptons in the final state are the decays of W, Z and Higgs bosons. The visible tau decay products of such events accumulate in the low momentum region, below 40 GeV. Therefore, the focus of the tau ID algorithms is on this momentum region. Nevertheless, the algorithms need to perform well at any tau momentum range, in order to allow searches for physics beyond the SM.

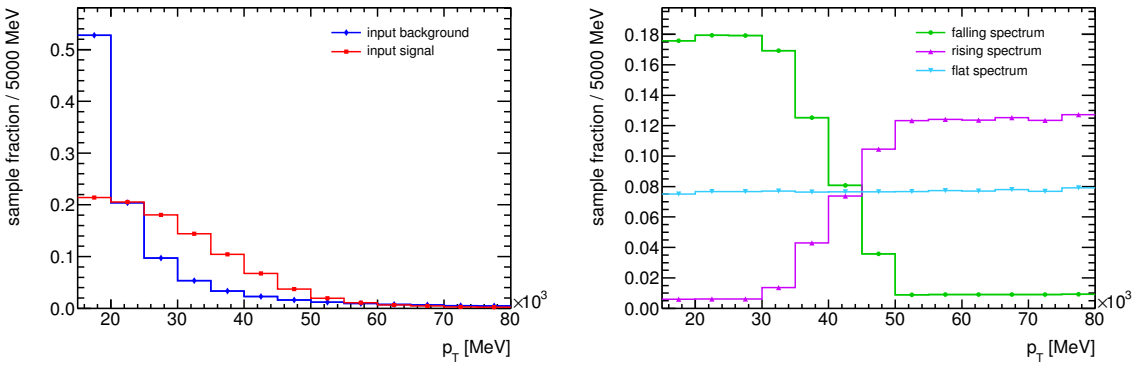
Besides including simulated Z' decays in the training sample, no optimisation of the identification of high-p_{T} tau leptons is done. For instance, the BDT is trained inclusively in momentum and has no direct access to $p_{\mathrm{T}}(\tau_h)$. In this section, the following aspects of the momentum treatment are investigated: the reweighting with respect to the p_{T} spectrum and possible gains of a p_{T}-flattening of variables.

5.5.1. Samples and setting

The studies are based on 8 TeV simulated $Z \rightarrow \tau\tau$ samples (signal) and multi-jet events from data (background), following the procedure and object definition as described in Sec. 5.3.1. The BDT training options were optimised as part of the introduction of substructure variables to the BDT tau ID (Sec. 5.4.1). Accordingly, the new BDT setting is used for the following studies. The major change is the introduction of *pruning* in order to reduce the tendency to overtraining. The pruning step is applied at the very end, after the forest has grown to its maximum size. Starting at the bottom, the trees are cut back by removing all statistically insignificant nodes [107]. The pruning algorithm in use is 'CostComplexity' with a strength of 60. With the exception of the number of total trees, which is increased to 70, other settings are not changed. Since pruning is discussed controversially [107], the results are cross checked using a setting without pruning. In order to avoid overtraining, the tree depth is limited to 8 or smaller, depending on the size of the sample and the number of variables. Even though small performance differences are seen comparing the two training settings, the conclusions are identical with both settings, i.e. comparing different variable sets within one setting shows the same behaviour (improvement / degradation / changes of momentum dependence / ...) in both settings. Therefore, only results obtained with the pruning-setting are shown here.

5.5.2. Reweighting

If the background and signal p_{T}-spectra deviate significantly, a reweighting of one sample is performed so that the spectra are equalised. This is done in order to prevent a separation of signal and background based on momentum, since the momentum as such is not a feature of the decay. In order to understand the impact of the reweighting, BDTs are trained with varying p_{T} spectra, shown in Fig. 5.24. The default training setup takes the 'input background' spectrum as a basis and the signal sample is reweighted to match it. The BDT is then trained for different scenarios, in which background and signal samples are either following the same spectra ('default', 'both flat' and 'both



(a) (b)

Figure 5.24: Momentum spectra for the 1-prong BDT training. (a) Spectra of the unweighted background (squares) and signal (diamonds) τ_h obtained from multi-jet events in data and simulated $Z \to \tau\tau$ events, respectively. (b) Artificial p_{T} spectra.

rising') or different spectra ('sig rising, bgr falling', 'bgr rising, sig falling', 'no reweighting'). As shown in Fig. 5.25, the latter setups increase the momentum dependence of the BDT score. The matching spectra show approximately the same dependence, which is especially not influenced by the spectrum's tilt. For the tau ID working points, the momentum dependence is removed by defining p_{T} dependent score thresholds. A flat score is nevertheless preferable to allow for a simple definition of control regions based on the BDT score in an analysis[14]. The influence of different input spectra on the background rejection power is shown in Fig. 5.26. The trainings done with opposite spectra for background and signal perform worse in each of the small p_{T} ranges. The default and the unweighted spectra perform equally well, and show the best background rejection in the low momentum region. This meets the requirements as analyses suffer from a high background in this region.

5.5.3. Flattening

The pile-up dependence of the BDT score could be significantly reduced by applying a pile-up correction to only two variables. The same approach is investigated for the p_{T} dependence. Furthermore, it is studied whether the performance in any momentum region can be improved by using p_{T} independent discriminating variables. The starting point is again the final 2012 variable set (including the substructure related variables). The default reweighting options are used, i.e. the simulated $Z \to \tau\tau$ signal candidates are reweighted to match the p_{T} spectrum of the background candidates from multi-jet events in data.

For the investigation of the momentum dependence of the BDT ID, flat cuts on the score are used in order to define the identification working points. The identification efficiency using these untuned WP as a function of momentum is shown in Fig. 5.27. After a rising phase the signal efficiency curves start leveling out around 40 GeV. The behaviour resembles the momentum profile of the variables central energy fraction f_{cent} and track radius R_{track}, which are therefore natural candidates to start with. The signal profile of each variable is plotted versus momentum, and a function is chosen to fit it. Since the study is intended to primarily explore the potential of a flattening procedure, the fit is not optimised and an arbitrary function is chosen that fits well within the training p_{T} range of 15 GeV to 80 GeV. This is illustrated in Fig. 5.28(a) for f_{cent} and in Fig. 5.28(b) for R_{track}. Besides the change

[14] See for example the tau identification efficiency measurement in Ref. [109].

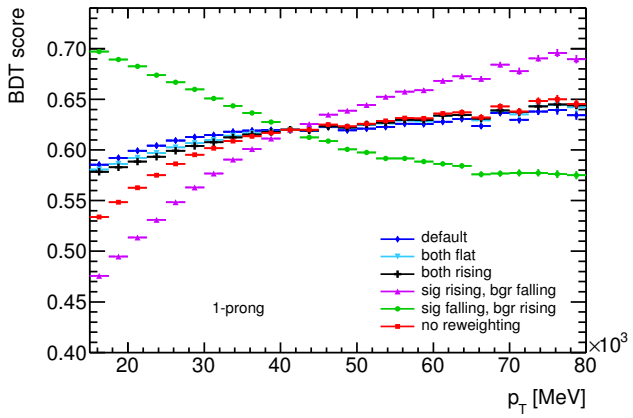


Figure 5.25: Mean BDT score as a function of τ_h momentum for samples reweighted to different p_{T} input spectra but otherwise identical training options. The profiles are shifted up or down for a better comparison of the momentum trend. In the default setup (diamonds), the signal is reweighted to match the unweighted background spectrum, compared to a training with no reweighting (squares). Both, background and signal candidates are reweighted for the other setups. They either follow the same flat (up-side-down triangle) or rising (square) spectrum, or opposite spectra with rising signal and falling background (triangle) or vice-versa (circle). The spectra are shown in Fig. 5.24.

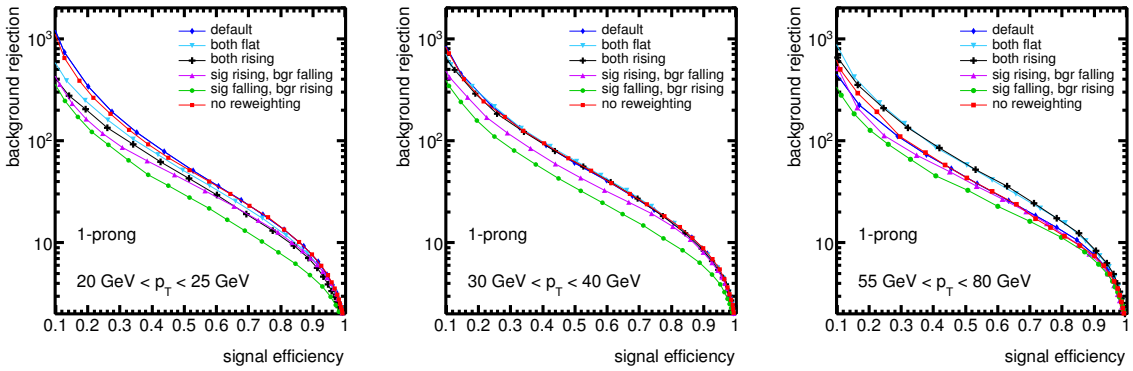
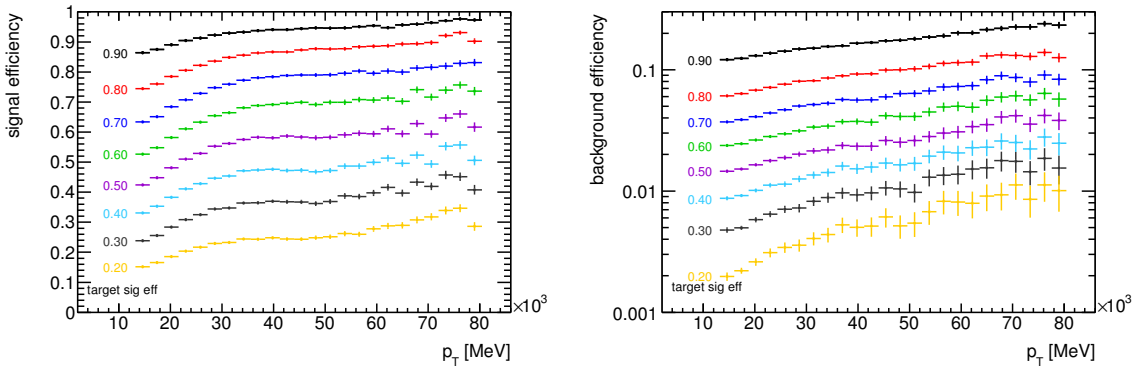


Figure 5.26: Background rejection versus signal efficiency for three momentum regions and for samples reweighted to different p_{T} input spectra but otherwise identical training options. The same (unweighted) spectra have been used for the efficiency curve, even so they have little influence here due to the small p_{T} ranges used. The increasing p_{T} ranges reflect the decreasing number of events available.



(a) (b)

Figure 5.27: Signal (a) and background (b) efficiency $\varepsilon_{\mathrm{ID}}$ as a function of τ_h momentum. The identified candidates pass a flat cut on the BDT score, targeting overall signal efficiencies of 90%, 80%, 70%, 60%, 50%, 40%, 30% and 20%. The tau candidates are 1-prong τ_h from a simulated $Z \to \tau\tau$ sample (signal) and multi-jet events from data (background).

of the mean value with p_{T}, also the width of the distribution changes, here defined as the (symmetric) standard deviation in each bin. Therefore, both the mean and the width are fitted:

$$\text{fitted mean of pile-up corr. } f_{\mathrm{cent}}: \quad m(p_{\mathrm{T}}) = -\frac{2.87\times10^3}{p_{\mathrm{T}}} - 1.10\times10^{-7}p_{\mathrm{T}} + 1.00\times10^{-2}\ , \tag{5.6}$$

$$\text{fitted spread of pile-up corr. } f_{\mathrm{cent}}: \quad s(p_{\mathrm{T}}) = -\frac{5.23\times10^3}{p_{\mathrm{T}}} - 3.36\times10^{-7}p_{\mathrm{T}} - 0.98\times10^{-2}\ , \tag{5.7}$$

$$\text{fitted mean of } R_{\mathrm{track}}: \quad m(p_{\mathrm{T}}) = \frac{0.76\times10^3}{p_{\mathrm{T}}} - 1.71\times10^{-7}p_{\mathrm{T}} + 2.63\times10^{-2}\ , \tag{5.8}$$

$$\text{fitted spread of } R_{\mathrm{track}}: \quad s(p_{\mathrm{T}}) = \frac{0.51\times10^3}{p_{\mathrm{T}}} + 0.19\times10^{-7}p_{\mathrm{T}} - 1.00\times10^{-2}\ . \tag{5.9}$$

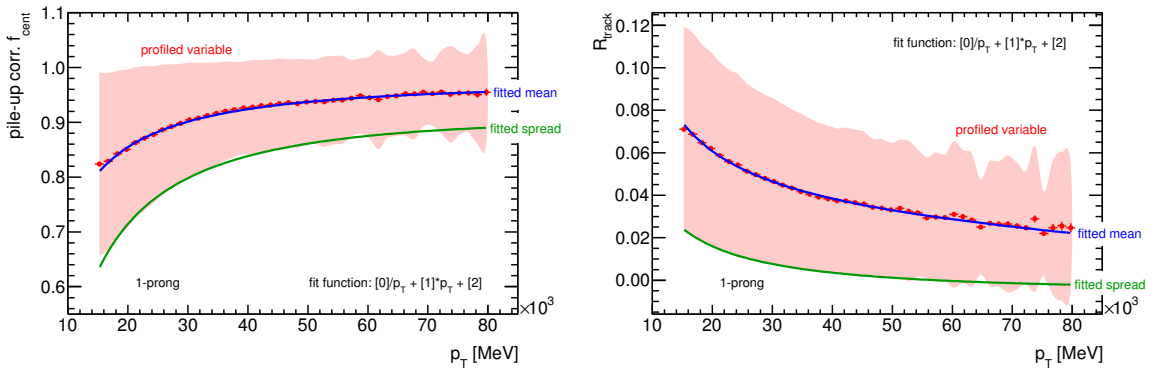
Two ways of variable flattening are tested: a simple flattening of the mean value by subtraction of the fitted function, and a variable transformation that p_{T}-flattens the width along with the mean. The idea behind the latter is inspired by the fact that the decision tree consists of simple cuts, which are not necessarily applied at the variable's mean value. A highly varying width can therefore lead to a momentum dependent identification efficiency. The two transformations are defined as follows, using the fitted functions $m(p_{\mathrm{T}}$ and $s(p_{\mathrm{T}})$:

$$\text{Flattened variable } x: \quad x^{\mathrm{flat}} = x - m(p_{\mathrm{T}}) \tag{5.10}$$

$$\text{Spread-flattened variable } x: \quad x^{\mathrm{spread-flat}} = \frac{\alpha}{|s(p_{\mathrm{T}}) - m(p_{\mathrm{T}})|} \times x^{\mathrm{flat}}\ . \tag{5.11}$$

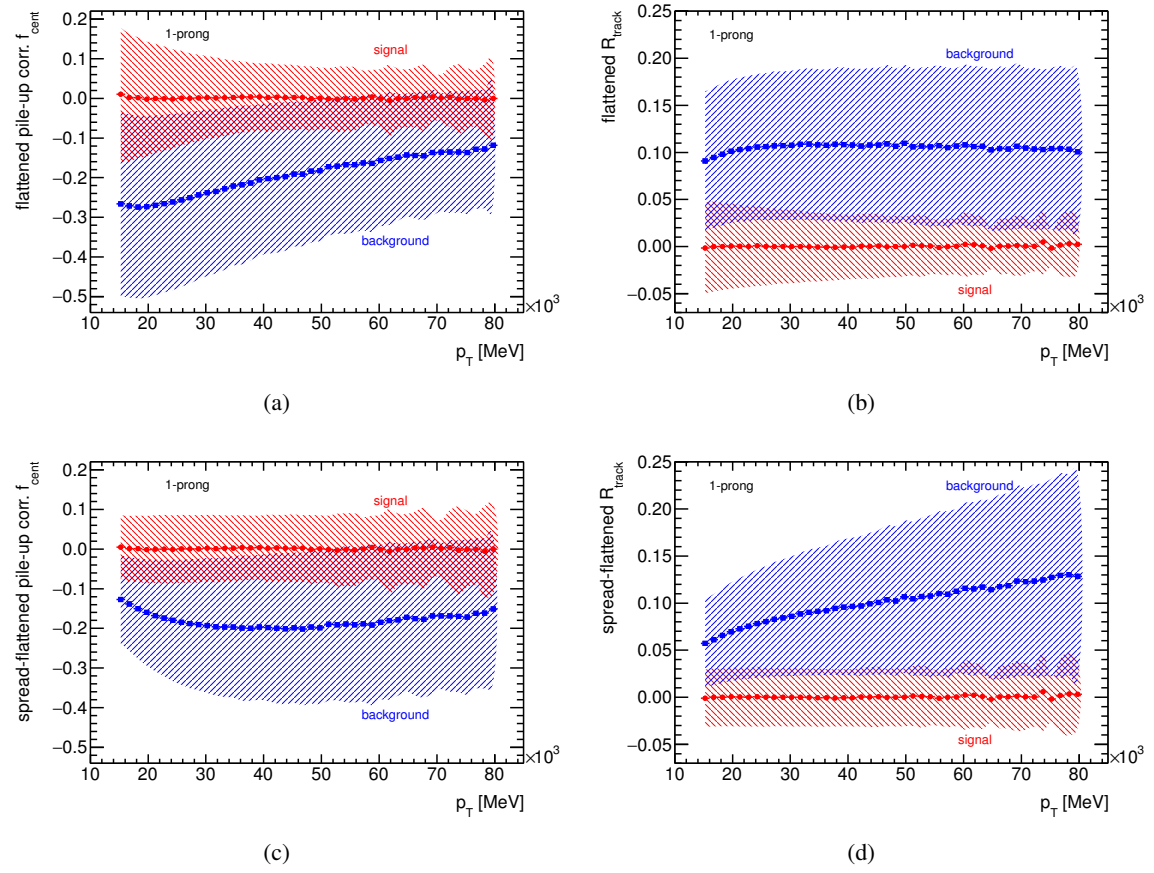
The factor α can be arbitrarily chosen and is set to the average width over all p_{T} bins, $\alpha = \mathrm{average}(|s(p_{\mathrm{T}}) - m(p_{\mathrm{T}})|)$. For simplicity reasons, no off-set correction is applied and the distribution is shifted towards zero. The values as such are hard to interpret, but the BDT training is not affected by the global shift. The flattened and spread-flattened variables f_{cent} and R_{track} are shown in Fig. 5.29. A BDT is then trained with the default variable set with replacing both f_{cent} and R_{track} with their flattened or spread-flattened version.

The result is shown in Fig. 5.30 for flat BDT score thresholds targeting different signal efficiencies. Both trainings show a much smaller p_{T} dependence compared to the default set as shown in Fig. 5.27. The use of flattened variables leads to a BDT score which is on average p_{T} independent (corresponding here to the 50% WP). The WPs further away from the average show an increasing dependence.



(a) (b)

Figure 5.28: Momentum dependence of (a) the pile-up corrected central energy fraction f_{cent} and (b) the track radius R_{track} for 1-prong signal candidates. The points show the mean value per p_{T} bin and the band indicates the standard deviation. The solid curves show the fitted mean and width, as given in Eq. (5.6) to (5.9).



(a) (b)

(c) (d)

Figure 5.29: The flattened (top) and spread-flattened (bottom) variables pile-up corrected central energy fraction f_{cent} and track radius R_{track} as a function of momentum. The points show the mean value per p_{T} bin and the band indicates the standard deviation. The 1-prong tau candidates are from a simulated $Z \to \tau\tau$ sample (signal) and multi-jet events from data (background).

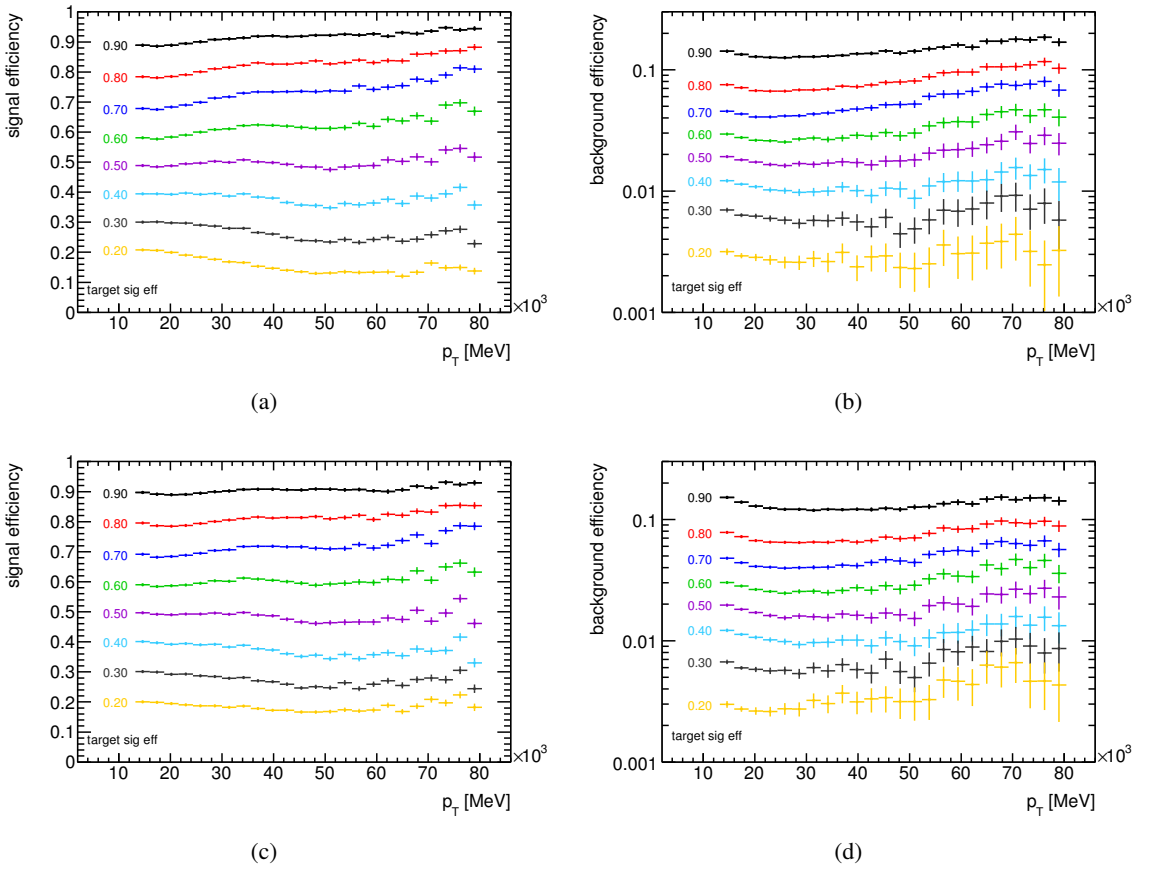


Figure 5.30: Signal (left) and background (right) efficiency $\varepsilon_{\mathrm{ID}}$ as a function of τ_h momentum for a BDT trained with the default variable set replacing f_{cent} and R_{track} with their flattened (top) or spread-flattened (bottom) version. The identified candidates pass a flat cut on the BDT score, targeting overall signal efficiencies of 90%, 80%, 70%, 60%, 50%, 40%, 30% and 20%. The tau candidates are 1-prong τ_h from a simulated $Z \rightarrow \tau\tau$ sample (signal) and multi-jet events from data (background).

This observation points towards an influence of the momentum dependent width of the input variable. Indeed, the training with the spread-flattened variables shows an improved behaviour. Over the full range of working points, the signal efficiency is notably flattened, despite the fact that only two of in total eight variables were modified. The behaviour of the background is also of importance. Figure 5.30 shows that the background efficiency flattens along with the signal efficiency, which is important in order to maintain a high rejection power. A closer look at the discrimination power in bins of p_{T} (Fig. 5.31) shows indeed that the background rejection is not significantly influenced by the transformation of the variables. While this (unfortunately) means that the procedure cannot positively influence the background rejection at high τ_h momenta, it indicates that current tau ID does not suffer from the momentum dependence of the input variables.

Also a flattening of other variables than f_{cent} and R_{track} is tested along with different combinations of altered and unaltered variables. The flattening of additional variables does however not further improve the momentum dependence. In addition, there is no straight forward way to transform integer valued variables as for example the number of isolation tracks.

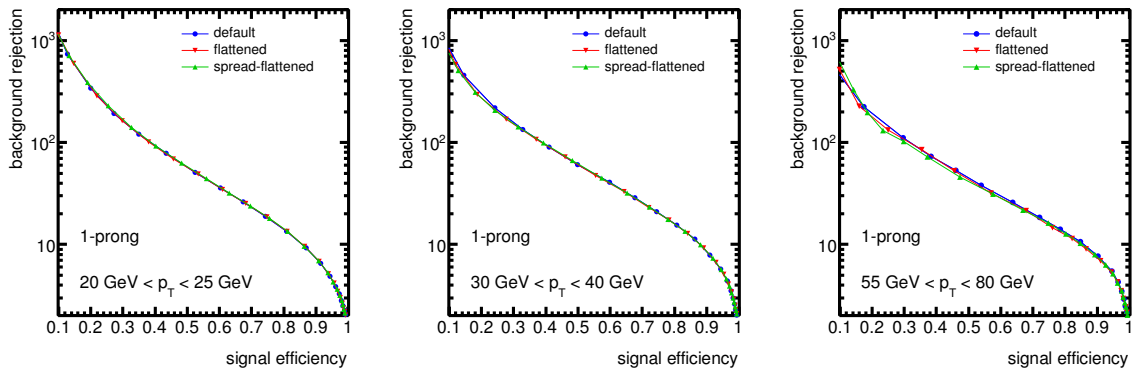


Figure 5.31: Background rejection versus signal efficiency for the three BDT trainings using different versions of f_{cent} and R_{track}: the default version, the flattened-version and the spread-flattened version. The performance of the 1-prong candidates is shown for three momentum regions.

5.5.4. Conclusion

The reweighting study underlines the importance of equalised momentum spectra for signal and background candidates. The rapidly falling background p_{T} spectrum puts the focus of the BDT training on the low momentum region, which matches the current analysis' focus.

The momentum dependence of the BDT score can be reduced by a flattening of only a few input variables. However, this does not lead to an improved background rejection in any momentum region under investigation. While an intrinsically p_{T} independent BDT score is favourable, flattened (and even more so a spread-flattened) variables add to the complexity of the identification method. At this point, the study was not continued, due to the availability of new substructure algorithms. The focus of the algorithm development switched towards studies potentially leading to an increased background rejection. These studies are described in the next section.

Not explicitly mentioned before is the influence of the momentum dependence on the performance plots showing the background rejection versus the signal efficiency. For an optimal comparison of different BDTs the influence of the score's p_{T} dependence needs to be eliminated. This can be done by creating the performance plots in bins of p_{T} as done in Fig. 5.31. Another possibility is to apply p_{T} dependent cuts on the score, as it is done in the construction of the default working points.

5.6. Substructure

During 2013 new algorithms exploring the tau decay substructure became available. The new algorithms focus on the energy flow within the decay in order to specifically identify the charged and neutral pions [118]. Only one of the new algorithms is discussed here, which was in the focus of the author's investigations. Others focussed on the remaining algorithms and came to similar conclusions. The algorithm under study is an improvement of the version described in Ref. [119]. The ECAL energy deposit of the charged pions is determined from the track and HCAL measurement. The expected energy deposition is then subtracted cell-by-cell from the calorimeter clusters, based on average charged pion shower shapes. The remaining ECAL cells are reclustered, using the topoclustering algorithm. Neutral pions are identified within the group of neutral clusters with the help of a BDT. Contributions from pile-up, noise and an imperfect subtraction are suppressed as part of the neutral pion identification. All steps are optimised to correctly identifiy the neutral pions within a

true hadronic tau decay, so that the reconstructed neutral and charged pions can be used to build the tau candidate. While the new algorithm is expected to greatly increase the tau momentum resolution and position measurement, the effect on the background rejection is not obvious. The algorithm is not tuned to differentiate between a QCD-jet and a true tau decay. It is therefore an interesting question whether to base the future tau ID entirely on the new substructure or to keep the current approach. Another question is whether the current tau ID can profit from the new substructure information. The development of the new algorithms as well as the investigations regarding tau ID are still ongoing within the ATLAS collaboration. Only the author's contributions are presented here, with a focus on the more general observations that are unlikely to be influenced by future developments of the substructure algorithms. Several ways of implementing the new substructure information and algorithm are investigated:

- Replacement of the substructure variables in the 2012 tau ID by the corresponding variables based on the new algorithm.
  → The replacement does not lead to significant differences in performance. However, small observed differences turn out to be sensitive to the version of the new algorithm used. Since the algorithm is still under development no final conclusion can be drawn at this point.

- Development of additional substructure related variables.
  → The additional variables are based entirely on substructure objects. They are inspired by existing variables and describe for example average radii, maximum distances, multiplicities or energy ratios. Different aspects and findings are discussed in sections 5.6.1, 5.6.2 and 5.6.4.

- Replacement of the currently cell-based variables by cluster-based variables.
  → Driven by the idea to base the tau reconstruction and ID entirely on substructure objects, replacements for the current variables are searched for. Especially, this requires the transition from a cell-based to a cluster-based calculation of variables. The implications are discussed in Sec. 5.6.3.

- Replacement of the currently calorimeter-based tau axis to an axis given by the identified charged and neutral pions.
  → First trainings using the new axis show no significant differences in performance, but more detailed studies are necessary to understand the impact on the various discriminating variables. Here, this part of the study is not discussed further.

Finally, in Sec. 5.6.5 a variable set is discussed which is not dependent on a specific substructure algorithm. Rather, the set includes variables that are very sensitive to the substructure of the decay.

The following studies are based on 8 TeV simulated samples (signal) and 2012 data (background), using a version of the substructure algorithm from summer and autumn 2013. The same training options and object definitions are used as in the previously described study (see Sec. 5.5.1).

5.6.1. Cluster reconstruction

Figure 5.32 shows the number of neutral clusters in comparison to the number of identified π^0. While the latter is a pile-up robust variable, the former is not. However, the discrimination of background and signal is much more distinct in the number of neutral clusters, especially for 1-prong candidates. This reveals a critical point: in the π^0 identification process, the algorithm rejects the soft clusters that appear numerously in a QCD-jet. What remains are τ-like objects, and the shapes of signal and background variables become similar. A tau identification exclusively based on the identified charged and neutral pions is therefore unlikely to be strong. A major ingredient for the discrimination of QCD-jets are the clusters that fail the pion reconstruction. A good example for this circumstance

is the variable $p_{\mathrm{T}}^{\pi^0+\mathrm{tracks}}/p_{\mathrm{T}}$ (Fig. 5.21). It provides separation power based on the rejection of low energetic clusters in a QCD-jet, causing the substructure-based calculation of the tau momentum to differ from the calorimeter-based calculation (which considers all core clusters).

5.6.2. Isolation cone clusters

The reclustering of ECAL cells is done within the full cone of $\Delta R < 0.4$, but the energy subtraction from charged pions is performed in the smaller cone of $\Delta R < 0.2$ only. The reclustered ECAL cells in the isolation annulus $0.2 < \Delta R < 0.4$ represent a new group of clusters, called *outer ECAL clusters* here. An interesting observation is the fact that a BDT training including the number of outer ECAL clusters or their summed up energy leads to additional background rejection. This seems to contradict earlier findings, in which the isolation energy $E_{\mathrm{T}}^{\mathrm{iso}}$ did not contribute significantly. A closer inspection reveals the key difference: while $E_{\mathrm{T}}^{\mathrm{iso}}$ takes into account the energy in both ECAL and HCAL, the outer ECAL clusters are based on the ECAL only. Unfortunately, the pile-up dependence of the number of outer ECAL clusters can not be ignored. Using the cluster energy instead of the (integer) number of clusters offers the possibility to apply a pile-up correction. The variable works equally well in terms of performance.

5.6.3. Clusters and cells

Attempts to replace the cell-based calorimeter variables f_{cent} and f_{track} by cluster-based versions did not lead to satisfactory results. Especially the central energy fraction profits from being based on cells, rather than clusters. Summing up cells in various cones provides the fraction of energy close to the tau barycenter independent of the cluster sizes and shapes. If instead entire clusters are summed up, this can not be ensured. A cluster can have its barycenter within the innermost cone, but can have tails far out in the outer cone. Especially for candidates with only a few clusters, all clusters can have their barycenter either within or outside the central cone, which leads to the extrem values of 1 or 0 if the clusters are taken to calculate the central energy fraction. Since the number of neutral clusters and even more so the number of reconstructed neutral pions is small, a cluster-based calculation of f_{cent} weakens the variable considerably. The fact that f_{cent} is based on cells provides separation power even for the extrem cases with just a very few topocluster, as shown in Fig. 5.33.

5.6.4. Variable sets for the BDT tau ID

BDTs trained with different variable sets indicate that the strength of the current tau ID lies in the variety of variables used. The current variables include all objects within the jet and make use of both the fine detector granularity and the larger scale entities. BDTs trained with additional variables based on substructure information indicate that the rejection power can be increased. By the end of the author's studies, the additional rejection could however not be traced back to a single feature. Rather, it seems to be the result of many additional variables each contributing a little bit. Unfortunately, the additional rejection is associated with a decreased pile-up robustness, and it remains the task to find a good balance between a minimalistic, robust tau ID and maximal background rejection.

5.6.5. Substructure algorithm independent tau ID

Several substructure algorithms are available and the development of all of them is still ongoing. Therefore, a BDT independent of a specific algorithm (or version) can be an advantage. It offers the flexibility to choose the algorithm at a later stage and simplifies an unbiased comparison of the

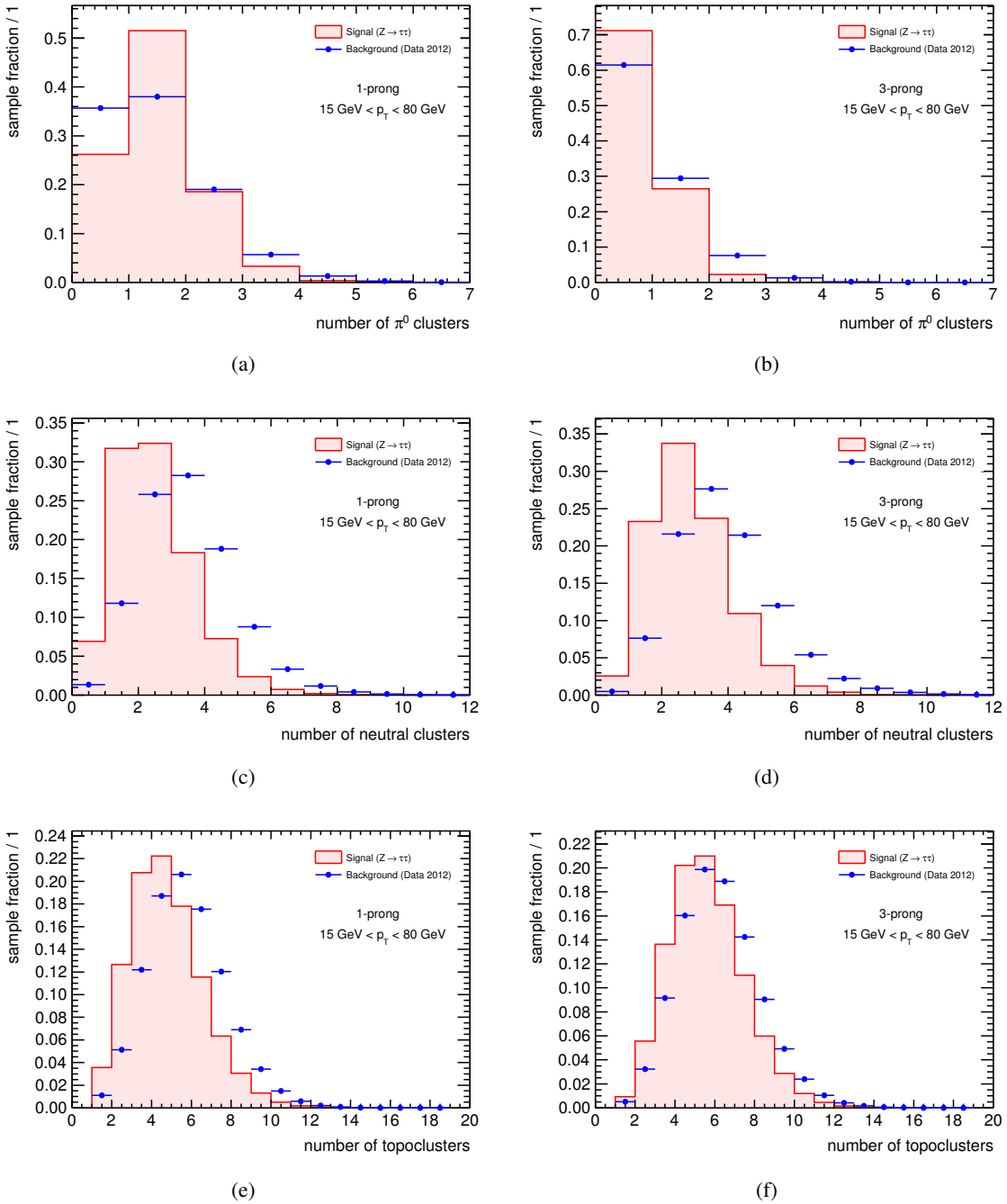


Figure 5.32: Number of identified neutral pions (top), number of reclustered neutral clusters (middle) and number of topoclusters (bottom) in the cone $\Delta R < 0.2$, for 1-prong (left) and 3-prong (right) candidates.

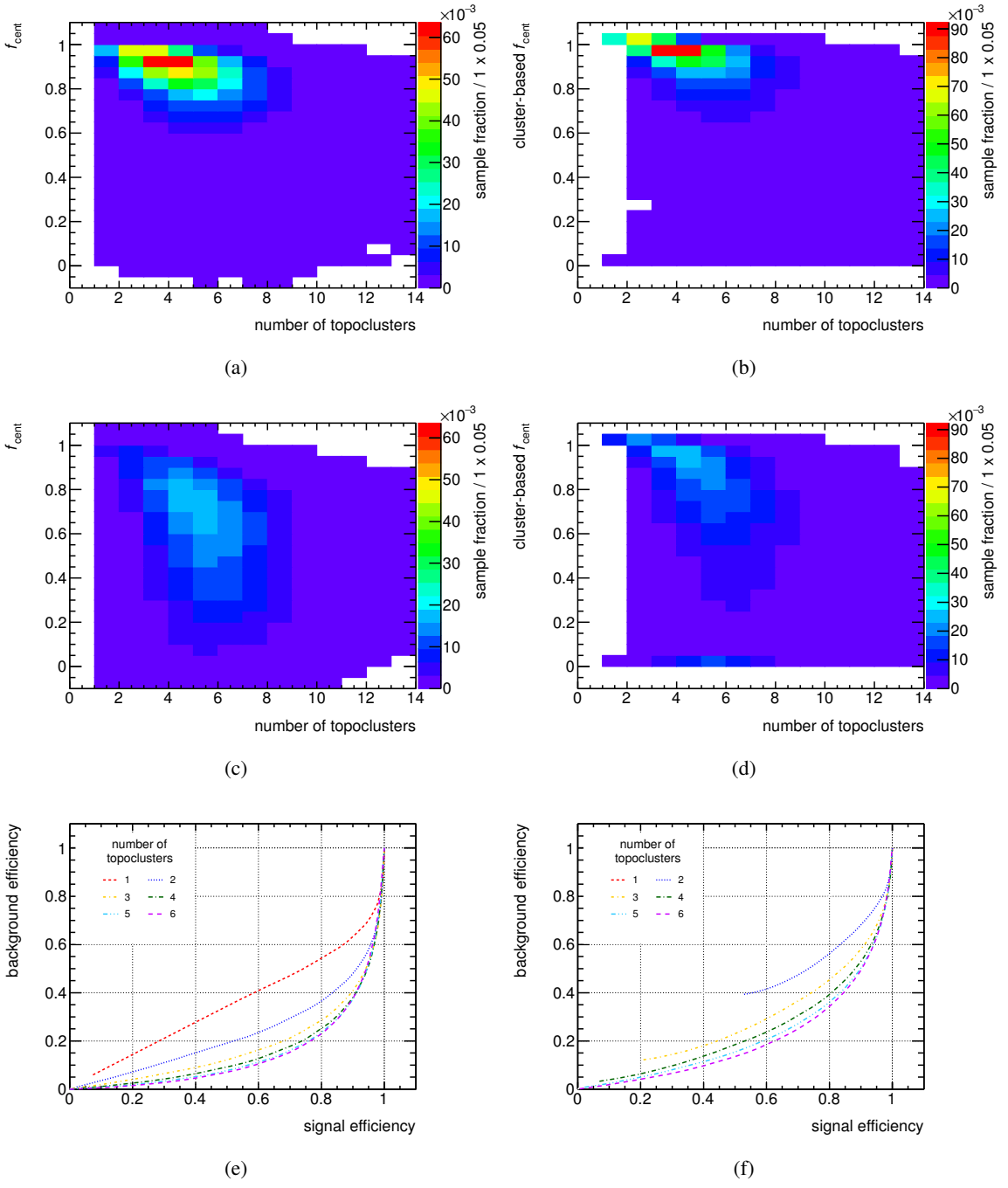


Figure 5.33: Cell-based (left) or topocluster-based calculation of the variable central energy fraction f_{cent} for 1-prong tau candidates. (a) – (d) shows a 2-dimensional illustration of the variable versus the number of topoclusters in $\Delta R < 0.2$, for signal (a,b) and background (c,d) candidates. Figure (e) and (f) show the rejection capability of f_{cent} per topocluster bin, by plotting the background efficiency as a function of the signal efficiency for flat cuts on f_{cent} for a fixed number of topoclusters.

performance of the different algorithms[15]. Furthermore, the substructure algorithms are (currently) not available at trigger level. Therefore, only a BDT that is based on relatively simple discriminating variables can be synchronised with the tau ID at EF trigger level. A second step of background discrimination could be applied on top of the substructure independent tau ID. The step could be optimised for each of the algorithms, so that the best method for a specific purpose can be chosen. The goal of this part of the study is therefore searching for replacements for the three π^0-related variables used in the final 2012 tau ID, without risking the discrimination power. As described in Sec. 5.4.1, the π^0 counting algorithm used in the 2012 tau ID is based on shower shape variables. Similar to the other (not substructure-related) variables in the BDT tau ID, they are calculated per tau candidate and are sums over all clusters or cells within a specific detector layer. The variables do not rely on the substructure algorithm and can therefore be used in the tau ID BDT. The π^0-related variables $m_{\pi^0+\text{tracks}}$ and $p_{\text{T}}^{\pi^0+\text{tracks}}/p_{\text{T}}$ cannot be easily replaced, because they rely on the reconstruction of a π^0 cluster. The two variables add mainly to the additional rejection of the 1-prong BDT tau ID, while N_{π^0} contributes to a much smaller extend. The 3-prong case behaves the opposite way, with N_{π^0} giving the highest additional rejection and $m_{\pi^0+\text{tracks}}$ contributing least [120].

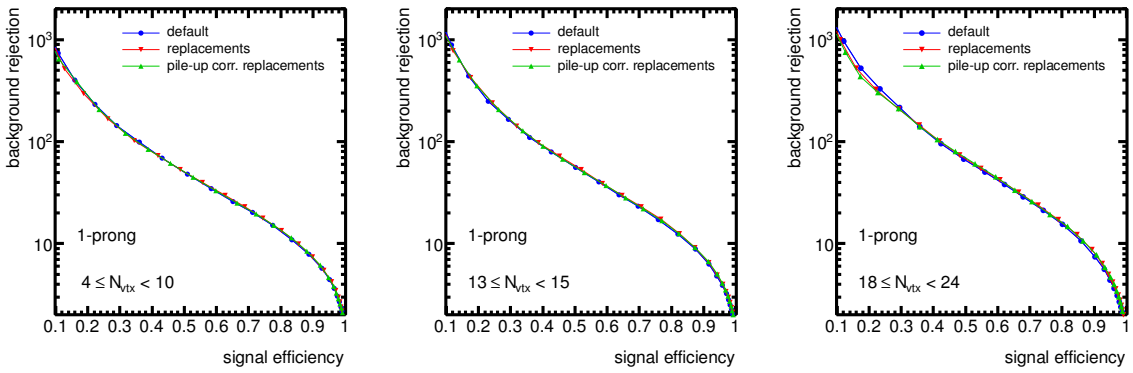
BDTs are trained by replacing the substructure-based variables N_{π^0}, $m_{\pi^0+\text{tracks}}$ and $p_{\text{T}}^{\pi^0+\text{tracks}}/p_{\text{T}}$ by variables which are input variables to the π^0-counting algorithm. Figure 5.34 shows the result of a training with the replacements f_{PSS}, $E_{\text{Ecal}}/p_{\text{tracks}}$ and $f_{\text{Ecal}}^{\pi^\pm}$. As expected, the 3-prong case shows better results, because the additional rejection was caused by N_{π^0} in the first place. Nevertheless, also the 1-prong case gives very similar performance compared to the original variable set. The disadvantage of the replacement set is the pile-up dependence. As shown in Fig. 5.35, the dependence is increased significantly. The π^0 reconstruction includes several steps to make it robust against pile-up or noise, which are missing when the input variables are used directly. However, the background efficiency plotted versus signal efficiency in bins of N_{vtx} shows a good performance in all bins. The increased pile-up dependence is therefore not necessarily a show-stopper, and it is rather a matter of finding a way to correct for it. A quick test of applying a linear pile-up correction on the three variables shows good potential (Fig. 5.35) and does not change the discrimination power (Fig. 5.34).

5.6.6. Conclusion

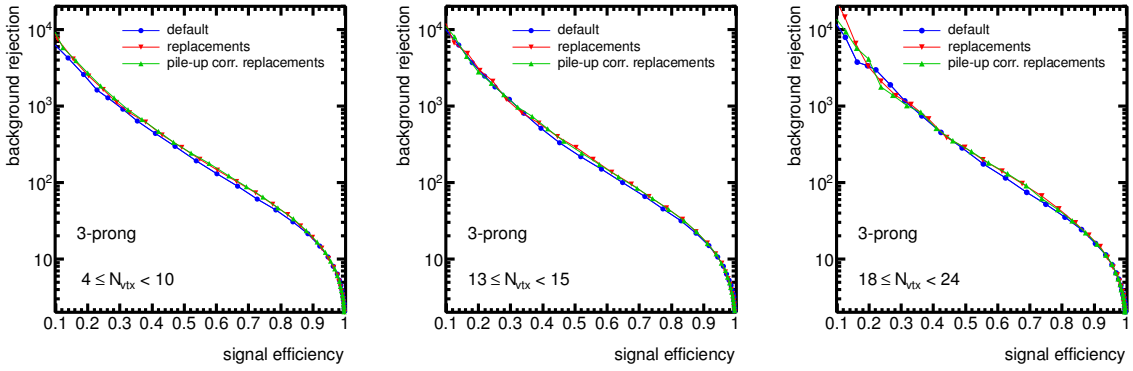
A variable set which is independent of a specific substructure algorithm is proposed. It is achieved by replacing N_{π^0}, $m_{\pi^0+\text{tracks}}$ and $p_{\text{T}}^{\pi^0+\text{tracks}}/p_{\text{T}}$ by f_{PSS}, $E_{\text{Ecal}}/p_{\text{tracks}}$ and $f_{\text{Ecal}}^{\pi^\pm}$. The high background rejection is maintained by using variables that are sensitive to the tau decay mode. However, the pile-up dependence of the BDT is signifiantly increased. Therefore, pile-up correction terms need to be introduced for each new variable.

The investigations concerning the discriminating variables in general show that the strength of the currect ID is the usage of cell-based variables, which can map small-scale differences of QCD-jets and tau objects. Cell-based variables are powerful when small cone sizes are considered, which is also important for high momentum tau leptons. Furthermore, the current tau ID considers all objects within the reconstructed jet object and a large fraction of the background rejection is based on the fact that the disciminating variables are altered by soft clusters. The strength of the ID method is hence based on objects that fail the substructure reconstruction algorithms. Variables that are based on the comparison of quantities calculated with all objects and with substructure objects (like $p_{\text{T}}^{\pi^0+\text{tracks}}/p_{\text{T}}$) are therefore expected to be more powerful than variables exclusively based on substructure objects. There was no time to develop additional variables of such kind, yet this option could be interesting for future studies.

[15]Performance indicators like the success rate of the decay classification or kinematic plots are often created for signal candidates passing a medium tau ID working point.

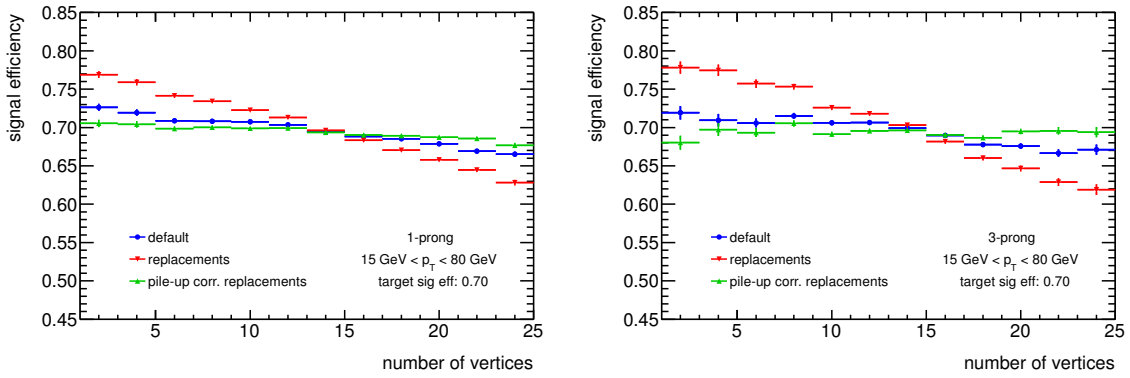


(a)



(b)

Figure 5.34: Background rejection versus signal efficiency for (a) 1-prong and (b) 3-prong candidates and for different pile-up conditions. The default variable set with the three substructure-based variables N_{π^0}, $m_{\pi^0+\text{tracks}}$ and $p_{\text{T}}^{\pi^0+\text{tracks}}/p_{\text{T}}$ (circles), is compared to new set with f_{PSS}, $E_{\text{Ecal}}/p_{\text{tracks}}$ and $f_{\text{Ecal}}^{\pi^\pm}$as a replacement for the three substructure variables (up-side-down triangle) and the new set with an additional pile-up correction applied on each of the three variables (triangle).



(a) (b)

Figure 5.35: Signal efficiency as a function of the number of vertices for (a) 1-prong and (b) 3-prong candidates. The default variable set with the three substructure-based variables N_{π^0}, $m_{\pi^0+\text{tracks}}$ and $p_{\text{T}}^{\pi^0+\text{tracks}}/p_{\text{T}}$ (circles), is compared to new set with f_{PSS}, $E_{\text{Ecal}}/p_{\text{tracks}}$ and $f_{\text{Ecal}}^{\pi^\pm}$ as a replacement for the three substructure variables (up-side-down triangle) and the new set with an additional pile-up correction applied on each of the three variables (triangle).

6. Search for $Z(\rightarrow ee/\mu\mu)H(\rightarrow \tau\tau)$

This chapter describes the search for associated Higgs boson productions with tau leptons in the final state. The analysis has been carried out by the author single-handed. The object quality requirements are inspired by a variety of other ATLAS analyses, especially Z boson analyses and Higgs boson searches. While the fake factor method and the background fit are also commonly used methods as such, the specific implementation and analysis flow has been developed by the author.

The chapter is structered as follows. Section 6.1 gives an introduction to the search for the Higgs boson, and the status of the measurements of the Higgs boson properties. The describtion of the analysis as developed and carried out by the author of this thesis starts in Sec. 6.2 with the search strategy. It is followed by overview of the data and simulated samples in Sec. 6.3. The different steps of the analysis are described in Sec. 6.4 to Sec. 6.9. The discussion of uncertainties is interleaved with the description of the analysis and only summarised in Sec. 6.10. Finally, the result is presented in Sec. 6.11.

6.1. Searches for the Higgs boson

From the very start, one of the main tasks of the multi-purpose LHC experiments CMS [65] and ATLAS [64] was the search for the Higgs boson. Before the start of the LHC, the Higgs boson was being searched for at LEP[1] at CERN and at the Tevatron at Fermilab[2]. The LEP experiments set a lower limit on the Higgs boson mass of 114.4 GeV at 95% confidence level [123]. A global fit to electroweak precision data including the lower mass limit from LEP restricted the mass to be smaller 182 GeV at 95% confidence limit (status 2009 [124]). The high luminosity and high energy of the LHC allowed the ATLAS and CMS collaborations to put tighter constrains on the mass after the first year of data taking at 7 TeV centre-of-mass energy [125, 126]. In July 2012, nearly 50 years after its prediction, the experiments announced the discovery of a Higgs-like boson with a mass around 125 GeV in a joined seminar at CERN [127–129].

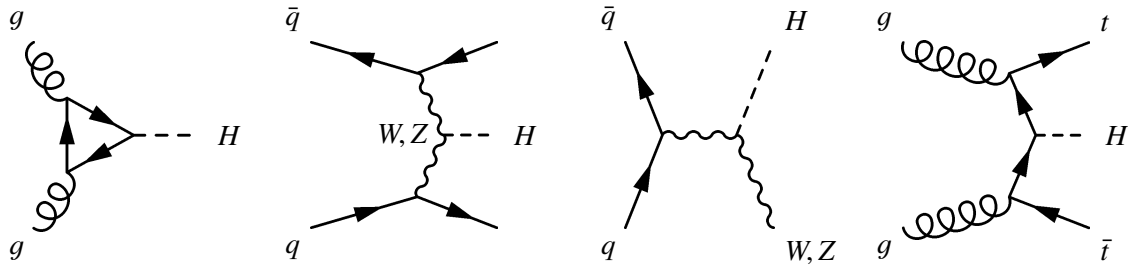
6.1.1. Production of the Higgs boson in pp-collision at the LHC

The Higgs boson can be produced via four leading processes (Fig. 6.1), ordered by cross section for pp-collisions at the LHC (Fig. 6.2):

Gluon-gluon fusion (ggF) is the dominant production mode at the LHC. The interaction of two gluons is mediated by a triangular quark loop to produce a Higgs boson. Since the Higgs boson coupling is proportional to the fermion mass, the quark loop is dominated by top quarks.

[1] The Large Electron-Positron (LEP) collider accelerated and collided electrons and positron between 1989 and 2000. LEP was first operated at a centre-of-mass energy of approximately the Z boson mass. In a second operation phase, the energy was raised to the WW threshold region. Eventually, the energy was raised to a maximum of 209 GeV in order to search for the Higgs boson. The experiments were ALEPH, DELPHI, L3 and OPAL [35, 121].

[2] Fermi National Accelerator Laboratory (Fermilab) is a large accelerator centre in Batavia, Illinois, USA. The Tevatron was operated from 1985 to 2011, bringing protons and antiprotons to collision at energies per beam of nearly 1 TeV. Main experiments were CDF and D0 (or DØ, pronounced D-Zero) [122].



(a) ggF (b) VBF (c) VH (d) ttH

Figure 6.1: Dominating Higgs boson production modes at the LHC: (a) gluon-gluon fusion, (b) weak vector boson fusion, (c) associated production with weak vector bosons, and (d) $t\bar{t}$ fusion.

Vector boson fusion (VBF) is an order of magnitude less likely than the gluon-fusion mode. The two quarks (qq, $\bar{q}\bar{q}$, $q\bar{q}$) in the initial state both radiate off a W or Z boson. The vector bosons then fuse to produce the Higgs boson. The distinct signature of two forward jets in opposite ends of the detector helps to identify the process over the background.

Associated production or Higgs-strahlung (VH) is the production of a Higgs boson in association with a vector boson (W or Z boson). The additional boson can be used to trigger these events, independent of the Higgs decay mode. This production mode is investigated in this thesis.

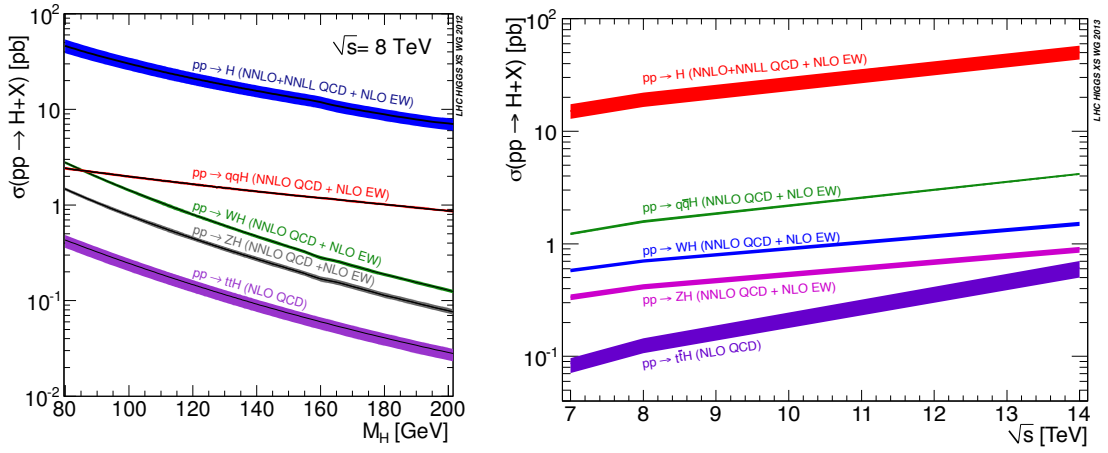
Top quark fusion (ttH) is the production of a Higgs boson in association with two top quarks ($t\bar{t}$). At LHC energies, this production mode has the lowest cross section. It will however profit the most from the increased collision energy in the next data-taking period.

6.1.2. Decay modes of the Higgs boson

The Higgs boson decay branching ratios are largely driven by the mass of the final state particles, due to the mass dependence of the coupling strength of the Higgs field, as shown in Fig. 6.3. A Higgs boson with a mass > 160 GeV decays primarily to $W^{\pm}W^{\mp}$ and ZZ, favouring the decay into W bosons due to the extra degree of freedom from the electromagnetic charge. Below this mass, one of the weak bosons has to be produced off-shell, which suppresses these decay modes. The b quarks are the heaviest fermions that can be produced on-shell here and the decay into $b\bar{b}$ is therefore dominating. This decay mode, in the same way as $H \rightarrow gg$ and $H \rightarrow c\bar{c}$, is hard to separate from the dominat background of multi-jet events at the LHC. The decay into a pair of tau leptons is consequently an important mode for the study of the Yukawa coupling of the Higgs boson to fermions. Furthermore, it is the only channel currently accessible to measure the coupling to leptons. The decay $H \rightarrow \gamma\gamma$ can only proceed via a W boson or (top) quark loop, and is greatly suppressed. Nevertheless, this decay mode played an important role in the discovery of the Higgs boson, due to a clear detector signature, the very well modelled background, and the high mass resolution.

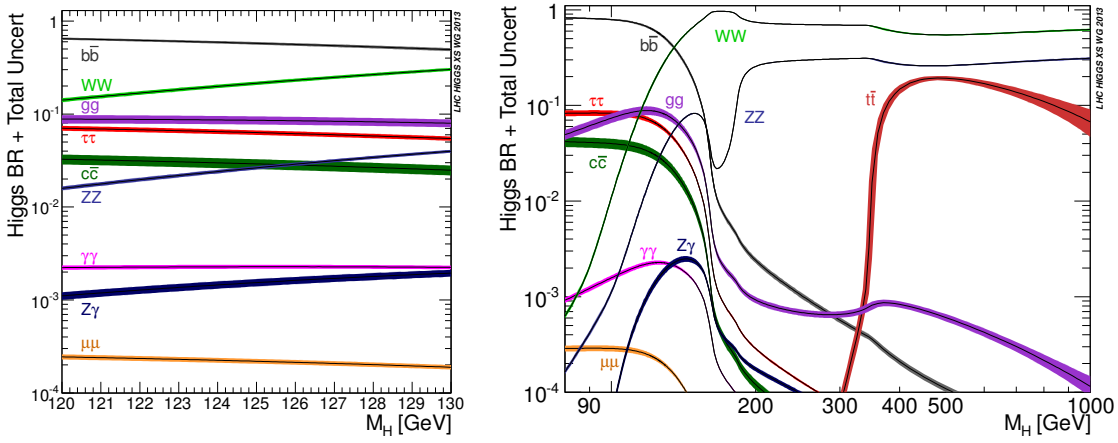
6.1.3. Status of the searches for the Higgs

The announcement of the discovery of a new Higgs-like boson in July 2012 was based on the full data set collected with the ATLAS and CMS experiment in 2011 at $\sqrt{s} = 7$ GeV and the first months of the 2012 data taking period at $\sqrt{s} = 8$ GeV. It was based on a combined measurement of the Higgs



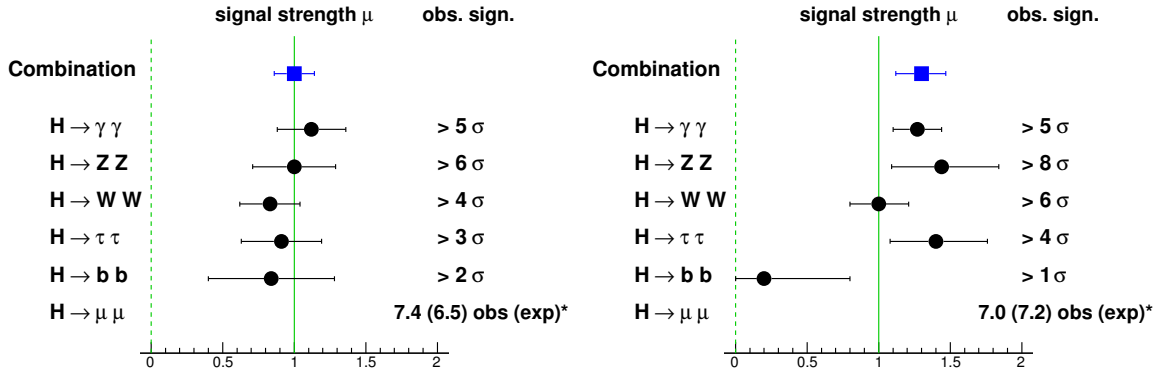
(a) (b)

Figure 6.2: Cross sections of the Higgs boson production modes for the LHC (pp-collisions), (a) as a function of the Higgs boson mass for a centre-of-mass energy of 8 TeV, and (b) as a function of centre-of-mass energy for $m_H = 125$ GeV. The bands indicate the theoretical uncertainties [8, 130].



(a) (b)

Figure 6.3: Higgs boson decay branching ratios as a function of the Higgs boson mass for (a) the mass range (120 – 130) GeV and (b) an extended range up to 1 TeV. The bands indicate the theoretical uncertainties, modified [130].



(a) CMS (b) ATLAS

Figure 6.4: Summary of Higgs boson searches with the best fit signal strength per channel and for the global fit ('Combination'), followed by the observed significance of the excess in each channel. The di-muon analyses are not yet sensitive to the signal. The upper observed (expected) limit at the 95% confidence level for the signal strength is given instead [131, 132].
(a) Measurements by the CMS collaboration. All numbers are calculated for $m_H = 125\,\text{GeV}$ [133].
(b) Measurements by the ATLAS collaboration. The values for the signal strength are from Ref. [134], assuming $m_H = 125.5\,\text{GeV}$, with the exception of the di-photon analysis, which have been updated recently [135]. The combined signal strength is based on the previous study, with $\mu = 1.57^{+0.33}_{-0.28}$. The significances are from Refs. [136–139] and based on the best fit mass.

boson decaying to $\gamma\gamma$, ZZ and WW. Since then, more data has been collected, the analyses have been refined and measurements of properties and additional channels have been carried out. Also the Tevatron experiments published the results of their full dataset, seeing hints of the newly discovered boson. An overview of the current status of the SM Higgs boson searches is given here.

With the full 2011 and 2012 dataset available, CMS and ATLAS have released searches for the decay channels $H \to \gamma\gamma, ZZ, WW, \tau\tau, bb, \mu\mu$ (Fig. 6.4) [131–139]. Even though the first two decay modes are relatively rare, they have the greatest sensitivity. The decay signature involves photons, electrons and muons only[3], with a good mass resolution and low uncertainties. The excess in data over the background expectation exceeds the 5σ level in both channels individually. The signal strength is defined as the observed cross section relative to the SM cross section, $\mu = \sigma_{\text{obs}}/\sigma_{\text{SM}}$. Combining all measured decay chains, the best fit signal strength is [133, 134]:

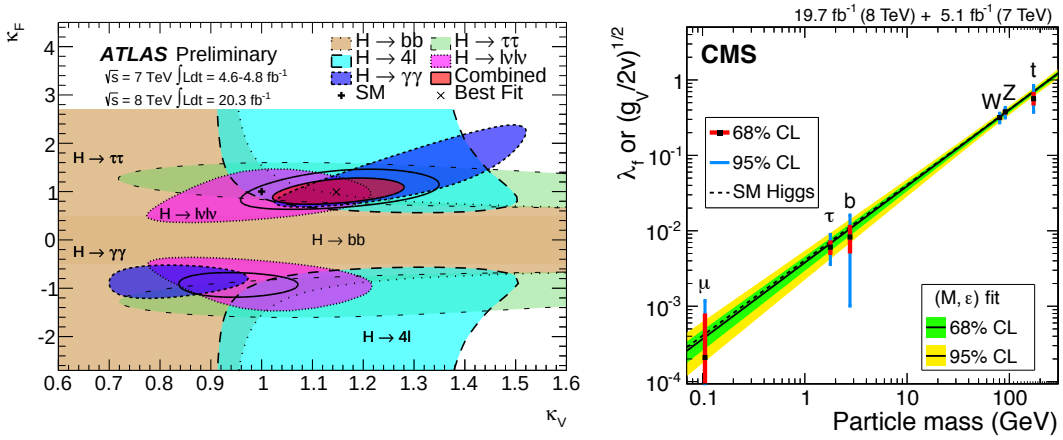
$$\text{CMS}: \quad \mu = 1.00^{+0.14}_{-0.13} \,,$$
$$\text{ATLAS}: \quad \mu = 1.30^{+0.18}_{-0.17} \,.$$

This is in good agreement with the SM expectation $\mu = 1$.

The decay channels most sensitive to the mass are $H \to ZZ$ and $H \to \gamma\gamma$, which are consequently used to determine the mass [133, 140]:

$$\text{CMS}: \quad m_H = (\,125.02\ ^{+0.26}_{-0.27}\ (\text{stat})\ ^{+0.14}_{-0.15}\ (\text{syst})\,)\ \text{GeV} \,,$$
$$\text{ATLAS}: \quad m_H = (\,125.36\ \pm 0.37\ (\text{stat})\ \pm 0.18\ (\text{syst})\,)\ \text{GeV} \,.$$

[3]The $H \to ZZ$ analysis considers $ZZ \to 4l$, $l = e, \mu$ only.



(a) (b)

Figure 6.5: a) Result for a model probing different coupling strength scale factors for fermions (κ_F) and for bosons (κ_V). The plot shows the correlation, derived for the individual decay channels and their combination. For more details on the model, see Ref. [134].
b) Result for a model probing different coupling strength scale factors for each particle type analysed. The y-axis is specific for fermions and bosons, taking into account the difference of Yukawa and weak boson couplings. For more details and the model, see Ref. [133].

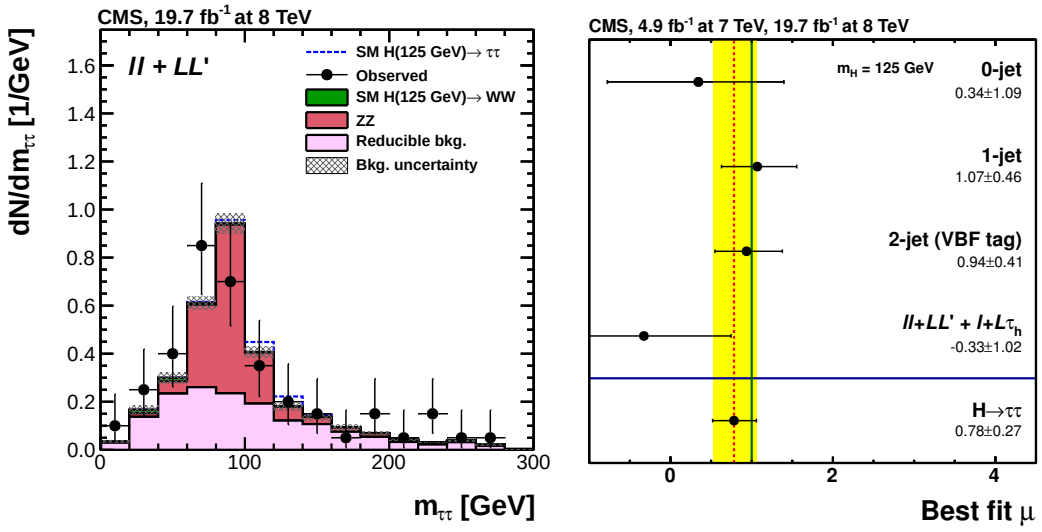
Spin and parity measurements done in CMS, ATLAS, and at the Tevatron have rejected various spin hypotheses and did not find any deviations from the SM expectation [141–143]. $J^{PC} = 0^{++}$ is clearly favoured. By combining measurements from all observed decay modes, the Higgs boson couplings are probed in multiple ways and models, finding no deviation from the SM expectation [133, 134]. An example is shown in Fig. 6.5(a), testing models with deviations of the coupling strength to fermions or vector bosons. Here, the importance of the di-photon channel becomes clear, being sensitive to the relative sign due to the interference of W and top quark loops in this decay mode. Figure 6.5(b) illustrates a fit to a model in which a different signal strength is allowed for each of the studied decay channels. The deviation is translated into a variable scaling with the particle mass, taking into account the difference in Yukawa and weak boson couplings. Even though the uncertainties for some decay channels are still high, the plot nicely demonstrates the scaling of the Higgs boson coupling with particle mass, for leptons, quarks and weak bosons.

The total Higgs boson decay width Γ_{tot} of around 4 MeV [144] cannot be determined directly at the LHC. However, the relative on-shell and off-shell measurements of $H \to ZZ$ give access to the total width, assuming no energy-scale dependence of the couplings, i.e. no physics beyond the Standard Model. The analysis has been performed by both CMS and ATLAS, achieving compatible limits for the 95% CL [145, 146]:

$$\begin{aligned} \text{CMS}: \quad & \Gamma_{tot}/\Gamma_{SM} < 5.4\ (8.0)\ \text{observed (expected)}\ , \\ \text{ATLAS}: \quad & \Gamma_{tot}/\Gamma_{SM} < 5.7\ (8.5)\ \text{observed (expected)}\ . \end{aligned}$$

The associated production of a Higgs boson and a weak vector boson (VH) is an important channel at the Tevatron. The collision of protons and antiprotons at $\sqrt{s} = 1.96$ TeV favours the VH production over the VBF production. Combining the results of both experiments, D0 and CDF see an excess of more than 3σ local significance at $m_H = 125$ GeV, dominated by the VH production mode [147, 148].

At the LHC, VH production is suppressed as opposed to ggF and VBF, as shown in Fig. 6.2. It



(a) (b)

Figure 6.6: Results of the CMS search for $H \to \tau\tau$ [150].
(a) Mass distribution $m_{\tau\tau}$ in the search region optimised for the ZH production mode for $\sqrt{s} = 8$ TeV. The background is normalised according to the global fit result.
(b) Signal strength μ for the best-fit result for the independent search categories, and the combined result from the global fit, for a SM Higgs mass of 125 GeV. The entries $ll + LL'$ and $l + L\tau_{\mathrm{h}}$ refer to analyses optimised for the VH production modes.

has therefore played a minor role in most LHC Higgs boson searches so far. An exception is the search for $H \to b\bar{b}$ [139, 149], which happens at relatively high rate, but is hard to separate from the enormous multi-jet background in the ggF and VBF production modes. The CMS collaboration sees a 2σ excess, based solely on VH events and being in very good agreement with the SM expectation. Other analyses suffer from large uncertainties but have started to add dedicated VH search regions for the decay modes ZZ [136], $\gamma\gamma$ [133, 135], WW [133] and $\tau\tau$ [150]. Of special interest here is the latter, which includes the channel under investigation in this thesis. Figure 6.6 shows the result of a CMS analysis of this channel. The signal expectation in the ZH search region is small compared to the statistical uncertainty. The best fit signal strength gives negative values, indicating a measured deficit. This is however not in contradiction to the SM expectation due to the large uncertainties.

In future data taking periods, the associated production is expected to play a more important role. Due to the increased center-of-mass energy and potentially higher number of simultaneous proton-proton interactions, the trigger p_{T} thresholds likely have to be raised and (relatively) fewer ggF events will be collected. VH events then provide a good signature for multi-object or topological triggers, which can be operated with lower momentum thresholds.

6.2. Search signature and strategy

The signature discussed in this thesis is the production of a Higgs boson in association with a Z boson. As illustrated in Fig. 6.7, only the leptonic decays $Z \to ee$ and $Z \to \mu\mu$ are considered, with the Higgs boson decaying into a pair of tau leptons. All tau decays are included, so that there are six

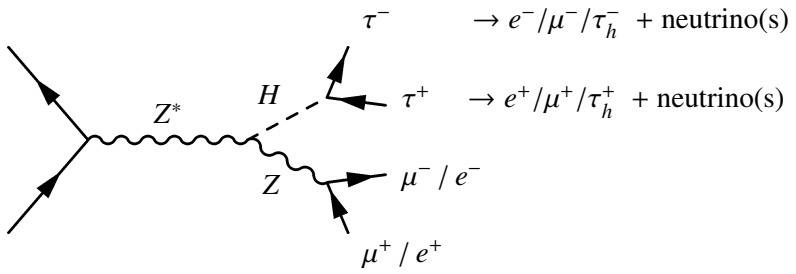


Figure 6.7: Diagram of the signature under study.

final states for the Higgs boson decay[4]: $\tau_e\tau_e$, $\tau_\mu\tau_\mu$, $\tau_e\tau_\mu$, $\tau_e\tau_{\text{had}}$, $\tau_\mu\tau_{\text{had}}$, $\tau_{\text{had}}\,\tau_{\text{had}}$. Due to the different probabilities for a leptonic or hadronic tau decay, the final states occur at different rate. The double hadronic final state and the mixed leptonic-hadronic modes constribute roughly equally, with about 42% for $\tau_{\text{had}}\tau_{\text{had}}$ and 46% for $\tau_{\text{lep}}\tau_{\text{had}}$. The double leptonic mode contributes with the remaining 12%.

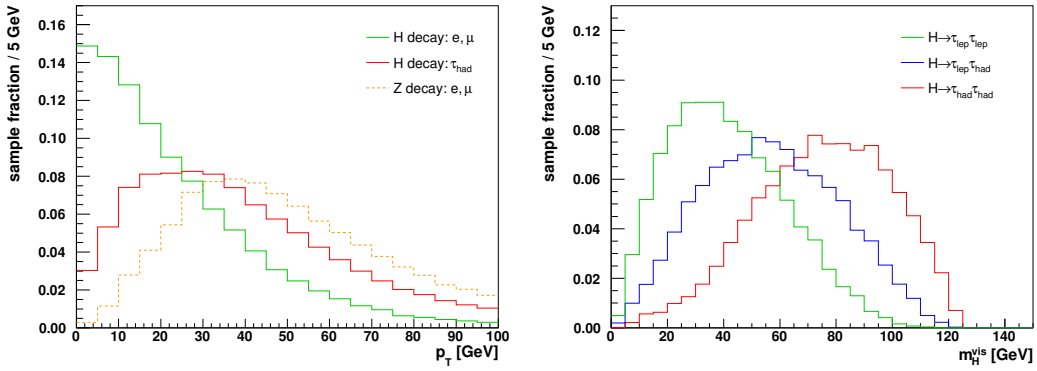
The SM cross section of the process $Z(\rightarrow ee/\mu\mu)H(\rightarrow \tau\tau)$ is small and only about 34 events are expected in the full 2012 data. Besides being one of the less probable production modes at the LHC, the requirement of the leptonic decay of the Z boson reduces the number of expected events drastically. The two considered modes, $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$, represent only 6.7% of all Z boson decays. Because the sensitivity towards the signal is not expected to be high, the analysis presented here is carried out with a focus on the exploration of the signature and the various backgrounds as well as finding background control regions and side bands. The background contribution is estimated with simulated samples. This approach was chosen in order to provide a cross check for the fully data-driven approach used by other members of the collaboration[5]. Furthermore, it provides the possibility to validate the MC samples and hence helps to decide to which extend simulated samples are useful for the signal search in the upcoming data taking period.

The analysis strategy is as follows: Events are selected by searching for the Z boson, i.e. a pair of electrons or muons with opposite electric charge and an invariant mass close to the mass of the Z boson (*Z-pair*). Tight identification requirements on the electrons and muons reduce the amount of events (backgrounds) without an actual Z boson decay. Among the selected events, events with additional e, μ or τ_h are considered in the analysis. Events with only one additional object (e, μ or τ_h) are used for background estimation measurements, for instance measurements of mis-identification probabilities. Events with (at least) two additional objects represent signal candidates if two objects carry opposite electrical charge and both are well identified (*H-pair*). The invariant mass of the pair is used for the signal extraction. The cross section for the process $Z(\rightarrow ee/\mu\mu)H(\rightarrow \tau\tau)$ can be derived, or a limit can be set, by comparing the observed distribution to the expected one.

Neglecting for the moment the production of additional jets, the Z boson and the Higgs boson recoil against each other and decay. In contrast to the Z-decay, the H-decay to tau leptons contains neutrinos in the final state, which carry away momentum. As a consequence, the momentum spectrum of the objects associated to the Higgs boson is softer than the spectrum of the leptons associated to the Z boson. Hence, it is of advantage to maintain low momentum thresholds for the H-pair objects, as illustrated in Fig. 6.8. The invariant mass distribution of the Z-pair peaks at the mass of the Z boson,

[4]The following symbolic notation is used here: τ_{lep}, τ_e and τ_μ for the leptonic decays, and τ_{had} for the hadronic decay. The detector object, i.e. the visible part of the hadronic decay, is denoted τ_h.

[5]The official ATLAS analysis is not yet publicly available, but sub-channels are described in References [151] and [120]. Unfortunately, the specific requirements used in each approach diverged stronger than anticipated, which currently complicates a direct comparison.



(a) (b)

Figure 6.8: The $ZH(\rightarrow \tau_{\text{had}}\tau_{\text{had}})$ signature: (a) transverse momentum spectrum of the electron, muon and τ_h associated to the Higgs boson, and (b) visible transverse mass of the Higgs boson decay products. The distributions are obtained from simulated ZH samples with a Higgs boson mass of 125 GeV.

which helps to select the ZH events and to suppress backgrounds, that do not contain a Z boson. The invariant mass distribution of the H-pair is shifted towards lower values relative to the mass of the Higgs boson and is broader, again caused by the presence of neutrinos from the tau decays. The leptonic tau decays include two neutrinos, as opposed to the hadronic decays with only one neutrino, so that the position of the invariant mass peak depends on the specific final state. As shown in Fig. 6.8, the invariant mass distributition peaks at ≈80 GeV for the τ_{had} τ_{had} final state, at ≈55 GeV for the $\tau_{\text{lep}}\tau_{\text{had}}$ final state and at ≈30 GeV for the $\tau_{\text{lep}}\tau_{\text{lep}}$ final state.

The number of events in the signal region is rather small, because there are not many Standard Model processes with a similar signature. Due to the requirement of a well-identified Z boson, the main background comes from events in which a Z boson is created along with other particles, such as Z + jets, ZW or ZZ events.

***ZZ* events** represent the irreducible background. The same particles appear in the final states of ZZ and ZH events, and a distinction is only possible by reconstructing the boson mass or spin. The reconstruction of the boson spin is currently not feasible, due to the low number of events and high uncertainties for the spin reconstruction. An exact reconstruction of the boson mass is also not possible due to the presence of neutrinos from the tau decays. Instead, the *invariant mass* of the visible tau decay products (H-pair) is used. A better mass discriminator can be obtained from the event topology and the missing transverse energy. The latter is then used as an estimator for the sum of the transverse momenta of all neutrinos in the event. Advanced methods, such as the MMC technique [152], evaluate the missing tranverse energy and the tau leptons in order to identify neutrinos that are most consistent with the Higgs boson-decay to tau leptons. However, such techniques require a very good modelling and understanding of the missing energy, so that the invariant mass is given priority here.

Because the Higgs boson couples to mass, the direct decays to ee and $\mu\mu$ are very heavily suppressed. The Z boson, on the other hand, decays with equal probability to ee, $\mu\mu$ and $\tau\tau$. The different final states of the ZH analysis therefore encounter a varying amount of ZZ background. The $\tau_e\tau_e$ and $\tau_\mu\tau_\mu$ states see the highest contribution. The invariant mass of the electron or muon pair in $Z \rightarrow ee$ or $Z \rightarrow \mu\mu$ decays peaks at the Z boson mass of 91 GeV, while the distribution of a pair from a tau decays ($H \rightarrow \tau_{\text{lep}}\tau_{\text{lep}}$ or $Z \rightarrow \tau_{\text{lep}}\tau_{\text{lep}}$) peaks below 50 GeV. A veto on the Z boson mass peak therefore greatly increases the sensitivity in these channels. Off-shell production of the second Z boson

contributes also to the low-mass spectrum. Requiring some minimal amount of $E_{\mathrm{T}}^{\mathrm{miss}}$ helps to cut down such events. The other final states are dominated by actual $Z \to \tau\tau$ decays. Due to the small difference of Z boson mass and Higgs boson mass, the invariant mass distributions overlap[6].

***WZ* events** contribute mainly to the control regions, where only one additional e, μ or τ_h is required. Nevertheless, if an additional jet appears in the event, this can be mistaken for a τ_h, or in rare cases for an electron or muon. WZ events therefore contribute also to the signal region. Because the H-pair does not originate from a resonant decay in this case, the invariant mass distribution is expected to be wider than for the signal.

Z + jet events only enter the signal region if at least two jets are mistaken for τ_h, e or μ. Due to the large cross section for Z production, this nevertheless happens. The dominant mis-identification is jet $\Rightarrow \tau_h$. Even though the mis-identification probability decreases with momentum, the fakes appear over a large p_{T} range. Jets can also fake electrons. The probability can however be drastically reduced by applying tight identification requirements on the electrons. Z + jet events are also a source of non-prompt electrons. Photons radiated off leptons from Z decays may convert early in the detector into a pair of electrons. Typically, such electrons appear close to one of the Z leptons. Even though these are actual electrons, they are considered 'fakes' here. Muons are not directly faked by jets, but appear as non-prompt muons from the jet. Charged pions, for example, are jet constituents which decay dominantly into muons, and are thus a source of real muons if they decay early. Such a muon is often produced in the direction of the jet and can therefore be rejected by requirements on the muon isolation. Another source of non-prompt muons and also electrons is the decay of b quarks. Such leptons stand out as being rather isolated and are thus harder to reject. Such leptons are often rejected by requirements on the transverse impact parameter, i.e. the fact that the track does not precisely point back to the primary vertex. However, this handle cannot be used here, because the signal displays the same behaviour.

$t\bar{t}$ events contribute to the analysis background, because of the high cross section and (relatively) great number of prompt and non-prompt leptons. Top quarks are extremely short-lived and decay dominantly to b quarks under the emission of a W boson. The b quarks can emit light leptons or their jets can be mistaken for a hadronic tau decay. The two W bosons are a source of well-isolated leptons, with a broad invariant mass spectrum. Therefore, some $t\bar{t}$ events are picked up despite the requirement on the Z boson mass. The amount of top quark events in the signal region is however small, as the two b quark decays have to result in at least two fake e, mu, or τ_h. In addition to $t\bar{t}$ events, other topologies with top quarks contribute to the top background, for instance single top events, tW, tZ or $t\bar{t}Z$ events. Due to the much lower cross section, these contribution are insignificant.

Other backgrounds such as W+jet, WW or tri-boson events contribute at a negligible level, either due to a low cross section or a very different even topology (missing Z boson, low number of prompt leptons). Multi-jet events do not contribute at a significant level either, despite the high cross-section.

Higgs boson events Other Higgs boson events can contribute to the signal region of the ZH search. The dominat contribution comes from the decay $H \to ZZ \to 4$ leptons, which is visible in

[6]The search discussed here is not restricted to the discovered Higgs boson with a mass of $\approx 125\,\mathrm{GeV}$. Rather, masses in the range $m_H[\,\mathrm{GeV}] \in \{100, 105, ..., 145, 150\}$ are considered. Due to the dropping branching ratio of $H \to \tau\tau$ with increasing mass, the problem of a dominating contribution from Z decays remains valid even for the higher masses under study.

Table 6.1: Fraction of good quality data delivered by the various detector subsystems in 2012 [153].

ATLAS p-p run: April-December 2012										
Inner Tracker			Calorimeters		Muon Spectrometer				Magnets	
Pixel	SCT	TRT	LAr	Tile	MDT	RPC	CSC	TGC	Solenoid	Toroid
99.9	99.1	99.8	99.1	99.6	99.6	99.8	100.	99.6	99.8	99.5
All good for physics: 95.5%										
Luminosity weighted relative detector uptime and good quality data delivery during 2012 stable beams in pp collisions at $\sqrt{s}$=8 TeV between April 4th and December 6th (in %) – corresponding to 21.3 fb^{-1} of recorded data.										

the $\tau_e\tau_e$ and $\tau_\mu\tau_\mu$ final states. Other decay modes contribute by negligible amounts. This background is taken into account assuming the SM predicted cross section for $m_H = 125\,\text{GeV}$.

6.3. Data sample and simulated samples

6.3.1. Data sample

The data analysed in this chapter is the recording of proton-proton collisions at a centre-of-mass energy of $\sqrt{s} = 8\,\text{TeV}$. The data was collected between April and December 2012 with the ATLAS detector. Collision events are only considered if they are recorded during stable beam conditions and when all detector subsystems were in operation. As shown in Table 6.1, the latter was the case for more than 95% of the recorded data. Noise bursts, hot calorimeter cells or other local detector failures can have negative impact on the data reconstruction quality. Therefore, such events are discarded. The high quality run numbers and luminosity blocks result in a data sample with an integrated luminosity of $20.3\,\text{fb}^{-1}$. Additional cleaning cuts are applied on data to remove rare detector failures which are not masked in the detector simulation and are not removed by the data quality requirements. With respect to the calculation of the integrated luminosity, the fraction of the data affected is in most cases insignificant. In the rare cases where it is not insignificant, L_{int} is scaled accordingly.

6.3.2. MC samples

The recorded data are compared to simulations. In order to limit the amount of simulated samples only processes that are expected to contribute to the analysis phase space are considered. These processes are in particular: Z + jets, W + jets, diboson events (WW, WZ, ZZ), $t\bar{t}$, single top, and Higgs boson events. Due to the pre-selection requirement of an opposite sign lepton pair (the Z-pair), processes without leptons in the final state are largely rejected. Therefore, the focus is on simulated samples with leptons in the final state. A detailed list of the processes used is shown in Appendix B.1. The samples are scaled according to the integrated luminosity L_{int} of the 2012 data sample. The weight w_i applied on each simulated event is dependent on the cross section σ_i of the process i and the number of events N_i generated:

$$w_i = L_{\text{int}} \times \frac{\sigma_i \times \epsilon_i \times k_i}{N_i} \quad . \tag{6.1}$$

The factor ϵ_i accounts for filter efficiencies during the MC production. For some processes, such filters are applied early in the event generation process in order to enhance the amount of events relevant for a physics analysis. Examples for such filters are momentum thresholds, pseudorapidity ranges or requirements of at least one lepton. The factor k_i is applied to correct the cross section for high order contributions. In particular, cross sections calculated by the event generators at leading order are brought to next-to-leading order level. Higher order contributions can significantly change the

Table 6.2: MC generators, parton shower model, parton distribution function (PDF) and generator tune for the production of the simulated samples used in the VH analysis.

Process	Generator, parton shower	PDF	Generator tune
Z + jets, W + jets	Alpgen + Pythia + Photos + Tauola	CTEQ6L1	Perugia2011C
DY Z + jets	Alpgen + Herwig Jimmy + Photos + Tauola	CTEQ6L1	AUET2 CTEQ6L1
$t\bar{t}$, single top, $gg \to ZZ$	McAtNlo + Herwig Jimmy + Photos + Tauola	CT10	AUET2 CT10
$t\bar{t}Z$, tZ ZZZ^*, ZWW^*	MadGraph + Pythia + Photos + Tauola	CTEQ6L1	AUET2B CTEQ6L1
WH+ZH: $H(\to \tau\tau)$ $H(\to ZZ)$ $H(\to WW)$	Pythia8 + Photospp	CTEQ6L1	AU2 CTEQ6L1
WZ, ZZ ggF+VBF: $H(\to \tau\tau)$ $H(\to ZZ)$ $H(\to WW)$	Powheg + Pythia8 + Photospp	CT10	AU2 CT10
WW, WZ, ZZ	Herwig Jimmy + Photos + Tauola	CTEQ6L1	AUET2 CTEQ6L1

calculated cross section. For example, the Z + jets samples used in the analysis receive a correction of about 20%[7]. The number of simulated events for each MC sample along with the product of cross section and k-factor and filter efficiencies are listed in Appendix B.1.

Different MC generators, parton shower models, and generator tunes are used for the various processes, as listed in Table 6.2. The Z + jets and W + jets samples are generated with ALPGEN [155]. The parton shower and hadronisation comes from PYTHIA [156], tuned according to the Perugia2011C [157] scheme[8]. CTEQ6L1 [158] provides for the parton distribution functions at leading order. Other combinations of generator and showering tool using the CTEQ6L1 PDF set are ALPGEN [155] with HERWIG [159] and JIMMY [160] (Drell Yan Z + jets) or MadGraph [161] with PYTHIA (top quark production in association with a vector boson and triboson events). Furthermore, standalone PYTHIA 8 [32] (VH Higgs events) and standalone HERWIG+JIMMY (WW events, and WZ, ZZ for cross checks) are used with this PDF set. The generator tunes come from the AU*T2 [162] family and are optimised for the CTEQ6L1 PDFs. The specific tune is chosen depending on the tool

[7]For more information on cross sections and k-factors see for example Ref. [154]. The numbers used here are obtained from ATLAS internal sources.

[8]The parameters of the event generation are tuned with data both from previous experiments (e.g. at LEP or Tevatron) and from LHC experiments, for example by comparing distributions in data and MC of charged particle multiplicities or kinematic quantities.

for parton showering and hadronisation, matching AUET2 and HERWIG+JIMMY, AUET2B and PYTHIA, and AU2 and PYTHIA 8. Processes simulated at next-to-leading order make use of the CT10 [163] PDF set. The events are generated with MC@NLO [164] and HERWIG+JIMMY (top quark events and $gg \rightarrow ZZ$) or POWHEG [165] and PYTHIA 8 (WZ, ZZ events, ggF and VBF Higgs events). The tunes come again from the AU*T2 family, but are optimised for the CT10 PDFs. Samples generated with HERWIG+JIMMY or PYTHIA make use of PHOTOS [33] for the simulation of QED radiative corrections and TAUOLA [166] for the simulation of the decay of polarised tau leptons. PHOTOS++ is also used to simulate the QED radiative corrections in combination with PYTHIA 8.

An important background in the analysis are $Z(\rightarrow ee, \mu\mu)$ + jets events because they have a high cross section and therefore dominate the analysis, especially at an early selection state. The events can enter the signal region if jets are mistaken for electrons, muons or τ_h. As mentioned before, leptons can also be emitted from jets, e.g. due to the semi-leptonic decay of a b quark. Such leptons are considered 'fakes' as well. It is therefore important to model and understand the misidentification probability correctly. Leptons are more likely to be emitted from heavy flavour jets, i.e. jets induced by c or b quarks, than from light flavour jets. Therefore, dedicated $Z+c\bar{c}$+jets and $Z+b\bar{b}$+jets samples are used, which provide a better description of the heavy flavour jets by considering their mass. In order to avoid an overestimation ('double counting'), events with heavy flavour jets are removed from the standard Z + jets sample. Furthermore, the $Z + c\bar{c}$ + jets sample is cleaned from light flavour and b jet events, and the $Z + b\bar{b}$ + jets sample is cleaned from light flavour and c jet events.

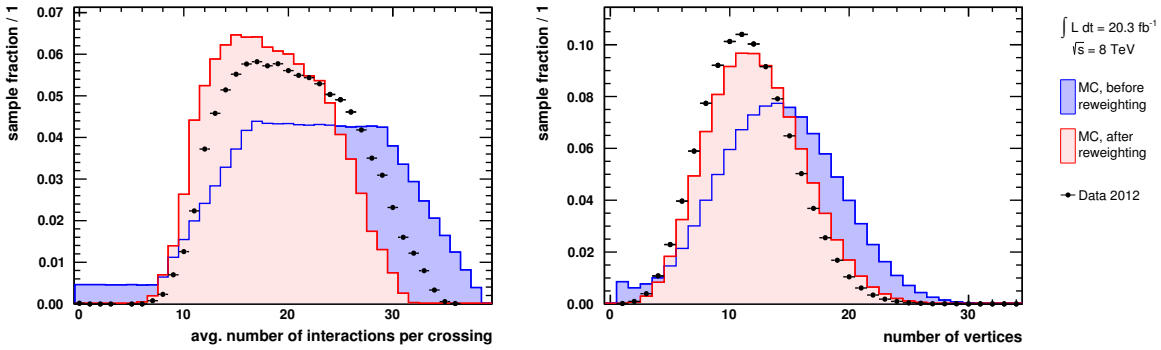
6.3.3. Reweighting of MC samples

Event activity

The MC samples were simulated before the actual data taking began and the pile-up conditions could only be estimated based on the foreseen beam parameters. The simulated distribution of the number of simultaneous proton-proton interactions does therefore not match the distribution observed in data. In order to allow for a reweighting, the MC samples were simulated with sufficient statistics in the low and high tails (see Fig. 6.9(a)). All simulated samples are reweighted with respect to the average number of interactions per bunch-crossing, $\langle\mu\rangle$. Measurements performed on minimum bias events show that an additional scaling of $\langle\mu\rangle$ is needed, in order to reproduce the average event activity as observed in data. The number of vertices in the event is a measure for the event activity, and illustrated in Fig. 6.9(b)), before and after the reweighting procedure. The fact that the distributions in data and MC do not match completely is due to a different beam spot size in the MC samples (σ_z = 66 mm versus σ_z = 47 mm in data). The smaller beam spot size in data leads to a lower vertex multiplicity, due to more merged vertices. The small mis-match in the number of vertices is therefore an artefact of the vertex reconstruction; the average event activity in MC and data matches after the reweighting. The scaling factor is provided by the respective working group.

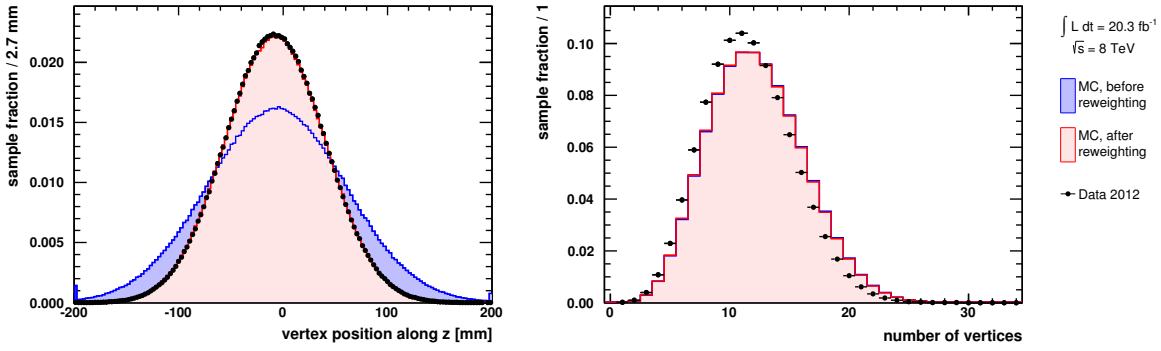
Vertex position

The MC samples are reweighted in order to match the beam spot size along the z-direction. The reweighting is done with respect to the generated z-position of the hard interaction. The result for the reconstructed vertex position is shown in Fig. 6.10(a) . The MC samples correctly reflect the beam spot size measured in data after the reweighting. The impact on the distribution of the number of vertices as observed in the analysis sample is however barely visible (Fig. 6.10(b)). The subsequent reweighting has no impact on the vertex reconstruction procedure as such and can therefore not mirror the probability of merged vertices.



(a) (b)

Figure 6.9: Distribution before and after reweighing of (a) the average number of interactions per bunch-crossing and (b) the number of reconstructed vertices N_{vtx}. The reweighting has the purpose of matching the average event activity in data and MC. All distributions are normalised to unit area.



(a) (b)

Figure 6.10: Distribution before and after reweighting of (a) the z-position of the first primary vertex in the event and (b) the number of reconstructed vertices N_{vtx}. The reweighting has the purpose to correct for the different beam spot size in data and MC. All distributions are normalised to unit area.

6.4. Pre-selection: Z boson decay

The first step of the analysis is the pre-selection. Events are selected that are of interest for the ZH search. The selection is very loose so that the sample is not biased by the presence of signal events. Consequently, the data-MC agreement can be validated, and the sample can be used to control mismodelled background contributions. The pre-selection consist primarily of the selection of events that are compatible with the decay of a Z boson to electrons or muons. Furthermore, all cleaning cuts are applied and basic objects, namely electrons, muons and hadronically decaying tau leptons, are selected.

6.4.1. Event cleaning

Trigger: the event has to pass either of four triggers, requiring a single muon of 24 GeV or 36 GeV or a single electron of 24 GeV or 60 GeV. The lower momentum triggers have an additional isolation requirement: the sum of track momenta in a cone of $\Delta R \leq 0.2$ around the muon (electron) must not exceed 12% (10%) of the lepton's momentum (see also Sec. 4.3).

Data streams: The selected data streams 'Muons' and 'Egamma' contain events passing the muon and electron triggers, respectively (see Sec. 3.3.6). Events passing both an electron and a muon trigger will be part of both streams. In order to avoid a double counting of events, a trigger hierarchy is used: events are taken from the stream 'Muons' if they pass any of the two muon triggers. If the events fail both muon triggers, but pass any of the two electron triggers, they are taken from the stream 'Egamma'. This hierarchy has no influence on the event selection as such, as both streams contain identical information and are treated alike in the analysis. Because simulated samples are not organised in streams, the trigger hierarchy is neither necessary nor applicable to MC.

Primary vertex: the first primary vertex in the event is required to have at least four associated tracks. This requirement removes non-collision events, that arise for example from cosmic muons passing through the detector or from beam background.

Jet cleaning: events are discarded if they contain jets which are not associated to real energy deposits in the calorimeter. Such jets originate from spikes in the HEC, calorimeter noise, non-collision background, or cosmics. The rejection is based on jet variables, such as energy fractions in various calorimeter parts, LAr pulse shapes, the number of cells containing the jet energy, and the signal timing.

6.4.2. Object selection

The object selection divides into several steps. Very loose objects are defined for an overlap removal procedure that prevents the repeated usage of the same detector object. For the analysis, basic muon, electron and τ_h candidates are defined. These satisfy basic quality criteria, such as charge, momentum and pseudorapidity requirements. Additional requirements are applied later on in the analysis, depending on whether the objects are associated to the Z boson, the Higgs boson, or background-enhanced control regions.

Loose objects for overlap removal: The same detector object can be reconstructed as multiple physics objects (see Sec. 4.2). An electron, for instance, appears as a track with an associated calorimeter deposit in the detector. It will therefore be reconstructed as an electron, as well as a τ_h

and a jet. To avoid using the same detector object multiple times, an overlap removal is performed. Priority is given to particle types that are reconstructed with higher purity, which results in the order $\mu > e > \tau_h >$ jet. Hence, electrons are rejected if they overlap with muons, τ_h are rejected if they overlap with muons or electrons, and so forth. The quality requirements are kept very loose for the overlap removal. This in done in order to reduce the amount of misidentified particles: a muon might deposit an unusually large amount of energy in the calorimeter and does therefore not pass tight muon identification criteria. If the overlap removal is performed with respect to tight muons, the muon is mis-classified as an electron or τ_h. If, instead, loose objects are used, the muon is considered and the electron and the τ_h are discarded. Consequently, the following loose objects are defined:

- μ_{olr}: all muon types (SA, ST and CB) are considered. A minimal transverse momentum of 2 GeV is required and $|\eta| < 2.5$.
- e_{olr}: the electrons are reconstructed with the standard cluster-based algorithm and have to pass the loose identification. A minimal transverse momentum of 7 GeV is required and $|\eta| < 2.47$.
- $\tau_{h,\text{olr}}$: the tau candidates have to pass a cut on the BDT score of ≥ 0.3, and have to be reconstructed with one or three core tracks and an absolute electromagnetic charge of 1. A minimal transverse momentum of 20 GeV is required and $|\eta| < 2.47$.

Muons have to pass the kinematic requirements $p_{\text{T}} \geq 7\,\text{GeV}$ and $|\eta| < 2.5$. Only CB muons are used, having an inner detector track satisfying basic quality criteria on the number of detector hits. Loose requirements on the absolute distances of closest approach in the transverse plane, d_0, and the longitudinal plane, z_0, limit them these to a maximum of 10 mm. The additional track momentum in a cone of $\Delta R < 0.4$ is not allowed to exceed five times the muon transverse momentum: $p_{\text{T}}^{\text{iso}(0.4)}/p_{\text{T}} < 5$. A similar requirement is placed on the additional energy in the calorimeter in a cone of size $\Delta R < 0.2$: $E_{\text{T}}^{\text{iso}(0.2)}/p_{\text{T}} < 5$. These isolation requirements are rather loose and will be tightened at a later analysis step. Finally, if two muon candidates are found that are closer than $\Delta R < 0.1$, only the muon with the higher transverse momentum is kept.

Electrons have to pass the kinematic requirements $p_{\text{T}} \geq 7\,\text{GeV}$ and $|\eta| < 2.47$. Any electrons in the 'crack' region of the detector, $1.37 \leq |\eta| < 1.52$, are rejected. The electrons have to be reconstructed by the standard cluster-based algorithm and pass the loose identification working point and be of good object quality. Loose requirements on the absolute distances of closest approach in the transverse plane, d_0, and the longitudinal plane, z_0, limit these to a maximum of 10 mm. The additional track momentum in a cone of $\Delta R < 0.4$ is not allowed to exceed five times the electron transverse momentum: $p_{\text{T}}^{\text{iso}(0.4)}/p_{\text{T}} < 5$. A similar requirement is placed on the additional energy in the calorimeter in a cone of size $\Delta R < 0.2$: $E_{\text{T}}^{\text{iso}(0.2)}/p_{\text{T}} < 5$. These isolation requirements are rather loose and will be tightend at a later analysis step. Furthermore, electrons are discarded if they geometrically overlap with loose muon candidates, which is the case if $\Delta R(\mu_{\text{olr}}, e) < 0.2$. Finally, if two electrons are found that are closer than $\Delta R < 0.1$, only the electron with the higher transverse momentum is kept.

Tau leptons have to pass the kinematic requirements $p_{\text{T}} \geq 20\,\text{GeV}$ and $|\eta| < 2.47$. The score of the BDT for the rejection of QCD-jets has to exceed 0.3. The number of associated core tracks is limited to one or three, and the electromagnetic charge has to add up to $\pm 1\,[q_e]$. In order to ensure that all objects (e, μ, τ_h) come from the same vertex, the tau vertex position has to be consistent with the first primary vertex in the event. Therefore, they are required to be less than 1 mm apart. Candidates are also rejected if they geometrically overlap with loose muon or electron candidates, i.e.

if $\Delta R(\mu_{\text{olr}}, \tau_h) < 0.2$ or $\Delta R(e_{\text{olr}}, \tau_h) < 0.2$. In addition, the tau candidate is required to pass the loose electron veto, so that the amount of electrons being mistaken for τ_h is minimised. Finally, if two τ_h are found that are closer than $\Delta R < 0.2$, only the tau lepton with the higher transverse momentum is kept.

Jets have to pass the kinematic requirements $p_{\text{T}} \geq 20\,\text{GeV}$ and $|\eta| < 4.5$. If the jet falls within $|\eta| < 2.4$ and $p_{\text{T}} < 50\,\text{GeV}$, the jet vertex fraction must be at least 0.5. The jet is rejected if it geometrically overlaps with loose muons, electrons or hadronic tau decays, so if either $\Delta R(\mu_{\text{olr}}, \text{jet}) < 0.2$, $\Delta R(e_{\text{olr}}, \text{jet}) < 0.2$ or $\Delta R(\tau_{h,\text{olr}}, \text{jet}) < 0.2$ is satisfied.

6.4.3. Selection of the Z-pair

For the selection of the Z-pair, the requirements on electrons and muons are tightened further. The minimum transverse momentum is raised to $p_{\text{T}} \geq 12\,\text{GeV}$ and the isolation is tightened to $p_{\text{T}}^{\text{iso(0.4)}}/p_{\text{T}} < 0.1$ and $E_{\text{T}}^{\text{iso(0.2)}}/p_{\text{T}} < 0.1$. In order to reject non-prompt leptons, an additional requirement on the impact parameter significance of $|d_0/\sigma(d_0)| < 3.5$ (muons) and $|d_0/\sigma(d_0)| < 6.5$ (electrons) is introduced. Furthermore, the electrons have to pass the tight identification working point.

SF OS Z-pair

Among the selected electrons and muons, the pair is searched which is most consistent with the decay of a Z boson. The pair has to be of same flavour (SF) and opposite sign[9] (OS): e^+e^- or $\mu^+\mu^-$. The leading lepton of the pair is required to have $p_{\text{T}} \geq 26\,\text{GeV}$ and to have fired the trigger[10]. The invariant mass of the lepton pair is required to be within $\pm 20\,\text{GeV}$ of the Z boson mass. If more than one pair exists, the pair with an invariant mass closest to 91 GeV is chosen.

The SF OS pairs are shown in Fig. 6.11. The selected pairs are dominated by $Z(\rightarrow ee, \mu\mu)$ + jets events, for which the contribution of light flavour jets (Z lf) and heavy flavour jets (Z hf) are shown separately. In terms of event counts, the second dominat contribution comes from $t\bar{t}$ events (top). They are shown together with other top events, i.e. single top events and vector boson-top events, which however contribute very little. Since $t\bar{t}$ events do not contain an actual Z boson, the distribution of the invariant mass is non-resonant. Other electroweak backgrounds are summarised as 'other EW'. This category includes the non-resonant W + jets and WW events, as well as tri-boson events. Furthermore, it includes $Z(\rightarrow \tau\tau)$ + jets events, which peak at a lower invariant mass due to the escaping neutrinos. The ZH signal is scaled up by a factor of 100, in order to illustrate its distribution. Other Higgs boson backgrounds are summarized as 'H125' and scaled to the SM predicted cross-section. Due to the Z-pair selection, the contribution is dominated by $H \rightarrow ZZ$ events.

The shape of the data distribution is generally well reproduced by the simulation. The deviation in the tails of the invariant mass distribution is caused by ee-pairs. This has been studies in more detail by other groups, who concluded that, at least partially, the observed effect is caused by small imperfections in the material description up to the calorimeter[11]. To a large extend, the deviation is covered by the systematic uncertainties associated to the energy calibration of electrons. The influence on the analysis has been studied and found to be negligible.

[9] The terms 'sign' and 'electric charge' are used interchangeable in this chapter. The abbreviations OS and SS are used for object pairs with opposite sign electric charge and same sign electric charge, respectively.

[10] An actual geometrical match to the trigger object is done in the case of muons. In the case of electrons, the necessary information to perform the match was missing. Therefore, tight requirements on ID level, isolation and momentum are imposed, in order to ensure that the leading electron satisfies the trigger criteria.

[11] Internal communication.

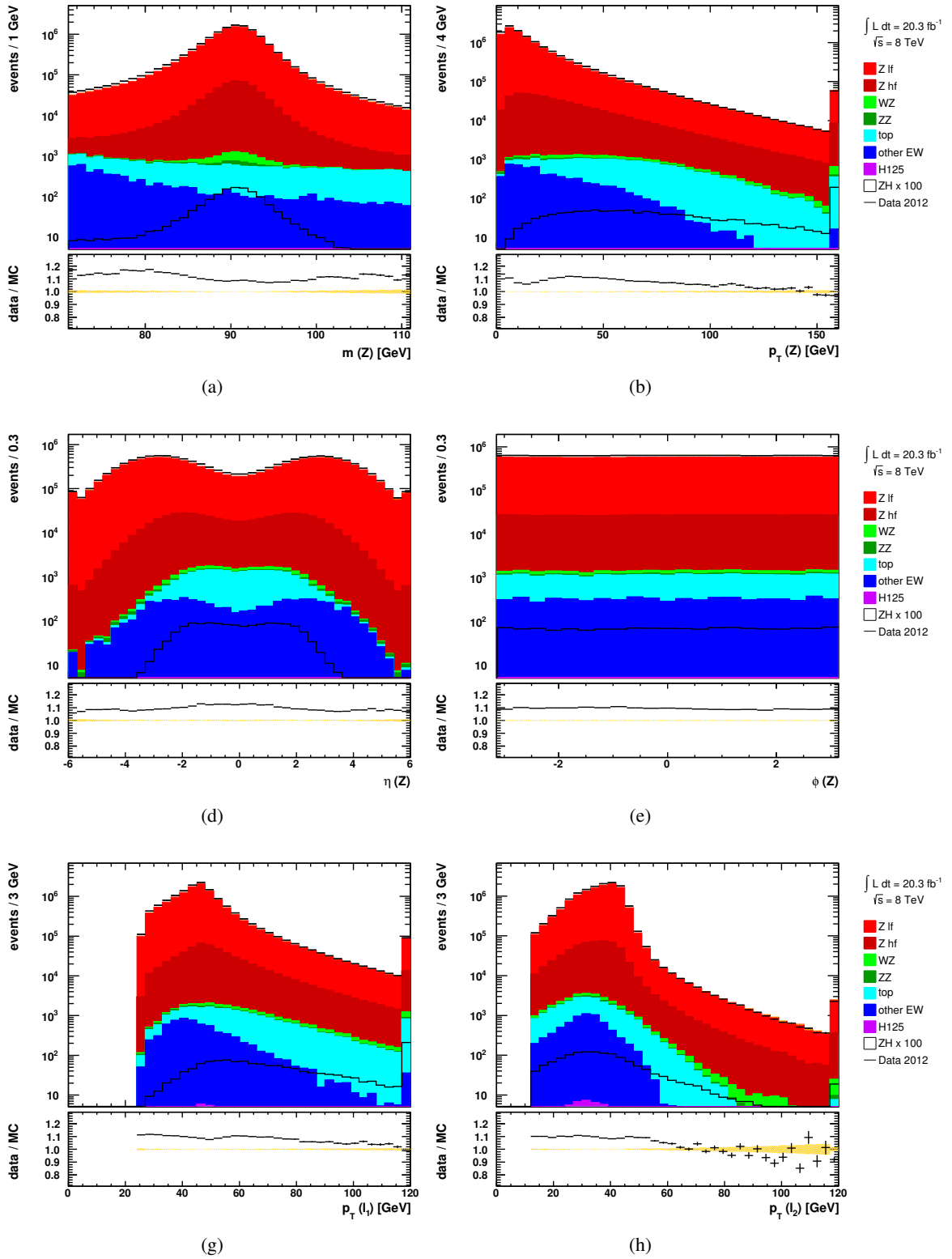


Figure 6.11: Distributions of the OS SF Z-pair: (a) invariant mass, (b) transverse momentum, (c) pseudorapidity, (d) azimuthal angle, (e) leading lepton p_{T}, and (f) subleading lepton p_{T}. The events include both di-electron and di-muon pairs. The first and last bins include all events in the under- and overflow of the x-axis range. The legend is explained in the main text. The yellow band shows the statistical uncertainty on the MC samples.

The total number of observed events is about 9% lower than the prediction. This is observed for both ee-pairs and $\mu\mu$-pairs. Several cross-checks have been performed (see App. B.2 for more details and plots):

- Comparison to events generated with POWHEG+PYTHIA8: For this sample, the prediction is about 5% lower than the observed number of events. A difference in the predicted number of events of about 5% is in good agreement with previous measurements [167].
- Measurement of scale factors for isolation cuts: The measured scale factors are found to be consistent with unity within errors.
- Missing backgrounds: Several backgrounds are not considered, for instance multi-jet events and hadronic decays of the Z boson. These contributions are however expected to be much smaller than the observed disagreement. Furthermore, these backgrounds are not expected to be resonant.

None of the studies can explain the large observed discrepancy, even though it remains possible that many smaller effects add up to a larger mis-modelling. Concerning this ZH search, the Z + jets background contributes to the signal region via jets faking (or emitting) electrons, muons and τ_h. The cross section uncertainty on Z+2jets events is significantly higher than the uncertainty on the inclusive cross section. This is even more true for the simulation of heavy flavour jets, which are an important source for fake muons. As a consequence, there is no alternative to taking the data to normalise the Z + jets background - which is done at a later stage of the analysis. The overall difference of the normalisation is therefore not considered critical for the ZH search.

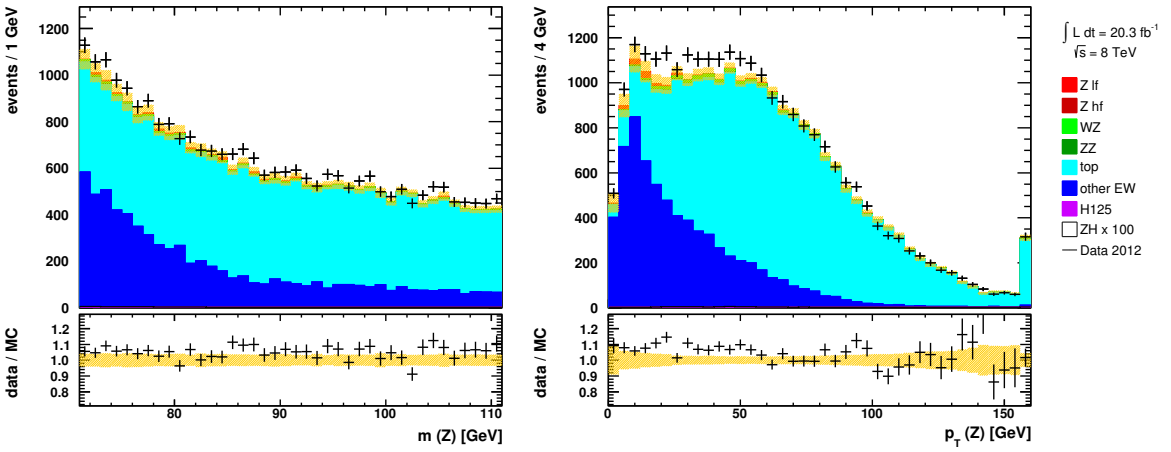
OF OS top quark pair

If no e^+e^- or $\mu^+\mu^-$ pair is found, a search for a $e^\pm\mu^\mp$ pair is performed. With the exception of the lepton flavour, the same criteria are used as in the search for the SF OS Z-pair. The selected events are dominated by W + jets events and $t\bar{t}$ events, as shown in Fig. 6.12. Also here, a small excess of data over MC is observed, of about 5%. However, no estimation of the QCD background has been performed here. Requiring additional e, μ or τ_h rejects most of the W + jets events (and potential contributions of multi-jet events), so that the sample of OF OS pairs can be utilised for the estimation of the top background. This selection will also be called the *alternative* pre-selection, in contrast to the SF OS Z-pairs, which are part of the *nominal* pre-selection.

6.5. Main selection: Higgs boson decay

For the ZH search, events with at least two additional objects of type e, μ, and τ_h are of interest. Using the basic definition given in Sec. 6.4.2, the objects can be grouped into candidates of good and bad quality, depending on whether they *pass* or *fail* additional requirements:

- *pass*:
  - μ: $p_{\mathrm{T}}^{\mathrm{iso(0.4)}}/p_{\mathrm{T}} < 0.08$ and $E_{\mathrm{T}}^{\mathrm{iso(0.2)}}/p_{\mathrm{T}} < 0.08$,
  - e: tight ID, $p_{\mathrm{T}}^{\mathrm{iso(0.4)}}/p_{\mathrm{T}} < 0.2$ and $E_{\mathrm{T}}^{\mathrm{iso(0.2)}}/p_{\mathrm{T}} < 0.2$,
  - τ_h: medium ID,
- *fail*:
  - μ: $p_{\mathrm{T}}^{\mathrm{iso(0.4)}}/p_{\mathrm{T}} > 0.2$ or $E_{\mathrm{T}}^{\mathrm{iso(0.2)}}/p_{\mathrm{T}} > 0.2$,



(a) (b)

Figure 6.12: Top control region: (a) invariant mass and (b) transverse momentum of the $e^{\pm}\mu^{\mp}$ pairs (OS OF pairs). The first and last bins include events in the under- and overflow of the x-axis range. The yellow band shows the statistical uncertainty.

- e: failing tight ID,
- τ_h: failing medium ID.

Objects that are neither part of *pass* nor *fail* are discarded. The 'gap' in the categorisation of muons, i.e. the disregard of muons with an relative isolation between 0.8 and 0.2, is introduced to minimise the contribution of prompt muons to the *fail* region. For electrons, the categorisation is obtained by inverting the requirement on the identification (tight versus not tight). Well-defined electrons have also to be isolated in both the inner detector and the calorimeter. In order to increase the number of objects in the *fail* region, the requirement on the isolation is released there.

The objects are then paired to match any of the six possible final states: ee, $\mu\mu$, $e\mu$, $e\tau_h$, $\mu\tau_h$ and $\tau_h\tau_h$. The pairing is performed on the joined set of *pass* and *fail* objects, considering all events with at least two additional objects. Only the pair that best satisfies the Higgs boson hypothesis is considered for further study. The pair is required to have

- opposite electric charge,
- an angular separation of $0.8 \leq \Delta R < 5$,
- an invariant mass between 12 GeV and 150 GeV.

If more than one pair is found, the pair with the highest sum p_{T} is selected. This happens in about 4% of all cases where an OS H-pair is found. The selection requirements on the pair are kept rather tight in order to select pairs compatible with a Higgs boson decay. The lower mass cut rejects possible backgrounds from J/ψ or Υ decays. The highest Higgs boson mass under study is $m(H) = 150$ GeV, so that the upper mass requirement does not restrict the signal distribution. The requirement on ΔR is an additional protection against picking the decay products of boosted hadrons or heavy quarks, which show a low anglular separation. Furthermore, it ensures that the isolation cones of the H-pair constituents are uncorrelated. This is important for the background estimation, which uses the previously defined *fail* objects to estimate the number of *pass* objects, via so called *fake factors* (see Sec. 6.7). Overlapping isolation regions would result in correlated fake factor. To avoid this

complication a lower limit of $\Delta R < 0.8$ is chosen, which, at the same time, does not discard significant amount of signal events.

In events without an H-pair of opposite electric charge, the search is repeated for pairs with the same electric charge. Events with SS pairs are background dominated, and are later used to increase the number of events available for the estimation of background events.

Based on the quality criteria defined above, each H-pair will appear in one of the exclusive categories *pass-pass*, *pass-fail*, *fail-pass*, or *fail-fail*. For each of the six possible final states, *pass-pass* defined the signal region. The other categories are used as side bands for the estimation and validation of the background modelling, because they are dominated by events with fake objects. The pairing procedure is special, because the H-pair is defined first, i.e. before the quality of the each object is determined. It ensures not only orthogonal signal regions, but also provides orthogonal side bands for each final state. This eases the use of fake factors at a later stage of the analysis.

The final states ee and $\mu\mu$ will not be considered in the final signal extraction. This is due to the very small cross section as well as the overwhelming ZZ background. Both final states are however carried along in the analysis flow. This offers the possibility to use them as control regions - or add them to the signal region as soon as tools are available to enhance the signal purity.

The tight requirements on the signal region minimise the contribution from Z+jets and $t\bar{t}$ events. In fact, the remaining amount of simulated events is too small for it to be used to model the background contribution. Instead, the events are taken from a background-enhanced region and transferred to the signal region. In order to minimise the dependence on the quality of the MC description, the procedure is driven by the transfer-probabilities and normalisation observed in data. The background-enhanced regions are given by the *pass-fail*, *fail-pass* and *fail-fail* side bands. The normalisation of the backgrounds in the *fail-fail* region are obtained from a fit to data. The backgrounds are then transferred to the signal region using fake factors measured in data. The two steps are described in the following sections.

6.6. Background estimation: normalisation

To correct for the normalisation of the Z + jets samples and mis-modelling of fakes, the background is fitted to data in the *fail-fail* side band. This side band is dominated by Z + jets and $t\bar{t}$ events, with negligible contributions from true e, μ or τ_h. The fit is performed separately for each of the six final states, allowing the magnitude of the mis-modelling to be different for each particle type or final state. The shape of the Z-pair invariant mass distribution is found to be independent of the additional fake objects in the event, and is therefore used for the fit. An exception are events in which a photon is radiated from one of the final state leptons and is converted in the detector material to an electron-positron pair. In such events, the occurrence of an additional electron is correlated with a lower mass of the Z-pair. This background contribution is fitted along with the other backgrounds.

In total, four templates are used in the fit as shown in Fig 6.13. The templates are obtained from simulated samples, right after the pre-selection step. At this stage of the analysis, many thousands of events are available so that smooth templates are obtained.

- resonant (Z shape): all backgrounds peaking at the Z boson mass are combined for the Z-shape template, i.e. Z + jets, ZZ, WZ, Zt and triboson events. The dominant contribution comes from Z + jets events.
- non-resonant (top shape): the template consists of $t\bar{t}$, single top and WW events. The contributions from the latter two are minor, so that the template is dominated by $t\bar{t}$ events.
- conversions: the template contains the previously mentioned conversion events. The events are coming from Z + jets events that contain a truth-matched conversion electron.

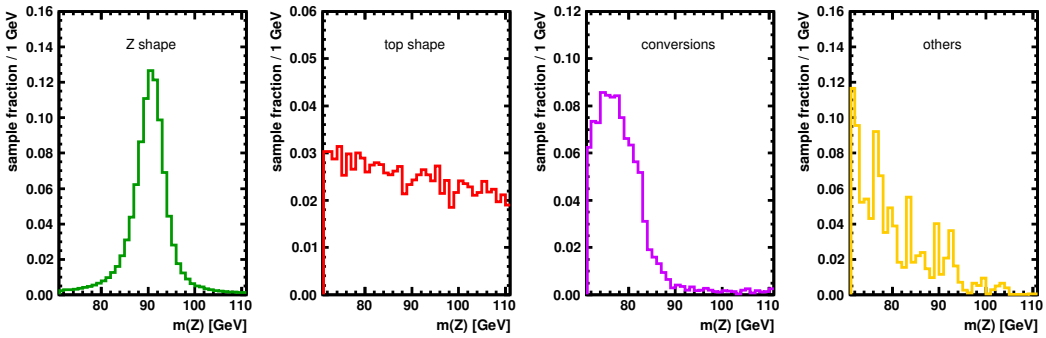


Figure 6.13: Templates used in the background fit. The variable shown in the invariant mass of the Z-pair constituents.

- others: all other backgrounds, i.e. non-signal Higgs boson events, W + jets and $Z(to\tau\tau)$ + jets events, are summarised in another template.

The Z + jets background dominates in all fit regions and prevents a proper estimation of the $t\bar{t}$ background. Therefore, the alternative pre-selection - the selection of the OF OS-pair - is used to constrain this background. As discussed in Sec. 6.4.3, the requirement of an $e\mu$ pair predominantly selects $t\bar{t}$ and W events. The requirement of two additional objects (the H-pair) rejects the W + jets events, so that the *fail-fail* region is largely dominated by $t\bar{t}$ events. The amount of $t\bar{t}$ events is expected to be identical for both pre-selections, OF OS and SF OS. The two leptons originate from the decay of the two opposite sign W bosons[12], and thus the probability to find same flavour pairs ($ee, \mu\mu$) is the same as to find opposite flavour pairs ($e\mu, \mu e$). Figure 6.14 shows the MC expectation of $t\bar{t}$ events for each fit region and both pre-selections, and confirms the hypothesis. Hence, the OF OS region can be used to directly constrain the $t\bar{t}$ background in SF OS events.

The fit is an extended likelihood fit and performed independently for each of the six final states. Events with H-pairs of opposite and same sign electric charge are considered, which improves the fit due to a higher number of events in each fit region. The distribution of the invariant mass of the Z-pair is used as input to the fit. SF OS and OF OS pairs are fitted simultaneously. The Z background has a negligible contribution to the OF OS region and is constrained by the fit in the SF OS region only. The top template is constrained to the same number of events in both regions[13]. The normalisation of the Z and top templates is allowed to float freely. Conversion events are only relevant for the final states containing electrons, and are here allowed to float freely as well. The normalisation of 'others' is bound to $\pm 10\%$ of the MC expectation[14]. A very few events with true leptons enter the signal region. Because the focus of the fit is on the contributions from fakes, and because the few events come with high statistical uncertainties, this contribution is fixed to its MC prediction. The templates are fitted in the *fail-fail* region for the observed data and the simulated data separately. Normalisation scale factors are then derived for each of the background types as the ratio of the the two fit results. This procedure reduces systematics uncertainties. The fit to data is shown for each of the final states in Figs. 6.15 and 6.16. The fit results are given in Table 6.3. The fit uncertainties are treated as systematic uncertainties for the normalisation of the different backgrounds.

[12] The top decays in predominantly via $t \to Wb$ [168].

[13] Requiring the same scaling factor instead of the same yield in both regions gives slightly different fit results for the fit to data and simulation. However, the final scaling, i.e. the ratio of the fit results, is identical within uncertainties.

[14] The choice of 10% is driven by the discrepancies seen in the pre-selection region. The fit itself has very little constraining power for the 'others' template, due to the small contribution of these backgrounds. Fixing the contribution does not change the fit result within uncertainties.

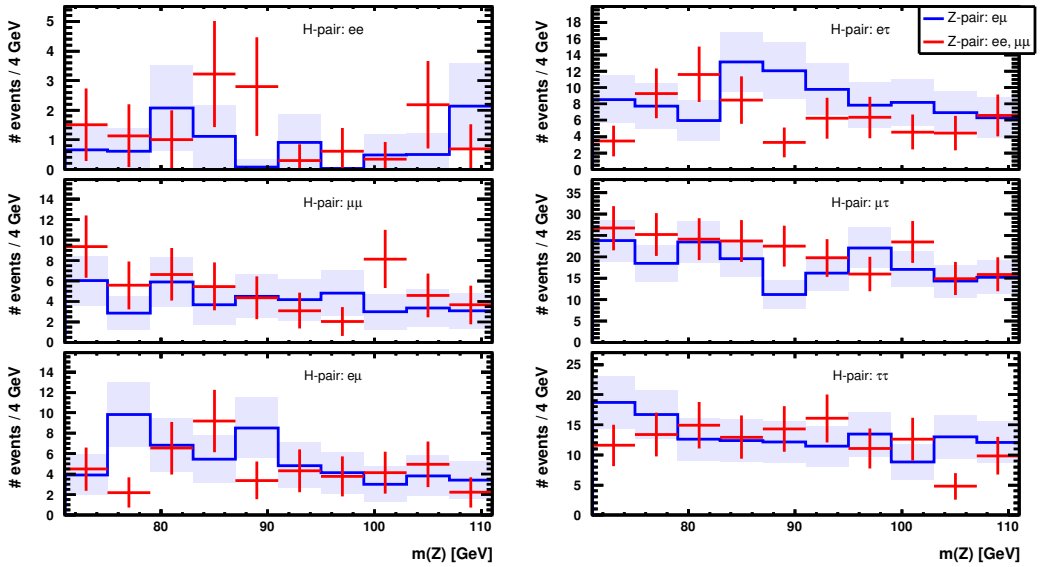


Figure 6.14: Expected $t\bar{t}$ background in the *fail-fail* control region for each of the six final states. The invariant mass distributions including statistical uncertainties are shown separately for the nominal SF OS pre-selection and the alternative OF OS pre-selection. The errors show the statistical uncertainty of the MC samples.

In order to test a possible bias due to the mis-modelling of the lower tail of the Z boson distribution, the fit is repeated for a fixed contribution of conversion events. The scale factors for the Z and top backgrounds are affected at a negligible level, and any deviation is covered by the fit uncertainty. Furthermore, the fit uncertainty on the number of conversion events is rather large, so that no additional uncertainty is asigned.

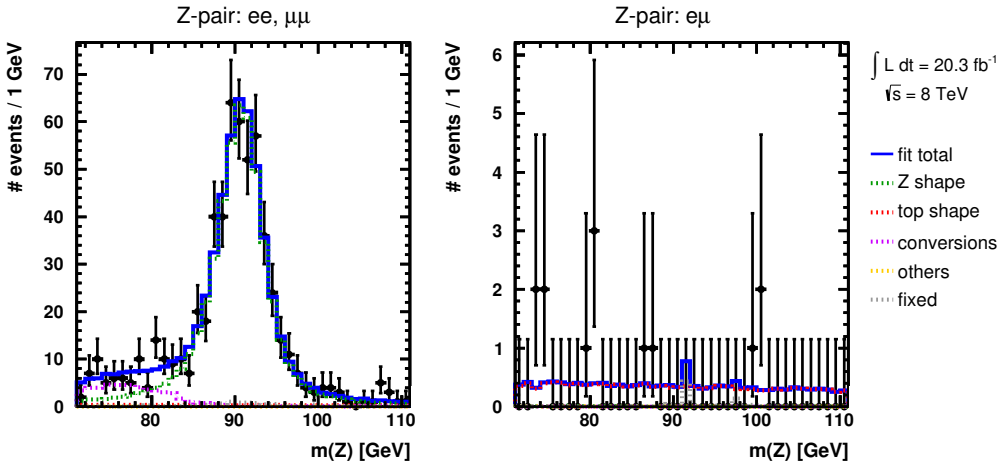
6.7. Background estimation: fake factors

The tight requirements on the signal region minimise the contribution from Z + jets and $t\bar{t}$ events. In fact, the remaining amount of simulated events is too small to give a reliable estimation of the background in the signal region. Instead, the events are taken from a background-enhanced region and transferred to the signal region. The transfer is carried out by weighting each object in the background-enhanced region with a fake factor, f_{F}, and thus obtain the background estimate in the signal region. The background enhancement is achieved by using the previously defined objects that satisfy the *fail* criteria. The fake factor is thus defined as

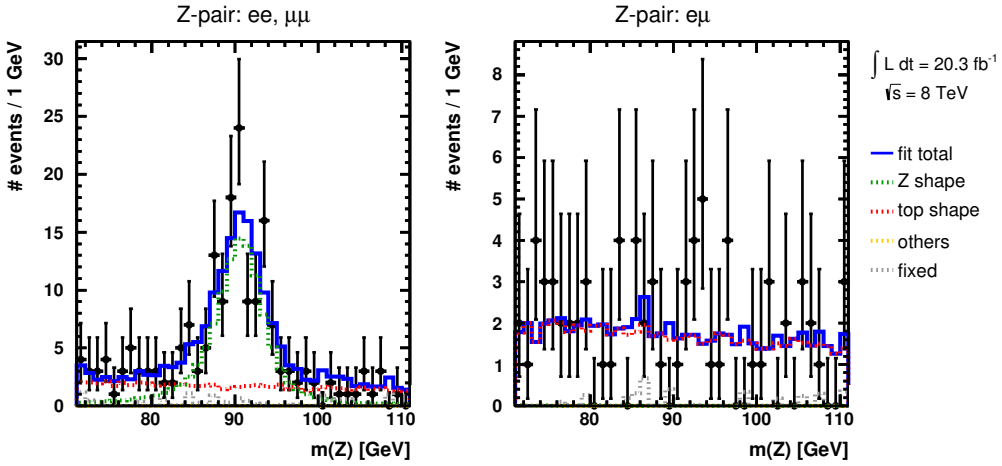
$$f_{\mathrm{F}} = \frac{\#\text{ probes satisfying the } \textit{pass} \text{ criteria}}{\#\text{ probes that satisfy the } \textit{fail} \text{ criteria}} \quad . \tag{6.2}$$

The fake factors are measured in a region which is orthogonal to the signal region, and largely independent of the *fail-fail* and *fail-pass* side bands, called the *control region*. The fake factors are measured individually for electrons, muons and hadronic tau decays. Ideally, the control region is well populated with *pass* and *fail* objects, and mirrors the conditions (i.e. background composition) of the side bands and the signal region. In order to account for possible difference in behaviour between the different regions, the fake factors are obtained as a function of several variables.

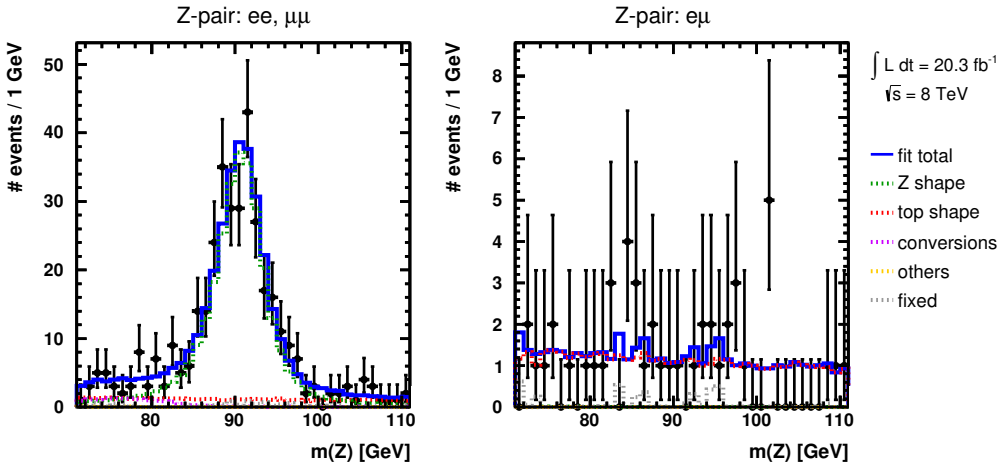
If the fake factor depends in an uncorrelated way on several variables, it can be measured indepen-



(a) H-pair: ee

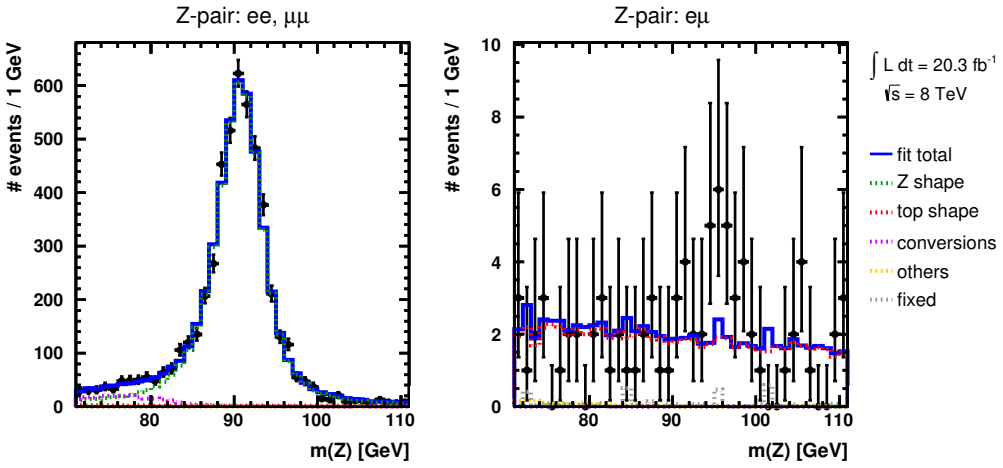


(b) H-pair: $\mu\mu$

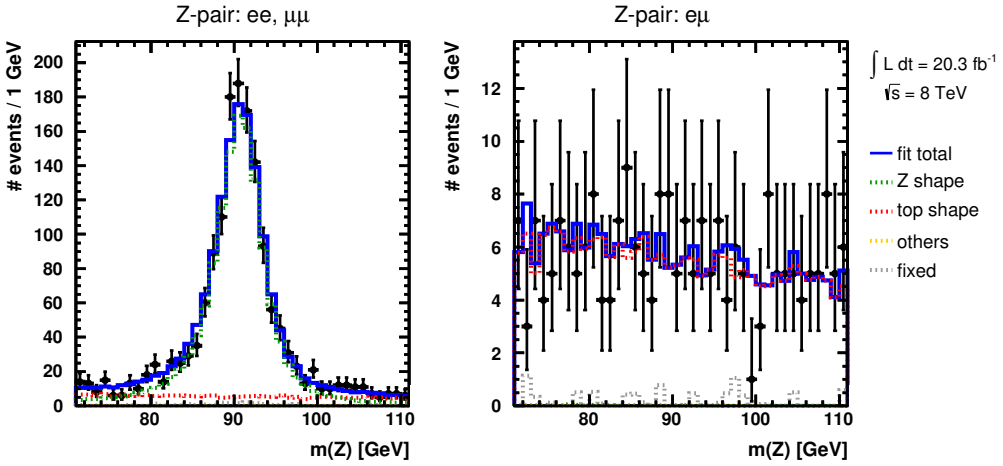


(c) H-pair: $e\mu$

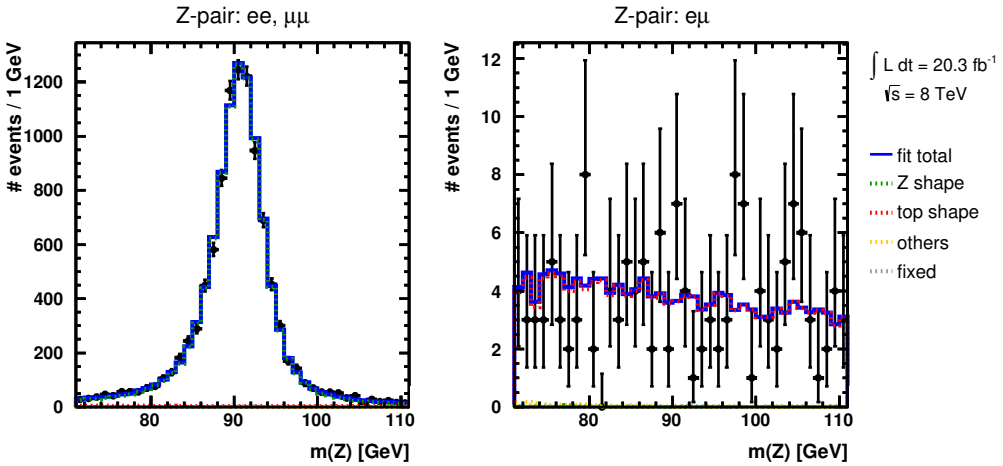
Figure 6.15: Normalisation of the background contributions in the *fail-fail* side band, for the fit to data and the final states (a) ee, (b) $\mu\mu$, and (c) $e\mu$. Both OS and SS H-pairs are considered. Events with SF Z-pairs are shown on the left, while the OF pairs are shown on the right.



(a) H-pair: $e\tau_h$



(b) H-pair: $\mu\tau_h$



(c) H-pair: $\tau_h\tau_h$

Figure 6.16: Normalisation of the background contributions in the *fail-fail* side band, for the fit to data and the final states (a) $e\tau_h$, (b) $\mu\tau_h$, and (c) $\tau_h\tau_h$. Both OS and SS H-pairs are considered. Events with SF Z-pairs are shown on the left, while the OF pairs are shown on the right.

Table 6.3: Normalisation of the background contributions in the *fail-fail* side band. Results (as numbers of events) of the fit for real and simulated data. Also shown is the prediction of the simulation as such. The last part labelled 'scaling' shows the ratios between the fit results in real and simulated data.

		ee	$\mu\mu$	$e\mu$	$e\tau_h$	$\mu\tau_h$	$\tau_h\tau_h$
fit to data	Z shape	501±26	114±19	293±25	4790±77	1335±57	10004±109
	top shape	13±4	67±9	45±6	74±10	218±14	149±15
	conversion	54±14	—	15±16	235±29	—	—
fit to MC	Z shape	452±22	74±11	224±17	4514±71	956±35	9305±98
	top shape	8±2	47±5	52±7	87±12	187±11	135±13
	conversion	58±14	—	4±4	190±22	—	—
MC pred.	Z shape	461±21	76±9	224±15	4508±67	940±31	9318±97
	top shape	14±4	53±7	45±7	66±8	212±15	123±11
	conversion	43±7	—	8±3	216±15	—	—
scaling	Z shape	1.1±0.1	1.5±0.3	1.3±0.1	1.1±0.02	1.4±0.1	1.1±0.02
	top shape	1.7±0.6	1.4±0.2	0.9±0.2	0.8±0.2	1.2±0.1	1.1±0.2
	conversion	0.9±0.3	—	3.9±6.1	1.2±0.2	—	—

dently for each variable. For n variables v_i, the combined (total) fake factor is the given by

$$f_{\mathrm{F}} = \frac{f^{v_1} f^{v_2} \dots f^{v_n}}{\left(\frac{1}{n}(\langle f^{v_1}\rangle + \langle f^{v_2}\rangle + \dots + \langle f^{v_n}\rangle)\right)^{n-1}} \quad . \tag{6.3}$$

Here, f^{v_i} is the fake factor obtained as function of variable v_i, and $\langle f^{v_i}\rangle$ is its average value. If all f^{v_i} are measured on the same sample, the $\langle f^{v_i}\rangle$ are equal.

If the fake factor depends in a correlated way on several variables, it is determined using a multi-dimensional binning in these variables. In particular, the fake factors measured for this analysis are observed to have a correlated dependence on p_{T} and η.

The fake factor approach is a major reason for not using the missing transverse energy in this analysis. The $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution is not correctly reproduced for fake factors that only depend on p_{T} and η. For utilising $E_{\mathrm{T}}^{\mathrm{miss}}$ in a mass estimator, both the direction and the magnitude are needed. Obtaining a reliable estimate in all signal regions was not possible within the scope of this study.

Each f_{F} bin has an associated uncertainty due to the statistical limitation of the control region. In order to estimate the uncertainty on a different sample (i.e. using a different binning), the per-bin uncertainties need to be combined. Uncertainties coming from the same bin are treated fully correlated and added linearly. Uncertainties coming from different bins are treated fully uncorrelated and added in quadrature. The total uncertainty due to the statistical limitation of the control region is thus evaluated as

$$\sigma_{\mathrm{tot}}^2 = (N_1\sigma_1)^2 + (N_2\sigma_2)^2 + \dots + (N_n\sigma_n)^2 \quad . \tag{6.4}$$

Here, σ_i is the statistical uncertainty on f_{F} in bin i, and N_i is the number of times the fake factor is obtained from bin i.

In many analyses a more conservative approach is chosen in which the f_F are varied up or down simultaneously in all bins. In other words, all bins are treated fully correlated. Due to the relatively fine-grained multi-dimensional binning used in this measurement, the conservative approach leads to very large uncertainties as soon as a coarser binning (compared to the f_F binning) is used. This is already the case when the dimensionality is reduced from 2-d to 1-d, i.e. when, for example, the momentum distribution is plotted. Therefore, the different approach is used here.

The modelling of fake objects is known to be poorly described in simulated samples. Therefore, the fake factors are measured in data. The pre-selection is used as described in Sec. 6.4. An additional muon, electron or τ_h is required, that functions as the probe object to measure the fake factor. The control region has to be free of signal events, which could otherwise bias the measurement. Therefore, events with more e, μ or τ_h that satisfy the *pass* criteria are rejected. The contribution of remaining true e, μ or τ_h is subtracted from the control regions. The amount of true objects is estimated using simulated samples and subtracted from the numerator and denominator of Eq. 6.2. If, in some bin, the subtraction exceeds the number of events in data, the fake factor is set to zero in the affected bin.

6.7.1. Muon fake factors

For the determination of the muon fake factor, a sample of *pass* and *fail* muons needs to be obtained, which is independent from the signal region. In addition to the pre-selection, the following requirements are applied:

- exactly one additional muon in the event,
- veto on events with additional electrons satisfying the *pass* criteria,
- veto on events with additional τ_h satisfying the *pass* criteria.

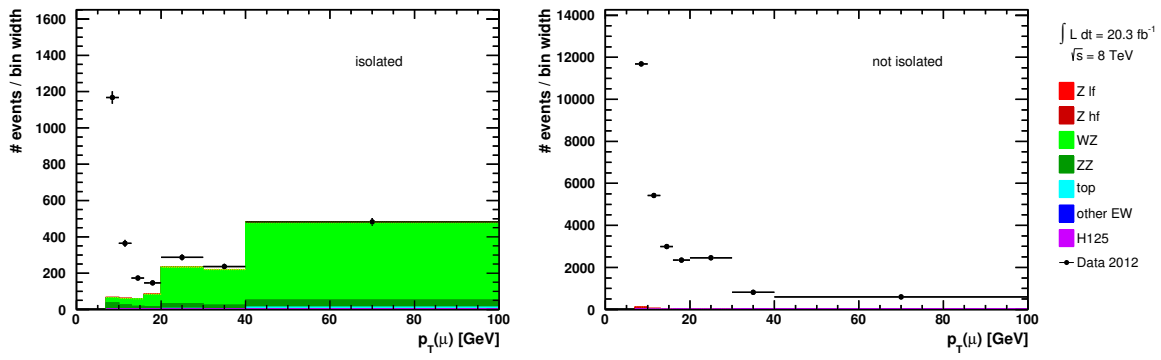
The samples of *pass* muons has a large contribution of true muons, mainly from WZ events, as shown in Fig. 6.17(a). All true muons are subtracted based on the simulation. This makes the measurement very dependent on the correct estimate of the true muon content. This is accounted for in the systematic uncertainties. The sample of *fail* muons (Fig. 6.17(b)) is dominated by fake objects over the full momentum range with no significant contribution from true muons.

The fake factor is measured as a function of p_T for various η ranges. For low momenta ($p_T < 20\,\text{GeV}$) three bins in η are used: $|\eta| < 0.5$, $0.5 \geq |\eta| < 1.5$, and $1.5 \geq |\eta|$. Above 20 GeV, the number of events is too small to allow for a binning in η, and the fake factor is binned in p_T only. In order to obtain a smoother behaviour of the fake factor as a function of p_T, a linear interpolation is applied between bin-centres. The interpolated fake factors for the three bins in η are shown in Fig. 6.18(a).

Finally, the fake factor is measured separately in events with or without jets. Figure 6.18(b) shows the fake factor binned for the two jet bins, $f_F(n)$, together with the average fake factor of the sample. The fake factors $f_F(n)$ and $f_F(p_T, \eta)$ are combined using the formula for uncorrelated variables (Eq. 6.3). The equation simplifies to a scaling of $f_F(p_T, \eta)$ by a factor 1.29 in events without jets, and by a factor 0.77 in events with at least one jet. The combined fake factor is validated in Fig. 6.19(a). The binning in η is not considered here, because the comparison would otherwise suffer from high statistical fluctuations. The combined fake factor $f_F(n) \times f_F(p_T, \eta)$ (using Eq. 6.3) is compared to a fully binned measurement $f_F(p_T, n)$. The two measurements agree within uncertainties, confirming that the variables are independent.

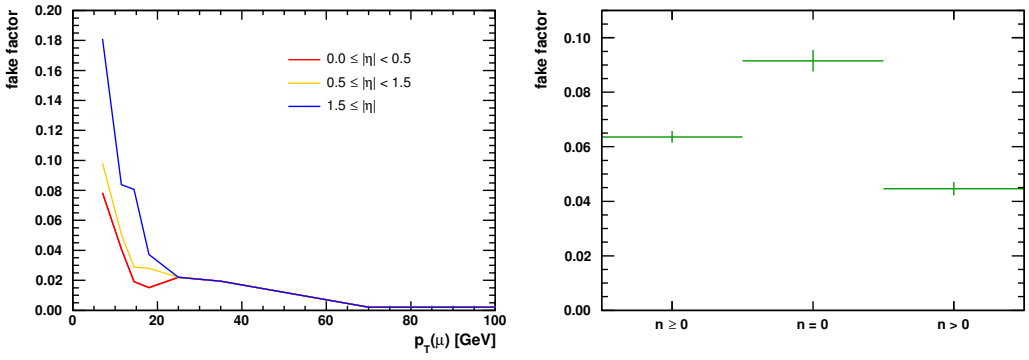
Uncertainties on the muon fake factors

The main uncertainties on the muon fake factors are due to the statistical limitation of the control region and the subtraction of true muons from the numerator.



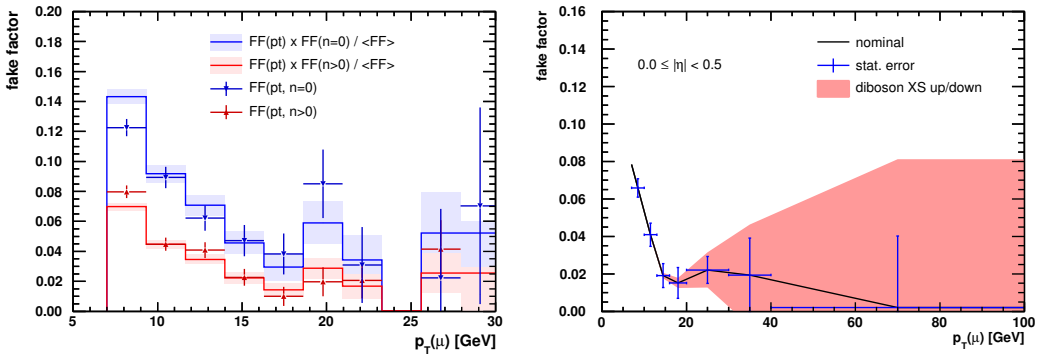
(a) (b)

Figure 6.17: (a) Numerator and (b) denominator for the measurement of the muon fake factor before the subtraction of true muons. The MC simulation (coloured area) shows true muons only. A varying bin width is used to account for the decreasing number of fake muons with p_{T}.



(a) (b)

Figure 6.18: Muon fake factors: (a) Interpolated fake factors $f_{\mathrm{F}}(p_{\mathrm{T}}, \eta)$ for the three η ranges as a function of p_{T}. (b) Average fake factor (first bin) and $f_{\mathrm{F}}(n)$.



(a) (b)

Figure 6.19: Muon fake factors: (a) Fake factor as a function of p_{T}, measured in events with (n>0) and without (n=0) jets, in comparison to the combined result from the combination of the individually measured factors. The errors are of statistical nature. (b) Interpolated fake factor for the most central η range. Also shown are the uncertainty due to the statistical limitation of the control region per bin (error bars) and the impact of the variation of the diboson cross section, i.e. the subtraction of true muons (shaded area).

- Statistical limitation: The number of events available for the f_{F} measurement is limited by the size of the data sample, in particular by the number of events in the numerator. Figure 6.19(b) show the statistical uncertainty for the most central η range in each momentum bin, which is the dominating uncertainty in the low momentum bins.
- Subtraction of true muons: The fake factor is highly dependent on the number of true muons in the *pass* region, which is determined from simulated samples. The highest contamination is coming from diboson WZ and ZZ events. Their cross section is varied by $\pm 10\%$, and the fake factor is recalculated. As shown in Fig. 6.19(b), this uncertainty dominates the high momentum region, allowing the f_{F} to vary between values from 0 to about 0.08.
- Sample composition: The largest difference between the average fake factor and the fake factor as a function of the number of jets is taken as an overall systematic uncertainty.

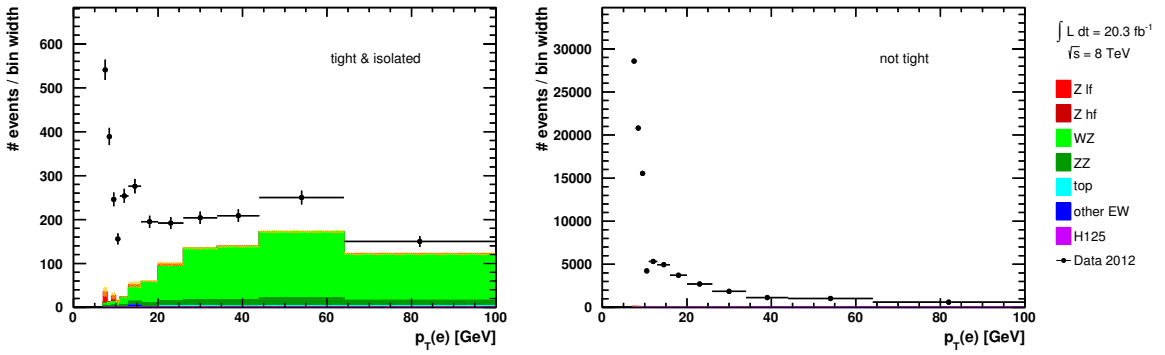
6.7.2. Electron fake factors

As described above, the requirement of an additional electron enhances the number of conversion events in the low tail of the invariant mass distribution of the Z-pair. Because conversion electrons have a different fake factor than other non-prompt electrons, a joint fake factor measurement would cause a dependence on the Z-pair mass. In order to avoid this dependence, the electron fake factor is measured separately for conversion electrons and other non-prompt electrons.

Electrons

In order to obtain a sample for the measurement of the electron fake factors, the following requirements are applied in addition to the pre-selection:

- exactly one additional electron,
- veto on events with additional muons satisfying the *pass* criteria,



(a) (b)

Figure 6.20: (a) Numerator and (b) denominator for the measurement of the electron fake factor before the subtraction of prompt and conversion electrons. The MC simulation (coloured area) shows the contribution of prompt and conversion electrons. A varying bin width is used to account for the decreasing number of fake electrons with p_{T}.

- veto on events with additional τ_h satisfying the *pass* criteria,
- $m(Z) \geq 85\,\mathrm{GeV}$.

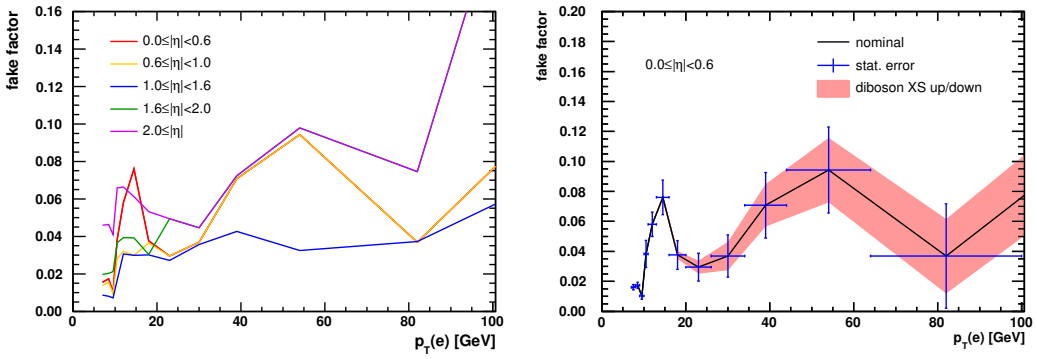
The last requirement leads to a sample of predominantly non-conversion electrons. Any residual electrons from conversions are subtracted based on the MC estimate. Furthermore, all prompt electrons are subtracted based on the simulation. Especially the *pass* region has a significant contribution from prompt electrons, mainly from *WZ* events, as shown in Fig. 6.20.

The fake factor is measured as a function of p_{T} for various η ranges. Below $p_{\mathrm{T}} = 20\,\mathrm{GeV}$ a fine binning in η is used, with five bins: $|\eta| < 0.6$, $0.6 \geq |\eta| < 1$, $1 \geq |\eta| < 1.6$, $1.6 \geq |\eta| < 2$, and $2 \geq |\eta|$. Above that threshold, the granularity is reduced to three bins: $|\eta| < 1$, $1 \geq |\eta| < 1.6$, and $1.6 \geq |\eta|$. In order to obtain a smoother behaviour of the fake factor as a function of p_{T}, a linear interpolation is applied between bin-centres. The interpolated fake factors for the different η ranges are shown in Fig. 6.21(a). It is apparent that the fake factors exhibit a more complicated momentum behaviour than the muon or τ_h fake factors. This is in particular caused by the tuning of the electron identification working points, which are tuned separately for momenta below and above 10 GeV [87].

Uncertainties on the electron fake factors

The systematic uncertainties of the electron fake factors are coming from the statistical limitation of the control region, the subtraction of true electrons and the sample composition.

- Statistical limitation: The number of events available for the f_{F} measurement is limited by the size of the data sample, which is in particular limited due to the requirement of additional μ and τ_h in the events. Figure 6.21(b) shows the uncertainty due to the statistical limitation per bin for the most central η region. It is the dominat uncertainty for all momenta.
- Subtraction of true electrons: The fake factor is dependent on the number of true electrons in the *pass* region, which is determined from simulated samples. The highest contamination is coming from diboson *WZ* events. The diboson cross section is varied by ±10%, and the fake factor is recalculated. The uncertainty comes into play above 20 GeV, but remains lower than the statistical uncertainty.



(a) (b)

Figure 6.21: Electron fake factors: (a) Fake factors as a funtion of p_{T} for the various bins in $|\eta|$. (b) Interpolated fake factor for the most central η range. Also shown are the uncertainty due to the statistical limitation of the control region per bin (error bars) and the impact of the variation of the diboson cross section, i.e. the subtraction of true electrons (shaded area).

- Sample composition: Based on observations in validation regions, especially in regions with additional electrons (see Sec. 6.8.2), an additional overall uncertainty of $^{-10\%}_{+50\%}$ is assigned to the electron fake factors.

Conversion electrons

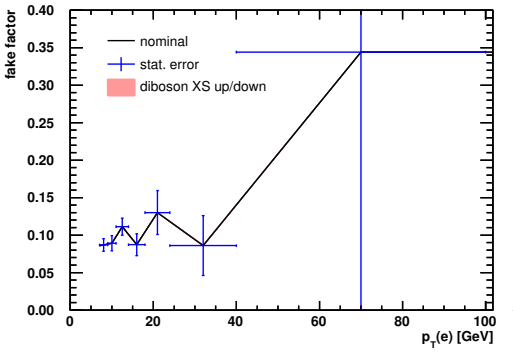
In order to collect a sample of conversion electrons, different mass requirements are applied. In contrast to the electron selection, a low invariant mass of the Z-pair is required, $m(Z) < 81\,\mathrm{GeV}$. To further increase the purity of the sample, the invariant mass of the Z-pair and the electron in the event has to satisfy $m(Z + e) < 100\,\mathrm{GeV}$. Any residual non-conversion fake electrons as well as prompt electrons are subtracted based on the simulation. The subtraction amounts to a negligible fraction of the sample.

The fake factor for conversion electrons is solely measured as a function of momentum, as shown in Fig. 6.22. Above 40 GeV, the tight selection criteria do not provide enough events for a meaningful measurement. However, these fake factors will only be applied on electrons that are tagged in the simulation as conversions originating from the Z-decay. The fraction of such electrons associated to the H-pair is negligible at high momenta. Therefore, the statistically limited fake factor for $p_{\mathrm{T}} > 40\,\mathrm{GeV}$ does not appear to be problematic. The systematic uncertainty due to the subtraction of true electron is negligible. The uncertainty of the fake factor is hence dominated by the statistical uncertainty.

6.7.3. Tau lepton fake factors

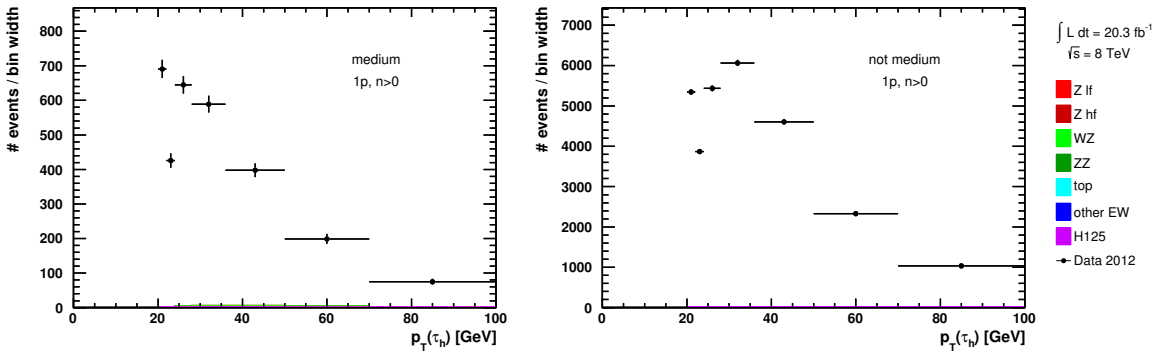
In order to select a sample of *pass* and *fail* τ_h, the following requirements are applied in addition to the pre-selection:

- exactly one additional τ_h in the event,
- veto on events with additional electrons satisfying the *pass* criteria,
- veto on events with additional muons satisfying the *pass* criteria.



(a)

Figure 6.22: Electron fake factors: interpolated fake factor for conversion electrons as a function of p_{T}, together with the statistical uncertainty (points). The variation due to the subtraction of true electrons (shaded area) is too small to be visible.



(a) (b)

Figure 6.23: (a) Numerator and (b) denominator for the measurement of the τ_h fake factors before the subtraction of true hadronic tau decays (coloured area), for 1-prong τ_h and $n > 0$. A varying bin width is used to account for the decreasing number of fake τ_h with p_{T}.

As shown in Fig. 6.23, both the numerator and denominator samples are dominated by fake τ_h. The simulation-based subtraction of true hadronic tau decays has no significant influence on the determination of the fake factor.

Because tau ID treats hadronic tau decays with one or three tracks independently, the fake factors are measured separately for the two cases. In simulated events, the fake factor is observed to depend on the number of τ_h candidates in the event, as shown in Fig. 6.24. The deviation is expected to come from a difference in the sample (jet type) composition. However, the data-driven measurement of the fake factor is performed on events with only one τ_h, in order to keep the signal region untouched. Therefore, the multi-τ_h region is emulated by requiring additional *τ_h-like objects* in events with only one τ_h. These τ_h-like objects are defined by reconstructed hadronic tau decays with two or more than three associated tracks and have a BDT score greater 0.2. Furthermore, they are required to satisfy standard kinematic cuts ($p_{\mathrm{T}} > 20\,\mathrm{GeV}$, $|\eta| < 2.47$), come from the same vertex, and are not allowed to overlap geometrically with any e, μ or τ_h objects. The use of reconstructed tau candidates leads to a very 'τ_h-like' samples of objects, while the contamination with true hadronic tau decays is minimised

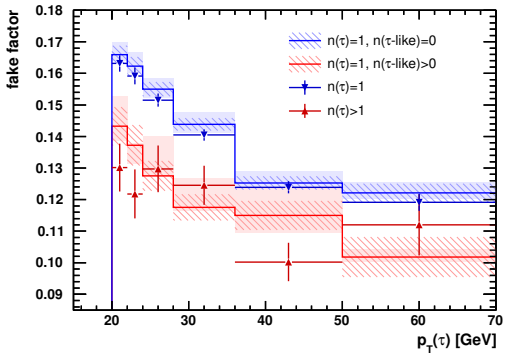
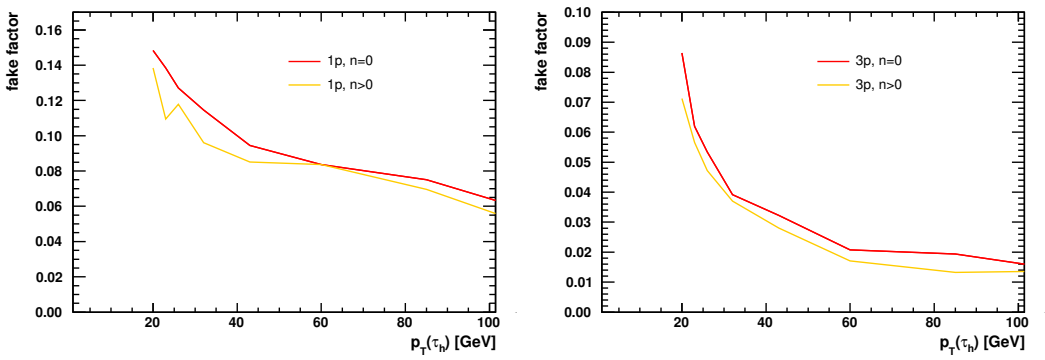


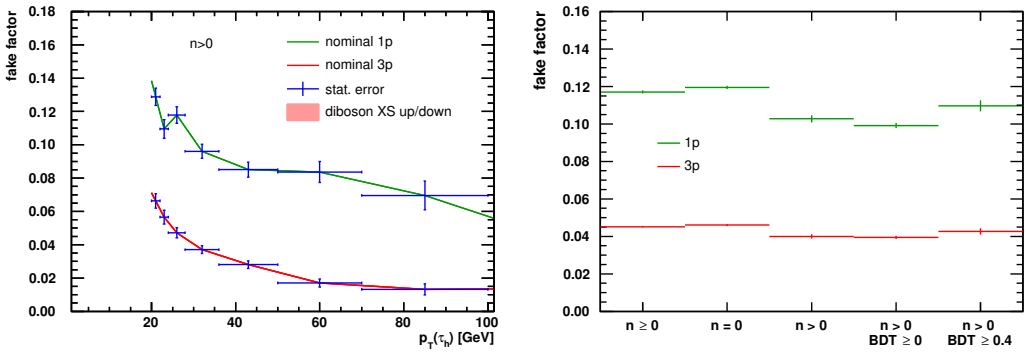
Figure 6.24: τ_h fake factors: fake factor for 1-prong τ_h as a function of p_{T}, for varying numbers of τ_h and τ_h-like objects. The shaded area illustrates the variation for varying the BDT score requirement on the τ_h-like objects between 0 and 0.4, while the striped area represents the statistical uncertainty. Only simulated events are used for this measurement.



(a) (b)

Figure 6.25: τ_h fake factors: Interpolated τ_h fake factors separately for (a) 1-prong and (b) 3-prong candidates, in each case for events with and without τ_h-like objects.

by the requirement on the number of tracks. Figure 6.24 shows the fake factor measured in simulated events with exactly one τ_h object and with or without additional τ_h-like objects. While the fake factor $f_{\mathrm{F}}(n(\tau_h) > 1)$ cannot be reproduced exactly, the trend of a lower fake factor is reproduced correctly. The shaded area shows the variation of the fake factor from varying the BDT score requirement of the τ_h-like objects between 0 and 0.4. While a lower threshold better reproduces $f_{\mathrm{F}}(n(\tau_h) > 0)$, the deviation from $f_{\mathrm{F}}(n(\tau_h) = 0)$ increases. This effect appears because the fake factor binned in the number of τ_h-like objects is derived from the same sample as $f_{\mathrm{F}}(n(\tau_h) = 0)$. Even though the fake factor measured in simulated samples is not exact, the trends are in general well reproduced. The approach using τ_h-like objects to determine the fake factor for events with many τ_h is therefore expected to work well also in data. Figure 6.25 shows measured fake factor in data, in events with and without τ_h-like object. The fake factor for $n > 0$ is then used also for events with more than one (*pass* or *fail*) τ_h. As for the muon and electron fake factors, a linear interpolation is applied between bin-centres in p_{T}.



(a) (b)

Figure 6.26: τ_h fake factors: (a) Interpolated fake factor for $n > 0$. Also shown are the uncertainty due to the statistical limitation of the control region per bin (error bars) and the impact of the variation of the diboson cross section (shaded area, not visible). (b) Average fake factor measured for a varying the number of τ_h-like objects, as well as for varying BDT score requirement for the τ_h-like objects.

Uncertainties on the tau lepton fake factors

The systematic uncertainties associated to the τ_h fake factors are coming from the statistical uncertainty as well as the sample composition

- Statistical limitation: The sample for the τ_h fake factor measurement is well populated and the dependence of the f_F on η weak, so that a fine binning was not necessary. The statistical uncertainties are therefore relatively small over the full momentum range, as shown in Fig. 6.26(a).

- Subtraction of true τ_h: The contamination of true τ_h has no consequence on the determination of the fake factor. The uncertainty due to the subtraction of true τ_h is therefore negligible.

- Definition of τ_h-like objects: Fig 6.26(b) shows the average factor for a varying number of τ_h-like objects as well as a varying BDT score threshold for their definition. The difference between the inclusive measurement $f_F(n \geq 0)$ and $f_F(n > 0)$ is taken as the systematic uncertainty, which covers the variations due to a different BDT score threshold.

6.8. Background estimation: validation and transfer model

6.8.1. Background transfer model

The signal region is defined by events with an H-pair, in which both objects satisfy the *pass* criteria. In terms of the true objects, the events are composed of three kinds of pairs:

- TT-pairs, with two true objects,

- TF-pairs, with one true and one fake object,

- FF-pairs, with two fake objects.

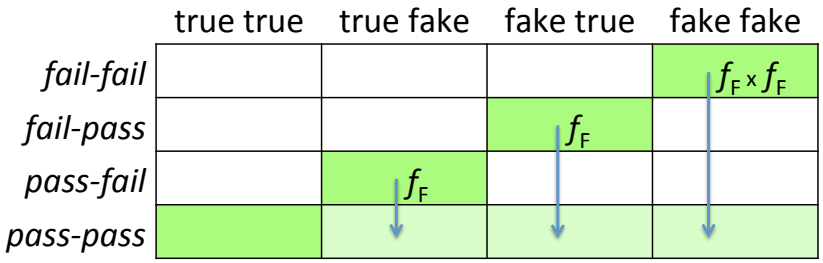


Figure 6.27: The background transfer model for the composition of the *pass-pass* signal region. The events in the light green areas are not well modelled and therefore replaced by events from the side bands. The latter are transferred to the signal region by multiplication with the corresponding fake factor(s) f_F. All events in the dark green areas therefore are used to model the signal region.

Table 6.4: Observed and expected number of events at different steps of the analysis. In relation to the figure legends, the category Z + jets summarises both events with light and heavy flavour jets (Z lf and Z hf). The category 'other' summarises the contribution from minorly contributing electroweak backgrounds and non-signal Higgs boson events (other EW and H125). Only the nominal selection is considered, i.e. SF OS Z-pairs and OS H-pairs.

	Data	Signal	Z + jets	WZ	ZZ	*top*	other	total bgr
Z-pair	13294241	14	12133321	4216	1539	17618	6869	12163563
H-pair	11114	4.5	10301.4	94.2	269.1	386.6	12.4	11063.7
fail-fail	9373	0.3	9077.7	16.7	11.8	336.6	3.5	9446.4
fail-pass	1482	1.8	1185.6	72.8	81.7	46.1	3.2	1389.5
pass-pass	259	2.5	38.0	4.7	175.6	3.8	5.6	227.8

As described above, the fakes of type *pass* are not well modelled by the simulation, and the amount of simulated events is not sufficient for a proper estimate of the amount of fakes in the signal region. Therefore, the TF and FF-pairs are transferred to the signal region from better populated regions. Figure 6.27 illustrates the process. FF-pairs are transferred from the *fail-fail* region, by making use of the fake factors twice (once for each object). TF and FT-pairs are transferred from the *pass-fail* and *fail-pass* regions, respectively, by applying the fake factor to the fake object that satisfies the *fail* criteria. Solely the TT-pairs are used directly and as predicted by the simulation. Because simulated samples are used in all regions it is ensured that each truth type combination (TT, TF, FT, FF) is used only once. Furthermore, since the pairing is carried out before the fake factors are applied, the presence of extra objects (in additional to the Z-pair and H-pair objects) can be disregarded. The background expectation in the signal region is hence composed as

$$\begin{aligned} N(\textit{pass-pass}) \;=\; & N(\mathrm{TT}, \textit{pass-pass}) + N(\mathrm{TF}, \textit{pass-fail}) \times f_F + \\ & N(\mathrm{FT}, \textit{fail-pass}) \times f_F + N(\mathrm{FF}, \textit{fail-fail}) \times f_F f_F \end{aligned} \tag{6.5}$$

The observed and expected number of events for the different steps of the analysis is listed in Table 6.4. The background normalisation as well as the background transfer model are used to estimate the number of contributing fakes.

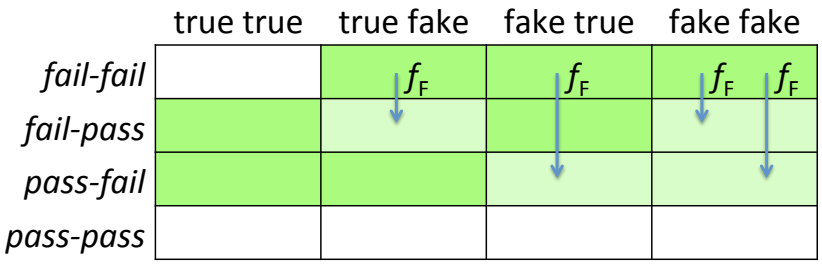


Figure 6.28: Application of the background transfer model on the *fail-pass* and *pass-fail* side bands. The events in the light green areas are not well modelled and therefore replaced by events from the *fail-fail* region. The latter are transferred by multiplication with the corresponding fake factor f_F. All events in the dark green areas therefore are used to model the *fail-pass* and *pass-fail* side bands.

6.8.2. Side bands

The intermediate step of going from *fail-fail* events to *fail-pass* events provides the possibility to validate the method and the fake factors[15]. The logic is illustrated in Fig. 6.28. Whenever a fake object is required to satisfy the *pass* criteria, the pair is taken from the region in which the fake objects satisfies the *fail* criteria and multiplied with the corresponding fake factor f_F. The *fail-pass* regions are then composed as follows[16]:

$$\begin{aligned} N(\textit{fail-pass}) &= N(\mathrm{TT}, \textit{fail-pass}) + N(\mathrm{TF}, \textit{fail-fail}) \times f_F + N(\mathrm{FT}, \textit{fail-pass}) + N(\mathrm{FF}, \textit{fail-fail}) \times f_F \\ N(\textit{pass-fail}) &= N(\mathrm{TT}, \textit{pass-fail}) + N(\mathrm{TF}, \textit{pass-fail}) + N(\mathrm{FT}, \textit{fail-fail}) \times f_F + N(\mathrm{FF}, \textit{fail-fail}) \times f_F \end{aligned} \tag{6.6}$$

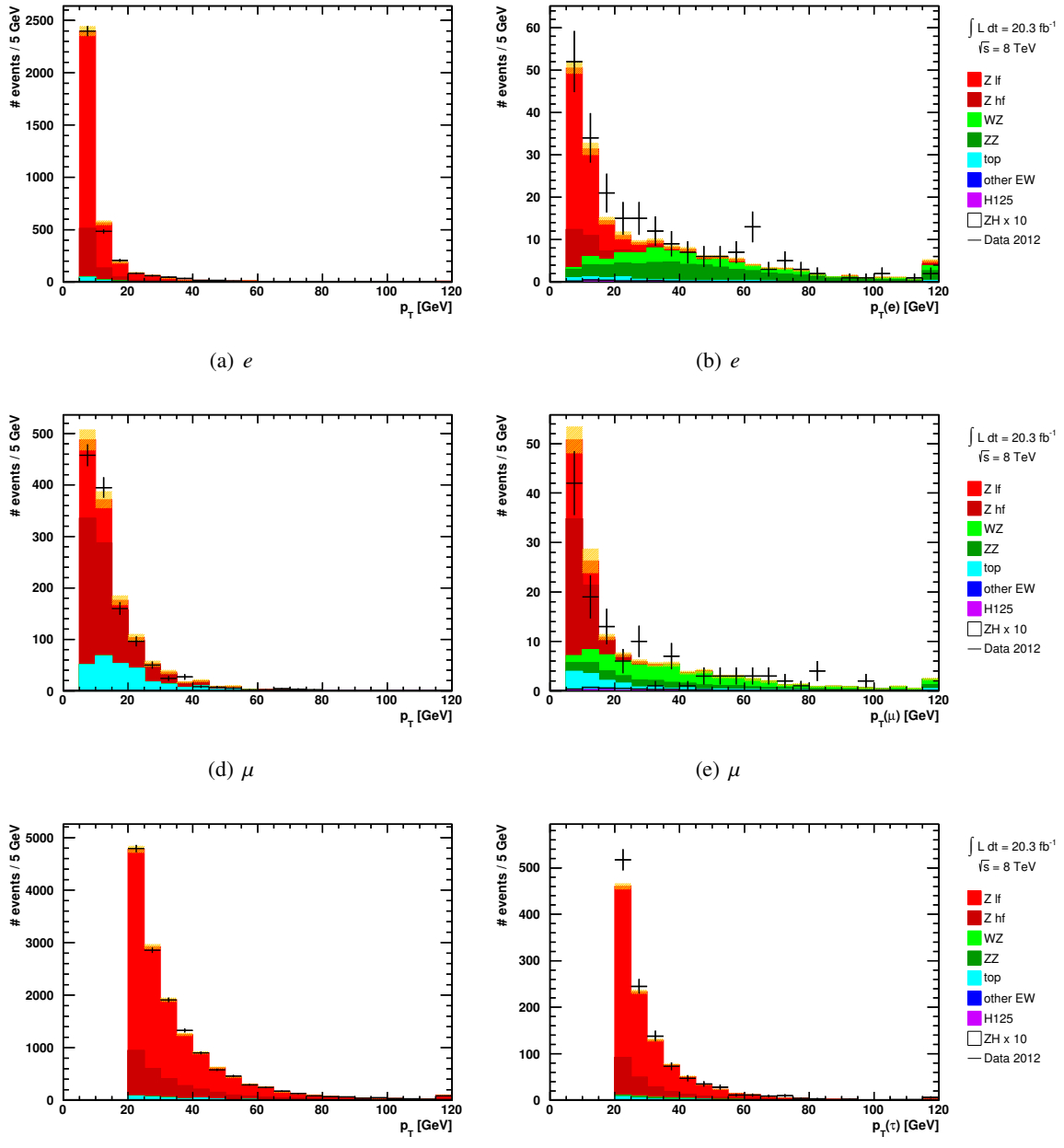
As a first test, the momentum distribution is evaluated for the three object types, e, μ, τ_h. Only objects associated to an OS H-pair are considered. The distributions on the left hand side of Fig. 6.29 shows the *fail-fail* region, sorted by object type. The MC prediction is normalised according to the fit result. Since the fit - in contrast to the figure - is performed on events with both OS and SS H-pairs, the very good agreement with the data also confirms that the balance of both types is correctly reproduced by the simulation. The right hand side of Fig. 6.29 shows the momentum distribution of the *pass* objects in the *fail-pass* and *pass-fail* regions. As given by Eq. 6.6, the fake objects that have to satisfy the *pass* criteria are transferred from the *fail-fail* regions. The true objects are taken directly from the simulation. The observed distributions (data) are not reproduced perfectly, but agree within uncertainties.

As a second test, distributions of the H-pair are evaluated. Figures 6.30 and 6.31 show the sum of the momenta of the H-pair objects and the invariant mass of the pair. The two variables are of interest to enhance the signal purity and to extract the signal at a later stage of the analysis. The purely leptonic modes (Fig. 6.30) show a worse modelling than the hadronic modes (Fig.6.31). The statistical uncertainties are high in these regions, both in the *pass-fail* region and in *fail-fail* region, where the fake leptons are transferred from[17]. Furthermore, the uncertainties from the fit are high for

[15] As a matter of fact, there is a small overlap between the regions used for the fake factor measurement and the side bands, because the fake factor control region is not confined to exactly one additional object. However, the number of events with at least two additional objects is small compared to the number of events with exactly one additional object, and the fake factors are not altered by including the events with many objects. Therefore, the side bands can nevertheless been used to validate the fake factors.

[16] It goes without saying that the composition of regions only applies for the MC prediction. The data is used as observed after applying the corresponding cuts.

[17] The distributions for the *fail-fail* region are shown in the Appendix B.3, Fig. B.5 and Fig. B.6.



(a) e (b) e

(d) μ (e) μ

(g) τ_h (h) τ_h

Figure 6.29: Momentum distribution of (left) all objects in the *fail-fail* region, and (right) the objects in the *pass-fail* and *fail-pass* regions that either satisfy the *pass* criteria directly (for data and true objects) or are transferred from the *fail-fail* region (fake objects). The events are sorted per particle type. The yellow shaded area shows the statistical uncertainty of the MC samples.

theses modes. Nevertheless, the deviation at low masses and low momenta as seen for the *ee* final state cannot be fully explained by the uncertainties. The cause of the deviation remains unclear, as it is not modelled in the simulation. This means in particular that a direct application of the *pass* requirements in the simulated samples (instead of using the fake factors) does not improve the data-MC agreement. Studies on additional MC samples (which are have otherwise not been included in this analysis), namely diboson events with hadronic final states and a higher-statistics $Z\gamma$ sample, are performed, in order to check for missing contributions from such events. However, the studied events do not contribute at a significant level and cannot explain the data-MC disagreement. The fact that the deviation is mainly seen in low populated regions complicates further investigations. As mentioned before, the *ee* and $\mu\mu$ modes are excluded from the signal region, so that the observed deviation does not directly influence the further analysis steps. It is however possible that similar deviations occur in the other electron final states in the signal region.

Based on the deviations seen in the validation plots, an additional asymmetric uncertainty is assigned to the electron fake factor, allowing for 10% less and 50% more electron fakes.

6.8.3. Alternative pre-selection

Another validation region free of signal events is provided by the alternative pre-selection, given by the OF OS-pairs (see Sec. 6.4.3). The region is statistically very limited, especially for the *fail-pass* and *pass-pass* regions. Nevertheless, the alternative pre-selection offers the possibility to validate the top background, as shown in Fig. 6.32. All final states are shown together, in order to increase the statistical significance. No significant difference between the simulation and the data is observed.

6.8.4. Diboson region

The *ee* and $\mu\mu$ final states are increasingly dominated by ZZ events when more objects of *pass* quality are required. The *pass-pass* region therefore provides a control region for the ZZ background, rather than a signal region[18]. Figure 6.33 shows the invariant mass of the H-pair jointly for the *ee* and $\mu\mu$ final states and *pass-pass* events. The estimated background contribution is very small compared to the ZZ events. Nevertheless, a disagreement between data and MC is observed at the order of 10%. This is the argument for the 10% systematic uncertainty assigned to the diboson background normalisaton, which was already introduced and used as part of the fake factor measurement.

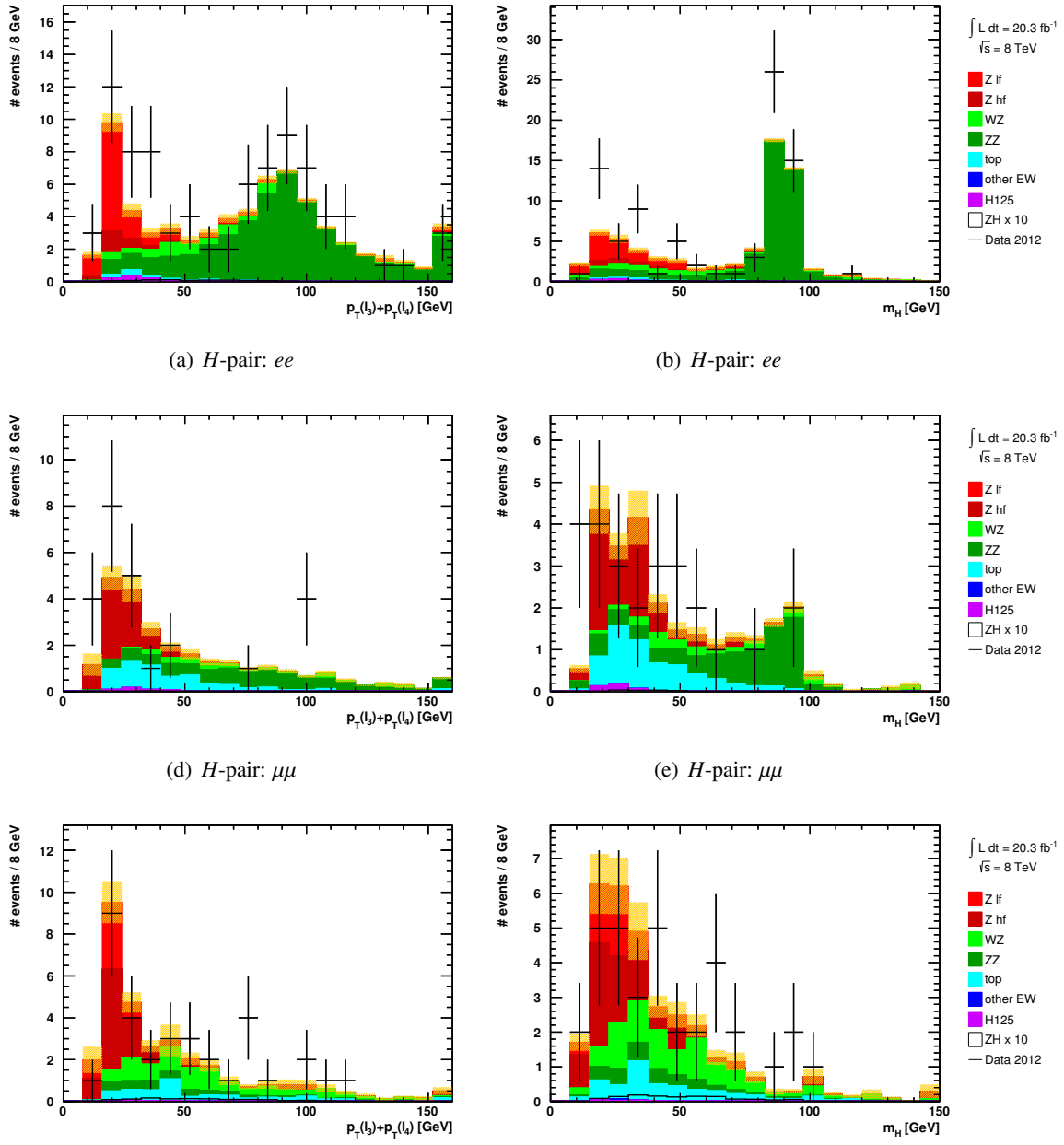
6.9. Signal region

6.9.1. Requirements

The signal region is given by the OS H-pairs in which both constituents satisfy the *pass* criteria. In order to further enhance the signal purity, a final requirement on the sum of the transverse momenta of the two constituents is made. The average momentum sum peaks at higher values for signal events than for the reducible backgrounds. Compared to requirements on the single momenta, a cut on the momentum sum keeps more signal events while rejecting a large fraction of the Z + jets background. Due to the different momentum spectra (see Sec. 6.2, Fig. 6.8) and the varying amount of background, the cut threshold is individually optimised for each final state by evaluating the expected contribution from all backgrounds:

- H-pair: $e\mu$: $p_{\mathrm{T}}(e) + p_{\mathrm{T}}(\mu) \geq 25\,\mathrm{GeV}$,

[18] The signal purity can be increased by a veto on the Z-mass peak and on the missing transverse energy in the event, in order to transform this control region to a signal region. This was however not done in this analysis.



(a) H-pair: ee

(b) H-pair: ee

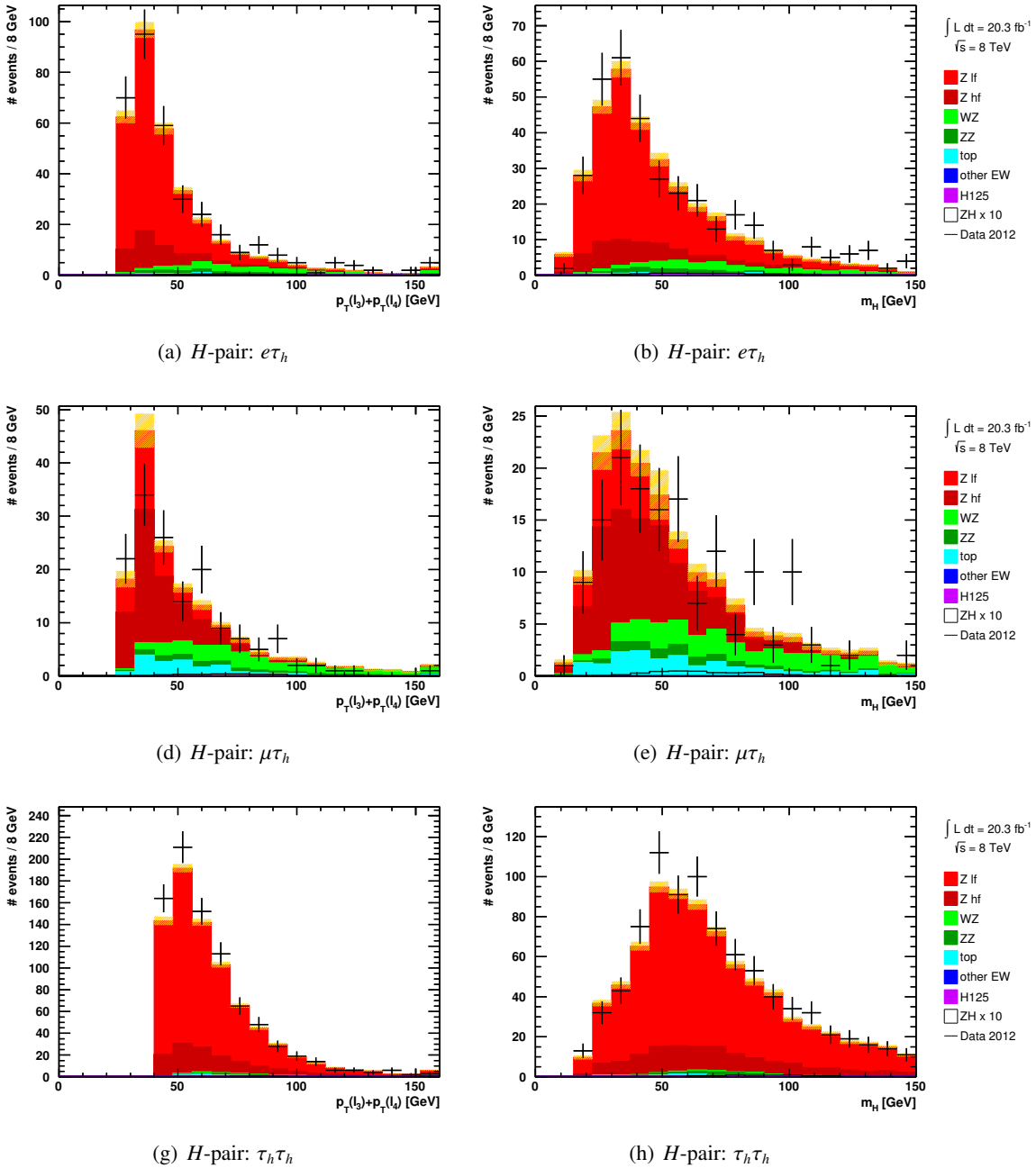
(d) H-pair: $\mu\mu$

(e) H-pair: $\mu\mu$

(g) H-pair: $e\mu$

(h) H-pair: $e\mu$

Figure 6.30: Distribution of (left) the momentum sum of the H-pair constituents, and (right) the invariant mass of H-pair, for the *fail-pass* and *pass-fail* regions and the final states (top) ee, (centre) $\mu\mu$, and (bottom) $e\mu$. The yellow shaded area shows the statistical uncertainty of the MC samples.

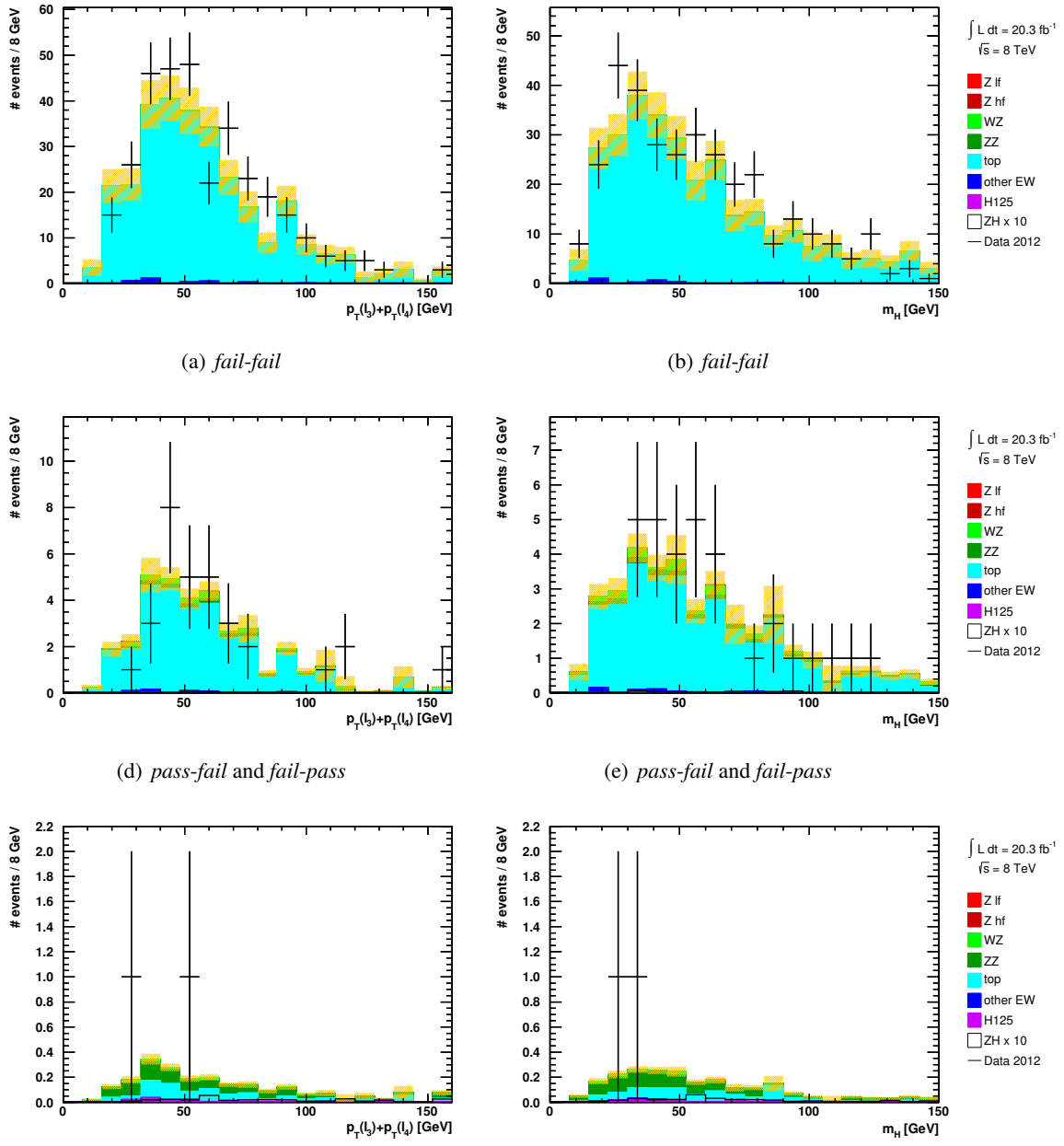


(a) H-pair: $e\tau_h$ (b) H-pair: $e\tau_h$

(d) H-pair: $\mu\tau_h$ (e) H-pair: $\mu\tau_h$

(g) H-pair: $\tau_h\tau_h$ (h) H-pair: $\tau_h\tau_h$

Figure 6.31: Distribution of the (left) the momentum sum of the H-pair constituents, and (right) invariant mass of H-pair, for the *fail-pass* and *pass-fail* regions and the final states (top) $e\tau_h$, (centre) $\mu\tau_h$, and (bottom) $\tau_h\tau_h$. The yellow shaded area shows the statistical uncertainty of the MC samples.

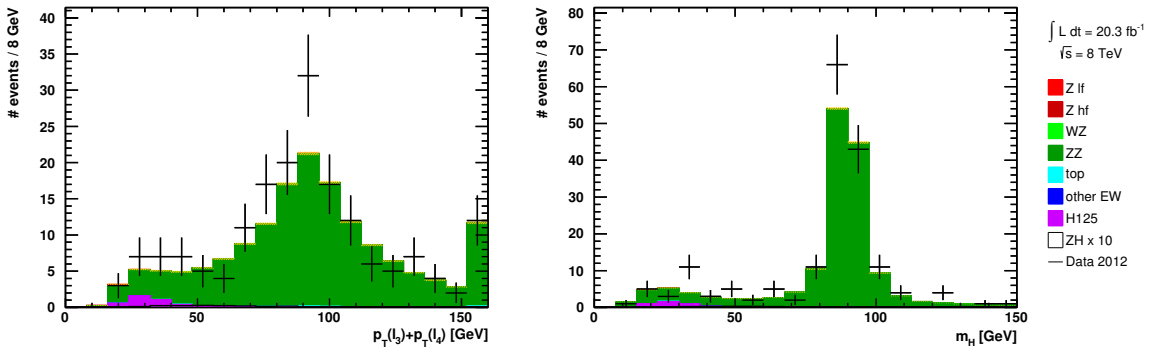


(a) *fail-fail* (b) *fail-fail*

(d) *pass-fail* and *fail-pass* (e) *pass-fail* and *fail-pass*

(g) *pass-pass* (h) *pass-pass*

Figure 6.32: Alternative pre-selection: distribution of (left) the momentum sum of the H-pair constituents, and (right) the invariant mass of the H-pair, for the (top) *fail-fail*, (centre) *fail-pass* and *pass-fail*, and (bottom) *pass-pass* regions. All final states are added together. The yellow band shows the statistical uncertainty of the MC samples.



(a) H-pair: $ee + \mu\mu$

(b) H-pair: $ee + \mu\mu$

Figure 6.33: ZZ control region: distribution of (left) the momentum sum of the H-pair constituents, and (right) the invariant mass of the H-pair, for *pass-pass* events.

- H-pair: $e\tau_h$: $p_T(e) + p_T(\tau_h) \geq 45\,\text{GeV}$,
- H-pair: $\mu\tau_h$: $p_T(\mu) + p_T(\tau_h) \geq 45\,\text{GeV}$,
- H-pair: $\tau_h\tau_h$: $p_T(\tau_h) + p_T(\tau_h) \geq 80\,\text{GeV}$,

The cuts are chosen so that a most a big part of the reducible background is discarded, while keeping most signal events, as shown in Fig. 6.34 and in Table 6.5. In terms of the ratio $S/\sqrt{S+B}$, where S is the number of signal events and B is the number of the background events, the optimal cuts are at higher thresholds. Using the optimal cuts reduces the number of expected background events below 5 in each channel. In order to allow a better judgement and control of the background modelling in the signal region, the slightly looser cuts are chosen. The expected limit evaluated on an Asimov dataset[19] is decreased by about 4% compared to the optimal set of cuts.

The remaining events are used to probe the data for the ZH signal. Essentially, the presence of the signal is tested by fitting the predicted background and signal contributions to the data under consideration of the statistical and systematic uncertainties. For the fit, the different background processes are summed up; all events with two true objects are taken together, as well as all other backgrounds. The first contribution is composed of mainly ZZ events and taken directly from the simulation, as described in Sec. 6.8.1. The second contribution contains fake objects, and is therefore a composite sample of events from all side bands. The four final states are evaluated in a combined fit.

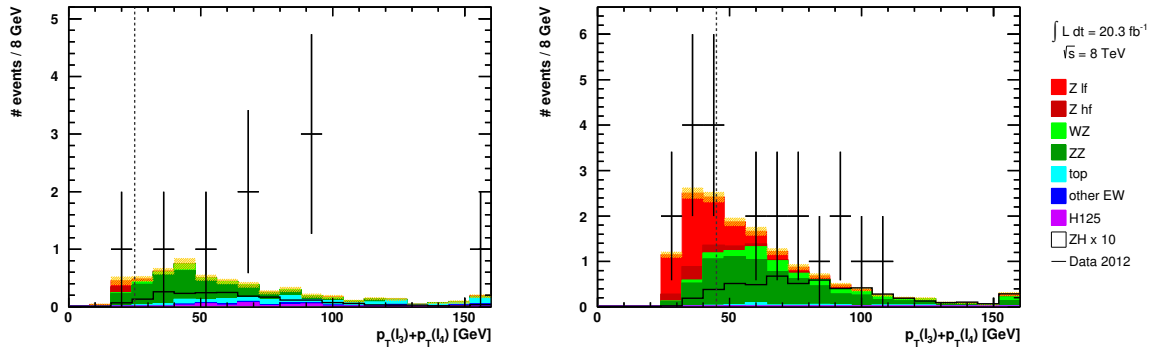
6.9.2. Limit setting

The goal of this analysis is to test the presence of associated Higgs boson production, within the collision data. This is done by testing the data against two hypotheses, the null hypothesis and the signal hypothesis. The null hypothesis is given by the background-only case, meaning that no ZH signal is present, and the data can be described by the background processes only. The signal hypothesis states the presence of the signal in addition to the background. Here, the signal is given by the process $Z(\rightarrow ee/\mu\mu)H(\rightarrow \tau\tau)$ and appears with the *signal strength* μ, with $\mu > 0$. The signal strength is

[19]The Asimov data set is a single representative data set, created from the signal and background expectation. It can be used to execute the limit setting procedure (see Sec. 6.9.2), while keeping the signal region *blinded*, i.e. without using the data events in the signal region.

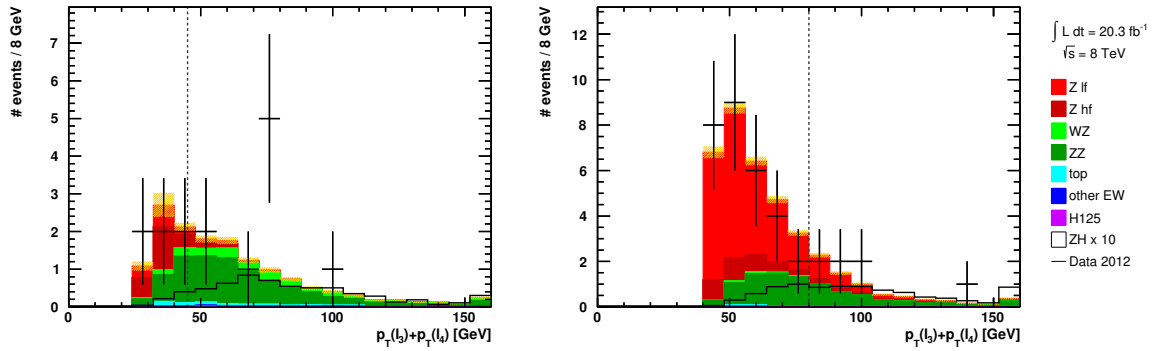
Table 6.5: Number of expected events in the signal region, before and after the requirement on the momentum sum of the H-pair constituents is imposed, separately for each final state. The background is listed as contributions from true-true events (True bgr) and all other events (Fake bgr), which are also the categories that are used in the final fit (see next section).

H-pair	Cut	Signal	Fake bgr	True bgr
$e\mu$	*pass-pass*	0.21	0.89	4.44
	sum $p_{\mathrm{T}} > 25$ GeV	0.20	0.61	4.15
$e\tau_h$	*pass-pass*	0.54	8.13	6.25
	sum $p_{\mathrm{T}} > 45$ GeV	0.49	4.14	5.13
$\mu\tau_h$	*pass-pass*	0.63	6.72	7.85
	sum $p_{\mathrm{T}} > 45$ GeV	0.58	3.17	6.32
$\tau_h\tau_h$	*pass-pass*	0.90	28.31	8.64
	sum $p_{\mathrm{T}} > 80$ GeV	0.62	3.62	3.37



(a) H-pair: $e\mu$

(b) H-pair: $e\tau_h$



(d) H-pair: $\mu\tau_h$

(e) H-pair: $\tau_h\tau_h$

Figure 6.34: Distribution of momentum sum of the H-pair constituents, for the four final states considered in the signal region. The dashed line indicates the chosen cut value (see text). The cuts are optimised without looking at the data (blind analysis). The yellow shaded area shows the statistical uncertainty of the MC samples.

normalised to the Standard Model prediction, so that $\mu = 1$ represents the case in which the process appears with the cross section predicted by the Standard Model. For the hypothesis testing, the mass of the Higgs boson is not a priori fixed. Mass values between 100 GeV and 150 GeV are tested, in steps of 5 GeV. Nevertheless, the focus is on a mass of 125 GeV, which is the mass point closest to the mass of the discovered Higgs boson. The signal strength μ is not fixed either. Instead, a fit is done in order to find the value of μ that best fits the observed data.

For the limit setting, the CL_s approach [169] is used, which is the preferred method for searches within the LHC community [170]. The CL_s method has the advantage of taking into account the measurement sensitivity. In particular, it means that observed data inconsistent with both the null and signal hypothesis will not result in the rejection of the signal hypothesis.

The (binned) invariant mass distribution of the H-pair is used for the hypothesis testing. For each bin i the observed number of events is d_i, and the expected number of events $E_i = \mu s_i + b_i$. The latter is composed of the signal contribution s_i appearing with the signal strength μ and the background contribution b_i. The compatibility in each bin i is evaluated using a Poisson probability,

$$L_i = \frac{(\mu s_i + b_i)^{d_i}}{d_i!} e^{-(\mu s_i + b_i)} \quad . \tag{6.7}$$

A Likelihood function $L(\mu)$ is constructed as the product over all bins. Due to the systematic uncertainties, the expected signal and background yields are not exactly known. The uncertainties enter the likelihood function as *nuisance parameters* θ, where θ stands for all the different nuisance parameters, $\theta = \theta_1, \theta_2, ...$ With these parameters, the full likelihood function appears as:

$$L(\mu, \theta) = \prod_{i=1}^{N} \frac{(\mu s_i(\theta) + b_i(\theta))^{d_i}}{d_i!} e^{-(\mu s_i(\theta) + b_i(\theta))} \quad . \tag{6.8}$$

Each nuisance parameter slightly changes the expected signal and/or background yield, either by effecting the overall normalisation or the shape of the input histogram. The nuisance parameters are constrained along with μ, by searching for the set of parameters that maximises the likelihood function. This is done in two ways. The *unconditional maximum likelihood* $L(\hat{\mu}, \hat{\theta})$ describes the global maximum, with the best fit values $\hat{\mu}$ and $\hat{\theta}$. The best fit signal strength is also called the *observed signal strength*. The *conditional maximum likelihood* $L(\mu, \hat{\hat{\theta}})$ is a function of μ with values $\hat{\hat{\theta}}$ for the nuisance parameters that maximise the likelihood function for a given μ. From the two maximum likelihoods, a profile likelihood ratio $\lambda(\mu)$ is constructed:

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\theta}})}{L(\hat{\mu}, \hat{\theta})} \quad . \tag{6.9}$$

The profile likelihood ratio is broadend by the nuisance parameters, which are not fixed and vary as a function of μ. The width of $\lambda(\mu)$ therefore reflects the degraded measurement precision due to systematic uncertainties.

The test statistic t_μ takes the general form $t_\mu = -\ln(\lambda(\mu))$. Low values of t_μ correspond to a good agreement of data and model for a given μ, while large values represent increasing incompatibility. The test statistic t_μ is slightly modified to take into account the specific requirements of the hypothesis testing. This encompasses especially the requirement of $\mu \geq 0$ (the ZH signal leads to an increase of the number of events in the signal region) and the goal to set an upper limit on the signal strength

(one-sided confidence interval). The modified test statistic q_μ is constructed as follows:

$$q_\mu = \begin{cases} -2\ln\left(\frac{L(\mu,\hat{\hat{\theta}}(\mu))}{L(0,\hat{\hat{\theta}}(0))}\right) & \hat{\mu} < 0 \ , \\ -2\ln\left(\frac{L(\mu,\hat{\hat{\theta}}(\mu))}{L(\hat{\mu},\hat{\theta})}\right) & 0 \leq \hat{\mu} \leq \mu \ , \\ 0 & \hat{\mu} > \mu \ . \end{cases} \tag{6.10}$$

The definition of the test statistic considers different cases for $\hat{\mu}$. First of all, a downward fluctuation in the data, i.e. the observation of fewer events than predicted by the background-only model, should not lead to a rejection of the null (background-only) hypothesis. If $\hat{\mu} < 0$, the best agreement of data and prediction is then achieved for $\mu = 0$. The converse case of a high observed signal strength ($\hat{\mu} > \mu$) should on the other hand not count against the signal hypothesis. It is therefore assigned a value of 0 for the test statistic. The probability p_μ to observe the value $\hat{\mu}$ or more extreme values for a given μ is calculated as

$$p_\mu = \int_{q_{\hat{\mu}}}^{\infty} f(q_\mu|\mu)\, \mathrm{d}q_\mu \ . \tag{6.11}$$

The probability density function $f(q_\mu|\mu)$ describes the probability to find the value q_μ for a given μ. The function $f(q_\mu|\mu)$ needs to be approximated, for example by running pseudo-experiments. The advantage of the profile likelihood ratio is that it allows to approximate $f(q_\mu|\mu)$ also with an analytic function depending on a single representative data set[20].

Following the similar logic, the probability for the background-only case is constructed as

$$p_b = \int_{-\infty}^{q_{\hat{\mu}}} f(q_\mu|0)\, \mathrm{d}q_\mu \ . \tag{6.12}$$

The CL_s methods sets both the signal and background-only probabilities in relation:

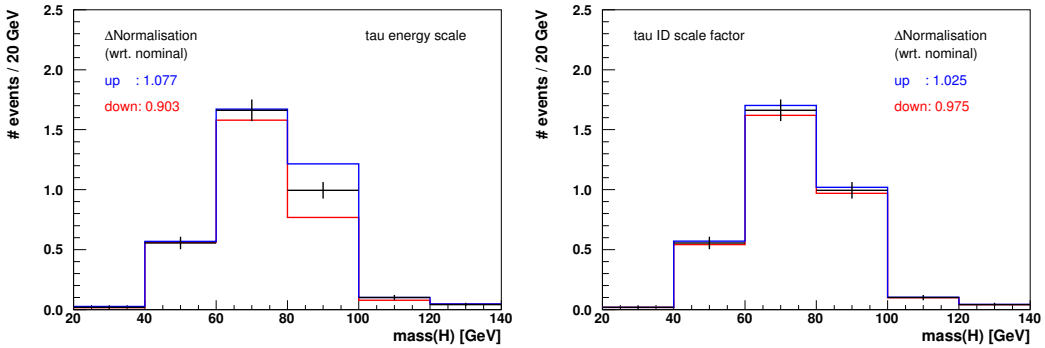
$$CL_s = \frac{p_\mu}{1 - p_b} \ . \tag{6.13}$$

If CL_s falls below a threshold α, the corresponding signal strength or more extreme values are rejected at a confidence level of $1-\alpha$. In this analysis, the 95% confidence level is used. Hence, for setting the upper limit on the signal strength, the value of μ is searched for for which $CL_s < 0.05$. Effectively, the signal probability p_μ is penalised by the background-only probability p_b. If the pdfs are well separated, p_b will be small and the CL_s limit will approach the direct limit on p_μ (commonly called the CL_{s+b} limit [169]). If the pdfs overlap to a large extend, the denominator $1 - p_b$ becomes small. A small sensitivity to the signal is therefore reflected in a weaker CL_s limit.

6.10. Systematic Uncertainties

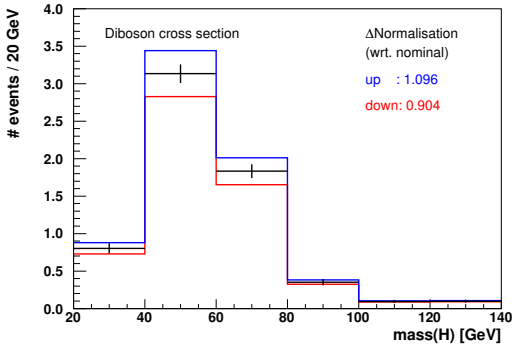
Systematic uncertainties enter the signal region from different sources. The scale factors associated to reconstruction and identification efficiencies as well as energy calibration corrections come with associated uncertainties. Furthermore, the simulated samples are weighted using calculated (theory) cross sections, which are known within certain bounds only. The largest systematic uncertainties are associated to the background estimation developed during the course of this analysis, both from the background normalisation fit and from the fake factor determination. The uncertainties enter the fit as variations of the shape and normalisation of the signal and background events. The systematic uncertainties are treated fully correlated between final states.

[20] Reference [170] provides a detailed discussion of the test statistic and the pdfs used here, including mathematic approximations and simplifications, as well as an explanation of the single representative data set, the so called Asimov data set.

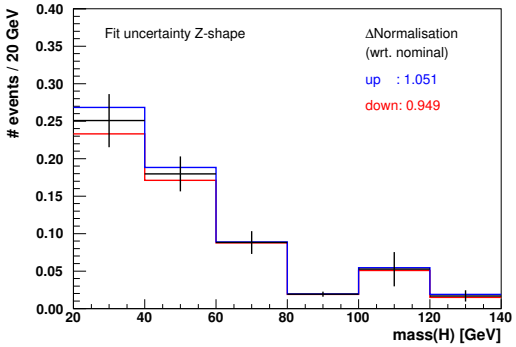


(a) H-pair: $\tau_h\tau_h$

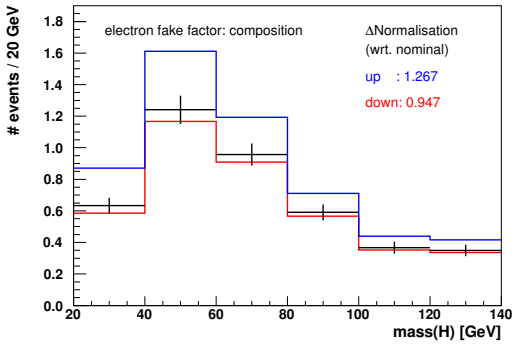
(b) H-pair: $\tau_h\tau_h$



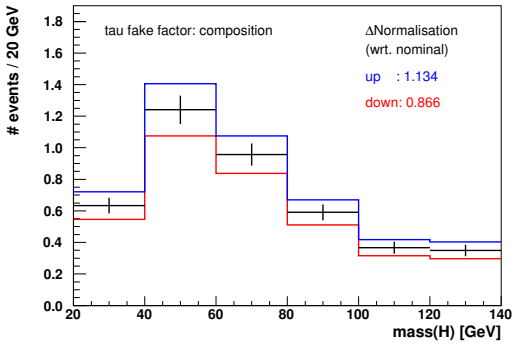
(c) H-pair: $\mu\tau_h$



(d) H-pair: $e\mu$



(e) H-pair: $e\tau_h$



(f) H-pair: $e\tau_h$

Figure 6.35: Size of representative systematic uncertainties for different final states. The nominal distribution of the invariant mass of the H-pair in the signal region is shown (black), as well as the up and down variation due to the following uncertainties: (a) tau energy scale, (b) tau ID scale factor, (c) diboson cross section, (d) Z-shape fit uncertainty, (e) electron fake factor, (f) τ_h fake factor. (a-c) show the background contribution from pairs of two true objects, while (d-f) show the contribution from all other backgrounds (fakes). The final state is indicated beneath each figure.

Luminosity: The uncertainty on the luminosity measurement amounts to 2.8%. The normalisation of all samples is varied up and down by that amount. An exception are the background events that are transferred from the *fail-fail* region, since these events obtain their normalisation from the fit.

Electron reconstruction: The uncertainties related to the energy calibration, the trigger, and the reconstruction and identification efficiencies are provided by the respective ATLAS working groups. They are given as recommendations for variations of the energy calibration and of the scale factors. These variations are carried through the entire analysis. The uncertainties due to the reconstruction and identification give normalisation as well as per bin variations (i.e. variations in the shape) of at maximum 2%. The calibration uncertainties mainly influence the shape of the mass distribution, with per bin shifts of mostly 2%-4%.

Muon reconstruction: As in the case of electrons, the uncertainties related to the energy calibration, the trigger, and the reconstruction efficiencies are provided by the respective ATLAS working groups. They are given as variations to the calibration or scale factors and carried through the entire analysis. The effect is at the sub-percent level.

Tau lepton reconstruction: The uncertainties related to hadronic tau decays are provided by the respective working group and are carried through the analysis. Variations on the tau energy calibration as well as the tau identification scale factors are provided. Due to the relative larger size of the uncertainties, the variations can have a notable effect on the mass distribution, as shown in Fig. 6.35(a) for the most affected final state, $\tau_h\tau_h$. The energy scale uncertainty causes variations of up to 10% or more in individual bins. In the mixed final states ($e\tau_h$, $\mu\tau_h$) the uncertainty is at the 5% level. The ID uncertainty typically is 2%-3%, as shown in Fig. 6.35(b).

Signal: Additional uncertainties on the ZH signal are due to theoretically calculated cross sections and branching rations. The size of the uncertainties is dependent on the mass point. The uncertainty is applied as up and down variations of the normalisation of the signal samples. The following variations are considered: The uncertainty on the branching ratio to tau leptons of 7% to 4% (for increasing mass), the QCD scale uncertainties of 2% to 4%, and the uncertainty on the parton distribution functions and parton scattering of 2% to 3%. The uncertainties are provided by the LHC cross section working group; details are given in Ref. [171] and [130].

Diboson events: As mentioned above, the cross section of the WZ and ZZ events is varied by $\pm 10\%$. The value exceeds the theoretical cross section uncertainty, and is motivated by the discrepancy seen in the ee and $\mu\mu$ final states. The variation of the diboson cross section does not only affect the true backgrounds, but also the fake backgrounds, due to the influence on the fake factor measurement. The latter is notable in the final states with muons and negligible for the di-τ_h state. The effect on the true backgrounds in the $\mu\tau_h$ final state is shown in Fig. 6.35(c).

Background estimation - fit: The largest uncertainties are associated with the data driven background estimation. Because the remaining top background is small, solely the fit uncertainty of the Z boson background enters the signal region at a notable level. Since each final state is fitted independently, the fit uncertainties are independent and treated fully uncorrelated. The variation on the mass distribution for the $e\mu$ final state is shown in Fig. 6.35(d).

Background estimation - fake factors: The largest uncertainties are connected to the fake factors. Especially the variations due to the sample composition are significant. The largest uncertainty is associated with the fake electrons, due to the discrepancy observed in the di-electron validation region. The influence of the fake factor uncertainty due to the sample composition is shown in Fig. 6.35(e) and Fig. 6.35(f) for $e\tau_h$ final state.

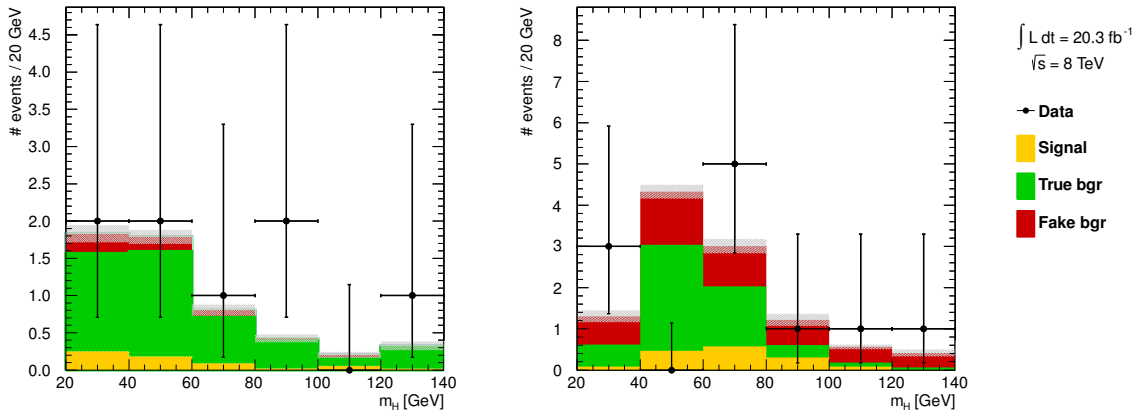
6.11. Results

The post-fit distributions of the invariant mass of the H-pair are shown in Fig. 6.36. The signal regions are dominated by the contribution of the ZZ background (the background contribution with two true objects), but have a remaining contribution of backgrounds with fakes. The background-only contributions tend to underestimate the data, so that the observed signal strength is at $\hat{\mu} = 3.1^{+2.9}_{-2.9}$. Nevertheless, the statistical limitation of the signal region is apparent, and within the large statistical and systematic uncertainties, the background-only hypothesis is in good agreement with the data. Figure 6.37 shows the upper limits set on the signal strength at 95% confidence level, per channel and combined. The limit is evaluated for Higgs boson masses between $m_H = 100\,\text{GeV}$ and $m_H = 150\,\text{GeV}$, in steps of 5 GeV. The sensitivity of the analysis is best for small Higgs masses and degrades for higher masses. Towards higher masses, the difference between observed and expected limit increases, showing a slight excess in the data. The behaviour is driven by the $e\mu$ and $e\tau_h$ final states. However, the excess remains within the uncertainty bands and is therefore in no contradiction to the expectation. For a Higgs boson mass of 125 GeV, the upper observed and expected limits are:

$$\text{observed}: \quad \mu_{95\%} = 8.96\ , \tag{6.14}$$

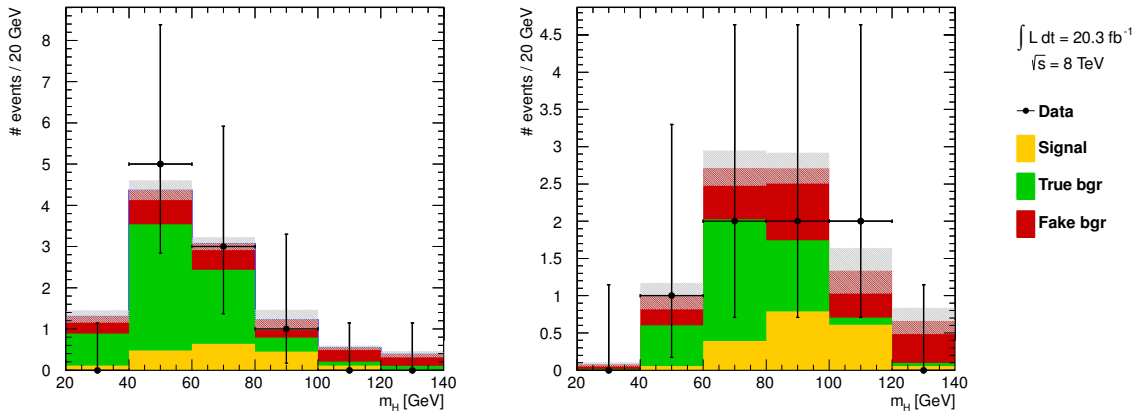
$$\text{expected}: \quad \mu_{95\%} = 6.18^{+8.94}_{-4.45}\ . \tag{6.15}$$

As expected, the data collected in 2012 is not sufficient to observe the signal process. An obvious improvement to increase the sensitivity is therefore the collection of more data - which has just started with the LHC Run II. The additional data would not only help to decrease the statistical uncertainties, but also help to understand the four-object side bands. In particular, it would be possible to understand which discrepancies are of statistical and which are of physical nature. The high uncertainties assigned to the fake factors, especially the electron fake factor, could then potentially be decreased. Increased MC statistics would also benefit the analysis. In depth studies of the composition of the signal region, and even the *pass-fail* regions, were not possible due to the low number of MC events as soon as the actual cuts were applied (instead of using the fake factor method). On the basis of the current dataset, the analysis could be improved by the usage of a better Higgs boson mass reconstruction. More advanced mass reconstruction methods take into account the missing transverse energy, and can narrow the width of the reconstructed mass distribution considerably. As a consequence, the irreducible ZZ background and the ZH signal would be more readily distinguishable. Finally, a more careful treatement of the true backgrounds (i.e. ZZ events) could potentially decrease systematic uncertainties. Even though the MC expectation for prompt leptons is much more reliable, any mismodelling of the ZZ background directly influences the limit setting, because these events represent the dominant background in the signal region. A very good modelling of the diboson background is also essential for the inclusion of the missing two final states (ee and $\mu\mu$).



(a) H-pair: $e\mu$

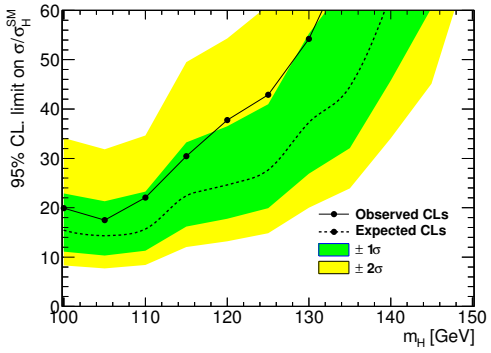
(b) H-pair: $e\tau_h$



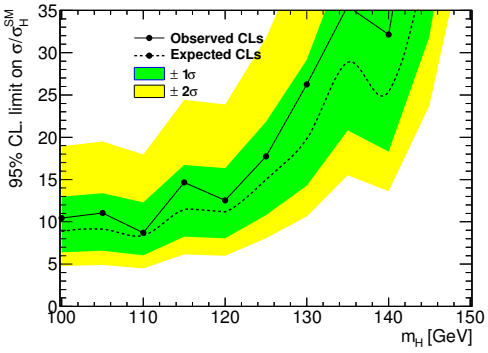
(d) H-pair: $\mu\tau_h$

(e) H-pair: $\tau_h\tau_h$

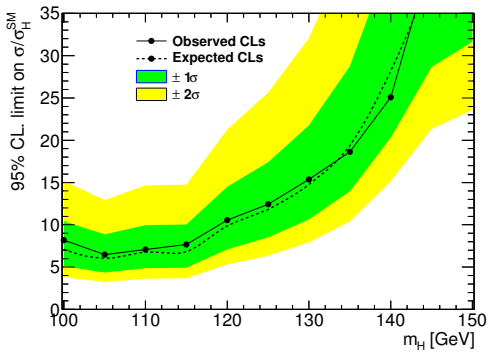
Figure 6.36: Invariant mass distribution of the H-pair for the global maximum of the combined fit for the four final states, and a Higgs boson mass of 125 GeV. The signal distribution is scaled to the observed signal strength of $\hat{\mu} = 3.1$. The gray area shows the fit uncertainty.



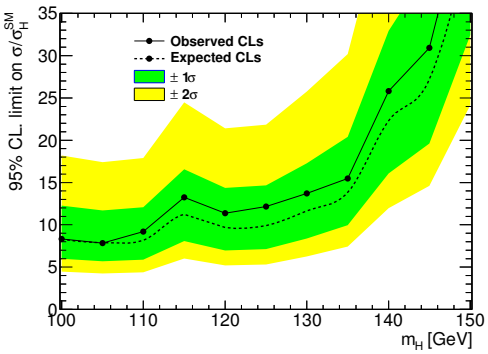
(a) H-pair: $e\mu$



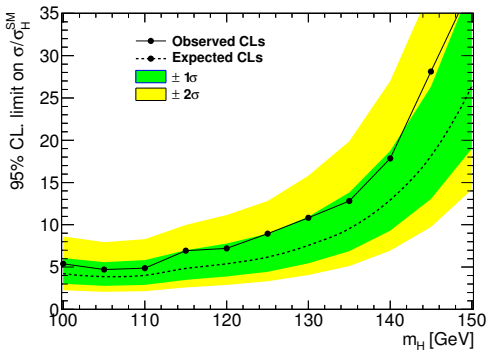
(b) H-pair: $e\tau_h$



(c) H-pair: $\mu\tau_h$



(d) H-pair: $\tau_h\tau_h$



(e) combination

Figure 6.37: The upper 95% confidence level on the signal strength for mass hypothesis ranging from $m_H = 100\,\text{GeV}$ to $m_H = 150\,\text{GeV}$ in steps of 5 GeV. The limit is evaluated for each mass point individually and then interpolated between mass points. The dashed and continous lines represents the expected and observed upper limit, respectively. The uncertainty on the expectation is shown by the green and yellow bands.

7. Conclusion

This thesis addresses the identification of hadronically decaying tau leptons and associated Higgs boson production with tau leptons in the final state. The work was performed within the ATLAS collaboration using proton-proton collisions from the LHC from the data-taking periods in 2011 at $\sqrt{s} = 7$ TeV and 2012 at $\sqrt{s} = 8$ TeV.

The identification of hadronically decaying tau leptons is a challenging task due to the enourmous background of multi-jet events at the LHC. Multivariate techniques are used to reject this background, which have to be maintained continously to ensure a stable performance despite the increasing pile-up and centre-of-mass energy over the years. The studies presented here target the optimisation of the multivariate-based tau identification with boosted decision trees. In particluar, the pile-up robustness of the algorithm is studied, which was a major concern before the start of the 2012 data-taking period with much higher expected pile-up conditions. As presented in this thesis, the dependence of the tau identification efficiency on pile-up was reduced considerably. Three steps contribute to the improvement. The set of discriminating variables is revised, discarding the ones that are very pile-up dependent and are not contributing significantly to the background rejection. Calorimeter-based quantities are redefined to use a smaller cone of $\Delta R < 0.2$ around the tau axis. Finally, the impact of the remaining pile-up dependence is minimised by adding a correction term to certain variables, and thus 'subtract' energy contributions from pile-up events. The improvements were implemented and applied as a basis for the tau identification algorithms used during the 2012 data-taking period.

Further studies of tau identification concern the momentum dependence and the substructure of the decay. In the default training procedure of the BDT, the momentum distributions of background and tau signal events are required to follow the same distribution. The studies confirm that this requirement is essential for a good performance over a wide range of momenta. Besides that, no critical impact of the momentum dependence is found on the performance of the BDT tau ID. Furthermore, it is tested whether a simple transformation of the discriminating variables can be used to obtain a momentum independent ID. While the transformation does indeed decreases the dependence, complete independence is not reached. Finally, the use of advanced substructure algorithms is investigated. These algorithms explore the substructure of the tau detector object, in order to identify the individual hadrons of the decay. The comparison of many variable sets show that a combined use of the currently used variables and substructure-based variables achieves the highest background rejection. In parallel, a variable set is developed which does not rely on a specific substructure algorithm.

The Higgs boson was discovered in 2012 by the CMS and ATLAS collaborations, and has since been studied in detail. Most studies are based on the ggF and VBF production modes, which are the main production modes at the LHC. In this thesis, a search is carried out for events in which the Higgs boson is produced in association with a Z boson, and decays into a pair of tau leptons. The search region is characterised by the presence of four objects, namely two electrons or two muons from the decay of the Z boson and two electrons, muons or τ_h from the subsequent decay of the tau leptons. The requirement of four objects leads not only to a statistically limited signal region, but turns out to be a limiting factor for the definition of four-object side bands as well. Simulated samples are used for the estimation of all backgrounds. Because the modelling of fakes is poorly described in MC samples, a data-driven approach is used to minimise the dependence on the mis-modelling. Nevertheless, the background model does not work in all validation regions equally well. Due to an insufficient amount of simulated events and the statistical limitation of the affected regions, the

cause of the discrepancy is not identified. As a consequence, large uncertainties are assigned on the backgrounds with fake objects. Nevertheless, the investigations show the importance of understanding the background composition of the signal region and the side bands. Especially, the mixture of light and heavy flavour jets seems to be of importance. Eventually, an upper limit is set on the signal strength of the process $Z(\rightarrow ee, \mu\mu)H(\rightarrow \tau\tau)$. Using a combined fit of the four final states $e\mu$, $e\tau_h$, $\mu\tau_h$, $\tau_h\tau_h$, the observed (expected) upper limit is $\mu_{95\%} = 8.96(6.18^{+8.94}_{-4.45})$ at the 95% confidence level, for a Higgs boson mass of 125 GeV.

Bibliography

[1] F. J. Litterst, *Atomismus und Kontinuum: Ein Streit der Vorsokratiker und seine Folgen*. Cramer, Braunschweig, 2010. http://www.digibib.tu-bs.de/?docid=00050433.

[2] M. Amoretti, et al., *Production and detection of cold antihydrogen atoms*, Nature **419** (2002)456–459, doi:10.1038/nature01096.

[3] BELLE Collaboration, *Observation of a resonance-like structure in the $\pi^{\pm}\psi'$ mass distribution in exclusive $B \to K\pi^{\pm}\psi'$ decays*, Phys. Rev. Lett. **100** (2008)142001, doi:10.1103/PhysRevLett.100.142001, arXiv:0708.1790 [hep-ex].

[4] BABAR Collaboration, *Search for the Z(4430)- at BABAR*, Phys. Rev. **D79** (2009)112001, doi:10.1103/PhysRevD.79.112001, arXiv:0811.0564 [hep-ex].

[5] LHCb Collaboration, *Observation of the resonant character of the $Z(4430)^{-}$ state*, Phys. Rev. Lett. **112** (2014)222002, doi:10.1103/PhysRevLett.112.222002, arXiv:1404.1903 [hep-ex].

[6] T. Y. Cao, *Conceptual developments of 20th century field theories*. Cambridge University Press, 1998.

[7] C. Quigg, *The state of the standard model*, AIP Conf.Proc. **542** (2000)3–28, doi:10.1063/1.1336236, arXiv:hep-ph/0001145 [hep-ph].

[8] K. A. Olive, et al. (Particle Data Group), *The review of particle physics*, Chin. Phys. C, **38**, 090001 (2014). http://pdg.lbl.gov.

[9] F. Halzen and A. D. Martin, *Quarks and leptons: An introductory course in modern particle physics*. John Wiley & Sons, 1984.

[10] E. Noether, *Invariante Variationsprobleme*, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse **1918** (1918)235–257.

[11] D. Griffiths, *Introduction to elementary particles*. WILEY-VCH Verlag, Weinheim, 2nd ed., 2008.

[12] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*. Perseus Book Publishing, L.L.C., Reading, Massachusetts, 1995.

[13] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13**, 9 (1964) 321–323, doi:10.1103/PhysRevLett.13.321.

[14] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13**, 16 (1964) 508–509, doi:10.1103/PhysRevLett.13.508.

[15] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global conservation laws and massless particles*, Phys. Rev. Lett. **13**, 20 (1964) 585–587, doi:10.1103/PhysRevLett.13.585. http://link.aps.org/doi/10.1103/PhysRevLett.13.585.

[16] S. L. Glashow, *Partial-symmetries of weak interactions*, Nucl. Phys. **22** (1961)579–588, doi:10.1016/0029-5582(61)90469-2.

[17] S. Weinberg, *A model of leptons*, Phys. Rev. Lett. **19** (1967)1264–1266, doi:10.1103/PhysRevLett.19.1264.

[18] A. Salam, *Weak and electromagnetic interactions*, in *Elementary particle theory. Relativistic groups and analyticity. Proceedings of the 8th Nobel Symposium*, N. Svartholm (ed.), pp. 367–377. Almqvist and Wiksell, Stockholm, 1968.

[19] J. Ellis, *Higgs physics*, CERN (Dec 2013), CERN-PH-TH-2013-315, arXiv:1312.5672.

[20] M. Drees and G. Gerbier, *Dark matter*, Sept 2013, in [8].

[21] S. P. Martin, *A supersymmetry primer*, Northern Illinois University and FNAL (Sept 2011), arXiv:hep-ph/9709356 [hep-ph].

[22] M. J. Mortonson, D. H. Weinberg, and M. White, *Dark energy*, Nov 2013, in [8].

[23] L. Canetti, M. Drewes, and M. Shaposhnikov, *Matter and antimatter in the universe*, New J. Phys. **14** (2012)095012, doi:10.1088/1367-2630/14/9/095012, arXiv:1204.4186 [hep-ph].

[24] K. Nakamura and S. T. Petcov, *Neutrino mass, mixing, and oscillations*, May 2014, in [8].

[25] J. L. Feng, *Naturalness and the status of supersymmetry*, Ann. Rev. Nucl. Part. Sci. **63** (2013)351–382, doi:10.1146/annurev-nucl-102010-130447, arXiv:1302.6587 [hep-ph].

[26] G. F. Giudice, *Naturally speaking: the naturalness criterion and physics at the LHC*, 2008, arXiv:0801.2562 [hep-ph].

[27] R. P. Woodard, *How far are we from the quantum theory of gravity?*, Rept. Prog. Phys. **72** (2009)126002, doi:10.1088/0034-4885/72/12/126002, arXiv:0907.4238 [gr-qc].

[28] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. **C63** (2009)189–285, doi:10.1140/epjc/s10052-009-1072-5, arXiv:0901.0002 [hep-ph].

[29] G. Watt, *private communication*, http://mstwpdf.hepforge.org/plots/plots.html. retrieved Jan 25 2015.

[30] A. Buckley, et al., *General-purpose event generators for LHC physics*, Phys. Rept. **504** (2011)145–233, doi:10.1016/j.physrep.2011.03.005, arXiv:1101.2599 [hep-ph].

[31] M. Bähr, et al., *Herwig++ physics and manual*, Eur. Phys. J. **C58** (2008)639–707, doi:10.1140/epjc/s10052-008-0798-9, arXiv:0803.0883 [hep-ph].

[32] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008)852–867, doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820 [hep-ph].

[33] P. Golonka and Z. Was, *PHOTOS Monte Carlo: a precision tool for QED corrections in Z and W decays*, Eur. Phys. J. **C45** (2006)97–107, doi:10.1140/epjc/s2005-02396-4, arXiv:hep-ph/0506026 [hep-ph].

[34] S. Agostinelli et al., *GEANT4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003)250–303, doi:10.1016/S0168-9002(03)01368-8.

[35] CERN, European Organization for Nuclear Research. www.cern.ch.

[36] C. D. Anderson, *The apparent existence of easily deflectable positives*, Science **76** (1932)238–239, doi:10.1126/science.76.1967.238.

[37] J. C. Street and E. C. Stevenson, *New evidence for the existence of a particle of mass intermediate between the proton and electron*, Phys. Rev. **52** (1937)1003–1004, doi:10.1103/PhysRev.52.1003.

[38] S. H. Neddermeyer and C. D. Anderson, *Note on the cature of cosmic-ray particles*, Phys. Rev. **51** (1937)884–886, doi:10.1103/PhysRev.51.884.

[39] C. M. G. Lattes, H. Muirhead, G. P. S. Occhialini, and C. F. Powell, *Processes involving charged mesons*, Nature **159** (1947)694–697, doi:10.1038/159694a0.

[40] C. F. Powell, D. H. Perkins, and P. H. Fowler, *The study of elementary particles by the photographic method : an account of the principal techniques and discoveries illustrated by an atlas of photomicrographs*. Pergamon Press Ltd., London, 1959.

[41] W. Pauli, *Offener Brief an die Gruppe der Radioaktiven bei der Gauvereins-Tagung zu Tübingen*, Dec 1930.

[42] E. Fermi, *Versuch einer Theorie der Betastrahlen*, Zeitschrift für Physik **Bd. 88** (1934)161–177, doi:10.1007/BF01351864.

[43] C. L. Cowan, F. Reines, F. B. Harrison, H. W. Kruse, and A. D. McGuire, *Detection of the free neutrino: a confirmation*, Science **124** (1956)103–104, doi:10.1126/science.124.3212.103.

[44] M. S. Livingston, *Early history of particle accelerator*, Adv. Electron El. Phys. **50** (1980)1–88, doi:10.1016/S0065-2539(08)61061-6.

[45] J. D. Cockcroft and E. T. S. Walton, *Experiments with high velocity positive ions. (I) Further developments in the method of obtaining high velocity positive ions*, Proc. Roy. Soc. Lond. **A136** (1932)619–630, doi:10.1098/rspa.1932.0107.

[46] J. D. Cockcroft and E. T. S. Walton, *Experiments with high velocity positive ions. (II) Disintegration of elements by high velocity protons*, Proc. Roy. Soc. Lond. **A137** (1932)229–242, doi:10.1098/rspa.1932.0133.

[47] J. Cockcroft and E. Walton, *Experiments with high velocity positive ions. (III) The disintegration of lithium, boron, and carbon by heavy hydrogen ions*, Proc. Roy. Soc. Lond. **A144** (1934)704–720, doi:10.1098/rspa.1934.0078.

[48] R. Wideröe, *Über ein neues Prinzip zur Herstellung hoher Spannungen*, Archiv für Elektrotechnik **21** (1928)387–406, doi:10.1007/BF01656341.

[49] E. O. Lawrence, *Method and apparatus for the acceleration of ions*, Feb 1934. http://www.google.com/patents/US1948384. US patent 1948384 A.

[50] E. J. N. Wilson, *Fifty years of synchrotrons*, Conf. Proc. **C960610** (1996)135–139.

[51] Brookhaven National Laboratory, *Cosmotron (1952-1966)*, http://www.bnl.gov/about/history/accelerators.php. retrieved Dec 18 2014.

[52] R. B. Neal, D. W. Dupen, H. A. Hogg, and G. A. Loew, *The Stanford two-mile accelerator*, 1968, W. A. Benjamin. http://www.slac.stanford.edu/spires/hep/HEPPDF/twomile/Stanford_Two_Mile_Complete.pdf.

[53] Stanford Linear Acceleration Center, *SLAC Linear Collider conceptual design report*, SLAC, Stanford University (June 1980), SLAC-R-229. http://www.slac.stanford.edu/pubs/slacreports/slac-r-229.html.

[54] J. Beringer (ed.), *High-energy collider parameters*, Sept 2013, in [8].

[55] T. Behnke, et al., *The international linear collider technical design report - volume 1: executive summary*, 2013, ILC-REPORT-2013-040, arXiv:1306.6327 [physics.acc-ph].

[56] M. Koratzinos, et al., *TLEP: a high-performance circular e^+e^- collider to study the Higgs boson*, 2013. arXiv:1305.6498 [physics.acc-ph].

[57] M. Aicheler, et al., *A multi-TeV linear collider based on CLIC technology: CLIC conceptual design report*, 2012, CERN-2012-007, SLAC-R-985, KEK-Report-2012-1, PSI-12-01, JAI-2012-001.

[58] M. Litos, et al., *High-efficiency acceleration of an electron beam in a plasma wakefield accelerator*, Nature **515** (2014)92–95, doi:10.1038/nature13882.

[59] C. Lefèvre, *The CERN accelerator complex. Complexe des accélérateurs du CERN*, Dec 2008. http://cds.cern.ch/record/1260465.

[60] CERN, *Experiments and accelerators*, http://home.web.cern.ch/about/experiments. retrieved Oct 29 2014.

[61] CERN, *The accelerator complex*, http://home.web.cern.ch/about/accelerators. retrieved Sept 2 2014.

[62] L. Evans and P. Bryant (eds.), *LHC machine*, JINST **3** (2008)S08001, doi:10.1088/1748-0221/3/08/S08001.

[63] J. Wenninger, *The CERN accelerator complex*, Aug 2008. http://www.teilchenphysik.de/multimedia/vortraege. retrieved Sept 2 2014.

[64] ATLAS Collaboration, *The ATLAS experiment at the CERN Large Hadron Collider*, JINST **3** (2008)S08003, doi:10.1088/1748-0221/3/08/S08003.

[65] CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST **3** (2008)S08004, doi:10.1088/1748-0221/3/08/S08004.

[66] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, JINST **3** (2008)S08002, doi:10.1088/1748-0221/3/08/S08002.

[67] LHCb Collaboration, *The LHCb detector at the LHC*, JINST **3** (2008)S08005, doi:10.1088/1748-0221/3/08/S08005.

[68] LHCf Collaboration, *The LHCf detector at the CERN Large Hadron Collider*, JINST **3** (2008)S08006, doi:10.1088/1748-0221/3/08/S08006.

[69] TOTEM Collaboration, *The TOTEM experiment at the CERN Large Hadron Collider*, JINST **3** (2008)S08007, doi:10.1088/1748-0221/3/08/S08007.

[70] ATLAS Outreach, *ATLAS fact sheet : to raise awareness of the ATLAS detector and collaboration on the LHC*, 2010. http://cds.cern.ch/record/1457044.

[71] L. Lapikás, M. Vreeswijk, and A.-P. Colijn (eds.), *Annual report 2003*, NIKEF ,. http://www.nikhef.nl/fileadmin/Doc/Docs%20&%20pdf/Annual_Report-2003%20vk.pdf.

[72] F. Bergsma, *Calibration of Hall sensors in three dimensions*, in *13th International Magnetic Measurement Workshop*, Z. Wolf (ed.). SLAC, Stanford, CA, 2003, http://www.slac.stanford.edu/pubs/slacwps/slac-wp-029.html.

[73] H. Bichsel, D. E. Groom, and S. R. Klein, *Passage of particles through matter*, Sept 2013, in [8].

[74] Particle Data Group, *Particle detectors at accelerators*, 2013, in [8].

[75] ATLAS Collaboration, *ATLAS experiment - public results: trigger operation public results*, http://twiki.cern.ch/twiki/bin/view/AtlasPublic/TriggerOperationPublicResults. retrieved Oct 29 2014.

[76] ATLAS Collaboration, *Expected performance of the ATLAS experiment - detector, trigger and physics*, CERN (2008/2009), CERN-OPEN-2008-020, arXiv:0901.0512 [hep-ex].

[77] ATLAS Collaboration, *Performance of the ATLAS trigger system in 2010*, Eur. Phys. J. **C72** (2012)1849, doi:10.1140/epjc/s10052-011-1849-1, arXiv:1110.1530 [hep-ex].

[78] W. J. Stirling, *private communication*, http://www.hep.ph.ic.ac.uk/~wstirlin/plots/plots.html. retrieved Oct 29 2014.

[79] S. van der Meer, *Calibration of the effective beam height in the ISR*, CERN (1968), CERN-ISR-PO-68-31. http://cds.cern.ch/record/296752.

[80] ATLAS Collaboration, *Improved luminosity determination in pp collisions at* $\sqrt{s}$ *= 7 TeV using the ATLAS detector at the LHC*, Eur.Phys.J. **C73** (2013)2518, doi:10.1140/epjc/s10052-013-2518-3, arXiv:1302.4393 [hep-ex].

[81] M. Hostettler and G. Papotti, *Luminosity lifetime at the LHC in 2012 proton physics operation*, in Dai and Petit-Jean-Genaz [172], p. 1403. TUPFI029.

[82] M. Lamont, *The first years of LHC operation for luminosity production*, in Dai and Petit-Jean-Genaz [172], p. 6. MOYAB101.

[83] ATLAS Collaboration, *ATLAS experiment - public results: luminosity public results*, http://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults. retrieved Jan 12 2015.

[84] ATLAS Collaboration, *Performance of the ATLAS inner detector track and vertex reconstruction in the high pile-up LHC environment*, CERN (Mar 2012), ATLAS-CONF-2012-042. http://cds.cern.ch/record/1435196.

[85] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. **A262** (1987)444–450, doi:10.1016/0168-9002(87)90887-4.

[86] R. Frühwirth, *Track fitting with non-Gaussian noise*, Comput. Phys. Commun. **100** (1997)1–16, doi:10.1016/S0010-4655(96)00155-5.

[87] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data*, CERN (Jun 2014), ATLAS-CONF-2014-032. http://cds.cern.ch/record/1706245.

[88] W. Lampl, et al., *Calorimeter clustering algorithms: description and performance*, CERN (Apr 2008), ATL-LARG-PUB-2008-002. http://cds.cern.ch/record/1099735.

[89] ATLAS Collaboration, *Jet energy measurement and its systematic uncertainty in proton-proton collisions at* $\sqrt{s} = 7$ *TeV with the ATLAS detector*, Eur. Phys. J. **C75** (2015)17, doi:10.1140/epjc/s10052-014-3190-y, arXiv:1406.0076 [hep-ex].

[90] T. Barillari, et al., *Local hadronic calibration*, CERN (Jun 2008), ATL-LARG-PUB-2009-001-2. https://cds.cern.ch/record/1112035.

[91] ATLAS Collaboration, *Jet energy measurement with the ATLAS detector in proton-proton collisions at* $\sqrt{s} = 7$ *TeV*, Eur. Phys. J. **C73** (2013)2304, doi:10.1140/epjc/s10052-013-2304-2, arXiv:1112.6426 [hep-ex].

[92] ATLAS Collaboration, *Expected photon performance in the ATLAS experiment*, CERN (Apr 2011), ATL-PHYS-PUB-2011-007. http://cds.cern.ch/record/1345329.

[93] ATLAS Collaboration, *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data*, Eur. Phys. J. **C74** (2014)3130, doi:10.1140/epjc/s10052-014-3130-x, arXiv:1407.3935 [hep-ex].

[94] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-kt jet clustering algorithm*, JHEP **04** (2008)063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189 [hep-ph].

[95] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, Phys. Rev. **D48** (1993)3160–3166, doi:10.1103/PhysRevD.48.3160, arXiv:hep-ph/9305266 [hep-ph].

[96] ATLAS Collaboration, *Characterisation and mitigation of beam-induced backgrounds observed in the ATLAS detector during the 2011 proton-proton run*, JINST **8** (2013)P07004, doi:10.1088/1748-0221/8/07/P07004, arXiv:1303.0223 [hep-ex].

[97] R. Abbasi et al., *The design and performance of IceCube DeepCore*, Astroparticle Physics **35** (2012)615 – 624, doi:10.1016/j.astropartphys.2012.01.004.

[98] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction in proton-proton collisions at* $\sqrt{s} = 7$ *TeV with ATLAS*, Eur. Phys. J. **C72** (2012)1844, doi:10.1140/epjc/s10052-011-1844-6, arXiv:1108.5602 [hep-ex].

[99] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction in ATLAS studied in proton-proton collisions recorded in 2012 at 8 TeV*, CERN (Aug 2013), ATLAS-CONF-2013-082. http://cds.cern.ch/record/1570993.

[100] ATLAS Collaboration, *Performance of the ATLAS muon trigger in pp collisions at* $\sqrt{s}$ = 8 *TeV*, Eur. Phys. J. **C75** (2015)120, doi:10.1140/epjc/s10052-015-3325-9, arXiv:1408.3179 [hep-ex].

[101] ATLAS Collaboration, *Performance of the electron and photon trigger in p-p collisions at* $\sqrt{s}$ = 7 *TeV with the ATLAS detector at the LHC*, CERN (Aug 2011), ATLAS-CONF-2011-114. http://cds.cern.ch/record/1375551.

[102] ATLAS Collaboration, *Performance of the ATLAS electron and photon trigger in p-p collisions at* $\sqrt{s}$ = 7 *TeV in 2011*, CERN (May 2012), ATLAS-CONF-2012-048. http://cds.cern.ch/record/1450089.

[103] A. Ezhilov, on behalf of the ATLAS Collaboration, *ATLAS electron and photon trigger performance in Run 1 and developments toward Run 2*, CERN (Jun 2014), ATL-DAQ-SLIDE-2014-371. http://cds.cern.ch/record/1712669.

[104] ATLAS Collaboration, *Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp collisions at* $\sqrt{s}$=8 *TeV*, CERN (2014), CERN-PH-EP-2014-227, arXiv:1412.7086 [hep-ex]. Submitted to Eur. Phys. J. C.

[105] M. L. Perl, et al., *Evidence for anomalous lepton production in* $e^+ - e^-$ *annihilation*, Phys. Rev. Lett. **35**, 22 (1975) 1489–1492, doi:10.1103/PhysRevLett.35.1489. http://link.aps.org/doi/10.1103/PhysRevLett.35.1489.

[106] ATLAS Collaboration, *Reconstruction, energy calibration, and identification of hadronically decaying tau leptons*, CERN (May 2011), ATLAS-CONF-2011-077. http://cds.cern.ch/record/1353226.

[107] A. Höcker, et al., *TMVA - Toolkit for Multivariate Data Analysis with ROOT: users guide*, CERN (Mar 2007),. http://tmva.sourceforge.net/.

[108] ATLAS Collaboration, *Determination of the tau energy scale and the associated systematic uncertainty in proton-proton collisions at* $\sqrt{s}$ = 7 *TeV with the ATLAS detector at the LHC in 2011*, CERN (Jun 2012), ATLAS-CONF-2012-054. http://cds.cern.ch/record/1453781.

[109] ATLAS Collaboration, *Performance of the reconstruction and identification of hadronic* τ *decays in ATLAS with 2011 Data*, CERN (Oct 2012), ATLAS-CONF-2012-142. http://cds.cern.ch/record/1485531.

[110] ATLAS Collaboration, *Performance of the reconstruction and identification of hadronic tau decays with ATLAS*, CERN (Nov 2011), ATLAS-CONF-2011-152. http://cds.cern.ch/record/1398195.

[111] Y. Freund and R. E. Schapire, *A decision-theoretic generalization of on-line learning and an application to boosting*, JCSS **55** (1997)119 – 139, doi:10.1006/jcss.1997.1504.

[112] R. E. Schapire and Y. Singer, *Improved boosting algorithms using confidence-rated predictions*, Machine Learning **37** (1999)297–336, doi:10.1023/A:1007614523901.

[113] M. Dam, 2011. private communication.

[114] ATLAS Collaboration, *Identification of the hadronic decays of tau leptons in 2012 data with the ATLAS detector*, CERN (Jul 2013), ATLAS-CONF-2013-064. http://cds.cern.ch/record/1562839.

[115] M. Trottier-McDonald, *Identifying hadronic tau decays using calorimeter topological clusters at ATLAS*, Master's thesis, Simon Fraser U., Burnaby, British Columbia, Canada, 2010. http://hep.phys.sfu.ca/theses/MTMcDonald_msc.pdf.

[116] M. Trottier-McDonald, *Evidence for the Standard Model Higgs boson decaying to one semi-hadronically decaying tau and one leptonically decaying tau at ATLAS*. PhD thesis, Simon Fraser U., Burnaby, British Columbia, Canada, Jun 2014. https://cds.cern.ch/record/1709068. Presented 29 05 2014.

[117] ATLAS Collaboration, *Determination of the tau energy scale and the associated systematic uncertainty in proton-proton collisions at* $\sqrt{s} = 8$ *TeV with the ATLAS detector at the LHC in 2012*, CERN (Apr 2013), ATLAS-CONF-2013-044. http://cds.cern.ch/record/1544036.

[118] C. Limbach, *Reconstruction and identification of tau leptons in ATLAS*, CERN (2014), ATL-PHYS-SLIDE-2014-680. http://cds.cern.ch/record/1951814.

[119] B. T. Winter, *Reconstruction of neutral pions in hadronic tau lepton decays in the ATLAS detector*, Master's thesis, Physikalisches Institut der Universität Bonn, Dec 2013. http://hep1.physik.uni-bonn.de/fileadmin/Publications/ATLAS_Analysis/Dipl/winter.pdf.

[120] L. A. Thomsen, *A search for associated production of a SM Higgs decaying into tau leptons with the ATLAS experiment*. PhD thesis, University of Copenhagen, Copenhagen, Denmark, Aug 2014. http://discoverycenter.nbi.ku.dk/teaching/thesis_page/LotteAThomsen_Thesis.pdf.

[121] R. Assmann, M. Lamont, and S. Myers, *A brief history of the LEP collider*, Nucl. Phys. Proc. Suppl. **109B** (2002)17–31, doi:10.1016/S0920-5632(02)90005-8.

[122] Fermi National Accelerator Laboratory. www.fnal.gov.

[123] ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration, The LEP Working Group for Higgs Boson Searches, *Search for the standard model Higgs boson at LEP*, Phys. Lett. **B565** (2003)61–75, doi:10.1016/S0370-2693(03)00614-2, arXiv:hep-ex/0306033 [hep-ex].

[124] G. Bernardi, M. Carena, and T. Junk, *Higgs boson: theory and searches*, Nov 2007, http://pdg.lbl.gov/2009/pdg_2009.html.

[125] ATLAS Collaboration, *Combined search for the Standard Model Higgs boson using up to 4.9* fb^{-1} *of pp collision data at* $\sqrt{s} = 7$ *TeV with the ATLAS detector at the LHC*, Phys. Lett. **B710** (2012)49–66, doi:10.1016/j.physletb.2012.02.044, arXiv:1202.1408 [hep-ex].

[126] CMS Collaboration, *Combined results of searches for the standard model Higgs boson in pp collisions at* $\sqrt{s} = 7$ *TeV*, Phys. Lett. **B710** (2012)26–48, doi:10.1016/j.physletb.2012.02.064, arXiv:1202.1488 [hep-ex].

[127] CERN, CMS Collaboration, ATLAS Collaboration, *Latest update in the search for the Higgs boson*, http://indico.cern.ch/event/197461/. retrieved Jan 2 2015.

[128] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012)1–29, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214 [hep-ex].

[129] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012)30–61, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235 [hep-ex].

[130] LHC Higgs Cross Section Working Group, *LHC Higgs cross section working group (2012-2013)*, http://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections. retrieved Jan 2 2015.

[131] CMS Collaboration, *Search for a standard model-like Higgs boson in the $\mu^+\mu^-$ and e^+e^- decay channels at the LHC*, 2014, CMS-HIG-13-007, CERN-PH-EP-2014-243, arXiv:1410.6679 [hep-ex].

[132] ATLAS Collaboration, *Search for the Standard Model Higgs boson decay to $\mu^+\mu^-$ with the ATLAS detector*, Phys. Lett. **B738** (2014)68–86, doi:10.1016/j.physletb.2014.09.008, arXiv:1406.7663 [hep-ex].

[133] CMS Collaboration, *Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV*, submitted to Eur. Phys. J. **C** (2014), arXiv:1412.8662 [hep-ex].

[134] ATLAS Collaboration, *Updated coupling measurements of the Higgs boson with the ATLAS detector using up to 25 fb^{-1} of proton-proton collision data*, CERN (Mar 2014), ATLAS-CONF-2014-009. http://cds.cern.ch/record/1670012.

[135] ATLAS Collaboration, *Measurement of Higgs boson production in the diphoton decay channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector*, Phys. Rev. **D90** (2014)112015, doi:10.1103/PhysRevD.90.112015, arXiv:1408.7084 [hep-ex].

[136] ATLAS Collaboration, *Measurements of Higgs boson production and couplings in the four-lepton channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector*, 2014, CERN-PH-EP-2014-170, arXiv:1408.5191 [hep-ex].

[137] ATLAS Collaboration, *Measurement of Higgs boson production in the diphoton decay channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector*, CERN (Aug 2014), CERN-PH-EP-2014-198, arXiv:1408.7084.

[138] ATLAS Collaboration, *Evidence for Higgs boson Yukawa couplings in the $H \to \tau\tau$ decay mode with the ATLAS detector*, CERN (Oct 2014), ATLAS-CONF-2014-061. http://cds.cern.ch/record/1954724.

[139] ATLAS Collaboration, *Search for the $b\bar{b}$ decay of the Standard Model Higgs boson in associated $(W/Z)H$ production with the ATLAS detector*, CERN (2014), CERN-PH-EP-2014-214, arXiv:1409.6212 [hep-ex].

[140] ATLAS Collaboration, *Measurement of the Higgs boson mass from the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector*, Phys. Rev. **D90** (2014)052004, doi:10.1103/PhysRevD.90.052004, arXiv:1406.3827 [hep-ex].

[141] CMS Collaboration, *Constraints on the spin-parity and anomalous HVV couplings of the Higgs boson in proton collisions at 7 and 8 TeV*, submitted to Phys. Rev. **D** (2014), arXiv:1411.3441 [hep-ex].

[142] ATLAS Collaboration, *Evidence for the spin-0 nature of the Higgs boson using ATLAS data*, Phys. Lett. **B726** (2013)120–144, doi:10.1016/j.physletb.2013.08.026, arXiv:1307.1432 [hep-ex].

[143] The TEVNPH Working Group Collaboration, *Combined CDF and D0 constraints on models for the Higgs boson with exotic spin and parity*, Aug 2014, FERMILAB-CONF-14-265-E.

[144] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs cross sections: 1. Inclusive observables*, 2011, CERN-2011-002, arXiv:1101.0593 [hep-ph].

[145] CMS Collaboration, *Constraints on the Higgs boson width from off-shell production and decay to Z-boson pairs*, Phys. Lett. **B736** (2014)64–85, doi:10.1016/j.physletb.2014.06.077, arXiv:1405.3455 [hep-ex].

[146] ATLAS Collaboration, *Determination of the off-shell Higgs boson signal strength in the high-mass ZZ final state with the ATLAS detector*, CERN (Jul 2014), ATLAS-CONF-2014-042. http://cds.cern.ch/record/1740973.

[147] CDF Collaboration, D0 Collaboration, *Higgs boson studies at the Tevatron*, Phys. Rev. **D88** (2013)052014, doi:10.1103/PhysRevD.88.052014, arXiv:1303.6346 [hep-ex].

[148] CDF Collaboration, D0 Collaboration, *Evidence for a particle produced in association with weak bosons and decaying to a bottom-antibottom quark pair in Higgs boson searches at the Tevatron*, Phys. Rev. Lett. **109** (2012)071804, doi:10.1103/PhysRevLett.109.071804, arXiv:1207.6436 [hep-ex].

[149] CMS Collaboration, *Evidence for the direct decay of the 125 GeV Higgs boson to fermions*, Nature Phys. **10** (2014)557–560, doi:10.1038/nphys3005, arXiv:1401.6527 [hep-ex].

[150] CMS Collaboration, *Evidence for the 125 GeV Higgs boson decaying to a pair of τ leptons*, JHEP **05** (2014)104, doi:10.1007/JHEP05(2014)104, arXiv:1401.5041 [hep-ex].

[151] E. A. Ideal, *A search for the Standard Model Higgs Boson produced in association with a vector boson and decaying to a hadronically-decaying tau pair at ATLAS*. PhD thesis, Yale University, 2015.

[152] A. Elagin, P. Murat, A. Pranko, and A. Safonov, *A new mass reconstruction technique for resonances decaying to di-tau*, Nucl. Instrum. Meth. **A654** (2011)481–489, doi:10.1016/j.nima.2011.07.009, arXiv:1012.4686 [hep-ex].

[153] ATLAS Collaboration, *ATLAS experiment - public results: data quality information*, https://twiki.cern.ch/twiki/bin/view/AtlasPublic/RunStatsPublicResults2010. retrieved Apr 15 2015.

[154] J. M. Campbell, R. K. Ellis, and C. Williams, *Vector boson pair production at the LHC*, JHEP **1107** (2011)018, doi:10.1007/JHEP07(2011)018, arXiv:1105.0020 [hep-ph].

[155] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **07** (2003)001, arXiv:hep-ph/0206293.

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

[157] P. Z. Skands, *Tuning Monte Carlo generators: the Perugia tunes*, Phys.Rev. **D82** (2010)074018, doi:10.1103/PhysRevD.82.074018, arXiv:1005.3457 [hep-ph].

[158] J. Pumplin, et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **0207** (2002)012, doi:10.1088/1126-6708/2002/07/012, arXiv:hep-ph/0201195 [hep-ph].

[159] G. Corcella, et al., *HERWIG 6: an event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, JHEP **0101** (2001)010, doi:10.1088/1126-6708/2001/01/010, arXiv:hep-ph/0011363 [hep-ph].

[160] J. Butterworth, J. R. Forshaw, and M. Seymour, *Multiparton interactions in photoproduction at HERA*, Z. Phys. **C72** (1996)637–646, doi:10.1007/s002880050286, arXiv:hep-ph/9601371 [hep-ph].

[161] J. Alwall, et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **1407** (2014)079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301 [hep-ph].

[162] ATLAS Collaboration, *ATLAS tunes of PYTHIA 6 and Pythia 8 for MC11*, CERN (Jul 2011), ATL-PHYS-PUB-2011-009. http://cds.cern.ch/record/1363300.

[163] H.-L. Lai, et al., *New parton distributions for collider physics*, Phys. Rev. **D82** (2010)074024, doi:10.1103/PhysRevD.82.074024, arXiv:1007.2241 [hep-ph].

[164] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **0206** (2002)029, doi:10.1088/1126-6708/2002/06/029, arXiv:hep-ph/0204244 [hep-ph].

[165] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **0411** (2004)040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146 [hep-ph].

[166] S. Jadach, Z. Was, R. Decker, and J. H. Kühn, *The tau decay library TAUOLA: version 2.4*, Comput. Phys. Commun. **76** (1993)361–380, doi:10.1016/0010-4655(93)90061-G.

[167] ATLAS Collaboration, *Measurement of the production cross section of jets in association with a Z boson in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector*, JHEP **1307** (2013)032, doi:10.1007/JHEP07(2013)032, arXiv:1304.7098 [hep-ex].

[168] A. Q. T. M. Liss, F. Maltoni (ed.), *The top quark*, Sept 2013, in [8].

[169] A. L. Read, *Presentation of search results: the CL(s) technique*, J. Phys. **G28** (2002)2693–2704, doi:10.1088/0954-3899/28/10/313.

[170] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. **C71** (Jun 2013)1554, doi:10.1140/epjc/s10052-013-2501-z, arXiv:1007.1727 [physics.data-an].

[171] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs cross sections: 3. Higgs properties*, 2013, CERN-2013-004, arXiv:1307.1347 [hep-ph].

[172] Z. Dai and Petit-Jean-Genaz (eds.), *Proceedings of 4th International Particle Accelerator Conference (IPAC 2013)*, May 13-17, 2013, Shanghai, China, . JACoW, 2013. http://accelconf.web.cern.ch/AccelConf/IPAC2013/index.htm.

List of Figures

List of Tables

Acronyms

AD Antiproton Decelerator.

ALFA Absolute Luminosity For ATLAS.

ALICE A Large Ion Collider Experiment.

ATLAS formerly for A Toroidal LHC ApparatuS, now proper name.

BCID Bunch Crossing Identifier.

BCM Beam Conditions Monitor.

BDT Boosted Decision Tree.

BEE Extra module of MDTs, in the transition region of barrel and end-cap.

BIL Inner layer of large barrel MDTs.

BIM Special inner MDT module, due to the support structure rails.

BIR Special inner MDT module, due to the support structure rails.

BIS Inner layer of small barrel MDTs.

BMF Modified BMS, due to support structure feet.

BML Middle layer of large barrel MDTs.

BMS Middle layer of small barrel MDTs.

BOF Modified BOS, due to support structure feet.

BOG Modified BOS, due to support structure feet.

BOL Outer layer of large barrel MDTs.

BOS Outer layer of small barrel MDTs.

CLIC Compact Linear Collider.

CMS Compact Muon Solenoid.

CNGS CERN Neutrinos to Gran Sasso.

CSC Cathode Strip Chamber.

CTF3 CLIC test facility.

ECAL Electromagnetic calorimeter.

EEL Extra module of large end-cap MDTs, placed between inner and middle layer.

EES Extra module of small end-cap MDTs, placed between inner and middle layer.

EF Event Filter.

EIL Inner layer of large end-cap MDTs.

EIS Inner layer of small end-cap MDTs.

EMEC Electromagnetic end-cap.

EML Middle layer of large end-cap MDTs.

EMS Middle layer of small end-cap MDTs.

EOL Outer layer of large end-cap MDTs.

EOS Outer layer of small end-cap MDTs.

FCal Forward calorimeter.

FSR Final State Radiation.

GRL Good Runs List.

HCAL Hadronic calorimeter.

HEC Hadronic end-cap.

ID Inner Detector.

ILC International Linear Collider.

IP Interaction point.

ISOLDE Isotope mass Separator On-Line facility.

ISR Initial State Radiation.

L1 Level 1.

L2 Level 2.

LAr Liquid Argon.

LB Luminosity Block.

LEIR Low Energy Ion Ring.

LHC Large Hadron Collider.

LHCb Large Hadron Collider beauty.

LHCf Large Hadron Collider forward.

LLH projective Likelihood.

LUCID LUminosity measurement using Cerenkov Integrating Detector.

MC Monte Carlo, refers to simulated samples.

MDT Monitored Drift Tube.

n-TOF neutron time-of-flight facility.

NMR Nuclear magnetic resonance.

OS Opposite sign, meaning opposite electric charge, $q_1 + q_2 = 0$.

PDF Parton Distribution Function.

pdf probability density functions.

PIX Pixel Detector.

PS Proton Sychrotron.

PSB Proton Synchrotron Booster.

QCD Quantum Chromo Dynamics.

QED Quantum Electro Dynamics.

QFT Quantum Field Theory.

RF Radio frequency.

RoI Region of Interest.

RPC Resistive Plate Chamber.

SCT Semiconductor Tracker.

SF Scale Factor.

SLAC Stanford Linear Accelerator Center, nowadays SLAC National Accelerator Laboratory.

SM Standard Model.

SPS Super Proton Synchrotron.

SS Same sign, meaning same electric charge, $q_1 - q_2 = 0$.

STVF Soft Term Vertex Fraction.

TES Tau Energy Scale.

TGC Thin Gap Chamber.

TOTEM TOTal cross section, Elastic scattering and diffraction dissociation Measurement at the LHC.

TRT Transition Radiation Tracker.

WP Working Point.

ZDC Zero-Degree Calorimeter.

Appendices

A. Discriminating Variables

The discriminating variables for the separation of hadronically decaying tau leptons from jets are defined as follows. The *jet seed axis* is the default axis in the 2011 tau reconstruction. It is given by the barycenter of all clusters that are associated to the tau candidate. The default axis in the 2012 tau reconstruction is the *intermediate axis*. It is calculated from the clusters that are within $\Delta R < 0.2$ around the jet seed axis. Their position is first recalculated with respect to the tau vertex (vertex with the highest jet vertex fraction). Then, the cluster sum is recalculated to define the intermediate axis.

The definitions are followed by the distributions of all variables, separately for 1-prong and 3-prong candidates. The plots illustrate the definition as used in the 2012 data taking period.

Track radius (R_{track}): p_{T}-weighted track width:

$$R_{\text{track}} = \frac{\sum_{i\in\{\text{core,iso}\}}^{\Delta R_i \leq 0.4} p_{\text{T},i}\, \Delta R_i}{\sum_{i\in\{\text{core,iso}\}}^{\Delta R_i \leq 0.4} p_{\text{T},i}},$$

where i runs over all core and isolation tracks of the τ_h candidate, within $\Delta R_i \leq 0.4$. ΔR_i is defined relative to the τ_h intermediate axis and $p_{\text{T},i}$ is the track transverse momentum.

Note that for candidates with only one track total in the core cone and isolation annulus, R_{track} simplifies to the ΔR between the track and the intermediate axis (2011: jet seed axis).

Number of tracks in isolation annulus ($N^{\text{iso}}_{\text{track}}$): number of tracks reconstructed within $0.2 < \Delta R \leq 0.4$ around the intermediate axis (2011: jet seed axis).

Leading track IP significance (S_{leadTrk}): impact parameter significance of the leading track in the core region.

$$S_{\text{leadTrk}} = \frac{d_0}{\delta d_0},$$

where d_0 is the distance of closest approach of the track to the tau vertex (2011: first primary vertex) in the transverse plane, and δd_0 is its estimated uncertainty.

Transverse flight path significance ($S_{\text{T}}^{\text{flight}}$): the decay length significance of the secondary vertex for multi-track τ_h candidates in the transverse plane:

$$S_{\text{T}}^{\text{flight}} = \frac{L_{\text{T}}^{\text{flight}}}{\delta L_{\text{T}}^{\text{flight}}},$$

where $L_{\text{T}}^{\text{flight}}$ is the reconstructed signed decay length, and $\delta L_{\text{T}}^{\text{flight}}$ is its estimated uncertainty. Only core tracks are used for the secondary vertex fit.

Maximum ΔR (ΔR_{max}): the maximal ΔR between a track associated to the τ_h candidate and the intermediate axis (2011: jet seed axis). Only core tracks are considered.

Track mass (m_{tracks}): invariant mass of the track system. Both core and isolation tracks are used for calculation of the invariant mass.

Leading track momentum fraction (f_{track}):

$$f_{\text{track}} = \frac{p_{\text{T}}^{\text{leadtrk}}}{\sum_{j\in\{\text{all}\}}^{\Delta R_j<0.2} E_{\text{T},j}^{\text{EM}}},$$

where $p_{\text{T}}^{\text{leadtrk}}$ is the transverse momentum of the leading p_{T} core track of the τ_h candidate. $E_{\text{T},j}^{\text{EM}}$ is the transverse energy, calibrated at the EM energy scale, deposited in cell j, and j runs over all cells of the calorimeter (Ecal + Hcal) in $\Delta R < 0.2$ (2011: $\Delta R < 0.4$) around the τ_h intermediate axis (2011: jet seed axis).

Note that for candidates with one track, f_{track} is the fraction of the candidate's momentum attributed to the track, compared to the total momentum of the candidate, which can have contributions from the calorimeter deposits from π^0s and other neutrals.

Pile-up-corrected leading track momentum fraction ($f_{\text{track}}^{\text{corr}}$):

$$f_{\text{track}}^{\text{corr}} = f_{\text{track}} + 0.003 * N_{\text{vtx}}$$

where N_{vtx} is the number of good vertices in the event, defined as the number of pile-up vertices with at least 2 tracks plus the primary vertex, which is required to have at least 4 tracks.

Central energy fraction (f_{cent}): fraction of transverse energy in the central region ($\Delta R < 0.1$) of the τ_h candidate:

$$f_{\text{cent}} = \frac{\sum_{i\in\{\text{all}\}}^{\Delta R_i<0.1} E_{\text{T},i}^{\text{EM}}}{\sum_{j\in\{\text{all}\}}^{\Delta R_j<0.2} E_{\text{T},j}^{\text{EM}}},$$

where $E_{\text{T},i}^{\text{EM}}$ ($E_{\text{T},j}^{\text{EM}}$) is the transverse energy, calibrated at the EM energy scale, deposited in cell i (j), and i runs over the cells in all layers associated with the τ_h candidate within $\Delta R < 0.1$ of the intermediate axis (2011: jet seed axis), while j runs over all cells in all layers within $\Delta R < 0.2$ (2011: $\Delta R < 0.4$).

Pile-up-corrected central energy fraction ($f_{\text{cent}}^{\text{corr}}$): Pile-up-corrected fraction of transverse energy in the central region ($\Delta R < 0.1$) of the τ_h candidate:

$$\begin{aligned} f_{\text{cent}}^{\text{corr}} &= f_{\text{cent}} + 0.003 * N_{\text{vtx}} && \text{for } p_{\text{T}} < 80\,\text{GeV} \\ f_{\text{cent}}^{\text{corr}} &= f_{\text{cent}} && \text{else} \end{aligned}$$

where N_{vtx} is the number of good vertices in the event, defined as the number of pile-up vertices with at least 2 tracks plus the primary vertex, which is required to have at least 4 tracks. p_{T} is the momentum of the τ_h candidate, calibrated at the tau energy scale [117].

Calorimeter radius (R_{cal}): transverse energy weighted shower width in the calorimeter:

$$R_{\text{cal}} = \frac{\sum_{i\in\{\text{all}\}}^{\Delta R_i<0.2} E_{\text{T},i}^{\text{EM}}\,\Delta R_i}{\sum_{i\in\{\text{all}\}}^{\Delta R_i<0.2} E_{\text{T},i}^{\text{EM}}},$$

where $E_{\text{T},i}^{\text{EM}}$ is the transverse energy, calibrated at the EM energy scale, deposited in cell i, j. The cells are within $\Delta R < 0.2$ (2011: $\Delta R < 0.4$) of any layer of the electromagnetic and hadronic calorimeter. ΔR_i is defined relative to the τ_h intermediate axis (2011: $\Delta R < 0.4$).

First two (three) leading cluster energy ratio ($f_{\text{2leadClus}}$ ($f_{\text{3leadClus}}$)): Ratio of the energy of the two (three) clusters with the highest energy over the total energy of all clusters associated to the τ_h candidate. The cluster energy is calibrated at LC scale.

Cluster mass (m_{clus}): invariant mass of the topoclusters associated to the τ_h candidate. All clusters are considered.

Effective cluster mass (m_{effClus}): invariant mass of the topoclusters with dominant contribution to the energy of the τ_h candidate. The dominatant clusters are selected based on the number of effective clusters N_{eff}. N_{eff} is equal to the total number of clusters if all clusters contribute equally to the total energy. It approaches 1, if the energy is concentrated in one cluster. N_{eff} is defined as

$$N_{\text{eff}} = \frac{\left(\sum_{i\in\{\text{all}\}} E_i^{\text{LC}}\right)^2}{\sum_{i\in\{\text{all}\}} \left(E_i^{\text{LC}}\right)^2}.$$

The number is rouned to its nearest larger integer, so that at least two clusters contribute to the calculation of m_{effClus} if more than one clusters is present. The cluster energy is calibrated at LC scale.

Cluster isolation energy ($E_{\text{T}}^{\text{iso}}$): the transverse energy of the isolation clusters:

$$E_{\text{T}}^{\text{iso}} = \sum_{i\in\{\text{all}\}}^{0.2\leq\Delta R<0.4} E_{\text{T},i}^{\text{EM}},$$

where l runs over all clusters in the isolation annulus $0.2 \leq \Delta R < 0.4$ and ΔR is defined with respect to the intermediate axis (2011: jet seed axis).

Corrected cluster isolation energy ($E_{\text{T,corr}}^{\text{iso}}$): the pile-up corrected transverse energy of isolated clusters:

$$E_{\text{T,corr}}^{\text{iso}} = E_{\text{T}}^{\text{iso}} - \min\left((1 - f_{\text{JVF}}) \sum_{i\in\{\text{jet}\}}^{\Delta R<0.2} p_{\text{T},i}^{\text{trk}} \;\;,\;\; 4\,\text{GeV}\right),$$

where f_{JVF} is the jet vertex fraction of the jet seed of the tau candidate calculated with respect to the primary vertex, and the sum runs over the transverse momenta of the tracks associated to that jet.

Number of neutral pions (N_{π^0}): output of the π^0 counting algorithm, which estimates the number of neutral pions based on two BDTs and the variables f_{PSS}, $f_{\text{Ecal}}^{\pi^\pm}$, $E_{\text{Ecal}}/p_{\text{tracks}}$, N_{strip}, and $p_{\text{T}}^{\text{leadtrk}}$ (see also Sec. 5.2). The number is limited to $\pi^0 = \{0, 1, 2\}$.

Invariant mass of the π^0-and-track-system ($m_{\pi^0+\text{tracks}}$): the π^0-and-track-system is given by the four-vector sum of the reconstructed π^0 cluster and the core tracks, assigning masses of 135 MeV and 140 MeV to the cluster and tracks, respectively. The π^0 cluster is given by the π^0 Finder, see Sec. 5.2. If the Finder returns a cluster-pair, the two (massless) clusters are combined first.

Ratio of tau momenta ($p_{\text{T}}^{\pi^0+\text{tracks}}/p_{\text{T}}$): the ratio of the momentum of the π^0-and-track-system and the momentum of the calorimeter-only measurement. The momentum of the π^0-and-track-system (numerator) is given by the output of the π^0 Finder and the core tracks. The momentum of the calorimeter-only measurement (denominator) is the fully calibrated momentum of the τ_h candidate, based on the topoclusters within $\Delta R < 0.2$ around the tau detector axis. See Sec. 5.2 for more information.

Presampler and strip layer energy fraction (f_{PSS}):

$$f_{\mathrm{PSS}} = \frac{\sum_{i\in\{\mathrm{all}\}} E_i^{\mathrm{PSS}}}{\sum_{i\in\{\mathrm{all}\}} E_i},$$

where i runs over all calorimeter clusters associated to the τ_h candidate, E_i^{PSS} denotes the part of cluster energy that is deposited in the presampler and strip layer (ECAL layer 1) and E_i is the total energy of a calorimeter cluster. The cluster energies are calibrated at the LC scale.

Estimated electromagnetic energy of charged pions over electromagnetic energy ($f_{\mathrm{Ecal}}^{\pi^\pm}$):

$$f_{\mathrm{Ecal}}^{\pi^\pm} = \frac{\sum_{i\in\{\mathrm{core}\}} p_i^{\mathrm{trk}} - \sum_{j\in\{\mathrm{all}\}} E_j^{\mathrm{Hcal}}}{\sum_{j\in\{\mathrm{all}\}} E_j^{\mathrm{Ecal}}}.$$

where j runs over all calorimeter clusters associated to τ_h candidate, E_j^{Hcal} denotes the part of cluster energy deposited in the hadronic calorimeter including the third layer of the electromagnetic calorimeter, E_j^{Ecal} is the part of cluster energy deposited in the electromagnetic calorimeter (presampler and first two layers) and i runs over the core tracks. All clusters are calibrated at the LC energy scale.

Electromagnetic energy over track momentum ($E_{\mathrm{Ecal}}/p_{\mathrm{tracks}}$):

$$E_{\mathrm{Ecal}}/p_{\mathrm{tracks}} = \frac{\sum_{i\in\{\mathrm{all}\}} E_i^{\mathrm{Ecal}}}{\sum_{j\in\{\mathrm{core}\}} p_j^{\mathrm{trk}}},$$

where i runs over all calorimeter clusters and E_i^{Ecal} is the part of cluster energy deposited in the electromagnetic calorimeter (presampler and first two layers), calibrated at LC energy scale. The index j runs over the core tracks.

Number of strip cells (N_{strip}): Number of cells in the strip layer (ECAL layer 1) with an energy of at least 200 MeV and within $\Delta R < 0.2$ (2011: $\Delta R < 0.4$) of the intermediate axis (2011: jet seed axis).

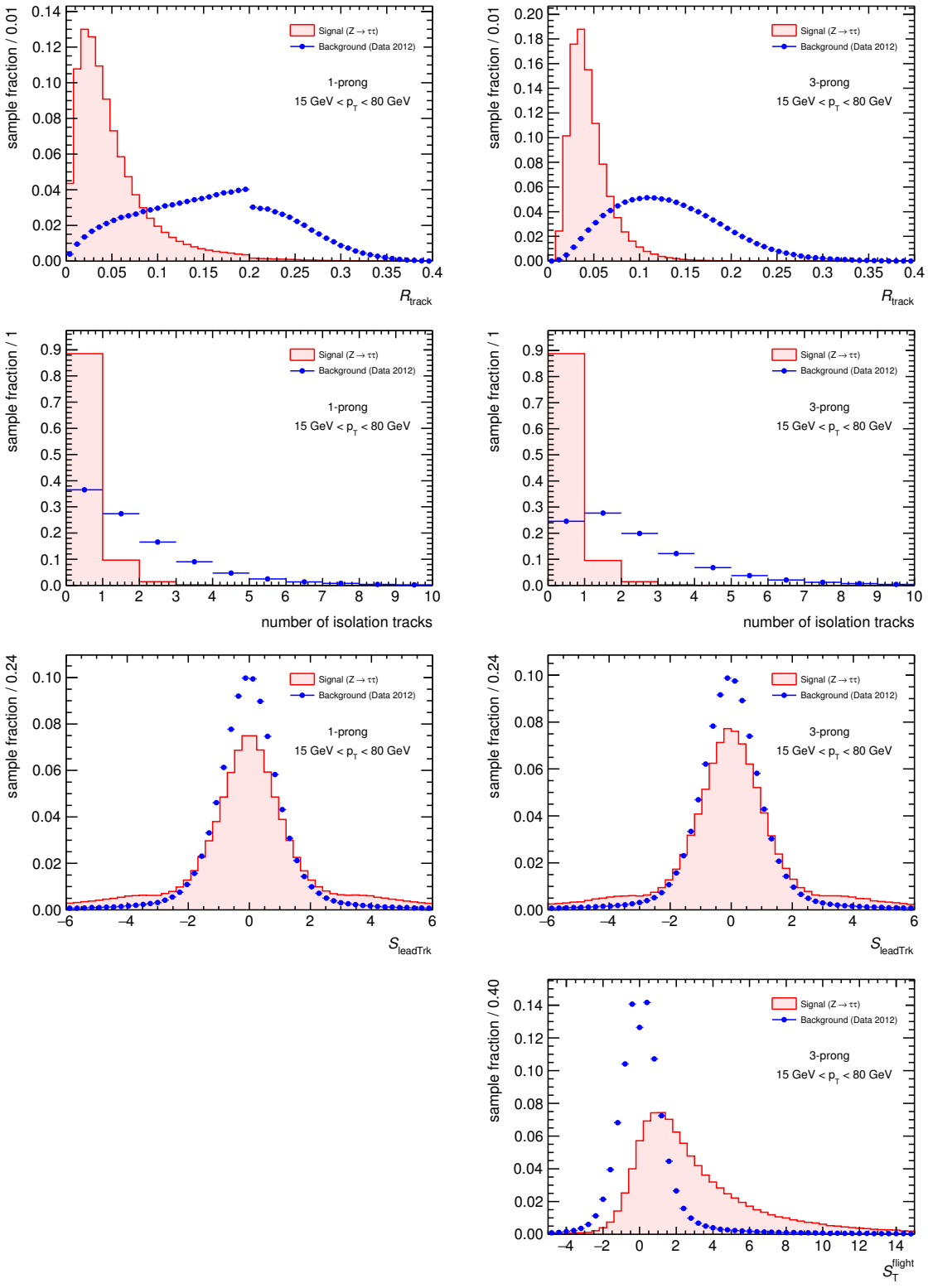


Figure A.1: Discriminating variables for 1-prong (left) and 3-prong (right) tau candidates. The variable S_{leadTrk} is not defined for 1-prong candidates. The signal candidates are from simulated $Z \to \tau\tau$ events matched to a true hadronic tau decay. The background candidates are from multi-jet events in 2012 data. All distributions are normalised to unity.

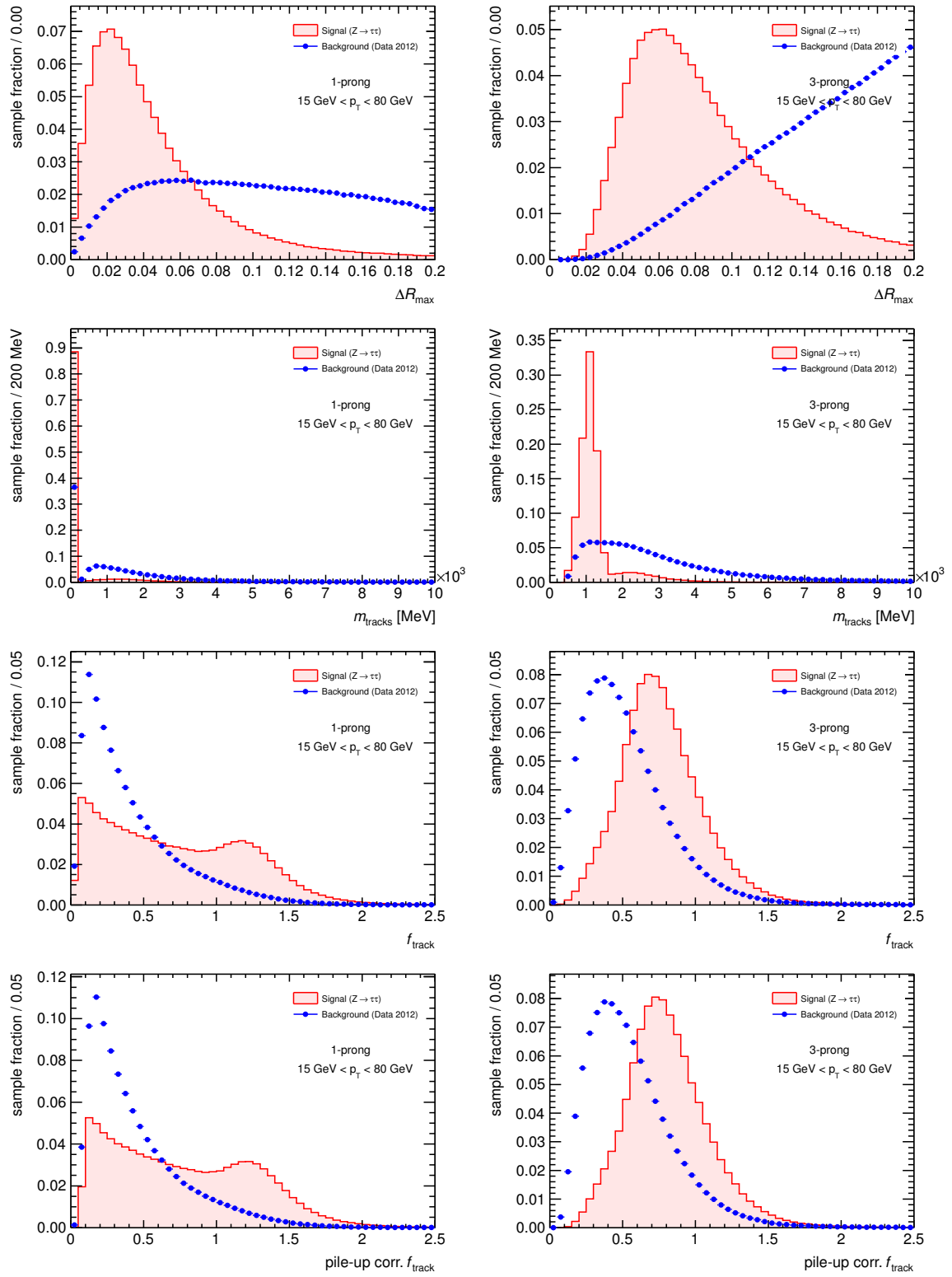


Figure A.2: Discriminating variables for 1-prong (left) and 3-prong (right) tau candidates. The signal candidates are from simulated $Z \rightarrow \tau\tau$ events matched to a true hadronic tau decay. The background candidates are from multi-jet events in 2012 data. All distributions are normalised to unity.

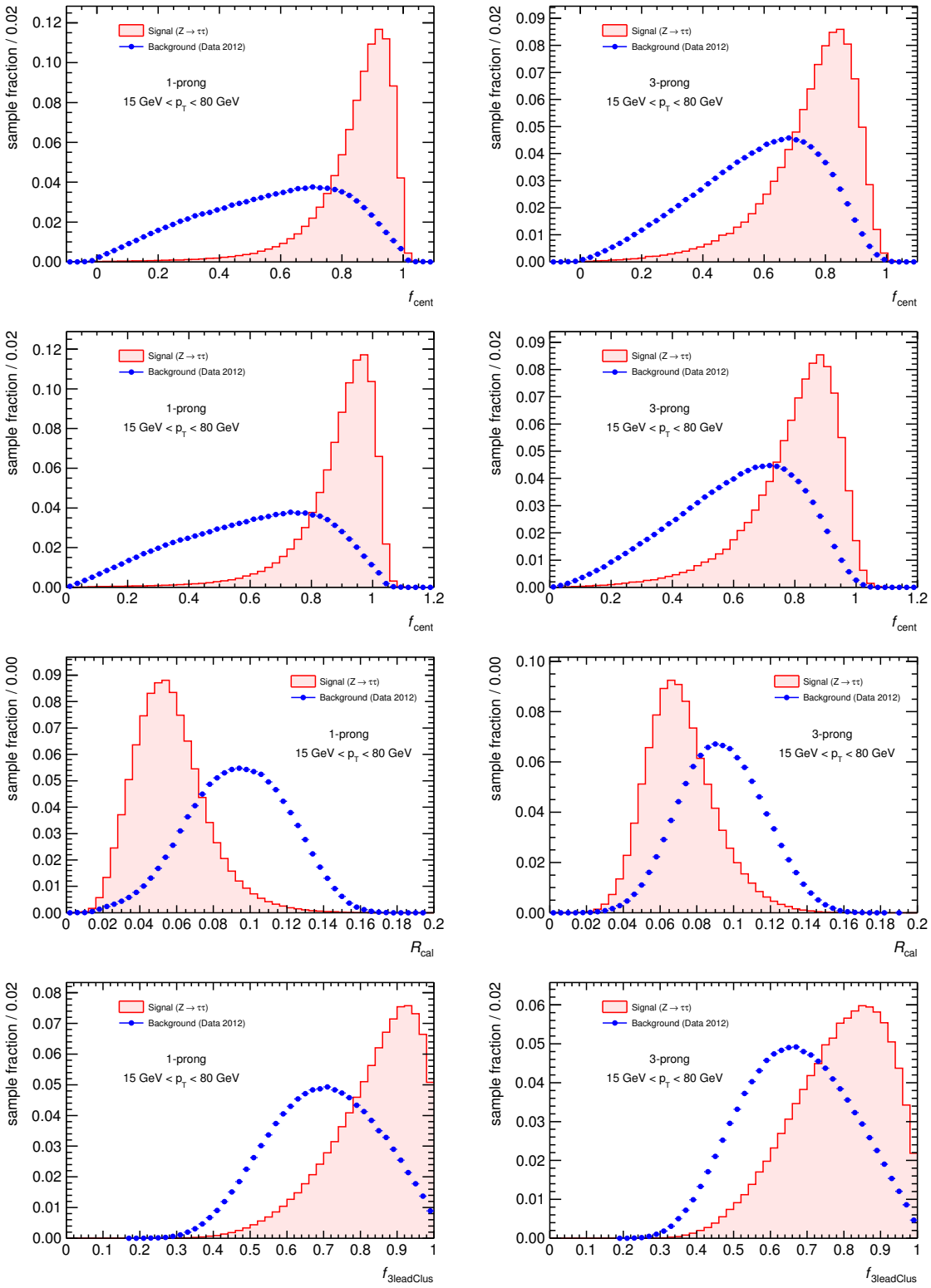


Figure A.3: Discriminating variables for 1-prong (left) and 3-prong (right) tau candidates. The signal candidates are from simulated $Z \rightarrow \tau\tau$ events matched to a true hadronic tau decay. The background candidates are from multi-jet events in 2012 data. All distributions are normalised to unity.

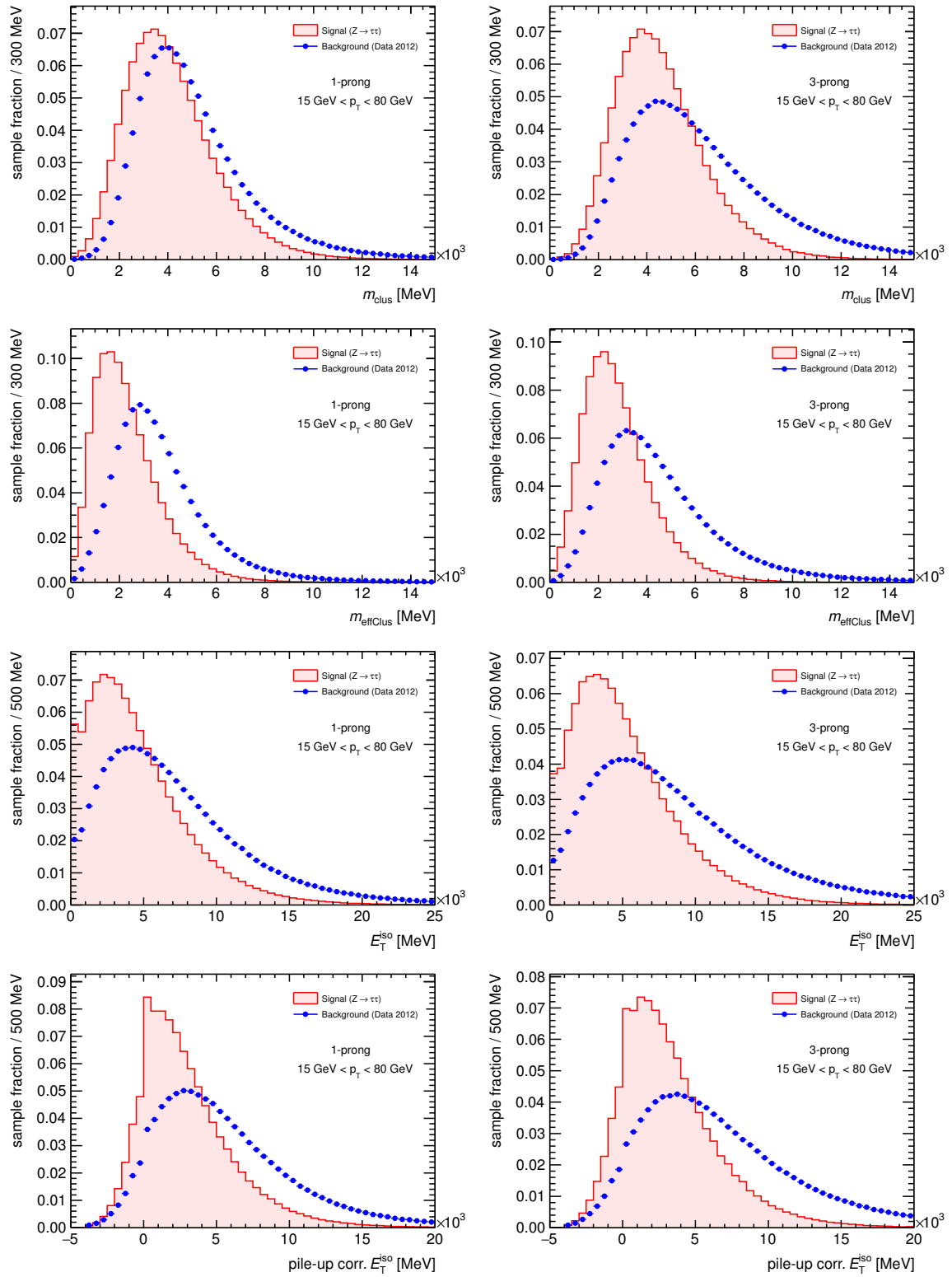


Figure A.4: Discriminating variables for 1-prong (left) and 3-prong (right) tau candidates. The signal candidates are from simulated $Z \to \tau\tau$ events matched to a true hadronic tau decay. The background candidates are from multi-jet events in 2012 data. All distributions are normalised to unity.

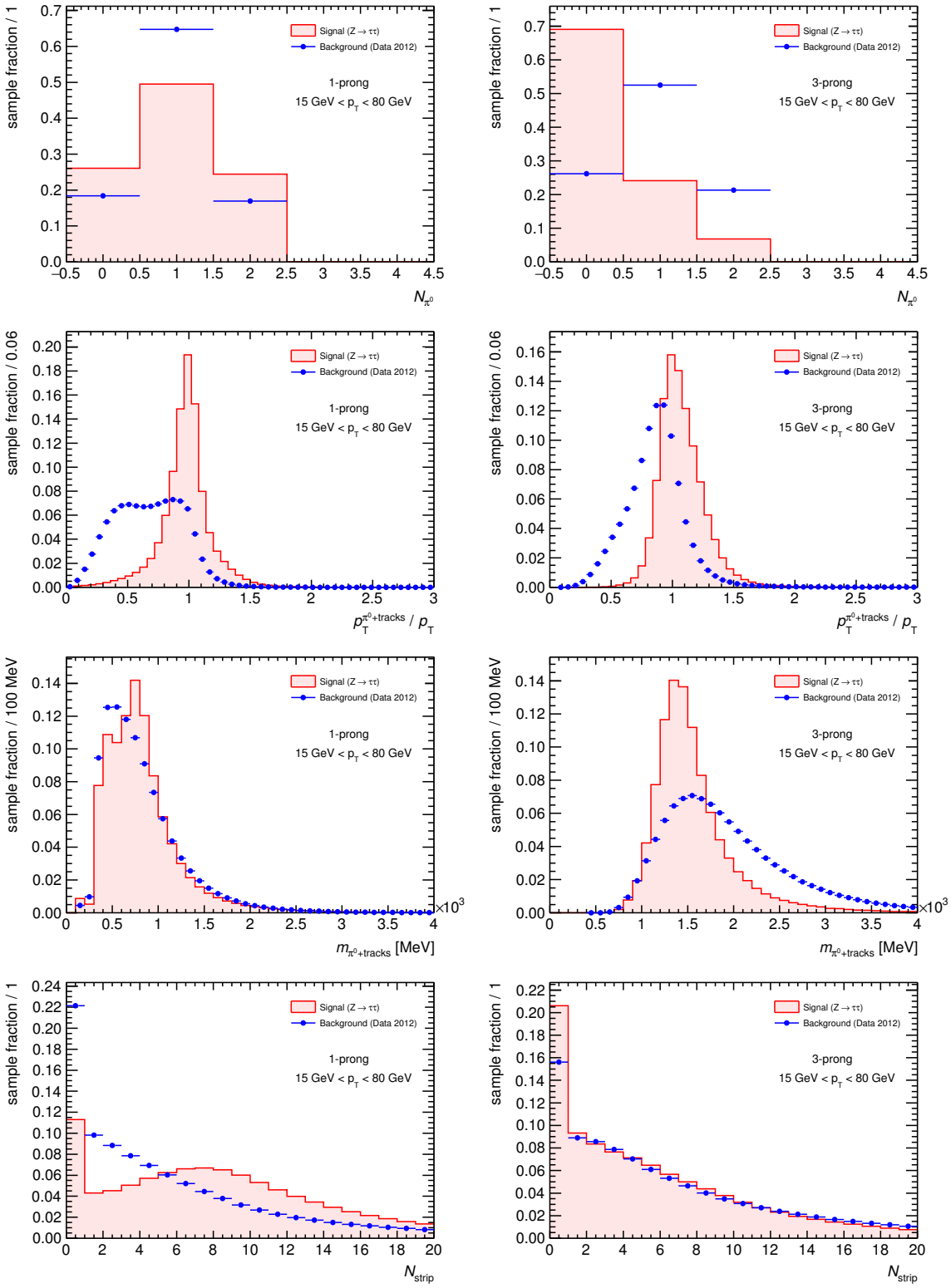


Figure A.5: Discriminating variables for 1-prong (left) and 3-prong (right) tau candidates. The signal candidates are from simulated $Z \rightarrow \tau\tau$ events matched to a true hadronic tau decay. The background candidates are from multi-jet events in 2012 data. All distributions are normalised to unity.

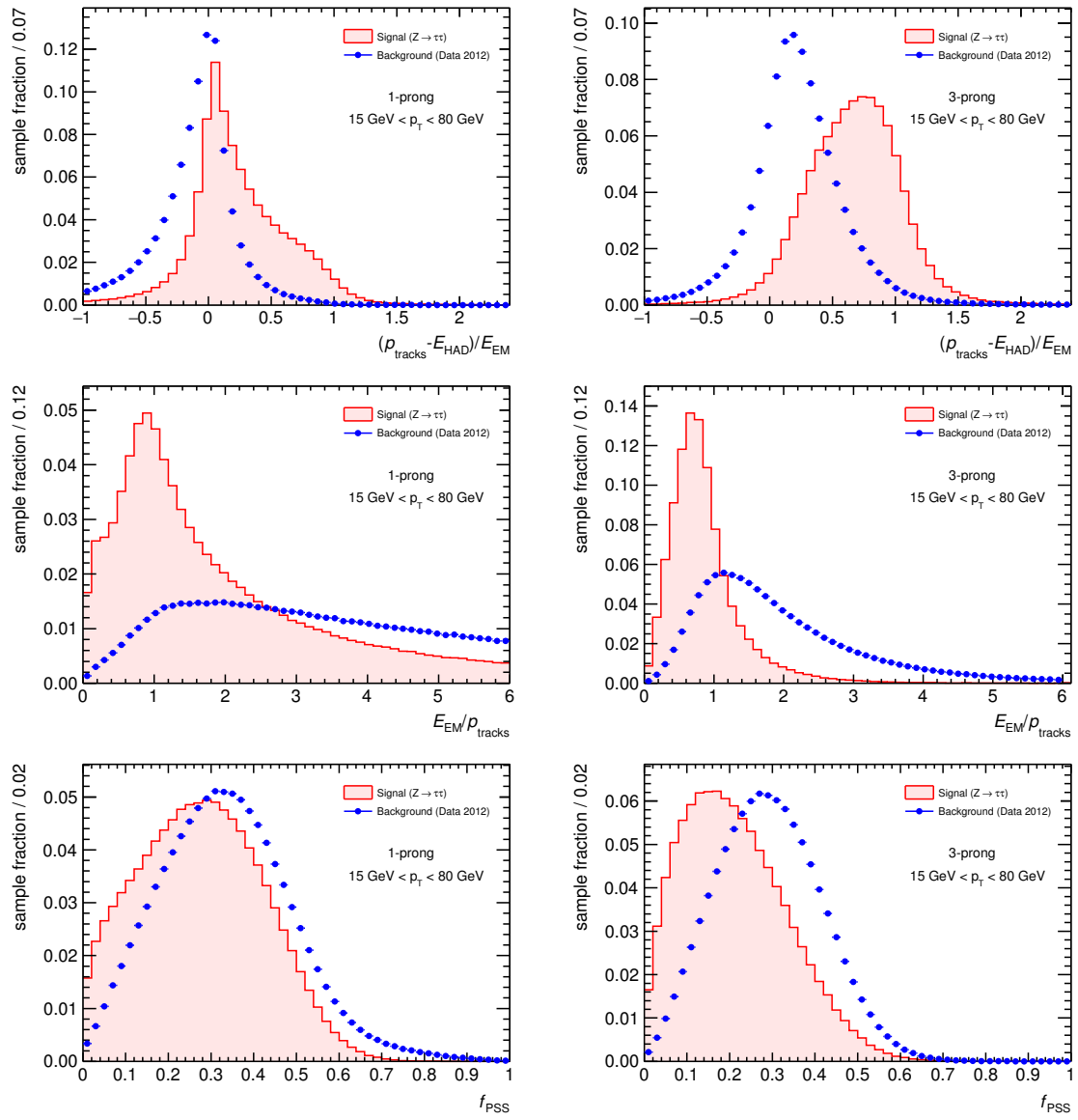


Figure A.6: Discriminating variables for 1-prong (left) and 3-prong (right) tau candidates. The signal candidates are from simulated $Z \rightarrow \tau\tau$ events matched to a true hadronic tau decay. The background candidates are from multi-jet events in 2012 data. All distributions are normalised to unity.

B. Appendices: ZH analysis

B.1. List of MC samples

Primary MC samples used in the ZH analysis. For each process, the following information is given: the filter applied during event generation, the MC generators, the totol number of events (before any cuts), and the cross section. The latter includes filter efficiencies and k-factors. The generators are abbreviated as follows: 'AlpJi' for Alpgen+Jimmy, 'AlpPy' for Alpgen+Pythia, 'HerJi' for Herwig+Jimmy, 'PowPy8' for Powheg+Pythia8, 'McNJi' for McAtNlo+Jimmy, 'MadPy' for MadGraph+Pythia and 'Py8' for Pythia8.

Process	Comment	Generator	# events	Cross section
$Z(\to \mu\mu)$ + 0 jets	$10 < M_{ll}$[GeV] < 60	AlpJi	999999	4156.0
$Z(\to \mu\mu)$ + 1 jet	$10 < M_{ll}$[GeV] < 60	AlpJi	299998	129.81
$Z(\to \mu\mu)$ + 2 jets	$10 < M_{ll}$[GeV] < 60	AlpJi	469998	62.947
$Z(\to \mu\mu)$ + 3 jets	$10 < M_{ll}$[GeV] < 60	AlpJi	144499	13.483
$Z(\to \mu\mu)$ + 4 jets	$10 < M_{ll}$[GeV] < 60	AlpJi	36300	3.0700
$Z(\to \mu\mu)+ \geq 5$ jets	$10 < M_{ll}$[GeV] < 60	AlpJi	79740	0.8296
$Z(\to \mu\mu)$ + 0 jets	$60 < M_{ll}$[GeV] < 2000	AlpPy	5918796	848.61
$Z(\to \mu\mu)$ + 1 jet	$60 < M_{ll}$[GeV] < 2000	AlpPy	7658384	89.373
$Z(\to \mu\mu)$ + 2 jets	$60 < M_{ll}$[GeV] < 2000	AlpPy	3055489	69.481
$Z(\to \mu\mu)$ + 3 jets	$60 < M_{ll}$[GeV] < 2000	AlpPy	734799	18.494
$Z(\to \mu\mu)$ + 4 jets	$60 < M_{ll}$[GeV] < 2000	AlpPy	398200	4.7267
$Z(\to \mu\mu)+ \geq 5$ jets	$60 < M_{ll}$[GeV] < 2000	AlpPy	179200	1.4802
$Z(\to \mu\mu) + c\bar{c}$ + 0 jets	$60 < M_{ll}$[GeV] < 2000	AlpPy	298998	13.918
$Z(\to \mu\mu) + c\bar{c}$ + 1 jet	$60 < M_{ll}$[GeV] < 2000	AlpPy	499799	8.3925
$Z(\to \mu\mu) + c\bar{c}$ + 2 jets	$60 < M_{ll}$[GeV] < 2000	AlpPy	499500	3.9775
$Z(\to \mu\mu) + c\bar{c}+ \geq 3$ jets	$60 < M_{ll}$[GeV] < 2000	AlpPy	443999	2.0130
$Z(\to \mu\mu) + b\bar{b}$ + 0 jets	$60 < M_{ll}$[GeV] < 2000	AlpPy	1799797	7.6766
$Z(\to \mu\mu) + b\bar{b}$ + 1 jet	$60 < M_{ll}$[GeV] < 2000	AlpPy	999897	3.8827
$Z(\to \mu\mu) + b\bar{b}$ + 2 jets	$60 < M_{ll}$[GeV] < 2000	AlpPy	999395	1.4869
$Z(\to \mu\mu) + b\bar{b}+ \geq 3$ jets	$60 < M_{ll}$[GeV] < 2000	AlpPy	885894	0.7302
$Z(\to ee)$ + 0 jets	$10 < M_{ll}$[GeV] < 60	AlpJi	999998	4156.1
$Z(\to ee)$ + 1 jet	$10 < M_{ll}$[GeV] < 60	AlpJi	299999	130.11

Process	Comment	Generator	# events	Cross section
$Z(\to ee)$ + 2 jets	$10 < M_{ll}$[GeV] < 60	AlpJi	469999	62.918
$Z(\to ee)$ + 3 jets	$10 < M_{ll}$[GeV] < 60	AlpJi	144500	13.514
$Z(\to ee)$ + 4 jets	$10 < M_{ll}$[GeV] < 60	AlpJi	36300	3.0763
$Z(\to ee)+ \geq 5$ jets	$10 < M_{ll}$[GeV] < 60	AlpJi	79619	0.8272
$Z(\to ee)$ + 0 jets	$60 < M_{ll}$[GeV] < 2000	AlpPy	6078988	848.38
$Z(\to ee)$ + 1 jet	$60 < M_{ll}$[GeV] < 2000	AlpPy	7879477	207.33
$Z(\to ee)$ + 2 jets	$60 < M_{ll}$[GeV] < 2000	AlpPy	3055991	69.473
$Z(\to ee)$ + 3 jets	$60 < M_{ll}$[GeV] < 2000	AlpPy	894995	18.451
$Z(\to ee)$ + 4 jets	$60 < M_{ll}$[GeV] < 2000	AlpPy	398597	4.7337
$Z(\to ee)+ \geq 5$ jets	$60 < M_{ll}$[GeV] < 2000	AlpPy	189700	1.4859
$Z(\to ee) + c\bar{c}$ + 0 jets	$60 < M_{ll}$[GeV] < 2000	AlpPy	224999	13.8803
$Z(\to ee) + c\bar{c}$ + 1 jet	$60 < M_{ll}$[GeV] < 2000	AlpPy	499500	8.4110
$Z(\to ee) + c\bar{c}$ + 2 jets	$60 < M_{ll}$[GeV] < 2000	AlpPy	498997	3.9652
$Z(\to ee) + c\bar{c}+ \geq 3$ jets	$60 < M_{ll}$[GeV] < 2000	AlpPy	443697	2.0185
$Z(\to ee) + b\bar{b}$ + 0 jets	$60 < M_{ll}$[GeV] < 2000	AlpPy	1799992	7.6798
$Z(\to ee) + b\bar{b}$ + 1 jet	$60 < M_{ll}$[GeV] < 2000	AlpPy	999896	3.8854
$Z(\to ee) + b\bar{b}$ + 2 jets	$60 < M_{ll}$[GeV] < 2000	AlpPy	999594	1.4802
$Z(\to ee) + b\bar{b}+ \geq 3$ jets	$60 < M_{ll}$[GeV] < 2000	AlpPy	885392	0.7282
$Z(\to \tau\tau)$ + 0 jets	$10 < M_{ll}$[GeV] < 60	AlpJi	999898	4156.3
$Z(\to \tau\tau)$ + 1 jet	$10 < M_{ll}$[GeV] < 60	AlpJi	299999	130.08
$Z(\to \tau\tau)$ + 2 jets	$10 < M_{ll}$[GeV] < 60	AlpJi	469999	63.069
$Z(\to \tau\tau)$ + 3 jets	$10 < M_{ll}$[GeV] < 60	AlpJi	119900	13.498
$Z(\to \tau\tau)$ + 4 jets	$10 < M_{ll}$[GeV] < 60	AlpJi	366100	3.0955
$Z(\to \tau\tau)+ \geq 5$ jets	$10 < M_{ll}$[GeV] < 60	AlpJi	79979	0.8250
$Z(\to \tau\tau)$ + 0 jets	$60 < M_{ll}$[GeV] < 2000	AlpPy	6299885	848.27
$Z(\to \tau\tau)$ + 1 jet	$60 < M_{ll}$[GeV] < 2000	AlpPy	8199883	207.40
$Z(\to \tau\tau)$ + 2 jets	$60 < M_{ll}$[GeV] < 2000	AlpPy	3174895	69.450
$Z(\to \tau\tau)$ + 3 jets	$60 < M_{ll}$[GeV] < 2000	AlpPy	894995	18.487
$Z(\to \tau\tau)$ + 4 jets	$60 < M_{ll}$[GeV] < 2000	AlpPy	398798	4.7343
$Z(\to \tau\tau)+ \geq 5$ jets	$60 < M_{ll}$[GeV] < 2000	AlpPy	229799	1.4821
$W(\to e\nu)$ + 0 jets		AlpPy	31894238	9208.2
$W(\to e\nu)$ + 1 jet		AlpPy	47465905	2031.1
$W(\to e\nu)$ + 2 jets		AlpPy	17559347	613.16

Process	Comment	Generator	# events	Cross section
$W(\to e\nu)$ + 3 jets		AlpPy	4985287	167.29
$W(\to e\nu)$ + 4 jets		AlpPy	2553792	42.755
$W(\to e\nu)+ \geq 5$ jets		AlpPy	769192	13.553
$W(\to \mu\nu)$ + 0 jets		AlpPy	31343557	9208.0
$W(\to \mu\nu)$ + 1 jet		AlpPy	44260419	2031.6
$W(\to \mu\nu)$ + 2 jets		AlpPy	17611454	614.36
$W(\to \mu\nu)$ + 3 jets		AlpPy	4966077	167.30
$W(\to \mu\nu)$ + 4 jets		AlpPy	2556595	42.765
$W(\to \mu\nu)+ \geq 5$ jets		AlpPy	788898	13.562
$W(\to \tau\nu)$ + 0 jets		AlpPy	29512260	9208.0
$W(\to \tau\nu)$ + 1 jet		AlpPy	48065178	2030.6
$W(\to \tau\nu)$ + 2 jets		AlpPy	17591943	614.39
$W(\to \tau\nu)$ + 3 jets		AlpPy	4977982	167.28
$W(\to \tau\nu)$ + 4 jets		AlpPy	2558295	42.806
$W(\to \tau\nu)+ \geq 5$ jets		AlpPy	789096	13.550
WW	has lepton ($p_{\mathrm{T}} > 10\,\mathrm{GeV}$)	HerJi	2494694	20.8999
WZ	has lepton ($p_{\mathrm{T}} > 10\,\mathrm{GeV}$)	HerJi	999998	6.9706
ZZ	has lepton ($p_{\mathrm{T}} > 10\,\mathrm{GeV}$)	HerJi	245000	1.5379
$W^-(\to e\nu)\, Z(\to ee)$	filter A	PowPy8	190000	0.4650
$W^-(\to e\nu)\, Z(\to\to \mu\mu)$	filter A	PowPy8	190000	0.3666
$W^-(\to e\nu)\, Z(\to \tau\tau)$	filter A	PowPy8	76000	0.0329
$W^-(\to \mu\nu)\, Z(\to ee)$	filter A	PowPy8	189999	0.4484
$W^-(\to \mu\nu)\, Z(\to \mu\mu)$	filter A	PowPy8	190000	0.3751
$W^-(\to \mu\nu)\, Z(\to \tau\tau)$	filter A	PowPy8	76000	0.0331
$W^-(\to \tau\nu)\, Z(\to ee)$	filter A	PowPy8	75400	0.2187
$W^-(\to \tau\nu)\, Z(\to \mu\mu)$	filter A	PowPy8	76000	0.1903
$W^-(\to \tau\nu)\, Z(\to \tau\tau)$	filter A	PowPy8	19000	0.0114
$W^+(\to e\nu)\, Z(\to ee)$	filter A	PowPy8	189899	0.3261
$W^+(\to e\nu)\, Z(\to \mu\mu)$	filter A	PowPy8	190000	0.2542
$W^+(\to e\nu)\, Z(\to \tau\tau)$	filter A	PowPy8	76000	0.0205
$W^+(\to \mu\nu)\, Z(\to ee)$	filter A	PowPy8	190000	0.3155
$W^+(\to \mu\nu)\, Z(\to \mu\mu)$	filter A	PowPy8	190000	0.2610
$W^+(\to \mu\nu)\, Z(\to \tau\tau)$	filter A	PowPy8	76000	0.0206

Process	Comment	Generator	# events	Cross section
$W^+(\to \tau\nu)\, Z(\to ee)$	filter A	PowPy8	76000	0.1570
$W^+(\to \tau\nu)\, Z(\to \mu\mu)$	filter A	PowPy8	76000	0.1346
$W^+(\to \tau\nu)\, Z(\to \tau\tau)$	filter A	PowPy8	19000	0.0072
$ZZ \to 4e$	filter B	PowPy8	1099997	0.0698
$ZZ \to 2e2\mu$	filter B	PowPy8	1599696	0.1454
$ZZ \to 2e2\tau$	filter B	PowPy8	599999	0.1021
$ZZ \to 4\mu$	filter B	PowPy8	1099798	0.0701
$ZZ \to 2\mu2\tau$	filter B	PowPy8	799900	0.1032
$ZZ \to 4\tau$	filter B	PowPy8	299999	0.0081
$gg \to ZZ \to 4e$		McNJi	90000	0.0007
$gg \to ZZ \to 4\mu$		McNJi	90000	0.0007
$gg \to ZZ \to 2e2\mu$		McNJi	90000	0.0014
$t\bar{t}$	has lepton	McNJi	9626750	137.31
$t\bar{t}$	no lepton	McNJi	923946	115.57
single top, $W \to e\nu$	s-channel	McNJi	169183	0.6059
single top, $W \to \mu\nu$	s-channel	McNJi	169100	0.6059
single top, $W \to \tau\nu$	s-channel	McNJi	169061	0.6059
single top, (inclusive)	t-channel	McNJi	1766958	22.371
$t\bar{t}Z$		MadPy	399996	0.0914
$t\bar{t}Zj$		MadPy	399995	0.0612
$t\bar{t}Zjj$	inclusive	MadPy	399798	0.0537
$tZ(\to ll)$	s+t-channel	MadPy	100000	0.0380
ZZZ^*		MadPy	50000	0.0003
ZWW^*		MadPy	50000	0.0016
$ggF : H \to ZZ \to 4l$	$m(H) = 125$ GeV	PowPy8	200000	0.0054
$VBF : H \to ZZ \to 4l$	$m(H) = 125$ GeV	PowPy8	199999	0.0004
$WH : H \to ZZ \to 4l$	$m(H) = 125$ GeV	Py8	100000	0.0002
$ZH : H \to ZZ \to 4l$	$m(H) = 125$ GeV	Py8	85000	0.0001
$ggF : H \to WW \to 2l2\nu$	$m(H) = 125$ GeV	PowPy8	500000	0.2136
$VBF : H \to WW \to 2l2\nu$	$m(H) = 125$ GeV	PowPy8	300000	0.0181
$WH : H \to WW \to 2l2\nu$	$m(H) = 125$ GeV	Py8	20000	0.0159
$ZH : H \to WW \to 2l2\nu$	$m(H) = 125$ GeV	Py8	20000	0.0094
$WH : H \to \tau_{\text{lep}}\tau_{\text{lep}}$	$m(H) = 125$ GeV	Py8	494953	0.0055

Process	Comment	Generator	# events	Cross section
$ggH : H \to \tau_{\text{lep}}\tau_{\text{had}}$	$m(H) = 125$ GeV	PowPy8	1994795	0.5628
$VBF : H \to \tau_{\text{lep}}\tau_{\text{had}}$	$m(H) = 125$ GeV	PowPy8	1099897	0.0455
$WH : H \to \tau_{\text{lep}}\tau_{\text{had}}$	$m(H) = 125$ GeV	Py8	230000	0.0201
$ggF : H \to \tau_{\text{had}}\tau_{\text{had}}$	$m(H) = 125$ GeV	PowPy8	1374899	0.5180
$VBF : H \to \tau_{\text{had}}\tau_{\text{had}}$	$m(H) = 125$ GeV	PowPy8	1848691	0.0419
$WH : H \to \tau_{\text{had}}\tau_{\text{had}}$	$m(H) = 125$ GeV	Py8	499362	0.0185
$ZH : H \to \tau_{\text{lep}}\tau_{\text{lep}}$	$m(H) = 100$ GeV	Py8	29999	0.0080
$ZH : H \to \tau_{\text{lep}}\tau_{\text{lep}}$	$m(H) = 105$ GeV	Py8	30000	0.0068
$ZH : H \to \tau_{\text{lep}}\tau_{\text{lep}}$	$m(H) = 110$ GeV	Py8	100000	0.0058
$ZH : H \to \tau_{\text{lep}}\tau_{\text{lep}}$	$m(H) = 115$ GeV	Py8	100000	0.0048
$ZH : H \to \tau_{\text{lep}}\tau_{\text{lep}}$	$m(H) = 120$ GeV	Py8	100000	0.0039
$ZH : H \to \tau_{\text{lep}}\tau_{\text{lep}}$	$m(H) = 125$ GeV	Py8	495000	0.0031
$ZH : H \to \tau_{\text{lep}}\tau_{\text{lep}}$	$m(H) = 130$ GeV	Py8	100000	0.0023
$ZH : H \to \tau_{\text{lep}}\tau_{\text{lep}}$	$m(H) = 135$ GeV	Py8	30000	0.0017
$ZH : H \to \tau_{\text{lep}}\tau_{\text{lep}}$	$m(H) = 140$ GeV	Py8	30000	0.0012
$ZH : H \to \tau_{\text{lep}}\tau_{\text{lep}}$	$m(H) = 145$ GeV	Py8	30000	0.0008
$ZH : H \to \tau_{\text{lep}}\tau_{\text{lep}}$	$m(H) = 150$ GeV	Py8	30000	0.0005
$ZH : H \to \tau_{\text{lep}}\tau_{\text{had}}$	$m(H) = 100$ GeV	Py8	30000	0.0295
$ZH : H \to \tau_{\text{lep}}\tau_{\text{had}}$	$m(H) = 105$ GeV	Py8	30000	0.0252
$ZH : H \to \tau_{\text{lep}}\tau_{\text{had}}$	$m(H) = 110$ GeV	Py8	30000	0.0213
$ZH : H \to \tau_{\text{lep}}\tau_{\text{had}}$	$m(H) = 115$ GeV	Py8	30000	0.0177
$ZH : H \to \tau_{\text{lep}}\tau_{\text{had}}$	$m(H) = 120$ GeV	Py8	30000	0.0144
$ZH : H \to \tau_{\text{lep}}\tau_{\text{had}}$	$m(H) = 125$ GeV	Py8	499998	0.0114
$ZH : H \to \tau_{\text{lep}}\tau_{\text{had}}$	$m(H) = 130$ GeV	Py8	30000	0.0086
$ZH : H \to \tau_{\text{lep}}\tau_{\text{had}}$	$m(H) = 135$ GeV	Py8	29998	0.0063
$ZH : H \to \tau_{\text{lep}}\tau_{\text{had}}$	$m(H) = 140$ GeV	Py8	30000	0.0044
$ZH : H \to \tau_{\text{lep}}\tau_{\text{had}}$	$m(H) = 145$ GeV	Py8	30000	0.0029
$ZH : H \to \tau_{\text{lep}}\tau_{\text{had}}$	$m(H) = 150$ GeV	Py8	30000	0.0018
$ZH : H \to \tau_{\text{had}}\tau_{\text{had}}$	$m(H) = 100$ GeV	Py8	100000	0.0271
$ZH : H \to \tau_{\text{had}}\tau_{\text{had}}$	$m(H) = 105$ GeV	Py8	99999	0.0232
$ZH : H \to \tau_{\text{had}}\tau_{\text{had}}$	$m(H) = 110$ GeV	Py8	100000	0.0196
$ZH : H \to \tau_{\text{had}}\tau_{\text{had}}$	$m(H) = 115$ GeV	Py8	100000	0.0163
$ZH : H \to \tau_{\text{had}}\tau_{\text{had}}$	$m(H) = 120$ GeV	Py8	100000	0.0133

Process	Comment	Generator	# events	Cross section
$ZH : H \to \tau_{\text{had}}\tau_{\text{had}}$	$m(H) = 125$ GeV	Py8	499997	0.0105
$ZH : H \to \tau_{\text{had}}\tau_{\text{had}}$	$m(H) = 130$ GeV	Py8	100000	0.0079
$ZH : H \to \tau_{\text{had}}\tau_{\text{had}}$	$m(H) = 135$ GeV	Py8	99999	0.0058
$ZH : H \to \tau_{\text{had}}\tau_{\text{had}}$	$m(H) = 140$ GeV	Py8	99998	0.0040
$ZH : H \to \tau_{\text{had}}\tau_{\text{had}}$	$m(H) = 145$ GeV	Py8	100000	0.0027
$ZH : H \to \tau_{\text{had}}\tau_{\text{had}}$	$m(H) = 150$ GeV	Py8	99999	0.0016

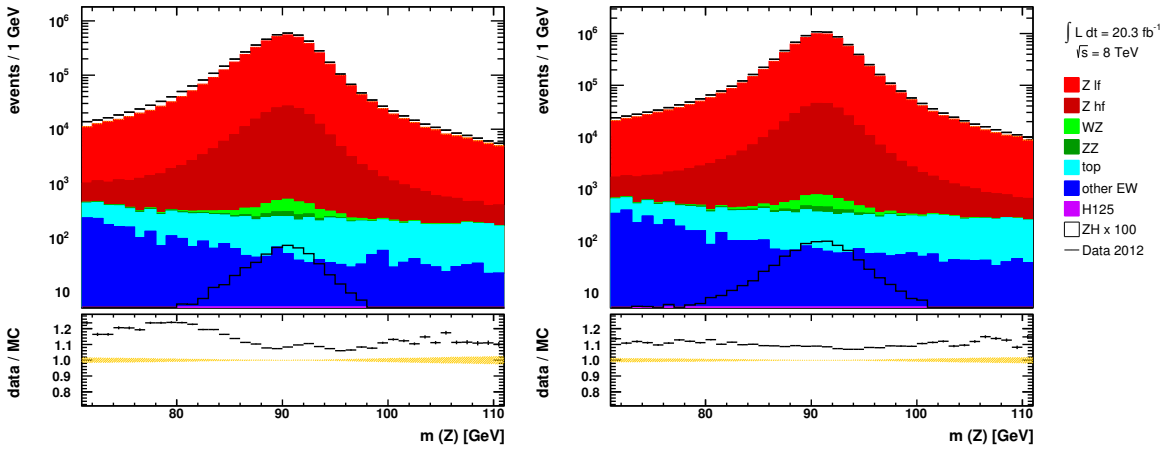
filter A: $m(ll) > 2m(l_Z) + 250\,\text{MeV}$, dilepton filter $p_{\text{T}} > 5\,\text{GeV}$, $|\eta| < 2.7$

filter B: $m(ll) > 4\,\text{GeV}$, dimuon filter $p_{\text{T}} > 5\,\text{GeV}$

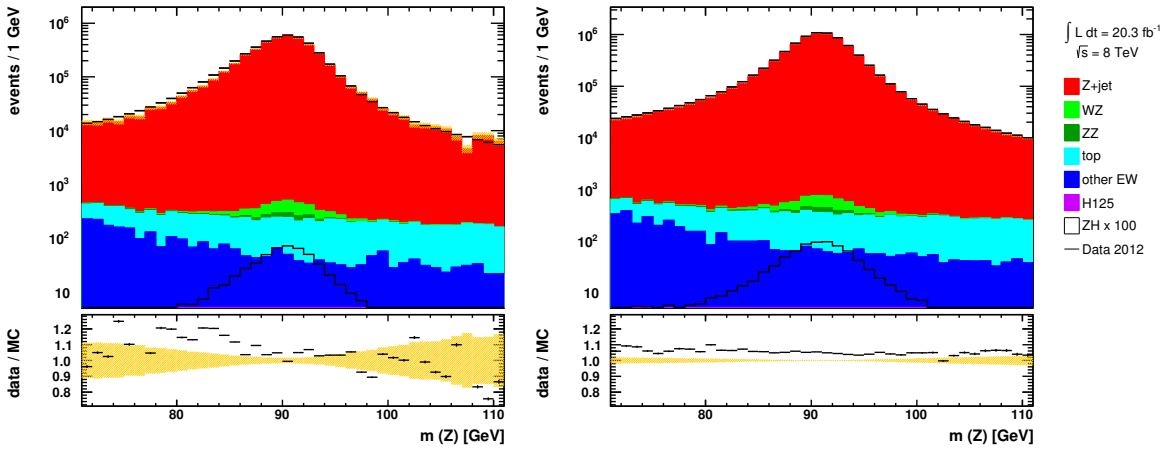
B.2. Control plots Z boson mass peak

Figure B.1 shows the invariant mass distribution of the Z-pair separately for ee-pairs (left) and $\mu\mu$-pairs (right). While the shape discrepancy in the tails is observed in for ee-pairs only, the underestimation of the predicted number of events appears in both channels. It is therefore unlikely to be caused by mis-modelled reconstruction, identification or trigger efficiencies, which are developed independently for electrons and muons. In the two central distributions the simulated $Z \to ee$ + jets and $Z \to \mu\mu$ + jets events are generated with POWHEG+PYTHIA8 instead of ALPGEN+PYTHIA (all other constributions are unchanged). The alternative sample does not distinguish between light and heavy flavour contributions. The tail for ee-pairs remains mis-modelled, which confirms the findings of other groups. The agreement in the predicted number of events improves to about 5%. The bottom two plots show again the nominal simulation, rescaled to match the event count measured in data. The systematic uncertainties cover most of the shape discrepancies in the tails of the invariant mass distribution.

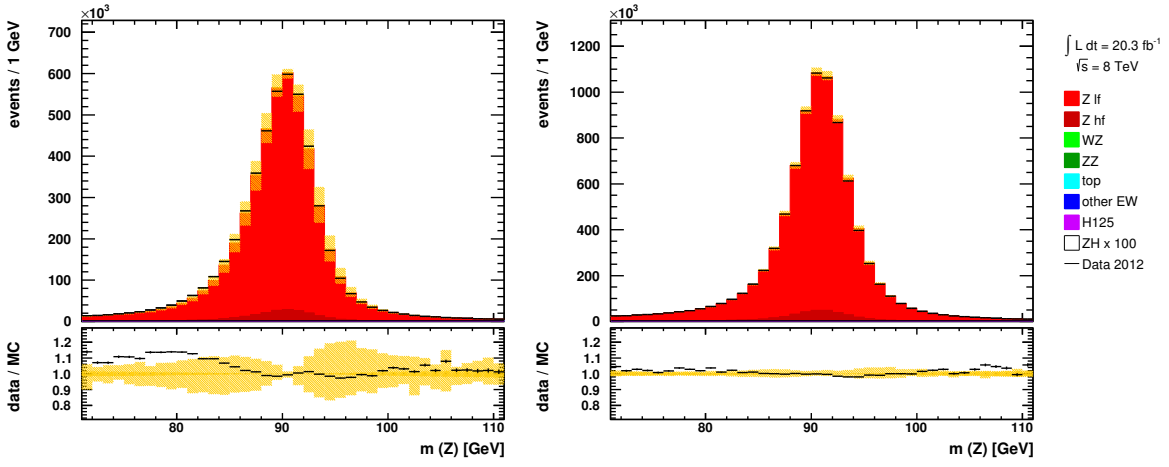
Figures B.2 to B.4 show the validation of the track and calorimeter isolation requirements as well as the cut on the impact parameter significance. The measurement is performed via a simplified tag-and-probe analysis. Electron or muon pairs consistent with the decay of a Z boson are selected: the pair has to be of same flavour and opposite electric charge and have an invariant mass of (91 ± 10) GeV. The leading lepton functions as the 'tag' and has to satisfy in particular $p_{\mathrm{T}} > 26$ GeV, tight identification, $p_{\mathrm{T}}^{\mathrm{iso}(0.4)}/p_{\mathrm{T}} < 0.1$, $E_{\mathrm{T}}^{\mathrm{iso}(0.4)}/p_{\mathrm{T}} < 0.1$ and $|d_0/\sigma(d_0)| < 3.5$ or 6.5 (muons or electrons). The subleading lepton functions as the 'probe' and satisfies basic criteria only (see 6.4.2). The probe lepton is then subjected to requirements imposed on the Z-pair leptons one at a time. That means that all cuts are applied apart from the one in question, i.e. for the measurement of the track isolation efficiency, the probe lepton is required to pass the cuts on ID, calorimeter isolation and $d_0/\sigma(d_0)$. This increases the purity of the probe sample. The cut efficiency measured in data is compared to the efficiency measured in data. To simplify the measurement, no dedicated background estimation or subtraction is performed. To estimated the contamination of fake leptons, the MC efficiencies are measured for all leptons or truth-matched leptons only. Figure B.2 shows this for the track isolation measurement. Apart from the first two (filled) bins, both MC efficiencies match, meaning that the measurement is not biased by the presence of fake leptons. Figure B.3 and B.4 show the efficiency ratios data/MC for the various requirements as a function of momentum and pseudorapidity. The cut efficiencies in data and simulation agree very well. Deviations larger than 1% are only observed for bins in which the measurement cannot be trusted due to increasing contamination with fake leptons. Due to the very good agreement, no scale factors are used in the analysis for requirememnts on the track isolation, calorimeter isolation or impact parameter significance.



(a) $Z \to ee$ (b) $Z \to \mu\mu$

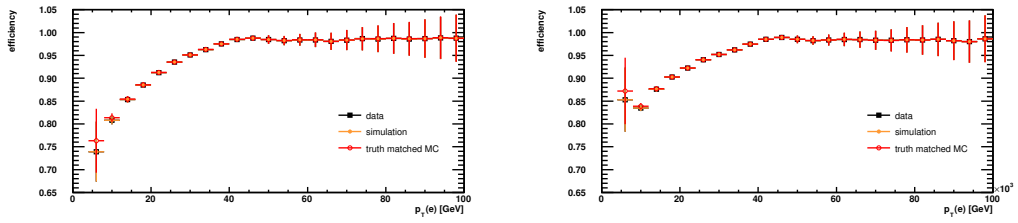


(d) $Z \to ee$ (e) $Z \to \mu\mu$



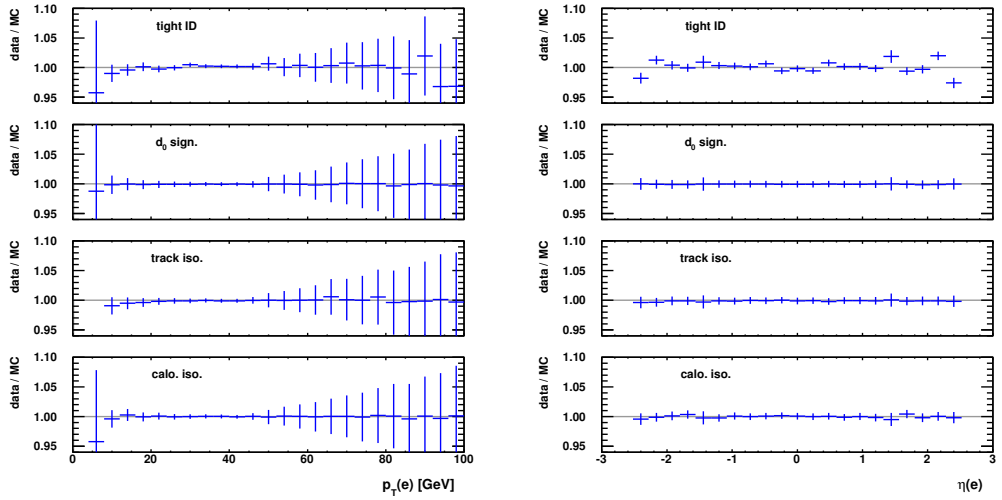
(g) $Z \to ee$ (h) $Z \to \mu\mu$

Figure B.1: Invariant mass of the Z-pair, separately for the electron pairs (left) and muon pairs (right). (a,b) Nominal distribution and (c,d) using Z+jets samples generated with POWHEG and PYTHIA8 instead of the nominal ALPGEN+PYTHIA sample. The yellow shaded area shows the statistical uncertainty. In (e,f), the nominal sample is used, but scaled up by 9%. The yellow shaded area shows here the systematic uncertainties connected to the electron and muon energy and position corrections, as well as trigger, reco and ID scale factors.



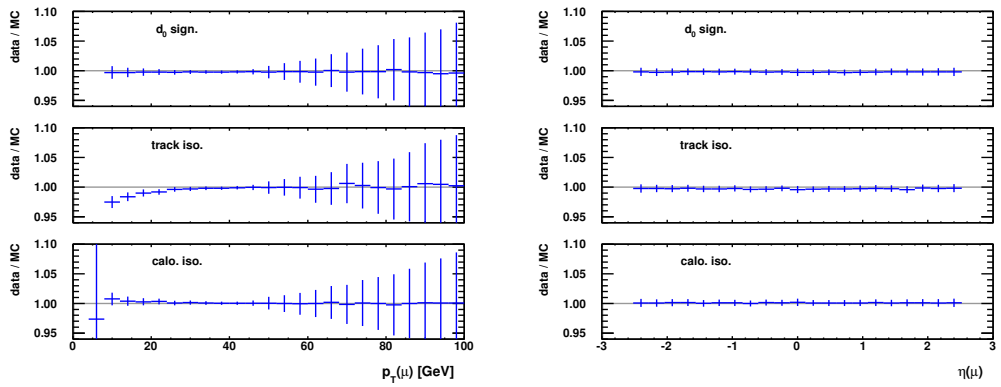
(a) electrons (b) muons

Figure B.2: Efficiency of the track isolation requirement $p_{\mathrm{T}}^{\mathrm{iso}(0.4)}/p_{\mathrm{T}} < 0.1$ as a function of p_{T} for (a) electrons and (b) muons. The efficiency is measured on all probe leptons for data and simulation, as well as on truth-matched probe leptons only.



(a) electrons (b) electrons

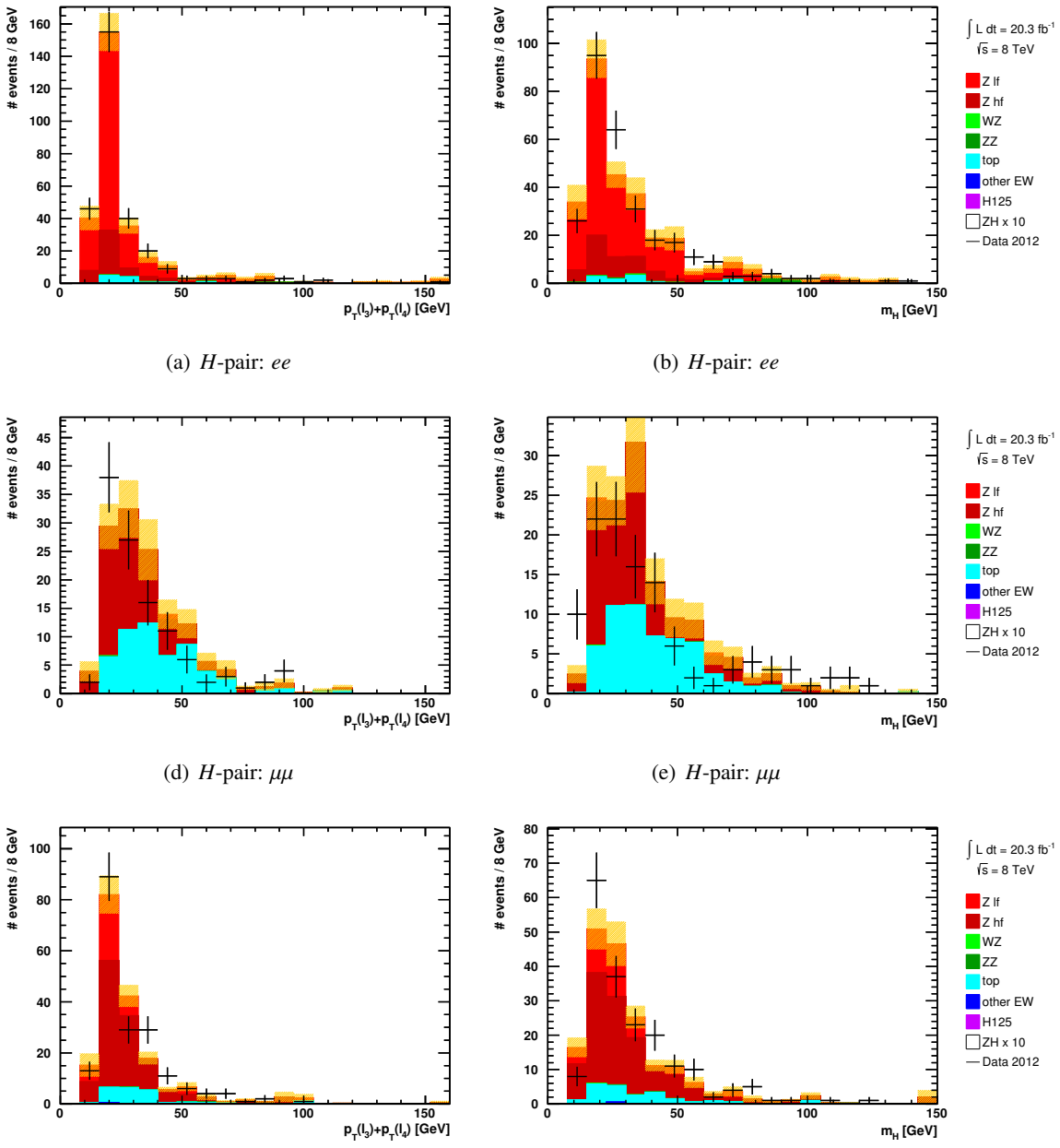
Figure B.3: Efficiency ratio for electrons as a function of (a) p_{T} and (b) η for the following requirements (from top to bottom): tight identification, $|d_0/\sigma(d_0)| < 6.5$, $p_{\mathrm{T}}^{\mathrm{iso}(0.4)}/p_{\mathrm{T}} < 0.1$, and $E_{\mathrm{T}}^{\mathrm{iso}(0.2)}/p_{\mathrm{T}} < 0.1$.



(a) muons (b) muons

Figure B.4: Efficiency ratio for muon as a function of (a) p_{T} and (b) η for the following requirements (from top to bottom): $|d_0/\sigma(d_0)| < 3.5$, $p_{\mathrm{T}}^{\mathrm{iso}(0.4)}/p_{\mathrm{T}} < 0.1$, and $E_{\mathrm{T}}^{\mathrm{iso}(0.2)}/p_{\mathrm{T}} < 0.1$.

B.3. Additional distributions

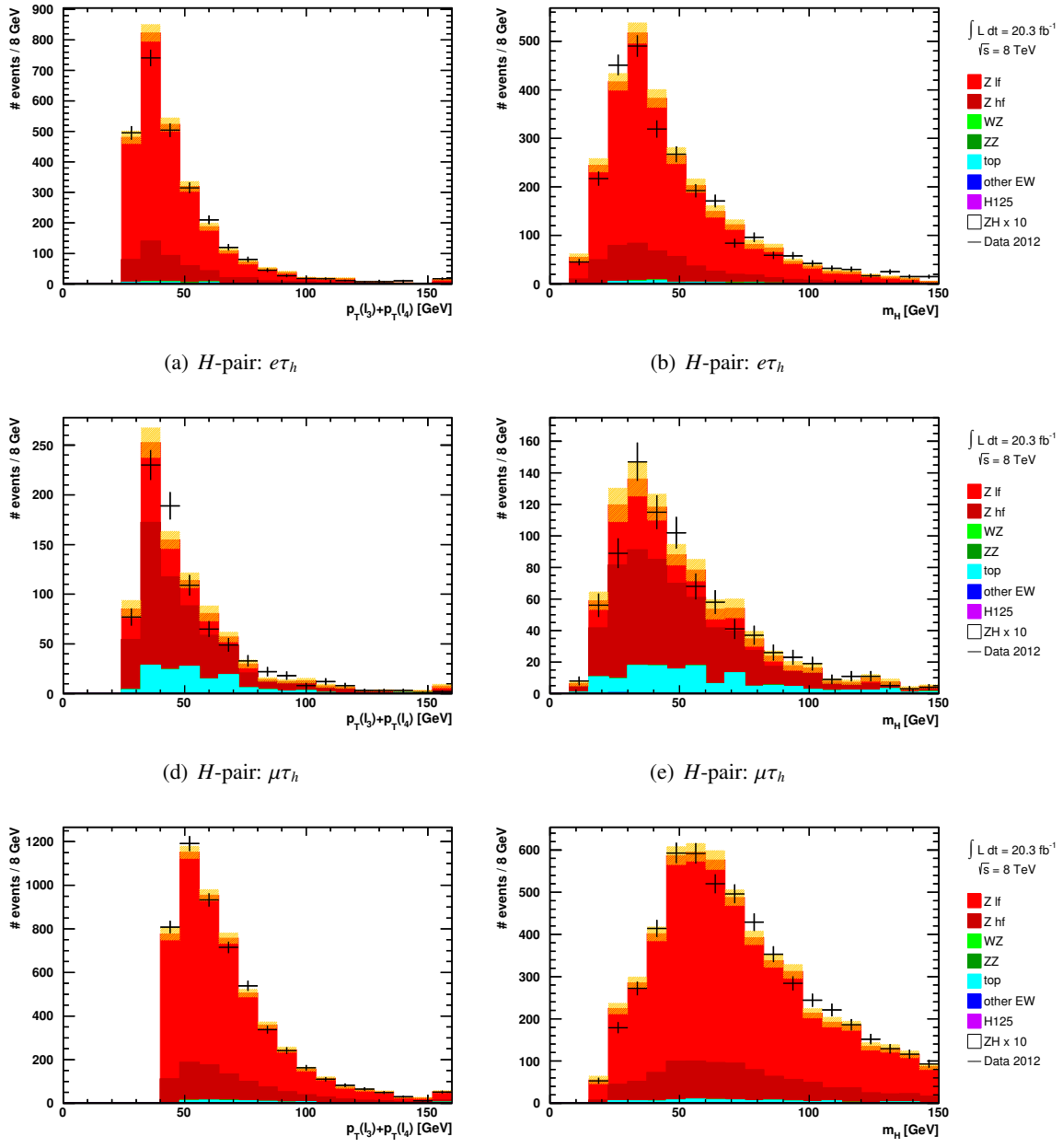


(a) H-pair: ee (b) H-pair: ee

(d) H-pair: $\mu\mu$ (e) H-pair: $\mu\mu$

(g) H-pair: $e\mu$ (h) H-pair: $e\mu$

Figure B.5: Distribution of (left) the momentum sum of the H-pair constituents, and (right) the invariant mass of H-pair, for the *fail-fail* regions and the final states (top) ee, (centre) $\mu\mu$, and (bottom) $e\mu$.



(a) H-pair: $e\tau_h$ (b) H-pair: $e\tau_h$

(d) H-pair: $\mu\tau_h$ (e) H-pair: $\mu\tau_h$

(g) H-pair: $\tau_h\tau_h$ (h) H-pair: $\tau_h\tau_h$

Figure B.6: Distribution of (left) the momentum sum of the H-pair constituents, and (right) the invariant mass of H-pair, for the *fail-fail* regions and the final states (top) $e\tau_h$, (centre) $\mu\tau_h$, and (bottom) $\tau_h\tau_h$.