

# **Kinematic Tau Reconstruction and Search For The Higgs Boson in Hadronic Tau Pair Decays with the CMS Experiment**

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der Rheinisch-Westfälischen Technischen Hochschule Aachen zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von

Diplom-Physiker Lars Perchalla

aus Lindlar

Berichter: Universitätsprofessor Dr. rer. nat. Achim Stahl  
Universitätsprofessor Dr. rer. nat. Stefan Schael

Tag der mündlichen Prüfung: 11.05.2011

Diese Dissertation ist auf den Internetseiten der Hochschulbibliothek online verfügbar.



## Zusammenfassung

Die vorliegende Dissertation behandelt die Suche nach dem Higgs-Boson, welches im Standardmodell der Elementarteilchenphysik vorhergesagt wird. Hierzu werden im Rahmen einer Monte Carlo-Simulation die Nachweismöglichkeiten des Zerfalls in Taupaaare mit dem CMS Experiment untersucht. Zwei separate Selektionen für die beiden dominanten Produktionskanäle, die Gluon- und die Vektorboson-Fusion, werden für leichte Higgs-Bosonen entwickelt.

Beide Selektionen berücksichtigen ausschließlich Tauzerfälle, in denen drei geladene Pionen entstehen. Effiziente Algorithmen werden entwickelt, um diese Zerfallsprodukte in den Hadron-Kollisionen am LHC zu identifizieren.

Mit Hilfe einer kinematischen Rekonstruktion, die die spezielle Topologie des Tauzerfalls berücksichtigt, lässt sich der vollständige Impuls des Tau-Leptons inklusive des Neutrinos berechnen. Die rekonstruierten Observablen werden Qualitätskriterien unterzogen, die für einen umfassenden Energiebereich der Tau-Leptonen gültig sind. Dies ermöglicht eine effiziente Unterdrückung von Quark- und Gluon-Jets, die Tauzerfälle imitieren.

Aus Paaren von Tau-Leptonen, die diese Anforderungen erfüllen, lassen sich Higgs-Bosonen rekonstruieren. Die unterschiedlichen kinematischen Eigenschaften der behandelten Produktionskanäle werden zur weiteren Abtrennung des Signals von Untergründen herangezogen. Die Massenverteilung des  $Z^0$ -Bosons weist weite Ausläufer bis in den Massenbereich des untersuchten Signals auf. Diese lassen sich im Falle der Gluon-Fusion nicht von Higgs-Bosonen unterscheiden. Bei der Erzeugung durch Vektorboson-Fusion entstehen zusätzliche Quark-Jets, die zusammen mit dem deutlichen Transversalimpuls des Higgs-Bosons eine klare Signatur zur Abtrennung aller Untergründe liefern.

Für die untersuchten Produktionsprozesse wird die Signifikanz der Selektionen diskutiert, die mit einer integrierten Luminosität von  $30 \text{ fb}^{-1}$  erreicht werden kann.



## Abstract

The thesis prepares a search for the Standard Model Higgs boson in the di-tau channel with the CMS experiment. Based on Monte Carlo simulations, two selections are developed for light Higgs bosons produced by the dominant processes, gluon fusion and vector-boson fusion.

Both selections exclusively consider the hadronic tau decay into three charged pions. They rely on an efficient algorithm to identify the tau decay products within the hadronic environment at the LHC.

A kinematic fit exploits the topology of the particular decay mode and enables the reconstruction of the entire tau momentum including the neutrino. A set of quality criteria is defined on the obtained observables, which is valid for a broad range of tau energies. This provides an efficient suppression of quark and gluon jets that fake tau decays.

The Higgs boson is reconstructed from pairs of tau leptons that pass the quality requirements. The selections derive further background suppression from the event kinematics. In case of the gluon fusion, the selected sample is dominated by off-shell  $Z^0$  bosons that decay into tau pairs. The vector-boson fusion involves two additional quark jets. Their signature and the significant transversal momentum of the Higgs boson allow for a background-free selection.

The significance of both selections is discussed for an integrated luminosity of  $30\text{ fb}^{-1}$ .



# Contents

<b>1. Introduction</b>	<b>1</b>
1.1. The Higgs Boson . . . . .	1
1.1.1. The Higgs Mechanism in the Standard Model . . . . .	2
1.1.2. Basic Properties of the Standard Model Higgs Boson . . . . .	4
1.1.3. The MSSM Higgs Boson . . . . .	7
1.1.4. Searches for the Higgs Boson . . . . .	8
1.2. The Large Hadron Collider . . . . .	12
1.3. The Compact Muon Solenoid . . . . .	14
1.3.1. The Tracking System . . . . .	16
1.3.2. Track Reconstruction . . . . .	16
1.3.3. Vertex Reconstruction . . . . .	18
1.3.4. Data Acquisition and Trigger . . . . .	19
1.4. Event Reconstruction Framework . . . . .	19
1.4.1. Monte Carlo Event Generation . . . . .	20
1.4.2. Detector Simulation and Digitization . . . . .	20
1.4.3. Event Reconstruction . . . . .	20
1.4.4. LHC Computing Grid . . . . .	21
<b>2. Kinematic Reconstruction of Tau Leptons</b>	<b>23</b>
2.1. Introduction . . . . .	23
2.2. Kinematics of Multi-Prong Tau Lepton Decays . . . . .	23
2.2.1. Tau Lepton Decay Topology . . . . .	23
2.2.2. Kinematic Calculation of the Tau Momentum . . . . .	25
2.2.3. Accessible Constraints . . . . .	31
2.3. Least Mean Squares Minimization with Lagrange Multipliers . . . . .	35
2.3.1. Basic Algorithm . . . . .	35
2.3.2. Iterations . . . . .	36
2.3.3. Unmeasured Parameters . . . . .	36
2.4. Kinematic Fit of Tau Leptons on Monte Carlo Level with CMS . . . . .	37
2.4.1. Standard Tau Reconstruction at CMS . . . . .	37
2.4.2. Kinematic Fit Workflow . . . . .	38
2.4.3. Quality Criteria . . . . .	44
2.4.4. Monte Carlo Truth Matching . . . . .	51
2.4.5. Efficiency of the Kinematic Tau Lepton Identification . . . . .	52

2.4.6. Resolution . . . . .	54
2.5. Technical Implementation in the CMS Software Framework . . . . .	58
2.5.1. The InputTrackSelector and the ThreeProngInputSelector . . . . .	60
2.5.2. The KinematicTauCreator and the ThreeProngTauCreator . . . . .	60
2.5.3. The KinematicTauProducer . . . . .	61
<b>3. Higgs Boson Reconstruction from Hadronic Tau Pair Decays</b>	<b>63</b>
3.1. Datasets and Expected Event Rates . . . . .	63
3.1.1. Signal-like Processes . . . . .	63
3.1.2. Background Processes . . . . .	64
3.2. Event Topology . . . . .	65
3.2.1. Higgs Kinematics . . . . .	66
3.2.2. Tau Pair Kinematics . . . . .	68
3.2.3. Kinematics of Single Tau Leptons . . . . .	68
3.2.4. Kinematics of Quarks from VBF . . . . .	70
3.2.5. Background Kinematics . . . . .	72
3.3. Selection Strategy . . . . .	73
3.4. Event Trigger . . . . .	73
3.5. Object Selection . . . . .	75
3.5.1. Tau Selection . . . . .	75
3.5.2. Assumptions on Backgrounds . . . . .	82
3.5.3. Jet Selection . . . . .	83
3.6. Event Selection . . . . .	91
3.6.1. Selection of Vector Boson Fusion . . . . .	91
3.6.2. Selection of Gluon Fusion . . . . .	96
3.6.3. Selection Summary . . . . .	101
3.7. Invariant Mass of the Tau Pair . . . . .	102
3.8. Systematic Uncertainties . . . . .	105
3.9. Signal Significance . . . . .	105
3.9.1. Possible Enhancement . . . . .	107
<b>4. Summary and Outlook</b>	<b>109</b>
<b>A. Appendix on Kinematic Tau Reconstruction</b>	<b>111</b>
A.1. Helix Propagation . . . . .	111
A.2. Vertex Link Significance . . . . .	112
<b>B. Appendix on Higgs Boson Reconstruction</b>	<b>113</b>
B.1. Jet Matching . . . . .	113
B.2. Basic Kinematic Distributions . . . . .	114
<b>References</b>	<b>117</b>

# Figures

1.1.	Higgs Potential . . . . .	3
1.2.	Single-vertex couplings of the Standard Model Higgs boson . . . . .	4
1.3.	Feynman Diagrams for Standard Model Higgs production from proton-proton collisions . . . . .	5
1.4.	Production cross sections of the Standard Model Higgs boson for proton-proton collisions . . . . .	6
1.5.	Branching fractions and the total decay width of the Standard Model Higgs boson . . . . .	7
1.6.	Most likely mass of the Standard Model Higgs Boson . . . . .	10
1.7.	The expected signal significance of the SM Higgs boson with CMS . . . . .	11
1.8.	Illustrations of the CERN Large Hadron Collider . . . . .	12
1.9.	Cross section for $p p$ and $p \bar{p}$ collisions at the Tevatron and the LHC . . . . .	13
1.10.	Illustration of the CMS detector . . . . .	14
1.11.	Total integrated luminosity of the CMS experiment . . . . .	16
1.12.	The CMS Inner Tracker . . . . .	17
1.13.	Track reconstruction performance . . . . .	18
2.1.	Invariant mass distribution of the three-pion system and the energy spectrum of the tau neutrino within the tau rest frame . . . . .	25
2.2.	Definition of the Gottfried-Jackson angle . . . . .	26
2.3.	Calculated tau momentum for typical $a_1$ parameters . . . . .	27
2.4.	Ambiguity for the magnitude of the tau momentum . . . . .	27
2.5.	Major dependencies of $\theta_{\text{GJ}}$ and $\theta_{\text{GJ}}^{\max}$ . . . . .	29
2.6.	Ratio of the actual Gottfried-Jackson angle $\theta_{\text{GJ}}$ and the maximal allowed one $\theta_{\text{GJ}}^{\max}$ in the laboratory system with uniformly distributed $\cos \theta_{\text{GJ}}^*$ in the tau rest frame . . . . .	30
2.7.	Dependence of $\theta_{\text{GJ}} / \theta_{\text{GJ}}^{\max}$ on the rest-frame angle $\theta_{\text{GJ}}^*$ . . . . .	30
2.8.	Dependencies of the maximal ambiguity width $w$ . . . . .	31
2.9.	Definition of the chosen angles $(\phi_p, \theta_p)$ and $(\phi_s, \theta_s)$ . . . . .	33
2.10.	Illustration of the correction derived from the helix propagation through the magnetic field . . . . .	35
2.11.	Illustration of the different cones within the particle-flow algorithm . . . . .	38
2.12.	Flowchart describing the workflow of the kinematic fit . . . . .	39
2.13.	Positional uncertainties of the primary and the secondary vertex, and mis-measurements of the track reconstruction . . . . .	41
2.14.	Variation of the Gottfried-Jackson angle due to the rotation of the primary vertex around the secondary vertex . . . . .	42
2.15.	Number of iterations per decay used by the kinematic fit in case of a pure $\tau \rightarrow 3\pi^\pm + \nu_\tau$ sample . . . . .	43

2.16. Quality criterion: $\chi^2$ -probability of the kinematic tau fit. . . . .	45
2.17. Quality criterion: The invariant mass of the hadronic part of the tau decay. . . . .	46
2.18. Quality criterion: Transverse-energy fraction of tau leptons reconstructed with the particle-flow algorithm and the kinematic fit. . . . .	47
2.19. Quality criterion: Number of tracks within the signal cone of the corresponding particle-flow tau candidate. . . . .	47
2.20. Quality criterion: Separation significance between the initial and the rotated primary vertex. . . . .	48
2.21. Quality criterion: Separation significance between the primary and the secondary vertex. . . . .	49
2.22. Overlap between the existing particle-flow discriminators and the newly proposed 3-prong tau discrimination. . . . .	49
2.23. Interrelation matrix for the set of quality requirements for signal-like 3-prong tau decays and background from QCD jets. . . . .	50
2.24. Monte Carlo truth matching based on $\Delta R$ assignment. . . . .	51
2.25. Purity gain from the application of tau quality requirements. . . . .	52
2.26. Angular resolution of the kinematically fitted tau-lepton momentum. . . . .	54
2.27. Momentum resolution of the kinematically fitted tau lepton. . . . .	55
2.28. Mass and transverse-momentum resolution of the $a_1$ resonance. . . . .	56
2.29. Spatial resolution of the kinematically fitted tau-decay vertex in Cartesian coordinates. . . . .	56
2.30. Dependency of the calculated tau momentum resolution w.r.t. $\cos \theta_{Gj}^*$ . . . . .	57
2.31. Dependencies of the expectation value of the tau-momentum resolution for uniformly distributed $\cos \theta_{Gj}^*$ . . . . .	58
2.32. Technical workflow of the kinematic fit. . . . .	59
 3.1. Illustration of the Higgs decay into a tau pair. . . . .	65
3.2. Distribution of the pseudorapidity for Higgs bosons from gluon fusion and vector-boson fusion. . . . .	66
3.3. Magnitude of the momentum vector for Higgs bosons from gluon fusion and vector-boson fusion. . . . .	67
3.4. Transversal momentum of Higgs bosons produced by gluon fusion and vector-boson fusion. . . . .	67
3.5. Angular difference between the two tau leptons from Higgs decays in the transversal plane. . . . .	68
3.6. Difference in $\Delta R$ between the two tau leptons from Higgs decays. . . . .	69
3.7. Balance of the transversal momentum between the two tau leptons from Higgs decays. . . . .	69
3.8. Transversal momentum of tau leptons from Higgs decays. . . . .	70
3.9. Pseudorapidity and separation of the quarks in vector-boson fusion. . . . .	71
3.10. Transversal momentum and the invariant mass of the quarks in vector-boson fusion. . . . .	71
3.11. Event topology of the $Z^0$ -boson decay into tau pairs. . . . .	72
3.12. Sketch of the selection strategy. . . . .	74
3.13. Selection criterion: $\chi^2$ -probability of the kinematic tau fit. . . . .	76
3.14. Selection criterion: Transverse-energy fraction of tau leptons reconstructed with the particle-flow algorithm and the kinematic fit $E_T^{\text{PF}} / E_T^{\text{kin}}$ . . . . .	77
3.15. Selection criterion: Transversal momentum of the $a_1$ . . . . .	78

---

3.16.	Mass dependence of the transversal $a_1$ momentum. . . . .	78
3.17.	Selection criterion: Track multiplicity within $\Delta R < 0.5$ around the direction of the kinematic tau. . . . .	79
3.18.	Number of tau pairs per event passing the selection. . . . .	81
3.19.	Kinematic correlation between both tau hemispheres. . . . .	83
3.20.	Transversal momentum resolution for jets in vector-boson fusion. . . . .	84
3.21.	Transversal momentum of reconstructed jets from vector-boson fusion. . .	85
3.22.	Separation between quark and tau jets in vector-boson fusion. . . . .	85
3.23.	Single jet selection in vector-boson fusion. . . . .	86
3.24.	Invariant mass of jet pairs. . . . .	88
3.25.	Acoplanarity between the di-jet and the di-tau systems. . . . .	88
3.26.	Number of jet pairs per event passing the selection. . . . .	90
3.27.	Event selection for vector-boson fusion: Vertex significance. . . . .	92
3.28.	Event selection for vector-boson fusion: Separation of the tau pair in the $(\eta, \phi)$ -plane. . . . .	93
3.29.	Event selection for vector-boson fusion: Pseudorapidity distribution of the di-tau system. . . . .	93
3.30.	Event selection for vector-boson fusion: Transversal momentum of the di-tau system. . . . .	94
3.31.	Event selection for vector-boson fusion: The invariant mass of the di-tau system. . . . .	95
3.32.	Event selection for the gluon-fusion process: Vertex significance. . . . .	96
3.33.	Event selection for the gluon-fusion process: Separation of the tau pair in the $(\eta, \phi)$ -plane. . . . .	97
3.34.	Event selection for the gluon-fusion process: The angle enclosed by the di-tau system. . . . .	98
3.35.	Event selection for the gluon-fusion process: Transversal momentum of the $p_T$ -leading $a_1$ . . . . .	98
3.36.	Event selection for the gluon-fusion process: Weighted transversal momentum of the di-tau system. . . . .	99
3.37.	Event selection for the gluon-fusion process: Invariant mass of the di-tau system. . . . .	100
3.38.	Invariant mass of tau pairs passing the event selection. . . . .	103
3.39.	Reconstructed di-tau mass from $Z^0$ decays. . . . .	104
3.40.	Fit of a gaussian distribution to the reconstructed di-tau mass. . . . .	104
3.41.	The signal significance $S_{cL}$ with $\mathcal{L} = 30 \text{ fb}^{-1}$ and the required luminosity for $S_{cL} = 1$ as a function of the Higgs boson mass. . . . .	106
3.42.	Illustration of the Higgs decay into a tau pair. One tau decays into a 3-prong, the other one into a 1-prong. . . . .	107
3.43.	Enhanced signal significance $S_{cL}$ with $\mathcal{L} = 30 \text{ fb}^{-1}$ as a function of the Higgs boson mass. . . . .	108
A.1.	Propagation of a charged particle in a solenoid magnetic field along the $z$ -direction. . . . .	111
B.1.	Distance between the reconstructed and generated jets. . . . .	113
B.2.	Transversal momentum of tau decays. . . . .	114
B.3.	Pseudorapidity of tau decays. . . . .	115
B.4.	Basic kinematics of the $Z^0$ boson. . . . .	115



# Tables

2.1. Overview of the degrees of freedom within the decay $\tau \rightarrow 3\pi^\pm + \nu_\tau$ . . . . .	31
2.2. Event-based efficiencies of the kinematic fit for several tau-production processes and fake tau leptons from the enriched QCD dataset. . . . .	52
2.3. Decay-based efficiencies of the quality criteria for several tau-production processes and fake tau leptons from QCD jets. . . . .	53
3.1. Number of generated events and the expected cross sections of the analyzed datasets. . . . .	64
3.2. Trigger requirements for single-tau and double-tau events. . . . .	73
3.3. Trigger efficiencies of the analyzed processes. . . . .	75
3.4. Decay-based efficiencies of the single tau selection. . . . .	79
3.5. Event-based efficiencies of the single tau selection. . . . .	80
3.6. Efficiency of the tau pair selection. . . . .	81
3.7. Efficiencies of the single jet selection. . . . .	87
3.8. Jet-based efficiencies of the jet pair selection. . . . .	89
3.9. Efficiency of the quality jet pair selection. . . . .	90
3.10. Efficiency of the custom jet pair selection. . . . .	90
3.11. Efficiency of the VBF selection. . . . .	95
3.12. Efficiency of the GF selection. . . . .	100
3.13. Summary of the VBF selection chain. . . . .	101
3.14. Summary of the GF selection chain. . . . .	101
3.15. Matching purity of the selected tau pairs. . . . .	102
3.16. Resolution of the reconstructed di-tau mass. . . . .	103



## Chapter 1

# Introduction

The knowledge of elementary particles and their interactions established the Standard Model of particle physics. Its predictions have been extensively tested with elaborated experiments. The model is very successful in describing the observed physics.

One of the unsolved fundamental questions is the origin of mass. A common answer is defined by the Higgs mechanism. A new undiscovered particle is introduced, the Higgs boson. The proof of its existence will be a milestone for particle physics and will confirm the validity of the Standard Model.

A new energy regime can be explored with the start of the Large Hadron Collider (LHC). Its experiments, like the Compact Muon Solenoid, aim to probe our understanding of nature at unexplored scales. One of the major tasks of the LHC experiments is the investigation of electroweak symmetry breaking. The LHC has the potential to discover the Higgs boson over the entire predicted mass range. The Standard Model is expected to be an effective theory up to an energy scale, which can be reached at the LHC. Physics beyond the current model is awaited.

The colliding protons with energies up to 14 TeV create a challenging hadronic environment. This complicates the identification and reconstruction of particles with hadronic decay modes. One example is the tau lepton, a particle that is expected to have a large coupling to the Higgs boson.

This thesis is divided into three major parts starting with an overall introduction in the current Chapter. The Higgs boson in the Standard Model and its prominent extensions are briefly discussed. The Large Hadron Collider and the experiments are introduced focussing on the Compact Muon Solenoid (CMS). The second part (Chapter 2) covers the kinematic reconstruction of tau leptons. Their unambiguous identification opens an important decay channel for light Higgs bosons, which is discussed in the last part, Chapter 3. It describes a Monte Carlo study to reconstruct Higgs bosons from hadronic tau decays with CMS.

The typical conventions of high energy physics are used in this thesis, considering  $\hbar = c = 1$ . Only explicitly denoted numeric values ignore this simplification to retain compatibility to the International System of Units (SI).

## 1.1. The Higgs Boson

The Standard Model of particle physics (SM) describes elementary particles and their interactions. It combines the electroweak and strong forces expressed as renormalized gauge field theories. The Glashow-Salam-Weinberg Model (GSW) [1-3] describes the interaction

of fermions mediated by gauge bosons as Yang-Mills theory [4]. Additionally, Quantum Chromodynamics (QCD) covers the interactions of colored quarks and gluons [5–9]. The corresponding Lagrangian is invariant under transformations of the weak isospin  $SU(2)_L$ , the weak hypercharge  $U(1)_Y$ , and the strong color rotation  $SU(3)$ .

The Lagrangian of the resulting  $SU(3)_C \times SU(2)_L \times U(1)_Y$  gauge group does not include particle masses. The Higgs mechanism described in Section 1.1.1 defines a framework of electroweak symmetry breaking to generate the masses of fermions and bosons while preserving the gauge invariance of the theory. The introduced Higgs field represents one additional massive particle, the Higgs boson, which has not yet been discovered.

The basic properties of the proposed Higgs boson are discussed in Section 1.1.2 and Section 1.1.3. Theoretical and experimental constraints on its mass, which is not predicted by the Standard Model, are summarized. An overview of the searches for the Higgs boson is given in Section 1.1.4.

### 1.1.1. The Higgs Mechanism in the Standard Model

The most frequently studied model of electroweak symmetry breaking is the Higgs mechanism. It was initially proposed in 1964 by Peter Higgs and independently by Francois Englert and Robert Brout [10–14].

The massive gauge bosons of the weak interaction cause divergent loop diagrams and destroy the renormalizability of the theory. Furthermore, the unitarity bounds of scattering amplitudes are not conserved up to high energies. The Higgs mechanism solves these problems by introducing the fundamental Higgs field  $\phi$ . In the Standard Model it consists of one self-interacting complex  $SU(2)$  doublet of two scalar fields

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_3 + i\phi_4 \\ \phi_1 + i\phi_2 \end{pmatrix} . \quad (1.1)$$

The potential  $V(\phi)$  is added to the SM Lagrangian and can be chosen as

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 . \quad (1.2)$$

It contains an imaginary parameter  $\mu$  and a real parameter  $\lambda$ , which has to be positive to achieve  $V(\phi \rightarrow \pm\infty) \rightarrow +\infty$ . The four real scalar fields  $\phi_i$  with  $i$  from 1 to 4 belong to a  $SU(2) \times U(1)$  multiplet to preserve gauge invariance.

The potential  $V(\phi)$  is illustrated in Figure 1.1 for three different choices of the parameter  $\mu$ . To simplify the depiction the fields  $\phi_3$  and  $\phi_4$  are ignored here. The potential  $V(\phi)$  satisfies the required symmetry under  $\phi \rightarrow -\phi$ . For  $\mu^2 \geq 0$  the ground state of the potential is zero. This corresponds to a vanishing vacuum expectation value. If the value of  $\mu^2$  decreases below zero, a set of non trivial minima satisfies the equation

$$\begin{aligned} \frac{\partial V(\phi)}{\partial \phi^\dagger \phi} &= \mu^2 + 2\lambda \phi^\dagger \phi \equiv 0 \\ \Rightarrow \phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2 &= -\frac{\mu^2}{\lambda} . \end{aligned} \quad (1.3)$$

This equation defines a circle with radius  $v \equiv \sqrt{-\mu^2/\lambda}$ , which confines the possible gauge.

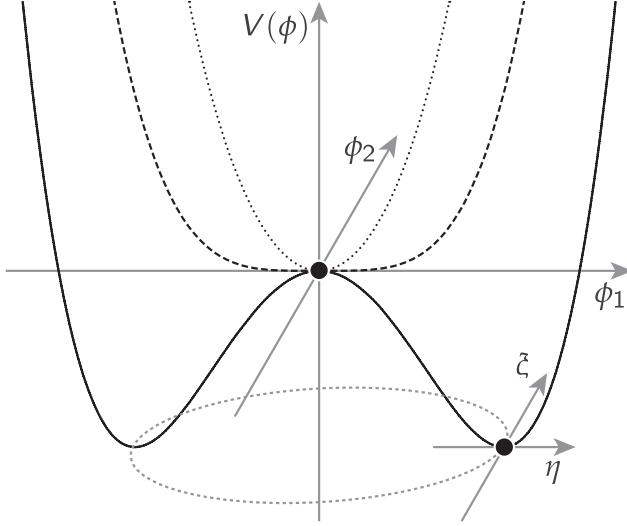


Figure 1.1.: Higgs Potential  $V(\phi)$  (ignoring  $\phi_3$  and  $\phi_4$ ) for three different values of  $\mu^2$ . The dashed line corresponds to the critical value  $\mu_{\text{crit}}^2 = 0$ . If  $\mu^2 < \mu_{\text{crit}}^2$ , the non trivial minima satisfy  $\phi_1^2 + \phi_2^2 = \mu^2 / \lambda^2$  depicted by the dashed circle.

One reasonable but arbitrary choice accounting for (1.3) is to set

$$\begin{aligned} \phi_2 &= \phi_3 = \phi_4 = 0 \\ \Rightarrow \phi_1^2 &= v^2 \quad . \end{aligned} \quad (1.4)$$

The phase symmetry is hidden, if a ground state is selected, but still preserved by the potential. As shown in [15], this spontaneous symmetry breaking leads to a renormalisable theory.

Translating the field to its minimum and expanding around the non vanishing vacuum expectation value yields

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h \end{pmatrix} \quad . \quad (1.5)$$

This introduces four new fields, which correspond to the real scalar Higgs field  $h$  and to three Goldstone scalars. These massless bosons occur according to the Goldstone theorem in every continuous symmetry of a physical system, which is spontaneously broken [16, 17]. They can be identified as excitations without resistance along the  $\xi$ -coordinate of Figure 1.1, whereas the Higgs field corresponds to the  $\eta$ -dimension.

Substituting these fields defined by the chosen gauge into the Lagrangian results in three massive gauge fields, the vector bosons with their mass terms, and one remaining massive scalar  $h$ . There is no mass term for the photon. It stays massless. The Goldstone bosons represent the longitudinal degrees of freedom of the vector bosons. The masses of fermions are generated through Yukawa interaction with the Higgs field.

The Higgs mechanism uses the described gauge in (1.4) and arranges the renormalizable interactions to lead to a vanishing charged component  $\phi^+$  of the initial doublet in (1.1) and a residual neutral part  $\phi^0$ . This field has a vacuum expectation value of

$$v = (\sqrt{2}G_F)^{-1/2} \approx 246 \text{ GeV} \quad . \quad (1.6)$$

It depends on the well-measured Fermi constant  $G_F$ . The mass of the corresponding Higgs boson is

$$m_H = v\sqrt{2\lambda} \quad . \quad (1.7)$$

As the self-coupling parameter  $\lambda$  is neither determined by the theory nor measured yet, the particle mass of the Higgs boson  $m_H$  remains unpredicted.

### 1.1.2. Basic Properties of the Standard Model Higgs Boson

In the SM the mass of the Higgs boson is only indirectly restricted by theory. One important constraint is the upper bound of  $m_{H^0} \sim 1 \text{ TeV}/c^2$  derived from unitarity arguments [18]. Further lower and upper bounds can be obtained, which depend on the cutoff energy scale  $\Lambda$ , at which new physics beyond the Standard Model becomes relevant. These bounds arise from the perturbativity of the Higgs self-coupling, the stability of the electroweak vacuum, and the fine-tuning in radiative corrections to the Higgs boson mass. A detailed summary of the theoretical mass restrictions can be found in [19].

The couplings of the Higgs boson to the particles of the Standard Model are determined by the theory. The Yukawa interaction to fermions  $f$  leads to a coupling  $g$  proportional to their masses. The coupling to bosons ( $V$  denotes  $W^\pm$  or  $Z^0$ ) depends on the squared boson mass. The Higgs mass sets the absolute scale of these couplings to fermions and bosons and determines the trilinear and quartic Higgs self couplings. According to the Feynman diagrams in Figure 1.2 this yields the possible single-vertex couplings [19]

$$g_{H\bar{f}f} = \frac{m_f}{v}, \quad g_{HVV} = \frac{2m_V^2}{v}, \quad g_{HHV} = \frac{2m_V^2}{v^2}, \quad g_{HHH} = \frac{3m_H^2}{v}, \quad g_{HHHH} = \frac{3m_H^2}{v^2} \quad . \quad (1.8)$$

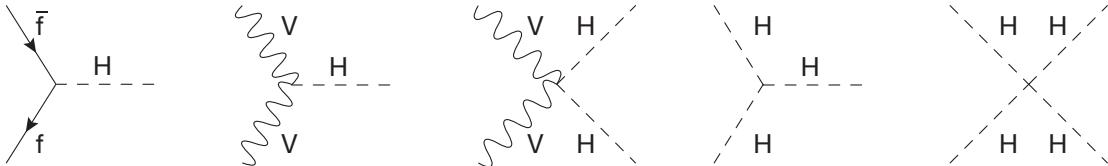


Figure 1.2.: Single-vertex couplings of the Standard Model Higgs boson.

The indices of the coupling  $g$  list the particles involved in the process. The parameter  $v$  denotes the vacuum expectation value from (1.6).

Due to the strong mass dependence of the coupling, the Higgs production is dominated by interactions with the heavy vector bosons  $W^\pm$  and  $Z^0$  and the third generation of quarks and leptons. At  $pp$ -colliders the Higgs boson is produced from gluon or quark pairs. The corresponding Feynman diagrams are depicted in Figure 1.3. The first diagram shows the gluon fusion ( $gg \rightarrow H$ ) mediated at lowest order by a virtual  $t\bar{t}$  loop. If only the depicted leading-order contribution is considered, the resulting Higgs Boson obtains no transversal momentum. The second process is the vector-boson fusion ( $qq \rightarrow qqH$ ) producing a Higgs boson with significant transversal momentum. It is accompanied by two quarks with small angles to the initial flight direction of the protons. The remaining two diagrams describe the associated production with either a weak gauge boson (Higgs Strahlung,  $q\bar{q} \rightarrow VH$ ) or with heavy quarks ( $gg \rightarrow t\bar{t}H$ ).

Figure 1.4 summarizes the Higgs boson cross sections for proton-proton collisions at the LHC (see Section 1.2). The major production processes are shown as function of the Higgs

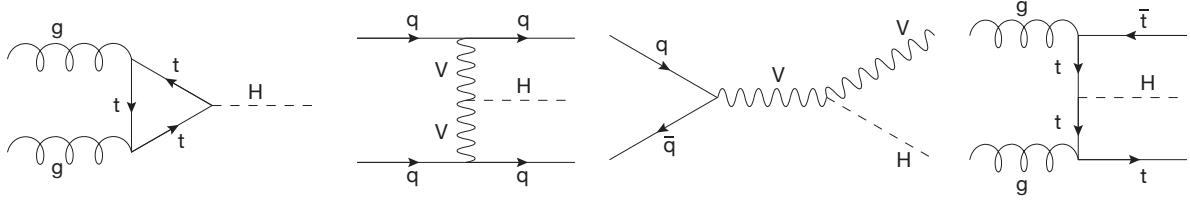


Figure 1.3.: Relevant Feynman diagrams for Standard Model Higgs production from proton-proton collisions at leading order. The vector bosons  $W^\pm$  and  $Z^0$  are denoted as  $V$ .

mass. They increase by a factor of about 5 from a center-of-mass energy of  $\sqrt{s} = 7 \text{ TeV}$  to  $\sqrt{s} = 14 \text{ TeV}$ . The relative contributions stay almost constant. Due to the large Yukawa coupling to the top quark and the gluon density, the depicted mass range is dominated by gluon fusion [21]. For small Higgs masses it has a cross section, which is a factor of 10 higher than the second highest from vector-boson fusion. The associated production processes have a considerable rate only for a light Higgs boson. All processes decrease in frequency for increasing Higgs masses.

The theoretical uncertainties are depicted in the error bands in Figure 1.4 labeled with the precision of the individual calculations for every process [20]. The cross section for gluon fusion depends on higher order QCD effects from heavy quark loops, which contribute up to 100 % at next-to-leading order (NLO). The major QCD corrections are known up to next-to-next-to-leading order (NNLO) in the heavy top-mass limit. They further increase the NLO result at the level of 10 – 20 %. Electroweak corrections give only small contributions at NLO calculation. The vector-boson fusion and the Higgs radiation from gauge bosons are known at NNLO of QCD and NLO of electroweak corrections. The associated production with top quarks is calculated at NLO of QCD.

The leading-order couplings of the Higgs boson in (1.8) determine its decay properties. The branching fractions and the total decay width of the Standard Model Higgs boson are shown in Figure 1.5. The depicted mass range corresponds to the allowed Higgs region. The calculations account for QCD and electroweak corrections.

The entire region below  $m_H = 135 \text{ GeV}/c^2$  is dominated by Higgs decays into  $b\bar{b}$  pairs and above into pairs of vector bosons, especially  $WW$ . Above the  $t\bar{t}$  threshold the Higgs decays also into top pairs. For lower Higgs masses various decay channels are possible. Below the dominant decay into  $b\bar{b}$ , decay modes with pairs of taus, gluons, and charm quarks are present. Decays into two photons, mediated by fermion and  $W$  loops, are suppressed to a few per mille. The branching ratio of the decay into tau pairs varies from nearly 10 % at  $m_H = 90 \text{ GeV}/c^2$  to about 1 % at  $m_H = 155 \text{ GeV}/c^2$  and rapidly drops for higher masses. This decay will be addressed in the analysis of Chapter 3.

Combining all possible decay channels, the total decay width of the Higgs Boson  $\Gamma_{H^0}$  is obtained (right plot in Figure 1.5). For light Higgs bosons the width is very small with  $\Gamma(H^0) \leq 10 \text{ MeV}$ . It rapidly raises up to a few hundred MeV at the  $ZZ$  threshold. At higher masses the width increases asymptotically and amounts to almost  $\Gamma_{H^0} = 1 \text{ TeV}$  at  $m_{H^0} = 1 \text{ TeV}/c^2$ .

A detailed study of the Standard Model Higgs production and decay at the LHC can be found in [20].

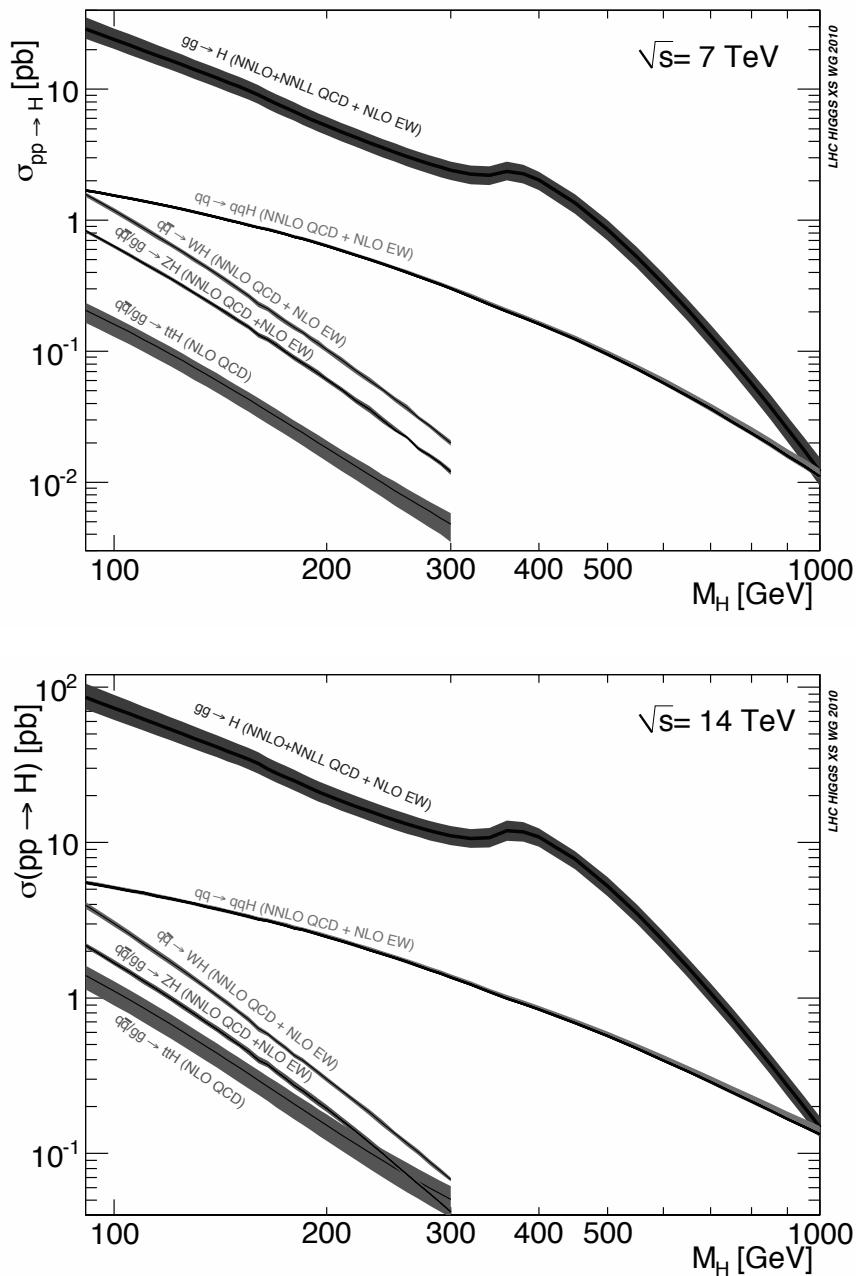


Figure 1.4.: Production cross sections of the Standard Model Higgs boson as function of the mass [20]. The error bands depict the theoretical uncertainties of the various production mechanisms. The upper plot considers proton-proton collisions at 7 TeV, the lower at 14 TeV.

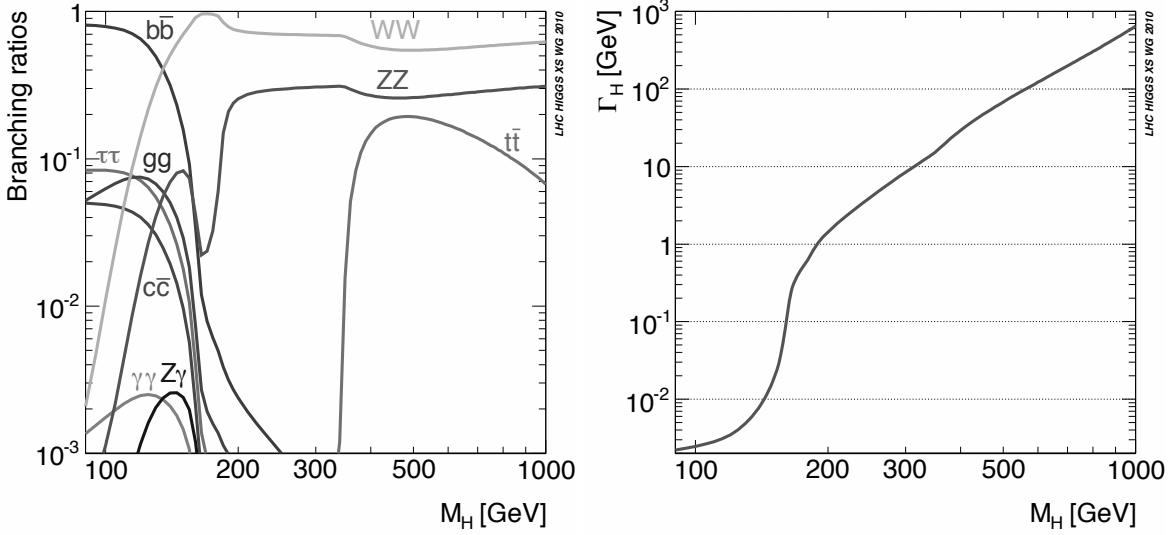


Figure 1.5.: Branching fractions (left) and the total decay width (right) of the Standard Model Higgs boson depending on its mass [20].

### 1.1.3. The MSSM Higgs Boson

The Standard Model Higgs mechanism discussed in Section 1.1.1 introduces a single SU(2) doublet, which results in one Higgs boson after spontaneous symmetry breaking. This mechanism causes quadratic divergencies arising from radiative corrections to the SM Higgs boson mass. There are several models exploring different Higgs mechanisms to cancel these divergencies. They extend the validity of the Standard Model beyond the TeV energy scale.

One famous model is the minimal supersymmetric extension of the Standard Model (MSSM) [22, 23]. Supersymmetry expands the Standard Model by proposing an additional symmetry between fermions and bosons. One supersymmetric partner, which differs in spin by 1/2, is assigned to each degree of freedom of the Standard Model. The MSSM assumes a minimal gauge group, a minimal particle content, and *R*-parity conservation, which leaves the lightest supersymmetric particle stable (details in [24]).

In the MSSM the electroweak gauge bosons acquire mass by a Higgs mechanism developed from two doublets with opposite hypercharges. It is a type of the general two-Higgs doublet model (2HDM). Starting from eight degrees of freedom of the scalar fields, three are absorbed by the gauge bosons  $W^\pm$  and  $Z^0$ . This leads to five residual Higgs bosons, two neutral CP-even states (the light  $h^0$ , the heavy  $H^0$ ), one neutral CP-odd state ( $A^0$ ), and the charged pair  $H^\pm$ .

The MSSM introduces a large set of new parameters. They have only small impact on the Higgs sector, which can be parameterized by the four Higgs masses and two angles. The mixing angle  $\alpha$  diagonalizes the neutral CP-even sector and the angle  $\beta$  expresses the ratio of the two vacuum expectation values  $v_u$  and  $v_d$  assigned to the Higgs doublets

$$\tan \beta = v_u / v_d . \quad (1.9)$$

The Higgs sector is restricted by its supersymmetric structure leaving only two unconstraint parameters at lowest order in terms of the gauge couplings. Usually one selects  $\tan \beta$  and  $m_{A^0}$ , the mass of the CP-odd  $A^0$  boson.

Once these two values are determined, all other parameters can be calculated up to higher-order corrections (complete derivation in [25]). The Higgs masses are given as

$$m_{H^\pm}^2 = m_{A^0}^2 + m_{W^\pm}^2 \quad (1.10)$$

$$m_{H^0, h^0}^2 = \frac{1}{2} \left[ m_{A^0}^2 + m_{Z^0}^2 \pm \sqrt{(m_{A^0}^2 + m_{Z^0}^2)^2 - (2m_{Z^0}m_{A^0}\cos 2\beta)^2} \right] \quad . \quad (1.11)$$

As a consequence, the mass of the lightest Higgs boson  $m_{h^0}$  is restricted already at tree level by the  $Z^0$  boson mass [26, 27]

$$m_{h^0} \leq m_{Z^0} |\cos 2\beta| \quad . \quad (1.12)$$

Considering important radiative and loop corrections in the calculation of the CP-even Higgs bosons [28], one obtains an upper bound of  $m_{h^0} \lesssim 135 \text{ GeV}/c^2$  [29].

In the decoupling limit [30], where  $m_{A^0} \gg m_{Z^0}$ , only one Higgs boson  $h^0$  remains light compared to the scale of the electroweak symmetry breaking. Its mass can be expressed as  $m_{h^0}^2 \simeq m_{Z^0}^2 \cos^2 2\beta$ . The properties of  $h^0$  become very similar to those of the SM Higgs boson, which complicates their distinction. The masses of the heavier Higgs bosons are nearly degenerate  $m_{H^\pm} \simeq m_{H^0} \simeq m_{A^0}$ .

The couplings of the MSSM Higgs bosons to gauge bosons and fermions depend strongly on  $\cos(\beta - \alpha)$  at tree level. At this order there is no coupling between a  $A^0$  or  $H^\pm$  Higgs boson to the electroweak bosons and the coupling to fermions distinguishes between down-type and up-type fermions. The couplings are strongly affected by radiative corrections. Their values are usually calculated for several benchmark scenarios of the supersymmetry breaking parameters [31]. An overview of the significant couplings and the impact of higher-order corrections can be found in [25].

Important decay modes of the neutral MSSM Higgs bosons are similar to those already discussed in Section 1.1.2. In the decoupling limit the light  $h^0$  decays with the same rates as the Standard Model Higgs boson. The branching fractions are only modified if the Higgs mass is heavy enough to open decay modes into supersymmetric particles [32]. As  $m_{h^0} \lesssim 135 \text{ GeV}/c^2$ , which is below the  $W^+W^-$  threshold, the decays into bottom quarks and tau pairs are dominant. Charged Higgs bosons decay into a tau lepton and a neutrino ( $H^\pm \rightarrow \tau^\pm \nu$ ) and into  $t b$  if the mass exceeds the threshold  $m_t + m_b$ . Decays into lighter quarks have small branching fractions. For small  $\tan \beta$  the heavy Higgs bosons significantly decay into lighter ones, according to

$$H^\pm \rightarrow W^\pm h^0, W^\pm A^0, \quad H^0 \rightarrow h^0 h^0, A^0 A^0, Z^0 A^0, \quad \text{and } A^0 \rightarrow Z^0 h^0 \quad . \quad (1.13)$$

The decay rates for the heavy Higgs bosons and also for  $h^0$  outside the decoupling limit can differ largely from the Standard Model branching fractions. Especially large  $\tan \beta$  increase the coupling to down-type fermions. This strongly enhances the decay rates into bottom-quark and tau-lepton pairs.

#### 1.1.4. Searches for the Higgs Boson

The Higgs boson is the last particle of the Standard Model, which has not yet been observed experimentally. Direct searches for its decay products are carried out at colliders, which have sufficient energy to enter the mass regime of the Higgs boson. Indirect measurements are sensitive to Higgs contributions through loop corrections.

A wide mass range is excluded with data collected at the LARGE ELECTRON-POSITRON COLLIDER (LEP) [33]. The dominant mechanism for SM Higgs production at LEP is the Higgs-Strahlung ( $e^+e^- \rightarrow H^0Z^0$ ). The explored mass range depends on the center-of-mass energy  $\sqrt{s}$  as the cross section for Higgs production drops rapidly above the threshold  $m_{H^0} = \sqrt{s} - m_{Z^0}$ . The data can be divided into collisions at the Z-pole (LEP1) and center-of-mass energies of up to  $\sqrt{s} = 209$  GeV (LEP2). Independent analyses at each collision energy and for various decay channels are combined by the four LEP experiments. The absence of an evident signal results in a lower bound on the Higgs mass of

$$m_{H^0} \geq 114.4 \text{ GeV}/c^2 \quad (1.14)$$

at 95 % confidence level [34]. The LEP measurements are also sensitive to the production of neutral Higgs bosons in the MSSM. The combined analysis of all four experiments sets upper bounds on the cross sections of Higgs-like event topologies [35]. For common benchmark scenarios it provides large exclusions in the parameter space spanned by  $\tan\beta$  and  $m_{h^0}$ . Charged Higgs bosons with masses below about  $75 \text{ GeV}/c^2$  are excluded in distinct decay channels [36].

The Tevatron at Fermilab [37] continues the Higgs search in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV. The benefit of a higher center-of-mass energy w.r.t. the LEP searches is diminished by a significantly smaller signal-to-background ratio and systematic uncertainties on the background rates in the hadronic environment. The experiments CDF and DØ combine their searches based on integrated luminosities up to  $5.9 \text{ fb}^{-1}$  at CDF and up to  $6.7 \text{ fb}^{-1}$  at DØ. They exclude a Standard Model Higgs boson within a mass range of

$$158 \text{ GeV}/c^2 < m_{H^0} < 175 \text{ GeV}/c^2 \quad (1.15)$$

at 95 % confidence level [38]. The analyses consider Higgs production in association with a vector boson ( $q\bar{q} \rightarrow W^\pm H^0$  or  $p\bar{p} \rightarrow Z^0 H^0$ ), through gluon fusion ( $gg \rightarrow H$ ), and through vector-boson fusion ( $q\bar{q} \rightarrow q\bar{q} H^0$ ). A summary of the MSSM Higgs boson searches at the Tevatron and the impact of different benchmark scenarios is discussed in [39].

Indirect searches for the Higgs boson add further mass constraints. A global fit of the Standard Model is performed [40]. It combines the Z-pole data measured at LEP and the STANFORD LINEAR COLLIDER (SLC) with collisions at higher energies recorded at LEP and the Tevatron. The leading contribution of the Higgs mass to the electroweak precision measurements through loop corrections is logarithmic. The fit results in a preferred Higgs mass of  $m_{H^0} = 89^{+35}_{-26} \text{ GeV}/c^2$ . The errors consider the experimental and theoretical uncertainties at 68 % confidence level. The fit yields an upper bound of  $m_{H^0} < 158 \text{ GeV}$  at 95 % C.L. [40] not taking the direct limits into account. These values strongly depend on the masses of the top quark and the  $W^\pm$  boson, which are obtained from direct measurements:  $m_{W^\pm} = 80.399 \pm 0.023 \text{ GeV}/c^2$  and  $m_t = 173.3 \pm 1.1 \text{ GeV}/c^2$ .

The current status of the experimental mass limits is summarized in Figure 1.6. It combines the results from the direct and indirect measurements. The vertical bands give the exclusion regions from direct searches at 95 % confidence level. The solid curve corresponds to the  $\Delta\chi^2 = \chi^2 - \chi^2_{min}$  of the global fit as function of the Higgs mass. It is enveloped by a band, which shows the theoretical uncertainties of the input parameters. The two other curves are obtained for modified fits. The dashed curve uses a different method to obtain the vacuum polarization of the photon field. The dotted curve incorporates also low energy measurements from atomic parity violation, Møller scattering and neutrino-nucleon scattering.

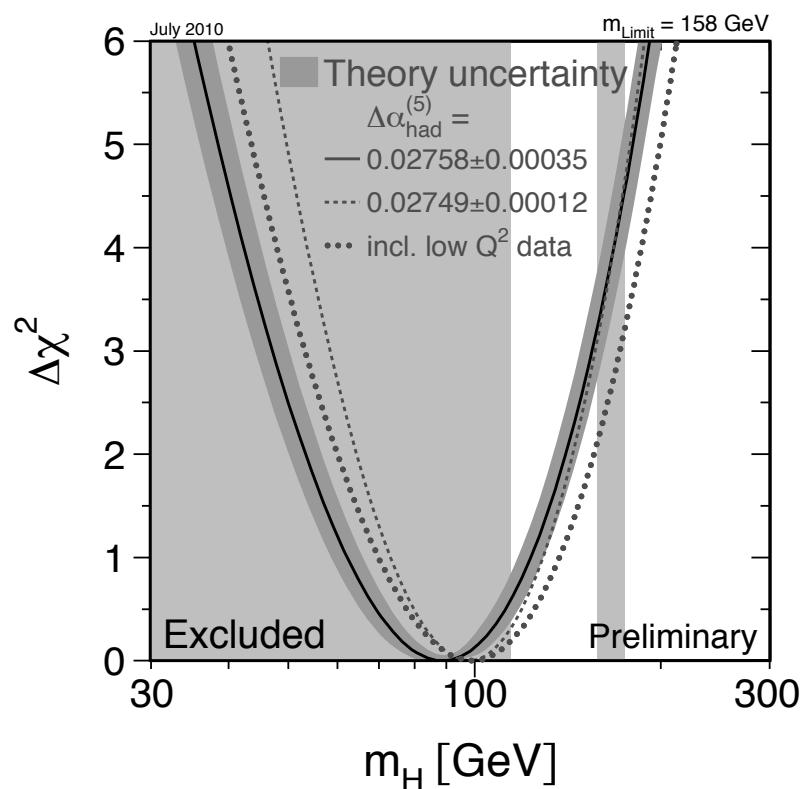


Figure 1.6.: Most likely mass of the Standard Model Higgs Boson derived from a global fit [40]. The vertical bands denote the excluded regions from direct measurements. See the text for further information.

A complete coverage of the allowed mass range will be provided at the Large Hadron Collider, which is introduced in Section 1.2. The experiments aim for a  $5\sigma$  discovery of the Standard Model Higgs boson combining the results of various decay modes. The expected discovery potential for MSSM Higgs bosons varies for the studied benchmark scenarios. Already with a moderate amount of data, neutral Higgs bosons can be observed or excluded for a large area within the parameter space of  $\tan \beta$  and  $m_{A^0}$ . The current limits on the charged Higgs bosons from Tevatron are expected to be improved at the LHC.

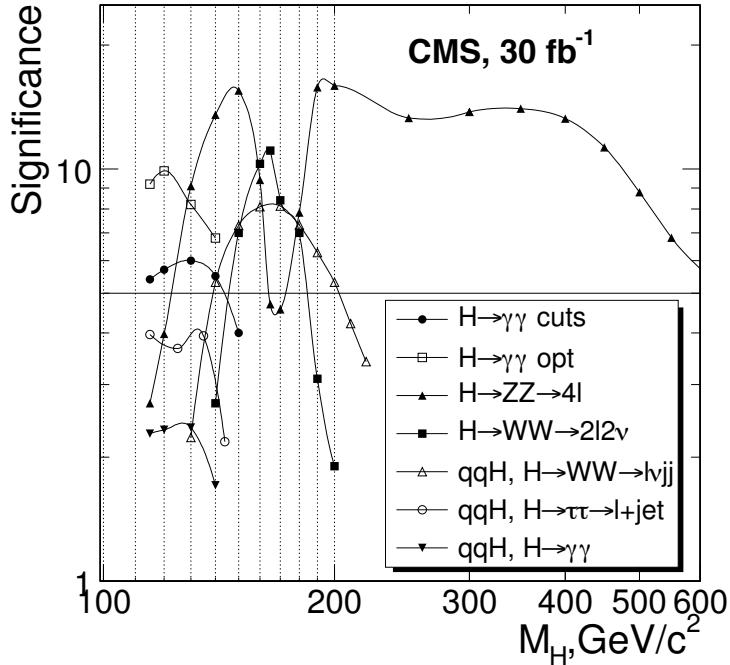


Figure 1.7.: The expected SM Higgs boson significance as a function of its mass [41]. The different curves for various Higgs boson production and decay channels correspond to  $30 \text{ fb}^{-1}$  of the integrated luminosity collected with CMS at a center-of-mass energy of  $14 \text{ TeV}$ .

The significance for the unambiguous identification of a Standard Model Higgs boson with the CMS detector (see Section 1.3) is depicted in Figure 1.7. The major discovery channels are shown independently as a function of the Higgs mass. A detailed summary of the analyzed decay channels and their sensitivity towards the Higgs searches (SM and MSSM) is given in [41, 42]. These studies consider the design center-of-mass energy of  $\sqrt{s} = 14 \text{ TeV}$ . The CMS sensitivities for SM Higgs boson searches at  $7 \text{ TeV}$  collisions are updated in [43].

The Higgs decay into tau pairs plays an important role. It is expected to provide a large significance for light Higgs bosons at the LHC exploiting the clean signature from isolated leptons in the final state. The various decay modes of the tau lepton are covered by independent analyses. The full hadronic mode, where both tau leptons decay into hadrons, is challenging in the hadronic environment at the LHC. It requires an efficient and clean tau identification. One approach is the kinematic reconstruction of hadronic tau decays introduced in Chapter 2.

## 1.2. The Large Hadron Collider

The LARGE HADRON COLLIDER (LHC) is the world's largest particle collider. This superconducting particle accelerator was built at CERN, the European Laboratory for Particle Physics [44], in the 27 km long circular tunnel, which formerly hosted the  $e^+e^-$  collider LEP. Located at the Franco-Swiss border near Geneva and about 100 m below the surface, a ring of 1232 dipole magnets stores two counter-rotating hadron beams - either protons or lead ions - and collides them at four distinct interaction points. Around these points the LHC houses four major detectors within huge underground caverns: ALICE (A LARGE ION COLLIDER EXPERIMENT [45]), ATLAS (A TOROIDAL LHC APPARATUS) [46, 47], CMS (COMPACT MUON SOLENOID [41, 48]), and LHCb (LARGE HADRON COLLIDER BEAUTY [49]). The large experiments ATLAS and CMS are general-purpose detectors designed to cover a broad range of physics, whereas ALICE and LHCb are specialized to particular phenomena. The first two are accompanied by smaller detectors - TOTEM (TOTAL ELASTIC AND DIFFRACTIVE CROSS SECTION MEASUREMENT [50]) and LHCf (Large Hadron Collider forward [51]) - to study forward particles, which are only slightly bent from their circular orbit. An illustration of the LHC location and the arrangement of the experiments around the ring is depicted in Figure 1.8. The construction of the storage ring and its experiments has been finished and the first beams circulated in 2008. The physics program started with collisions at modest energies in 2009.

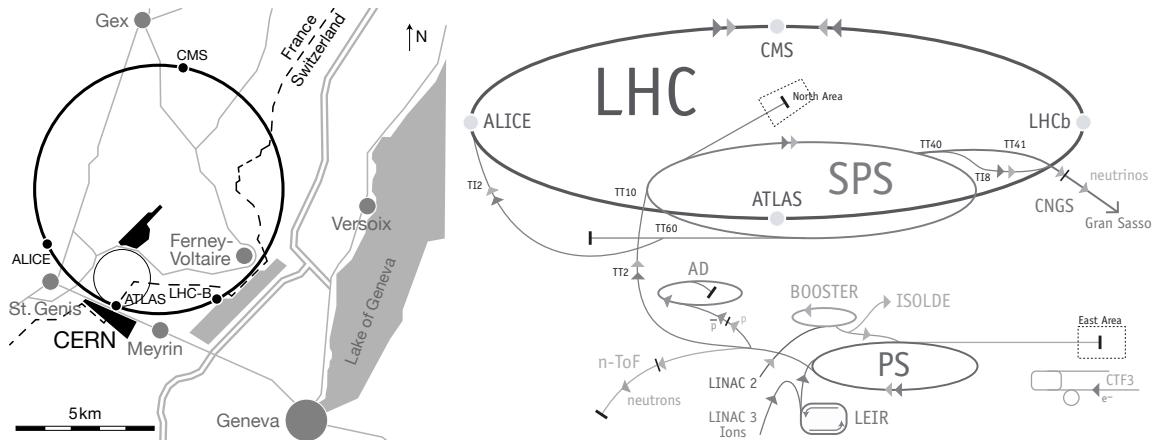


Figure 1.8.: Illustration of the CERN Large Hadron Collider. Left: Geographical location of the LHC and CERN sites near Geneva (based on [52]). Right: Schematic view of the LHC accelerator complex and its experiments (from [53]).

The LHC is designed to achieve proton-proton collisions with a center-of-mass energy of  $\sqrt{s} = 14 \text{ TeV}$  and a peak luminosity of  $\mathcal{L} = 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ . The instantaneous luminosity  $\mathcal{L}$  depends on the frequency  $f_{\text{rev}} = c/27 \text{ km} \approx 10^4 \text{ Hz}$ , the Lorentz factor  $\gamma_r$ , and  $A_T^{\text{eff}} = 4\pi\sigma_b^2$ , the effective transverse beam area with  $\sigma_b \approx 16 \mu\text{m}$

$$\mathcal{L} = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{A_T^{\text{eff}}} . \quad (1.16)$$

This determines a list of beam parameters, which is discussed in [54]. Each of the two rings are filled with  $n_b = 2808$  bunches containing  $N_b = 1.15 \cdot 10^{11}$  particles each. A nominal bunch spacing of 25 ns results in a beam current equal to 0.53 A and a stored energy per

nominal beam of 362 MJ. The beam energy of 7 TeV excludes the use of electrons (compared to LEP) due to synchrotron radiation. The implied beam intensities exclude the use of anti-protons as done at TEVATRON [37]. In order to bend the beams onto their circular orbit, a magnetic field of 8.4 Tesla is provided by superconducting magnets at a current of around 11 700 A operating at cryogenic temperatures of 1.9 K. These dipoles enclose two beam pipes consisting of continuous vacuum chambers at about  $10^{-10}$  Torr to avoid interactions between the beam and gas molecules. The stored energy in the magnets amounts to 11 GJ. Before entering the main LHC ring, the beams pass a segmented injector chain comprising Linac, Booster, PS and SPS (see Figure 1.8). Protons are finally injected into the LHC at 450 GeV, accelerated to their nominal energy, and stored up to 20 hours.

The LHC program is intended to analyze rare physics processes at unexplored energy scales. These events of interest are hidden by a huge rate of inelastic, non-diffractive collisions with a total cross section of about 60 mb. A comparison of the cross sections for different processes at the Tevatron and the LHC is depicted in Figure 1.9. The second  $y$ -axis represents the event

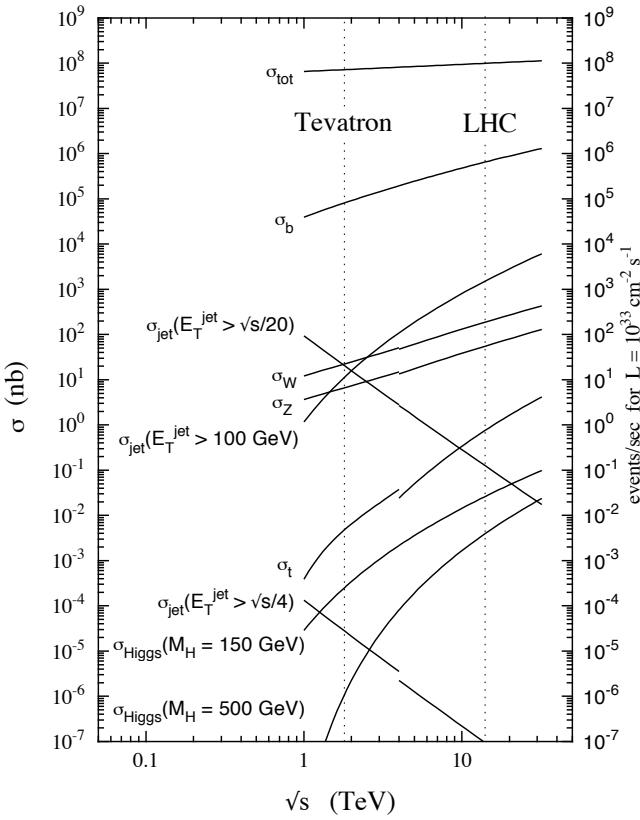


Figure 1.9.: Cross sections for various Standard Model processes depending on the center-of-mass energy  $\sqrt{s}$  of the  $p p$  or  $p \bar{p}$ -collisions [55]. The dashed, vertical lines indicate the design energies at the Tevatron and the LHC. The given event rates correspond to an instantaneous luminosity of  $\mathcal{L} = 2 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ .

rate  $N_{\text{event}}$  at an instantaneous luminosity of  $\mathcal{L} = 2 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ . The event rate is defined as the product of luminosity and cross section

$$N_{\text{event}} = \sigma \cdot \mathcal{L} . \quad (1.17)$$

The cross sections grow with rising center-of-mass energy  $\sqrt{s}$  of the primary hadron interaction as the parton density increases. At design luminosity the collision rate at the LHC will reach  $10^9$  Hz. Cross sections of interesting processes like  $\sigma_t$  are restrained by 8 orders of magnitude. One major challenge for the experiments is to identify such valuable events within this polluted environment on very short time scales.

The LHC has finished a period of data taking in 2010 and delivered  $pp$ -collisions with an integrated luminosity of more than  $47 \text{ pb}^{-1}$  at  $\sqrt{s} = 7 \text{ TeV}$  to the experiments. This enables the first physics analyses at an untested energy regime. The first results have been published.

### 1.3. The Compact Muon Solenoid

The hadronic environment at the LHC influences the design of the accompanying detectors. The large collision rate demands a fast response time of the detectors in the order of  $20 \text{ ns}$ . To resolve up to 1000 charged particles emerging from a single collision, a fine segmentation of the detectors is required. These detectors should cover the full solid angle around the interaction region. They have to withstand the high radiation doses.

The two general-purpose experiments at the LHC aim for a complete detection and reconstruction of rare events within this challenging environment. One of the two is CMS, the COMPACT MUON SOLENOID, located in Cessy, France. This thesis has been developed within the CMS collaboration. Thus, only CMS is described here.

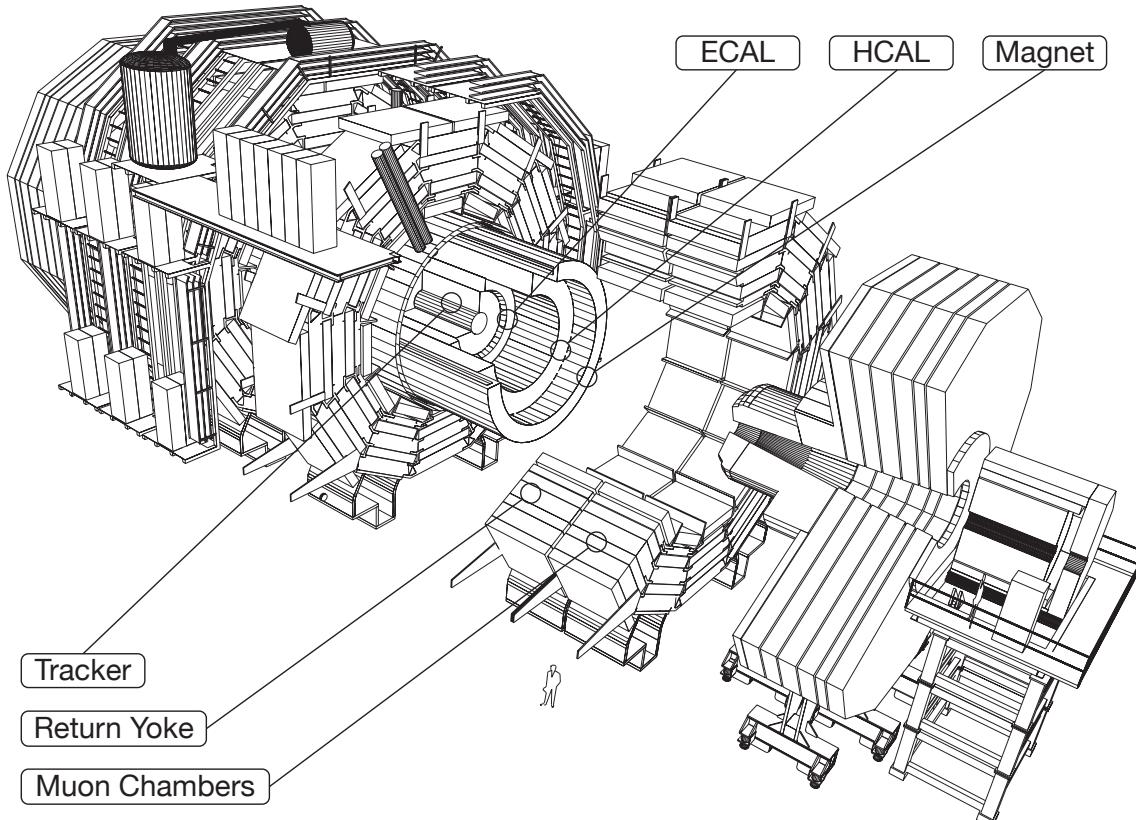


Figure 1.10.: Illustration of the CMS detector (based on [56]).

A schematic view of CMS is depicted in Figure 1.10. The detector is built around a superconducting magnet. This large solenoid ([57]) generates a field of up to  $4 \text{ T}$ . With its diameter of  $15 \text{ m}$ , a length of  $21 \text{ m}$ , and a weight of about  $14\,000 \text{ t}$ , CMS is a compact detector compared to ATLAS. The whole detector is centered around the nominal interaction point. It defines the origin of the CMS coordinate system with the  $x$ -axis pointing towards the center of the LHC and the  $y$ -axis vertically upwards. The  $z$ -axis completes a right-handed system pointing

westwards along the beam pipe, parallel to the magnetic field. The azimuth angle  $\phi$  lies in the  $x$ - $y$ -plane and the polar angle  $\theta$  is measured from the  $z$ -axis. It is often expressed in terms of the pseudorapidity defined as  $\eta = -\log [\tan \theta/2]$ . The cylindrical layout is divided into a barrel region for small values of  $|\eta|$  and endcaps that cover the forward region of angles up to  $|\eta| = 5$ .

Starting from the origin of this coordinate frame, the first sub-detector is the tracker. The fully silicon-based inner tracking system aims to measure the trajectories of charged particles bend by the magnetic field. The radius of the curvature determines their momenta. Moving further outwards, calorimeters measure the energy of particles by reconstructing their energy deposition in the traversed material. The scintillating electromagnetic calorimeter (ECAL, [58]) is optimized for detecting the electromagnetic showers of photons and electrons within its lead tungstate crystals ( $PbWO_4$ ). The subsequent hadron calorimeter (HCAL, [59]) detects hadronic showers. It is a sampling calorimeter, which consists of plastic scintillator tiles with embedded wavelength-shifting fibres and brass absorber. These sub-detectors are enclosed by the magnet coil. Outside, four layers of muon chambers are integrated into the iron return yoke of the magnet. Only muons are expected to reach the four layers of drift tubes, cathode strip chambers, and resistive plate chambers ([60]). A detailed description of CMS is given in [61].

All sub-detectors are linked to electronic read-out channels. To handle the enormous event rate at the LHC (see 1.2) and to reduce the amount of data to be stored, the observed events pass a complex trigger system. It will be discussed in 1.3.4.

The performance goals of CMS are driven by its primary physics tasks, which are extensively described in [48]. CMS is optimized to explore physics at the TeV scale and to analyze the mechanism of electroweak symmetry breaking. One major design goal is the complete coverage of Higgs bosons over the entire mass range. As muons are involved in important decay modes, this requires an unambiguous charge determination of muons, a relative momentum resolution of  $\Delta p/p \approx 10\%$  at a muon momentum of  $1\text{ TeV}/c$ , and a di-muon mass resolution of about 1% (at  $100\text{ GeV}/c^2$ ). The momentum resolution is achieved by combining the information from the tracking system and the muon stations. The energy loss of muons in the calorimeters is small. It is in the order of a few  $100\text{ MeV}$  for the ECAL and a few  $\text{GeV}$  for the HCAL. The ECAL has an energy resolution of  $\sigma_E/E < 0.5\%$  and a di-photon and di-electron mass resolution of 1% (at  $100\text{ GeV}/c^2$ ). The HCAL, located within the magnet coil, has a fine lateral segmentation of up to  $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$  providing a good di-jet mass resolution. Its hermetic coverage up to  $|\eta| < 5$  enables the reconstruction of the missing transversal energy  $E_T^{\text{miss}}$ .

The performance of the tracking system is discussed in 1.3.1. A description of the tau-lepton algorithm can be found in Section 2.4.1. The detailed reconstruction of other physics objects that are not essential for the current analysis are introduced in [48].

During the ongoing physics program, CMS has already collected  $pp$  and lead-ion collisions. The integrated luminosity of the initial proton run is depicted in Figure 1.11. CMS recorded  $43.17\text{ pb}^{-1}$  from  $47.03\text{ pb}^{-1}$  delivered by the LHC. Typical searches for the Standard Model Higgs boson need an integrated luminosity in the order of inverse femtobarns.

This thesis prepares the search for a Higgs boson via a pair of hadronically decaying tau leptons. The key aspects of CMS, on which this reconstruction is based are discussed in more detail in the subsequent sections.

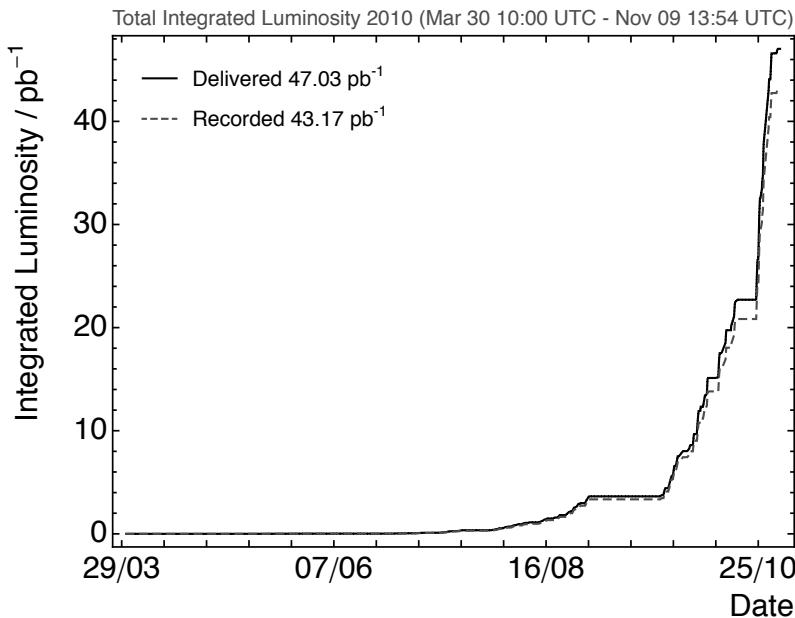


Figure 1.11.: Total integrated luminosity delivered to and recorded by CMS during stable beams at 7 TeV center-of-mass energy. Data obtained from [62].

### 1.3.1. The Tracking System

A vital task in the LHC physics program is the reconstruction of particle trajectories within the tracking system. The CMS tracker is a cylindrical detector with 5.8 m in length and a width of 2.6 m (see [63]). Its high granularity disentangles the large number of particles produced at the hadron collisions. The tracker provides an accurate momentum measurement of charged particles. Its design changes according to the charged particle flux, which decreases for larger radii. As shown in Figure 1.12, the detector is centered around the interaction point and consists of several subsystems (nomenclature explained in [64]). Divided into barrel and endcaps, it covers a polar angle up to  $|\eta| = 2.5$ .

The innermost part closest to the collision point has to withstand a charged particle flux of  $10^8 \text{ cm}^{-2} \text{ s}^{-1}$ . To measure the impact parameter of tracks and to resolve secondary vertices of primary particles with significant lifetime, it is located as close to the beam as possible. Starting at 4 cm from the beam axis, three layers of pixel detectors are placed. 66 million pixels of  $100 \times 150 \mu\text{m}^2$  each provide a spacial resolution of about  $10 \mu\text{m}$  in  $r\text{-}\phi$  and  $20 \mu\text{m}$  in  $z$ . At radii between 20 and 110 cm, 9.6 million silicon strips with sizes between  $10 \text{ cm} \times 80 \mu\text{m}$  and  $25 \text{ cm} \times 180 \mu\text{m}$  cover an active area of  $200 \text{ m}^2$ . This yields a single point resolution of up to  $23 \mu\text{m}$  in the  $r\text{-}\phi$  direction and  $23 \mu\text{m}$  in  $z$ . The material is minimized to reduce the interaction of particles in the tracker. A summary of the material budget is given in [48].

### 1.3.2. Track Reconstruction

From the signals collected by the CMS detector, particles are reconstructed (see 1.4.3 for a general description of the reconstruction chain). An important part is the track reconstruction. The algorithm can be divided into the following five processes.

The first step determines the position of hits and their uncertainty from charge clusters in the silicon sensors. The clusters are created from seed pixels and strips, which exceed a certain charge threshold.

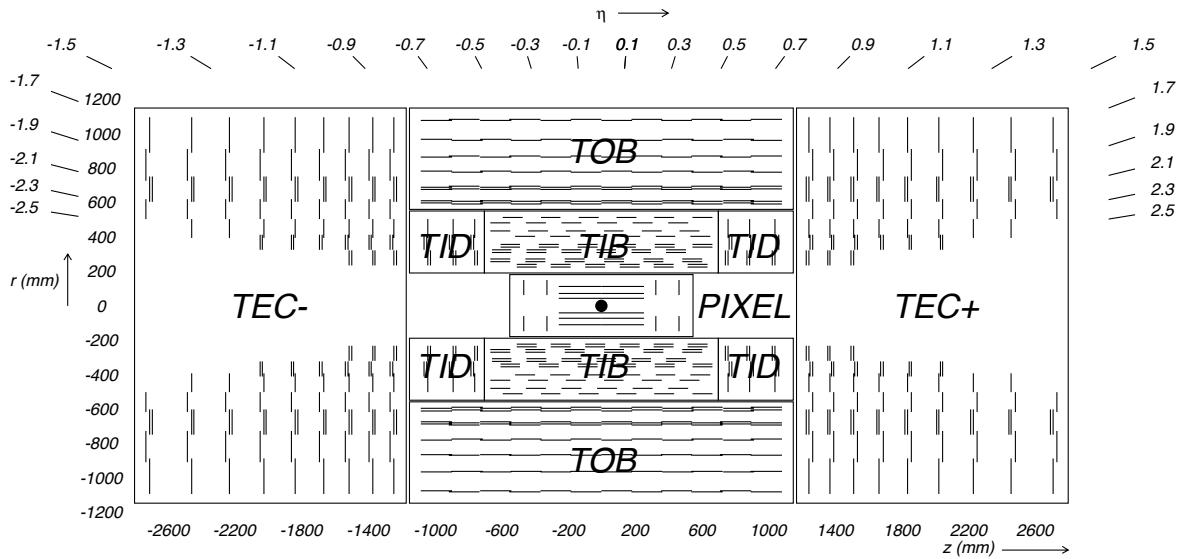


Figure 1.12.: The CMS Inner Tracker in  $r$ - $z$ -view (from [64]). Its layout in different subsystems and the pseudorapidity coverage are shown. Segments represent detector modules.

In the second step, track seeds are generated from the obtained hits or from other detector components. For a first estimate of all helix parameters of a track, a candidate has to have at least three points assigned. The number of hits can be enlarged by temporarily forcing the trajectory to originate from the beam line.

The pattern recognition defines the third step. It is based on a combinatorial Kalman filter method [65] and rebuilds trajectories by iteratively adding compatible hits from successive detection layers. The track is propagated through the whole tracker volume taking the magnetic field into account (the equations of motion are derived in Appendix A.1). The helix precision increases with each iteration.

Several tracks starting from the same seed or different seeds resulting into the same track can produce ambiguities. The fourth step resolves tracks sharing the same hits in favour of the best quality trajectory.

During the final step, each trajectory is refitted using a least-squares approach, combining a standard Kalman filter and smoother. First, a forward fit from the interaction point outwards removes approximations and biases of the seeding (e.g. possible beam line constraints) and accounts for the effects of energy loss and the deflection by multiple scattering. A second smoother fit starts from the outside and propagates back to the innermost layer. This determines the best final estimates of the track parameters at the surface associated with each hit and especially at the vertex.

This five-step tracking procedure is repeated iteratively. After each step, high quality tracks are stored and their associated hits are removed. A new step of track reconstruction starts with the remaining set of hits and looser requirements [64]. This iterative approach increases the reconstruction efficiency of the first iteration by up to 5 %.

More detailed information on the CMS track reconstruction can be found in [48, 66].

Figure 1.13 demonstrates the performance of the track reconstruction of the inner tracking system. On the left side the single track efficiency for pions is depicted as a function of pseu-

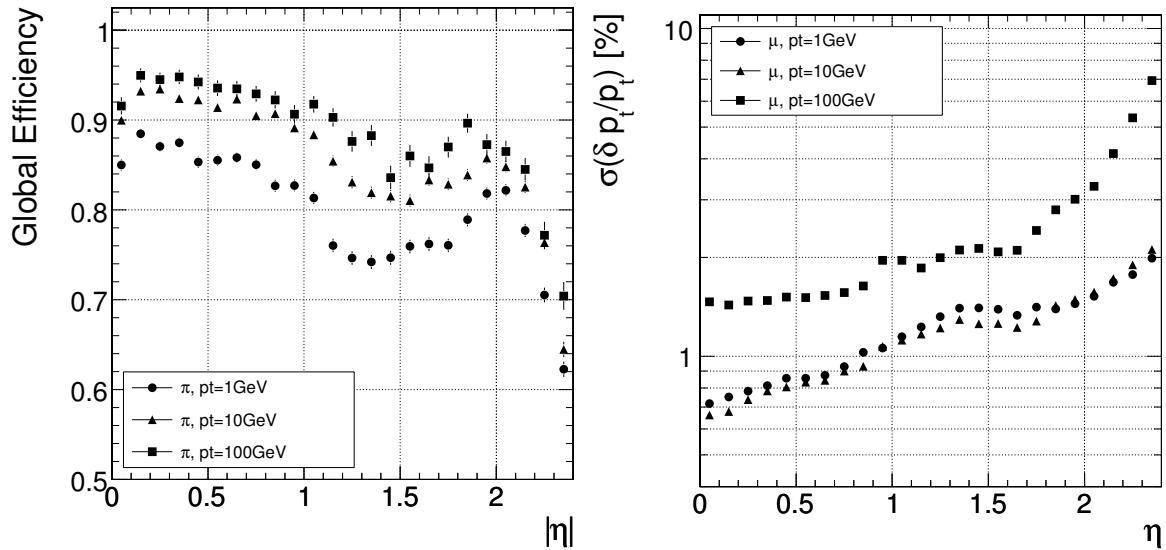


Figure 1.13.: Track reconstruction performance of the CMS detector (from [48]). Left: efficiency for pions. Right: Resolution of the transversal momentum for muons.

dorapidity. Depending on the particle momentum, one observes a rate up to 95 % over a broad  $\eta$ -region. For very large polar angles, the efficiency drops to about 70 %. The momentum resolution for isolated muons is shown in the right plot. It depends on the particle's momenta and their pseudorapidity and improves from about 2 % at  $|\eta| > 2.4$  up to 0.7 % at  $|\eta| = 0$ . The momentum resolution of pions involved in 3-prong tau decays is discussed in Figure 2.28 in Chapter 2.

The resolutions of the transverse and longitudinal impact parameters  $d_0$  and  $z_0$  for single tracks has been studied in [48]. For high momentum tracks, it is dominated by the hit resolution of the first hit in the pixel detector, whereas multiple scattering dominates at low momenta.

### 1.3.3. Vertex Reconstruction

Vertices are reconstructed from groups of tracks. In every collision at least one primary vertex should be identified. In addition, events may contain secondary vertices or tracks with significant impact parameters to the interaction point.

Different algorithms are used for different types of vertex reconstruction. Most algorithms include two steps, vertex finding and vertex fitting. The first combines vertex candidates from tracks. The second optimizes the vertex parameters by constraining the associated tracks to it. The CMS reconstruction uses two different vertex fitting methods. The first one is the Kalman Filter [65], which uses a global least-squares minimization. It works best for Gaussian distributed measurements and may include an update of the involved track parameters. The second method, the Adaptive Vertex Fitter [67], is an iterative re-weighted fit, which down-weights tracks according to their significant distance from the vertex. This enlarges the robustness by decreasing the sensitivity to outlying or mis-measured tracks. The performance of these techniques is compared in Table 6.5 in [48].

For the High Level Trigger (see 1.3.4), a primary vertex is created from re-weighting pixel standalone tracks [68], whereas the offline primary vertex reconstruction derives the best

parameter estimates from the fully reconstructed tracks. Its performance is summarized in Table 6.8 and 6.9 of [48]. The spacial resolution improves with the number of tracks associated to the vertex and with their transversal momentum. Depending on the physics process, it is in the order of 10 to 50  $\mu\text{m}$  for the transversal plane and slightly worse for the z-axis.

On top of the default vertex purposes, one can think of a specialized vertex reconstruction for exclusive decays. Chapter 2 introduces an approach to reconstruct tau-lepton decays based on a kinematic vertex fit.

### 1.3.4. Data Acquisition and Trigger

The data acquisition and trigger system of CMS handles the large bunch crossing rate of up to 40 MHz at the LHC (compare Section 1.2). It reduces the amount of data to a processable and storable size in two steps.

First, the Level-1 (L1) system, based on custom-made electronics, reduces the rate to approximately 100 kHz. This corresponds to a rejection factor of up to  $10^7$ . Then the High Level Trigger (HLT) system, implemented using commercial processors, further decreases the rate to  $\mathcal{O}(100)$  per second. In both steps vital decisions have to be made to discard a large fraction of data while retaining the few collisions of interest.

The bunch spacing of 25 ns is not sufficient to derive a trigger decision due to the required computation time and the signal transit from the detector to the trigger electronics. A buffer pipeline enlarges the time slot for event processing. The L1 decision considers primitives like electrons, muons, photons, jets and missing transversal energy. It is derived from limited detector information only and takes few microseconds.

The HLT algorithms run more sophisticated calculations accessing and combining the information stored by the L1 using more complex topological calculations. These computations performed in progressive series of HLT filters may take seconds and result in a reconstructed event. Based on the physics priorities, the HLT provides a classification into various data sets for further offline analyses. [69] gives a detailed review of the physics channels.

An entire description of the CMS trigger system and data acquisition can be found in [69–71]. An update of the performance of the trigger paths for the initial data taking is given in [72]. The efficiency of the paths required in the current analysis are listed in Table 3.3 in Section 3.4.

## 1.4. Event Reconstruction Framework

Millions of detector channels are read out electronically. The signals are passed to an extensive reconstruction chain to provide physics objects to the analyses. To estimate the performance of these algorithms and to study the physics sensitivity to certain processes, a complex detector simulation is needed. The simulation divides into the generation of particles emerging from high-energy collisions, the simulation of their path through the detector material, and the modeling of the detector response. The CMS collaboration integrates the full simulation chain and the event reconstruction into the CMS software framework (CMSSW [48]).

CMSSW is a C++ based framework designed around an event data model. It incorporates several independent libraries like the data analysis framework ROOT [73] and the class library

for high energy physics CLHEP [74]. Its primary goal is the development and deployment of reconstruction and analysis software.

### 1.4.1. Monte Carlo Event Generation

Event generators in high-energy physics interface theory and experiment. Based on the physics models, a collision between two incoming particles produces a set of outgoing particles. The rates and properties of events are generated with the Monte Carlo method [75].

CMSSW provides a large list of different event generators. General-purpose event generators, like PYTHIA [76], HERWIG [77], or SHERPA [78] include various models of the hard and soft interactions, parton distributions, initial and final state parton showers, multiple interactions, fragmentation, and decays. Matrix element calculators, such as ALPGEN [79], MADGRAPH [80], and MC@NLO [81], provide events at parton level as input to the multi-purpose generators. Particular particle decays can be described using distinct simulation packages like TAUOLA [82, 83] covering tau-lepton decays.

In CMSSW the generated events are stored in the HEPMC event record [84]. Its tree structure of particles and vertices reflects the evolution of the event. The convenient particle identifier scheme of the Particle Data Group [85] is used as unambiguous particle nomenclature.

### 1.4.2. Detector Simulation and Digitization

The interaction of generated particles penetrating the detector is evaluated within CMSSW using a full-scale GEANT4-based CMS detector simulation [86]. It combines a detailed description of the geometrical properties of every component and its material budget. The simulation provides a map of the magnetic field and accounts for the electromagnetic and hadronic interactions. The effects of primary vertices displaced from the nominal interaction point are covered by different vertex smearing methods.

The simulation accounts for luminosity dependent pile-up collisions. Multiple interactions within the same bunch and from the bunch crossing before and after a triggered event may cover the signal collision. This pile-up contribution is simulated separately and merged into the signal event [48].

The simulation step is complemented by the digitization. It simulates the detector's electronic response and generates a dataset, which is intended to be as close as possible to the real data detected with CMS.

### 1.4.3. Event Reconstruction

The same reconstruction algorithm is applied to the output of the Monte Carlo simulation chain as to the real data. It provides the reconstruction of events to the level of candidates, which can be used in the analyses. The independent algorithms for different types of particles or reconstruction objects are summarized in [48]. Vital parts for the current analysis are discussed in 1.3.2, 1.3.3, and 2.4.1. CMSSW can easily be extended to user-defined reconstruction modules as discussed in Chapter 2.

The resulting event format consists of three major data tiers [87]. The full event (FEVT) stores most of the intermediate products of simulation and reconstruction. The RECO format contains the reconstruction output and allows for possible event reprocessing. The amount of data can be reduced by extracting the information needed in particular analyses (AOD).

The displacements of detector modules from their expected position causes large uncertainties especially in the track reconstruction. These effects can be compensated with software corrections. They are defined in several scenarios for the alignment and the conditions of the detector. Their precision evolves in the course of the commissioning and is continuously optimized during data taking. The alignment of the silicon tracker is determined using optical survey measurements, a laser alignment system, and track-based alignment algorithms from cosmic rays [88, 89] until collision data is available. In CMSSW the scenarios are summarized within global tags to be specified during the event reconstruction. These tags also contain the current trigger menu for L1 and HLT.

#### 1.4.4. LHC Computing Grid

The vast amount of data produced at the LHC exceeds the storage and computation capacities of a single computation center. The CMS computing project relies on the Worldwide LHC Computing Grid (WLCG) to distribute the data around the world (see [48, 87]). WLCG is designed in several layers of data tiers providing resources for computation and storage.

The first layer is the Tier-0 at CERN, which stores the data collected by the CMS data acquisition (discussed in 1.3.4). It performs the first offline reconstruction and provides a copy to the subsequent Tier-1 centers. Here, physics analysis groups can access the reconstructed data streams. The Tier-1 centers provide data to and receive data from all Tier-2 sites, which are used for Monte Carlo production and end-user analyses.

CMSSW is distributed over the Tier's computing elements. This enables the submission of simulation or analysis tasks to the Grid to utilize its large resources. The private production and the further processing of official datasets used in the current analysis (see Section 3.1) are executed using the WLCG.



## Chapter 2

# Kinematic Reconstruction of Tau Leptons

The method described in this Chapter was developed in collaboration with Philip Sauerland. It is also part of his thesis [90]. This Chapter is published as an internal CMS note [91]. The kinematic tau algorithm is available to the CMS collaboration.

## 2.1. Introduction

The identification and reconstruction of tau leptons is a difficult task at the LHC. Due to their short lifetime, tau leptons decay before reaching the first detector layer. Therefore, they are solely detectable via their decay products. As tau-lepton decays incorporate neutrinos, which are not measurable with the CMS detector, the full information of the tau-lepton momentum is not accessible. A common technique is to ignore the neutrino energy fraction and approximate the tau flight direction by the visible decay products. Obviously, the precision of the tau-lepton reconstruction is limited in this case. In order to improve the experimental resolution and the reconstruction purity, one can utilize the specific decay topologies of the tau lepton and apply a kinematic fit to the measured tau-decay products.

In the following Sections the development of a kinematic fit for tau leptons is described. To begin with, the kinematics of multi-prong tau decays are introduced. This involves the kinematic calculation of the tau momentum as well as the derivation of kinematic constraints. Furthermore, the least mean squares minimization with Lagrange multipliers is discussed, which is the underlying approach of the kinematic fit. Afterwards, the newly developed kinematic fit for 3-prong tau decays is described in detail. Quality criteria are introduced, which aim to improve the tau-lepton identification and the suppression of background. Both the efficiency of the workflow and the achievable resolution of the kinematic tau-lepton parameters are tested with different tau-production processes via Monte Carlo simulated events. Finally, the technical implementation of the kinematic fit in the CMS software framework is described. The modularity of the whole procedure, its extensibility to additional tau-decay modes, and its flexible integration in physics analyses are pointed out.

## 2.2. Kinematics of Multi-Prong Tau Lepton Decays

### 2.2.1. Tau Lepton Decay Topology

The tau-lepton mass of  $M_\tau = 1776.82 \pm 0.16 \text{ MeV}/c^2$  enables a wide range of decay modes [85]. These modes can be categorized by the number of charged particles, usually called prongs, which are produced in a decay. Nearly 100 % of the total tau-decay width is made up of two main classes: 1-prong and 3-prong decays. The former provide the largest con-

tribution with a branching fraction of approximately 85 %. In nearly 15 % of all cases tau leptons decay into three charged particles. In contrast to muons, decay products of tau leptons are not limited to leptons only, but include hadrons as well.

In addition to their heavy mass, tau leptons feature a relatively long mean lifetime of  $\tau = (290.6 \pm 1.0) \times 10^{-15}$  sec. Considering for example electroweak processes like  $W^\pm$  and  $Z^0$  decays at the LHC, tau leptons are produced with high momenta  $|\vec{p}_\tau| = \mathcal{O}(100 \text{ GeV}/c)$ . This leads to large Lorentz factors  $\gamma$  and non-negligible mean tau flight distances

$$c\tau\beta\gamma = 87.11 \mu\text{m} \cdot \frac{|\vec{p}_\tau|}{M_\tau} = \mathcal{O}(5 \text{ mm}).$$

This is not sufficient to reach the first tracking detectors of the CMS experiment. As a result tau leptons are not directly detectable. Nevertheless, modern tracking detectors are able to resolve the tau flight distance. In case of the 1-prong decays, the distance is identified by a large impact parameter with regard to the primary vertex, which is formed by additional tracks deriving from the hard interaction. In 3-prong tau decays the particle tracks form a secondary vertex, which is displaced to the primary vertex. The link between the production and the decay vertex can be directly related to the tau flight distance.

Due to lepton-number conservation, at least one neutrino is produced in any tau decay. CMS is not able to detect neutrinos. The measurable tau energy is reduced by their energy. This affects measurements like invariant masses or missing transverse energy.

### 2.2.1.1. Tau Decay into Three Charged Pions and Neutrino

The largest contribution to the branching fraction of 3-prong tau decays comes from the decay into three charged pions and a tau neutrino. This mode incorporates three resonant processes, which add up to  $\mathcal{B}(\tau \rightarrow 3\pi^\pm + \nu_\tau) = \mathcal{B}_{a_1} + \mathcal{B}_{K^0} + \mathcal{B}_\omega = (9.32 \pm 0.07) \times 10^{-2}$  [85].

The dominant part is caused by the dynamic generation of an  $a_1(1260)$  resonance in the two-body decay  $\tau \rightarrow a_1(1260) + \nu_\tau$ . The  $a_1$  itself then decays into three charged pions. A possible decay chain is realized via an additional  $\rho$  resonance like  $a_1 \rightarrow \rho(770) + \pi^\pm \rightarrow 3\pi^\pm$ . The  $\rho(770)$  has a short lifetime in comparison to the tau. Thus the production vertices of the three charged pions are indistinguishable and are assumed to be identical.

The second process includes the production of a neutral kaon in the three-body tau decay  $\tau \rightarrow \pi^\pm + K^0 + \nu_\tau \rightarrow 3\pi^\pm + \nu_\tau$ . Because of the very long lifetime of the kaon, the two charged pions produced by its decay do not originate from the same vertex as the charged pion from the initial tau decay. Therefore, 3-prong tau decays via the  $a_1$  resonance or via kaons can in principle be distinguished by applying vertex fits to the three prongs.

The third and least dominant contribution is realized via an  $\omega(782)$  meson within the three-body tau decay  $\tau \rightarrow \pi^\pm + \omega(782) + \nu_\tau \rightarrow 3\pi^\pm + \nu_\tau$ . In contrast to the former process, the intermediate  $\omega(782)$  is short-lived. This results in a joint production vertex of the three prongs as in the  $a_1$  case.

Neglecting the contributions from kaons and  $\omega$  mesons, the process via the  $a_1$  leads to an exclusive branching ratio of  $\mathcal{B}(\tau \rightarrow 3\pi^\pm + \nu_\tau) = (9.00 \pm 0.06) \times 10^{-2}$  (excl.  $K^0, \omega$ ) [85]. This value corresponds to nearly 97 % of the inclusive branching fraction stated above. As further kinematic calculations assume that all three pions produced by the tau decay have to originate from a common secondary vertex, we ignore the small contribution of  $\mathcal{B}_{K^0}$  to

$\mathcal{B}(\tau \rightarrow 3\pi^\pm + \nu_\tau)$ . Furthermore, for reasons of simplification, we always identify the three-pion system with the  $a_1$  resonance and neglect the very small impact of the  $\mathcal{B}_\omega$  part, which is  $\mathcal{O}(10^{-4})$ .

One feature of the tau decay into three charged pions and neutrino is the absence of additional neutral particles besides the neutrino itself. That is to say, the visible energy of the tau lepton  $E_\tau^{\text{vis}}$  is directly correlated to the neutrino energy  $E_\nu$ . Figure 2.1(a) shows the invariant mass of the three-pion system  $m_{a_1}$ . The distribution is constrained via two bounds. The lower bound is three times the pion mass, whereas the upper bound is represented by the tau mass and thus  $3m_\pi < m_{a_1} < M_\tau$ . In case of the lower bound, the  $a_1$  has the highest momentum

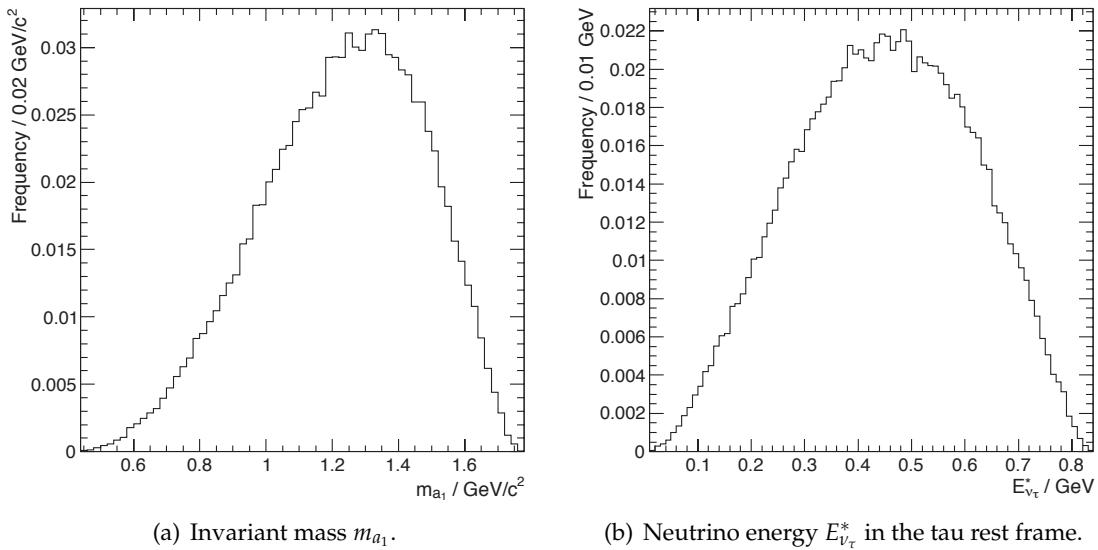


Figure 2.1.: The invariant mass distribution of the three-pion system  $m_{a_1}$  and the energy spectrum of the tau neutrino  $E_{\nu_\tau}^*$  within the tau rest frame.

within the tau rest frame. Following the basic kinematics of a two-body decay, the energy of the tau neutrino  $E_{\nu_\tau}^*$  is also maximal. Accordingly, the upper bound of the  $m_{a_1}$  spectrum corresponds to a very small neutrino energy within the tau rest frame. The resulting energy spectrum of the tau neutrino within the tau rest frame is shown in Figure 2.1(b).

## 2.2.2. Kinematic Calculation of the Tau Momentum

As shown above, the tau decay into three charged pions and neutrino ( $\tau \rightarrow 3\pi^\pm + \nu_\tau$ ) can be considered as a two-body decay into the  $a_1$  resonance, formed by the three pions, and the neutrino. The momentum of the tau lepton can be calculated from the  $a_1$  parameters only. The neutrino information is not needed in this case.

In the laboratory frame the decay can be written as

$$\begin{aligned} M_\tau^2 &= E_\tau^2 - \vec{p}_\tau^2 \\ M_\tau^2 &= (E_{a_1} + E_\nu)^2 - \vec{p}_\tau^2 \quad \text{neglecting the neutrino mass } m_\nu = 0 \\ M_\tau^2 &= (E_{a_1} + |\vec{p}_\tau - \vec{p}_{a_1}|)^2 - \vec{p}_\tau^2 \\ M_\tau^2 &= \left( E_{a_1} + \sqrt{\vec{p}_\tau^2 - 2|\vec{p}_\tau||\vec{p}_{a_1}|\cos\theta_{\text{GJ}} + \vec{p}_{a_1}^2} \right)^2 - \vec{p}_\tau^2 . \end{aligned}$$

The Gottfied-Jackson angle  $\theta_{GJ}$  is defined as the angle of the  $a_1$  w.r.t. the tau flight direction in the laboratory system (see Figure 2.2). Solving for the tau momentum one obtains the magnitude parametrized through the  $a_1$  parameters and the Gottfried-Jackson angle:

$$\begin{aligned} |\vec{p}_\tau| &= |\vec{p}_\tau|(\theta_{GJ}, m_{a_1}, |\vec{p}_{a_1}|) \\ &= \frac{(m_{a_1}^2 + M_\tau^2) |\vec{p}_{a_1}| \cos \theta_{GJ} \pm \sqrt{(m_{a_1}^2 + \vec{p}_{a_1}^2) \left[ (m_{a_1}^2 - M_\tau^2)^2 - 4M_\tau^2 \vec{p}_{a_1}^2 \sin \theta_{GJ}^2 \right]}}{2 (m_{a_1}^2 + \vec{p}_{a_1}^2 \sin \theta_{GJ}^2)} . \end{aligned} \quad (2.1)$$

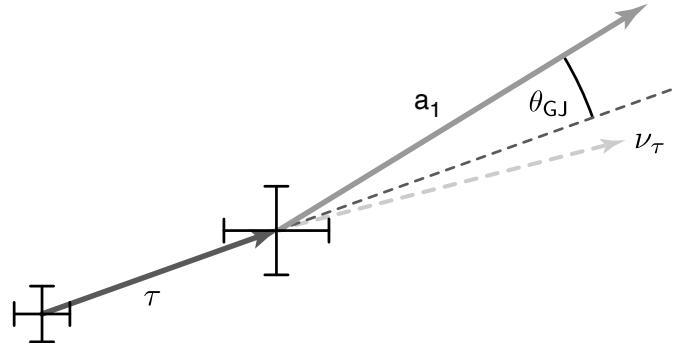


Figure 2.2.: The Gottfried-Jackson angle  $\theta_{GJ}$  is defined as the angle between the  $a_1$  momentum and the tau flight direction in the laboratory frame.

The same expression can be derived from the tau rest frame as follows (quantities referring to this frame are marked with an asterisk):

$$\begin{aligned} M_\tau^2 &= E_\tau^{*2} - \vec{p}_\tau^{*2} , \text{ with } \vec{p}_\tau^{*2} = 0 \text{ and } m_\nu = 0 \\ M_\tau &= E_{a_1}^* + |\vec{p}_{a_1}^*| = \sqrt{m_{a_1}^2 + \vec{p}_{a_1}^{*2}} + |\vec{p}_{a_1}^*| \\ \Rightarrow |\vec{p}_{a_1}^*| &= \frac{M_\tau^2 - m_{a_1}^2}{2M_\tau} \end{aligned} \quad (2.2)$$

$$\text{and } E_{a_1}^* = \frac{M_\tau^2 + m_{a_1}^2}{2M_\tau} . \quad (2.3)$$

A Lorentz transformation of (2.2) into the laboratory system yields again (2.1), which is plotted for typical  $a_1$  parameters in Figure 2.3.

### 2.2.2.1. Ambiguity

The square root in expression (2.1) leads to an ambiguity of the momentum magnitude. It arises from two different configurations in the tau rest frame, which are indistinguishable in the laboratory system. Either a lower energetic tau decays into an  $a_1$  under rather small angles w.r.t. the initial tau flight direction, or the  $a_1$  spans a larger angle towards a tau with higher momentum. In both cases one measures the same  $a_1$  momentum in the laboratory system (Figure 2.4).

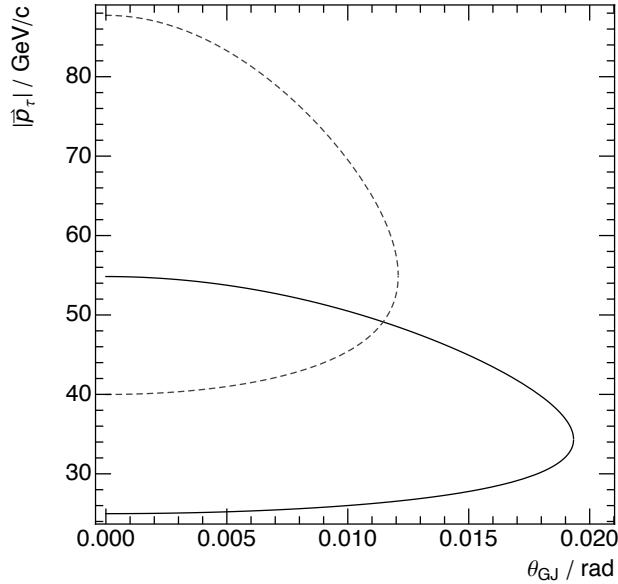


Figure 2.3.: The calculated tau momentum for typical  $a_1$  parameters ( $m_{a_1} = 1.2 \text{ GeV}/c^2$ , dashed:  $p_{a_1} = 40 \text{ GeV}/c$ , solid:  $p_{a_1} = 25 \text{ GeV}/c$ ). The range of allowed values for  $\theta_{GJ}$  is limited by the point where the ambiguity vanishes.

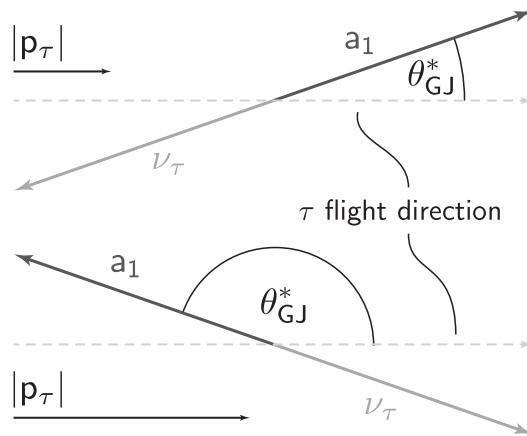


Figure 2.4.: Ambiguity for the magnitude of the tau momentum. An observer in a boosted system cannot distinguish between two different tau-decay configurations in the tau rest frame: 1. A tau lepton with a rather low momentum decays to an  $a_1$  under a small Gottfried-Jackson angle  $\theta_{GJ}^*$ . 2. A tau lepton with a higher momentum decays to an  $a_1$  under a larger Gottfried-Jackson angle  $\theta_{GJ}^*$ .

This ambiguity vanishes if the square root in (2.1) is zero. This determines an upper bound  $\theta_{\text{GJ}}^{\max}$  for the Gottfried-Jackson angle that only depends on the  $a_1$  mass and momentum:

$$\begin{aligned} (m_{a_1}^2 + \vec{p}_{a_1}^2) \left[ (m_{a_1}^2 - M_\tau^2)^2 - 4M_\tau^2 \vec{p}_{a_1}^2 \sin^2 \theta_{\text{GJ}} \right] &= 0 \\ \Rightarrow \theta_{\text{GJ}}^{\max}(m_{a_1}, |\vec{p}_{a_1}|) &= \arcsin \frac{M_\tau^2 - m_{a_1}^2}{2M_\tau |\vec{p}_{a_1}|} . \end{aligned} \quad (2.4)$$

Figure 2.5 shows some important dependencies of the Gottfried-Jackson angle and its maximal allowed value. All distributions show a common behavior of both variables over a wide range. Significant differences between  $\theta_{\text{GJ}}$  and  $\theta_{\text{GJ}}^{\max}$  only occur for very large  $|\cos \theta_{\text{GJ}}^*|$ . There are no important deviations in the dependency on the tau boost  $|\vec{p}_\tau|$  and the  $a_1$  mass.

One can show that the maximal allowed value  $\theta_{\text{GJ}}^{\max}$  is preferred. If one assumes the cosine of the angle in the rest frame  $\cos \theta_{\text{GJ}}^*$  to be equally distributed, the Lorentz transformation into the laboratory system results in a ratio  $r = \theta_{\text{GJ}} / \theta_{\text{GJ}}^{\max}$  shown in Figure 2.6. For small angles the ratio  $r$  is independent of the boost factor or the  $a_1$  mass.

This can be derived using (2.4) and (2.2):

$$\begin{aligned} \sin \theta_{\text{GJ}}^{\max}(m_{a_1}, |\vec{p}_{a_1}|) &= \frac{M_\tau^2 - m_{a_1}^2}{2M_\tau |\vec{p}_{a_1}|} = \frac{|\vec{p}_{a_1}^*|}{|\vec{p}_{a_1}|} \\ \text{and } p_{a_1\perp} &= |\vec{p}_{a_1}| \sin \theta_{\text{GJ}} . \end{aligned}$$

As  $p_{a_1\perp}$ , the orthogonal momentum component of the  $a_1$  w.r.t. the tau momentum, is not affected by the Lorentz boost and thus  $p_{a_1\perp} = p_{a_1\perp}^*$ , one obtains the ratio:

$$\begin{aligned} \frac{\sin \theta_{\text{GJ}}}{\sin \theta_{\text{GJ}}^{\max}} &= \frac{p_{a_1\perp}^*}{|\vec{p}_{a_1}^*|} \\ \Rightarrow \frac{\sin \theta_{\text{GJ}}}{\sin \theta_{\text{GJ}}^{\max}} &= \sin \theta_{\text{GJ}}^* \\ \text{and for small angles } r &= \theta_{\text{GJ}} / \theta_{\text{GJ}}^{\max} \approx \sin \theta_{\text{GJ}}^* . \end{aligned} \quad (2.5)$$

This dependency of the ratio  $r$  on the Gottfried-Jackson angle in the tau rest frame is shown in Figure 2.7. The sinus function gives small values of  $r$  for decays, where the  $a_1$  is emitted under  $|\cos \theta_{\text{GJ}}^*| \rightarrow 1$ .

### 2.2.2.2. Ambiguity Width

The deviation between the true tau momentum and the solution obtained by (2.1) is limited by the width of the ambiguity. As shown in Figure 2.3, the difference between both calculated solutions has a maximum for  $\theta_{\text{GJ}} = 0$ . The two solutions will be referred to as  $(+)$  or  $(-)$  depending on the sign of the square root term. This maximal width  $w$  can be calculated as

$$\begin{aligned} w &= |\vec{p}_\tau|(0, m_{a_1}, |\vec{p}_{a_1}|)^{(+)}) - |\vec{p}_\tau|(0, m_{a_1}, |\vec{p}_{a_1}|)^{(-)} \\ w &= \frac{(M_\tau^2 - m_{a_1}^2) \sqrt{m_{a_1}^2 + \vec{p}_{a_1}^2}}{m_{a_1}^2} = \frac{(M_\tau^2 - m_{a_1}^2) E_{a_1}}{m_{a_1}^2} . \end{aligned} \quad (2.6)$$

The maximal ambiguity width only depends on  $a_1$  parameters. Figure 2.8 shows the linear rise of  $w$  for increasing  $a_1$  momenta and the asymptotic decrease for larger  $a_1$  masses.

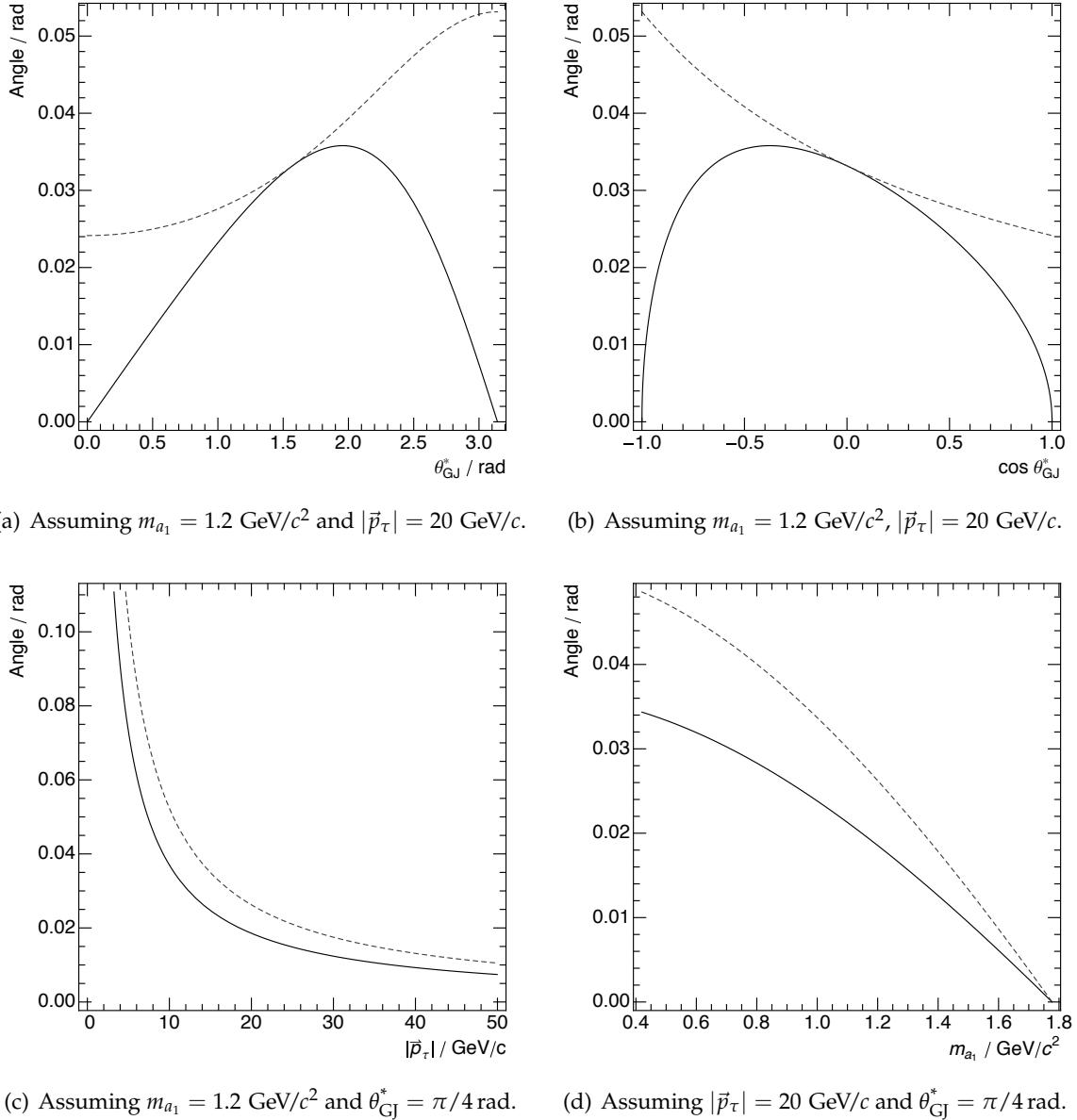
(c) Assuming  $m_{a_1} = 1.2 \text{ GeV}/c^2$  and  $\theta_{\text{GJ}}^* = \pi/4 \text{ rad.}$ (d) Assuming  $|\vec{p}_\tau| = 20 \text{ GeV}/c$  and  $\theta_{\text{GJ}}^* = \pi/4 \text{ rad.}$ 

Figure 2.5.: Major dependencies of the angles  $\theta_{\text{GJ}}$  (solid) and  $\theta_{\text{GJ}}^{\text{max}}$  (dashed) for typical decay parameters.

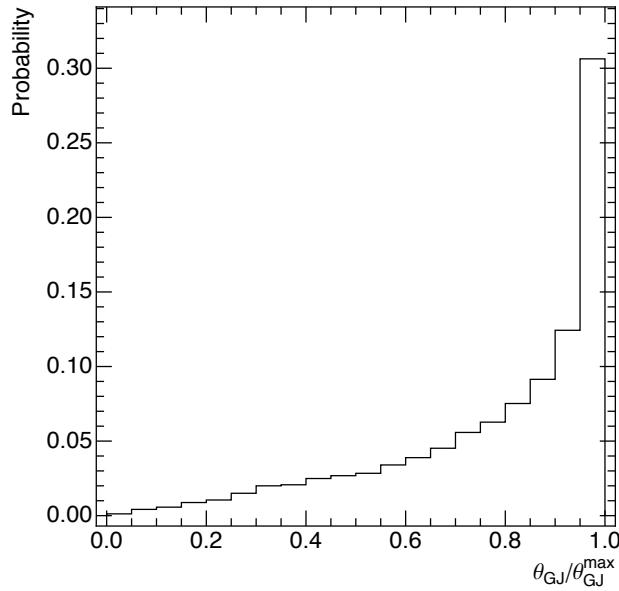


Figure 2.6.: Ratio of the actual Gottfried-Jackson angle  $\theta_{\text{GJ}}$  and the maximal allowed one  $\theta_{\text{GJ}}^{\text{max}}$  in the laboratory system. It is obtained from 1000 toy decays with uniformly distributed  $\cos \theta_{\text{GJ}}^*$  in the tau rest frame. The Lorentz transformation prefers the maximum value of  $\theta_{\text{GJ}}$ .

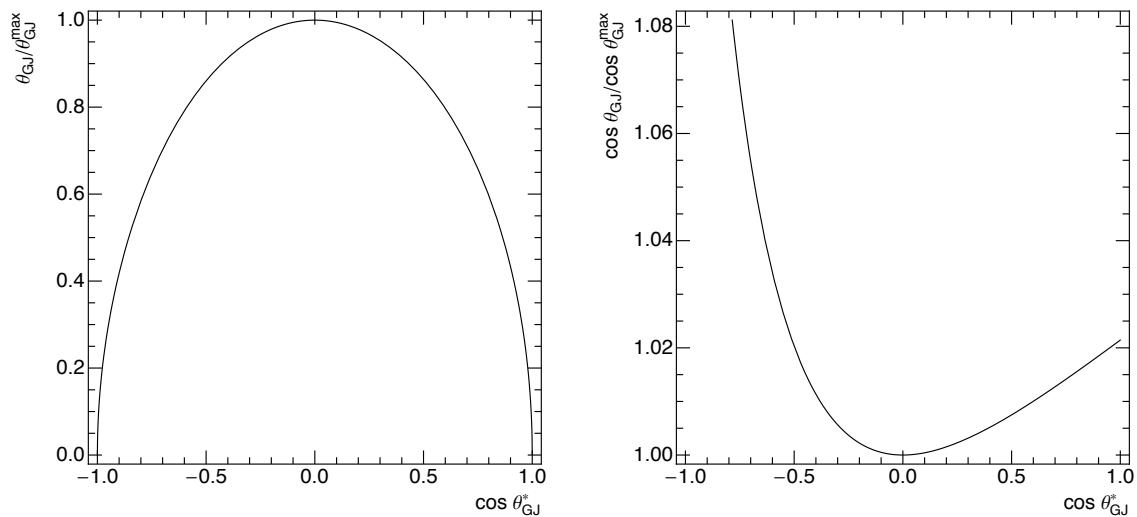


Figure 2.7.: Dependence of the ratio  $r = \theta_{\text{GJ}}/\theta_{\text{GJ}}^{\text{max}}$  on the rest-frame angle  $\theta_{\text{GJ}}^*$ . Small ratios  $r$  derive from large  $|\cos \theta_{\text{GJ}}^*|$ . The distributions are independent of the tau momentum and the  $a_1$  mass.

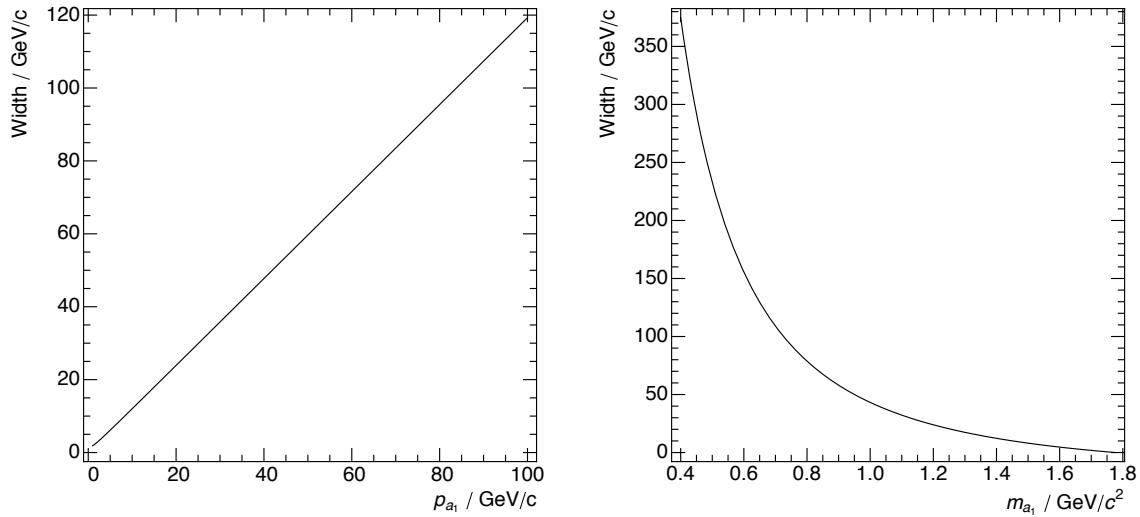


Figure 2.8.: Dependencies of the maximal ambiguity width  $w$ . Left: Linear increase ( $m_{a_1} = 1.2 \text{ GeV}/c^2$ ). Right: Asymptotic decrease ( $p_{a_1} = 20 \text{ GeV}/c$ ).

### 2.2.3. Accessible Constraints

Modern detectors in high energy physics can only measure few parts of the tau decay into three pions. Typically, the primary vertex, where the tau has been created, and the three charged tracks from the pions can be reconstructed. In order to fully determine the tau-lepton parameters and to improve the experimental resolution of the measured values, a set of constraints can be applied to the decay. These restrictions add additional knowledge to the measurement to minimize the number of unknown parameters.

Table 2.1.: Overview of the degrees of freedom (ndf) within the decay  $\tau \rightarrow 3\pi^\pm + \nu_\tau$ .

ndf	parameters	total
$4 \cdot 7$	track parameters + masses	+28
$3 \cdot 5$	pion reconstruction	-15
$4 \cdot 1$	particle masses	-4
$3 \cdot 1$	path length ( $n-1$ )	-3
$1 \cdot 1$	invariant mass constraint	-1
$4 \cdot 2 - 3$	vertex constraint ( $2n-3$ )	-5
$1 \cdot 2$	parallelism constraint	-2
		-2

Table 2.1 summarizes the number of parameters to determine in case of the 3-prong tau decay. An entirely determined decay requires the knowledge of 28 parameters in total. These are one mass and five helix parameters per particle involved. The helix parametrization only defines an infinite path without any start or end points. The absolute production vertex of the associated particle can be added using the path length  $s$  along the helix. This results in seven parameters for each of the four tau daughters.

The measurement of the three charged tracks provides five helix parameters each, and the mass of the tau daughters is set to their nominal values. As all daughters have to originate from

the same secondary vertex, the path lengths of these four particles are linked to each other. Therefore, assigning one path length will determine the others. The six missing parameters are obtained by applying three additional constraints:

The invariant mass of the four daughters is forced to be equal to the tau mass (Section 2.2.3.1). A vertex constraint (Section 2.2.3.2) ensures that all daughters origin from a common secondary vertex. The tau flight direction has to be in parallel to the link between the primary and the secondary vertex (Section 2.2.3.3). If all restrictions can be applied, one obtains a fully bound systems, where the number of constraints exceeds the degrees of freedom.

These three constraint classes are discussed in more detail in the following Sections, and their mathematical representations are given. According to the notation chosen in [92], constraints are described as functions  $H(\alpha) = 0$ , where  $\alpha$  is the full set of parameters of all particles. Each particle contributes with the parametrization  $\alpha_i = (x_i, y_i, z_i, p_{xi}, p_{yi}, p_{zi}, m_i)$ , where  $x_i, y_i, z_i$  defines the vertex position,  $p_j$  are the momentum components, and  $m_i$  is the mass of the particle. The calculation considers the equation of motion of charged particles in a solenoid magnetic field along the  $z$ -direction, which is derived in Appendix A.1.

### 2.2.3.1. Invariant Mass Constraint

A set of tracks can be forced to have a certain invariant mass  $M$ . From four-vector conservation one obtains the mass constraint  $H_M$  for the sum of  $n$  daughters of a decay:

$$\begin{aligned} E_{\text{sum}}^2 &= p_{\text{sum}}^2 + M^2 \\ H_M(\alpha, M) &\equiv E_{\text{sum}}^2 - p_{\text{sum}}^2 - M^2 = 0 \quad . \end{aligned} \quad (2.7)$$

In this case  $M$  corresponds to the tau-lepton mass.  $p_{\text{sum}}$  denotes the sum of momenta of all daughters propagated through the magnetic field (using equation A.2) to their common vertex  $(x^{(v)}, y^{(v)}, z^{(v)})$ , i.e. the three prongs plus the neutrino:

$$p_{\text{sum}} = |\vec{p}_{\text{sum}}| = \left| \sum_{i=1}^n \vec{p}_i \right| \quad .$$

This constraint is a so called 1C-constraint. This means it reduces the number of degrees of freedom by one.

### 2.2.3.2. Vertex Constraint

The vertex constraint  $H_V(\alpha)$  is a 2C-constraint for each particle  $i$  assigned to the vertex  $(x^{(v)}, y^{(v)}, z^{(v)})$ . These two equations correspond to the two planes defined by the bend of the helix caused by the Lorentz force. They can be derived from the equations of motion discussed in Appendix A.1. If the absolute reference point is unknown, the sum of the constraints is diminished by three. For  $n$  particles this results in a  $2n-3$  constraint.

Starting from (A.1) in Appendix A.1, the offset between the vertex and the reference point of the particle is obtained:

$$\begin{aligned} \Delta x &= x^{(v)} - x = \frac{p_x}{a} \sin \rho s - \frac{p_y}{a} (1 - \cos \rho s) \\ \Delta y &= y^{(v)} - y = \frac{p_y}{a} \sin \rho s + \frac{p_x}{a} (1 - \cos \rho s) \\ \Delta z &= z^{(v)} - z = \frac{p_z}{p} s = \frac{p_z}{a} \delta_T \quad . \end{aligned}$$

In order to eliminate terms of  $\rho_s$ , the following equations are used, where  $p_T$  is the transverse momentum:

$$\begin{aligned} p_x \Delta x &= \frac{p_x^2}{a} \sin \rho_s - \frac{p_x p_y}{a} (1 - \cos \rho_s) \\ p_y \Delta y &= \frac{p_y^2}{a} \sin \rho_s + \frac{p_x p_y}{a} (1 - \cos \rho_s) \\ p_x \Delta y &= \frac{p_x p_y}{a} \sin \rho_s + \frac{p_x^2}{a} (1 - \cos \rho_s) \\ p_y \Delta x &= \frac{p_x p_y}{a} \sin \rho_s - \frac{p_y^2}{a} (1 - \cos \rho_s) \\ \Rightarrow \quad \sin \rho_s &= \frac{a}{p_T^2} (p_{xi} \Delta x + p_{yi} \Delta y) \\ (1 - \cos \rho_s) &= \frac{a}{p_T^2} (p_x \Delta y - p_y \Delta x) . \end{aligned}$$

With these expressions the vertex constraints for the  $i$ th particle can be written as

$$H_{V\phi}(\alpha) \equiv p_{xi} \Delta y_i - p_{yi} \Delta x_i - \frac{a_i}{2} (\Delta x_i^2 + \Delta y_i^2) = 0 \quad (2.8)$$

$$H_{V\theta}(\alpha) \equiv \Delta z_i - \frac{p_{zi}}{a_i} \sin^{-1} \left( \frac{a_i}{p_{Ti}^2} (p_{xi} \Delta x_i + p_{yi} \Delta y_i) \right) = 0 . \quad (2.9)$$

### 2.2.3.3. Parallelism Constraint

The parallelism constraint  $H_P$  ensures that the vertex link between primary and secondary vertex is aligned with the resulting momentum vector  $\vec{p} = \sum_{i=1}^n \vec{p}_i$  of the decay daughters at the secondary vertex. It is a 2C-constraint as it restricts two orthogonal angles. Following

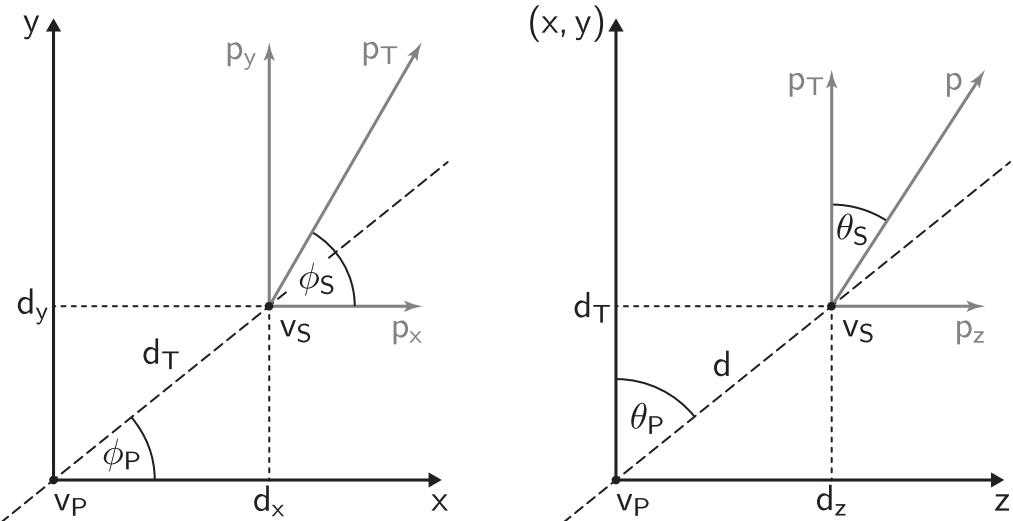


Figure 2.9.: Definition of the chosen angles  $(\phi_p, \theta_p)$  and  $(\phi_s, \theta_s)$ .

the parametrization in [93], the common azimuth angle  $\phi$  and the polar angle  $\theta$  defined from the transversal to the  $+z$ -direction are chosen. One should note that this is not the angle  $\theta_{\text{CMS}}$  according to CMS conventions, which is zero for tracks along the  $+z$ -direction,

but  $\theta = \pi/2 - \theta_{\text{CMS}}$ . To fulfill the constraint, the difference between the angles defined at the primary vertex  $(\phi_p, \theta_p)$  and the angles at the secondary vertex  $(\phi_s, \theta_s)$  have to be equal. As shown in Figure 2.9, one can express these angles in terms of the parametrization  $\alpha$ :

$$\begin{aligned}\cos \phi_s &= \frac{p_x}{p_T} & \cos \phi_p &= \frac{d_x}{d_T} \\ \cos \theta_s &= \frac{p_z}{p} & \cos \theta_p &= \frac{d_z}{d} .\end{aligned}$$

The distance between the primary vertex  $v_P = (x_p, y_p, z_p)$  and the secondary  $v_S = (x, y, z)$  is referred to as  $d_x = x - x_p$  (same for  $dy, dz$ ),  $d_T = \sqrt{d_x^2 + d_y^2}$ , and  $d = \sqrt{d_x^2 + d_y^2 + d_z^2}$ . To simplify further calculations, the difference is calculated in terms of  $\tan(\alpha/2)$  instead of comparing the angles itself. The tangent is well defined within the range of  $\theta$  and  $\phi$  and nearly linear for typical angles:

$$\begin{aligned}H_{P\phi}(\alpha, v_P) &\equiv \tan \frac{\phi_p}{2} - \tan \frac{\phi_s}{2} = 0 \\ \text{and } H_{P\theta}(\alpha, v_P) &\equiv \tan \frac{\theta_p}{2} - \tan \frac{\theta_s}{2} = 0 .\end{aligned}$$

Using  $\tan(\alpha/2) = (1 - \cos \alpha) / \sin \alpha$  and replacing the angles, one obtains the equations for the parallelism constraint:

$$H_{P\phi}(\alpha, v_P) = \frac{d_T - d_x}{d_y} + \frac{p_x - p_T}{p_y} = 0 \quad (2.10)$$

$$H_{P\theta}(\alpha, v_P) = \frac{d_s - d_T}{d_z} + \frac{p_T - p}{p_z} = 0 . \quad (2.11)$$

The first constraint equation assumes that the curvature of the path of the tau lepton due to the magnetic field is negligible for typical distances between the primary and the secondary vertex. As an example, in the solenoid field of CMS a tau lepton with a transverse momentum of  $30 \text{ GeV}/c$  will have a radius of curvature of  $26 \text{ m}$ . Assuming a typical tau lifetime, the helix bend causes an angle between the initial momentum measured at the primary vertex and the propagated momentum measured at the secondary vertex of  $\mathcal{O}(50 \mu\text{rad})$ . Compared to the scale defined by the maximal allowed Gottfried-Jackson angle, this is a correction at a percent level. According to the illustration in Figure 2.10 and with the equation of motion of charged particles in the magnetic field (Appendix A.1), the angular correction  $\alpha_S$  can be written as

$$\alpha_S = \delta_T/2 , \text{ where } \sin \delta_T/2 = \frac{d_T a}{2 p_T} . \quad (2.12)$$

Including  $\alpha_S$  within the parallelism constraint, the equation for the azimuth angle has to be modified. The polar angle is not affected by the helix bend.

$$\begin{aligned}H_{P\phi}(\alpha, v_P) &\equiv \tan \frac{\phi_p}{2} - \tan \frac{\phi_s}{2} - \tan \frac{\alpha_S}{2} = 0 \\ \Rightarrow H_{P\phi}(\alpha, v_P) &= \frac{d_T - d_x}{d_y} + \frac{p_x - p_T}{p_y} + \frac{\sqrt{4 p_T^2 - a^2 d_T^2} - 2 p_T}{a d_T} = 0\end{aligned} \quad (2.13)$$

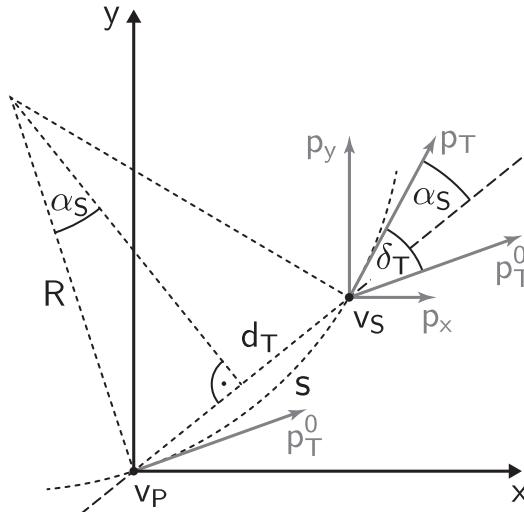


Figure 2.10.: Illustration of the correction derived from the helix propagation through the magnetic field.

## 2.3. Least Mean Squares Minimization with Lagrange Multipliers

A common mathematical approach to implement a kinematic fit with additional constraints is the Least Mean Squares minimization with Lagrange multipliers (LMS) [94]. This algorithm benefits from its sequential character. Each constraint is incorporated as a separated term within the  $\chi^2$  equation. This leads to a fast inversion due to factorized matrices and enables an implementation of the minimization process independent of the particular constraints applied by the user. LMS will result in an analytical solution, if the constraints are linear in the given parametrization. The application of LMS is limited to uncorrelated initial data and to hard constraints, where parameters are forced to match precise values. Soft constraint only requiring quantities to lie within certain bounds cannot be incorporated with LMS. A complete implementation of LMS into the CMS software framework is documented in [93].

### 2.3.1. Basic Algorithm

Given a set of unconstraint parameters  $\alpha_0$  from  $n$  measurements with their initial covariances  $V_{\alpha_0}$ , the  $\chi^2$  equation can be written as follows:

$$\chi^2 = (\alpha - \alpha_0)^T V_{\alpha_0}^{-1} (\alpha - \alpha_0) .$$

In addition, one can define  $r$  constraints  $H(\alpha) = 0$  with  $H = (H_1, H_2, \dots, H_r)$  (compare constraints defined in Section 2.2.3). If these constraints are not linear, they can be expanded about a linearization point  $\alpha_A$ :

$$\begin{aligned} 0 &= D(\alpha - \alpha_A) + d \\ &= D\delta\alpha + d , \end{aligned} \tag{2.14}$$

$$\text{where } D^{(r \times n)} = \begin{pmatrix} \frac{\partial H_1}{\partial \alpha_1} & \frac{\partial H_1}{\partial \alpha_2} & \dots & \frac{\partial H_1}{\partial \alpha_n} \\ \frac{\partial H_2}{\partial \alpha_1} & \frac{\partial H_2}{\partial \alpha_2} & \dots & \frac{\partial H_2}{\partial \alpha_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial H_r}{\partial \alpha_1} & \frac{\partial H_r}{\partial \alpha_2} & \dots & \frac{\partial H_r}{\partial \alpha_n} \end{pmatrix} \text{ and } d = \begin{pmatrix} H_1(\alpha_A) \\ H_2(\alpha_A) \\ \vdots \\ H_r(\alpha_A) \end{pmatrix} .$$

These linearized constraints (2.14) can be added to the  $\chi^2$  equation using a vector  $\lambda$  of  $r$  Lagrange multipliers:

$$\chi^2 = (\alpha - \alpha_0)^T V_{\alpha_0}^{-1} (\alpha - \alpha_0) + 2\lambda^T (D\delta\alpha + d) . \quad (2.15)$$

This equation can be minimized w.r.t. the parameters  $\alpha$  and the Lagrange multipliers  $\lambda$ :

$$\begin{aligned} \frac{\partial \chi^2}{\partial \alpha} &= 0 \Rightarrow V_{\alpha_0}^{-1} (\alpha - \alpha_0) + D^T \lambda = 0 \\ \text{and } \frac{\partial \chi^2}{\partial \lambda} &= 0 \Rightarrow D\delta\alpha + d = 0 . \end{aligned} \quad (2.16)$$

Solving these two equations yields a new set of constraint parameters, which depends on the initial values:

$$\begin{aligned} \alpha &= \alpha_0 - V_{\alpha_0} D^T \lambda & (2.17) \\ \text{where } \lambda &= V_D (D\delta\alpha_0 + d) \\ \text{and } V_D^{(r \times r)} &= (DV_{\alpha_0} D^T)^{-1} . \end{aligned}$$

The difference  $\delta\alpha_0 = \alpha_0 - \alpha_A$  is the offset between the unconstraint parameters and the linearization point. At this point the matrix  $D$  can be interpreted as a multidimensional gradient pointing towards the maximal modification in the parameter space and is therefore perpendicular to the equipotential surface generated by the combined constraints.

The minimized  $\chi^2$  equation can now be written as a sum of  $r$  terms:

$$\chi^2 = \lambda^T (D\delta\alpha_0 + d) . \quad (2.18)$$

This is one term per constraint. But, as the constraints are now correlated to each other through  $V_D$ , the assignment of each term to one particular constraint is no longer possible.

### 2.3.2. Iterations

The calculation above assumes that the Taylor expansion (2.14) around  $\alpha_A$  is suitable. The physical constraint has to be as linear as possible within the region of interest defined by  $\delta\alpha$ . Deviations from this linearity will lead to a result where the constraints are not completely fulfilled, thus  $H(\alpha) - \epsilon = 0$ . The second line in (2.16) only assures that the linearized version of the constraint is equal to zero, not the analytical one.

For sufficiently small  $\epsilon$  the recalculation of the whole algorithm using the last obtained set of parameters  $\alpha$  (see (2.17)) as a new linearization point can further optimize the minimization. Typically after a few iterations, the linearization point is well established and the LMS results in even smaller  $\epsilon$ . A criterion for how precise the analytical constraints are met by the obtained parameters is the sum of all remaining deviations:

$$c_{\text{sum}} = \sum_{i=1}^r |\epsilon_i| . \quad (2.19)$$

### 2.3.3. Unmeasured Parameters

The initial covariance  $V_{\alpha_0}$  represents the weight of each parameter involved in the LMS procedure. The larger the uncertainty of one parameter, the larger is the flexibility of the fit in modifying this quantity according to the given constraints.

Therefore, this algorithm can even deal with unmeasured parameters. If one can find enough restrictions to obtain a fully bound system where the degrees of freedom are at least compensated by the number of constraints, initially unknown quantities can be determined by LMS. As those new parameters have no finite prior uncertainties, they contribute to the covariance matrix using the Huge Error Method, described in [94]. According to this method each unknown parameter is treated as a measured one, which adds relatively large but finite errors  $L$  to the covariance matrix. The scale of these new errors has to be chosen w.r.t. the scale of the measured errors to avoid a loss in precision due to the subtraction of large numbers.

It can be shown that in the limit of  $L \rightarrow \infty$  the Huge Error Method leads to the same result as a direct but inconvenient calculation, which differs from the above algorithm and therefore would require a completely independent framework. However, applying the Huge Error Method allows to reuse the same calculation for both measured and unmeasured parameters.

## 2.4. Kinematic Fit of Tau Leptons on Monte Carlo Level with CMS

The standard tau-lepton reconstruction in CMS uses a generic approach to identify and reconstruct a large variety of tau-decay modes. In contrast to this, the kinematic fit of tau leptons needs to be specialized for specific decay topologies. The developed fit workflow is not intended to replace the standard tau-lepton reconstruction. It is designed as an add-on to the regular tau-reconstruction chain.

This Section gives a short overview of the standard tau reconstruction, which is used as an input for the kinematic fit. The logic of the kinematic fit and its application to 3-prong tau decays is described in detail. The efficiency of the method and the resolution of the kinematic tau parameters is shown. Quality criteria are introduced, which enhance the resolution and improve the suppression of background processes.

The kinematic fit has been developed and tested using Monte Carlo simulated events of the CMS detector. Throughout this Section the given numbers can depend on the version of the CMS software framework. Unless otherwise indicated the utilized release is CMS SW\_3\_6\_2.

### 2.4.1. Standard Tau Reconstruction at CMS

In the standard reconstruction chain there are currently two approaches for the identification of hadronically decaying tau leptons. One algorithm is based on calorimeter clusters and tracks (called CaloTau). The other one is based on the particles reconstructed by the particle-flow algorithm (called PFTau). Because of its potentially superior performance, only the latter will be discussed. Detailed information on both procedures and their comparison can be found in [95].

The particle-flow reconstruction collects the information derived from all sub-detectors of CMS and combines them to provide a global event description based on individually reconstructed particles. The obtained candidates (called PFJets) additionally require a minimal transverse momentum. At least one charged hadronic candidate with  $p_T > 5 \text{ GeV}/c$  has to be assigned to the jet inside a matching cone of  $\Delta R < 0.1$  around the jet axis.

As depicted in Figure 2.11, two additional regions are defined: A narrow signal cone around the leading candidate is intended to contain the collimated decay products from the boosted tau decay. In a broader isolation annulus low activity is expected for well separated tau candidates. In the latter region charged candidates with  $p_T > 1.0 \text{ GeV}/c$  and photon candidates with  $E_T > 1.5 \text{ GeV}$  are vetoed.

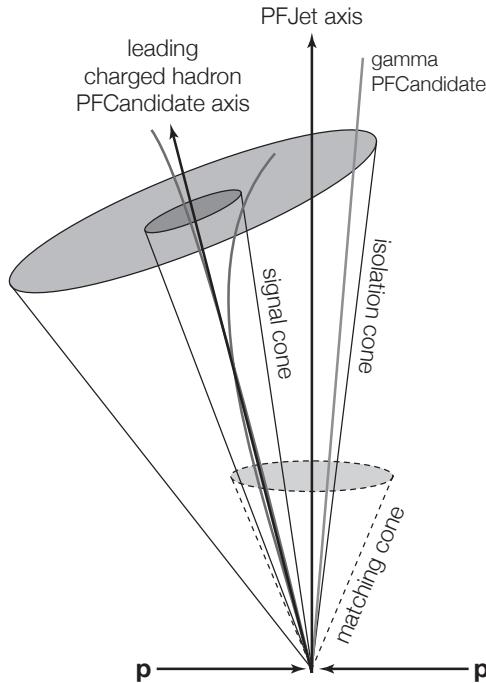


Figure 2.11.: Illustration of the defined cones within the particle-flow algorithm. The discrimination between hadronic tau-jets and jets from quarks and gluons is achieved by signal and isolation cones.

Depending on the size of the signal region, three different tau collections can be obtained. The first one, called `FixedCone`, uses a cone size of  $\Delta R < 0.07$ . A larger region of  $\Delta R < 0.15$  defines the second collection, called `FixedConeHighEfficiency`. In between of these two values the third one, called `ShrinkingCone`, varies its size according to the transverse energy of the reconstructed jet. The appropriate collection has to be selected for different tau boosts and decay modes. For example, 3-prong decays of soft tau leptons are less collinear and can exceed the signal region of the `FixedCone` tau algorithm. The larger cone sizes of the two other collections will lead to a larger background contribution.

On top of these tau-candidate collections, a set of discriminators is defined to enrich the reconstruction purity. Especially discriminators, which ensure the isolation of the particle-flow candidate, are useful to suppress quark and gluon jets. Additional discriminators cover the contamination of electrons and muons within the tau-lepton reconstruction. This higher-level discrimination is described in [96].

### 2.4.2. Kinematic Fit Workflow

The kinematic reconstruction of 3-prong tau decays is done in several steps. The basic idea is to take tau-lepton candidates, formed by the standard CMS particle-flow algorithm, as the starting point of the kinematic fit. Furthermore, the algorithm requires the existence of a primary vertex. It is crucial to note that the procedure solely depends on these two requirements and nothing else. It is independent of further event variables like the missing transverse energy. The application of the kinematic fit delivers a collection of kinematically refitted tau leptons, which is a sub-sample of the original particle-flow tau candidates. In contrast to the standard reconstruction, these refitted tau-lepton candidates incorporate the

full information on the tau momentum and its decay vertex. As described in Section 2.4.3, the fit quality can be evaluated via several unique variables.

Although not discussed here, the algorithm does not even have to start from particle-flow tau candidates. This requirement can be reduced to at least three reconstructed tracks per tau candidate. The user can utilize a custom tau-lepton identification and can easily deploy the kinematic fit in various steps of a physics analysis.

The flowchart in Figure 2.12 illustrates the intermediate steps needed to perform the kinematic fit. In the following Sections the workflow is described in detail.

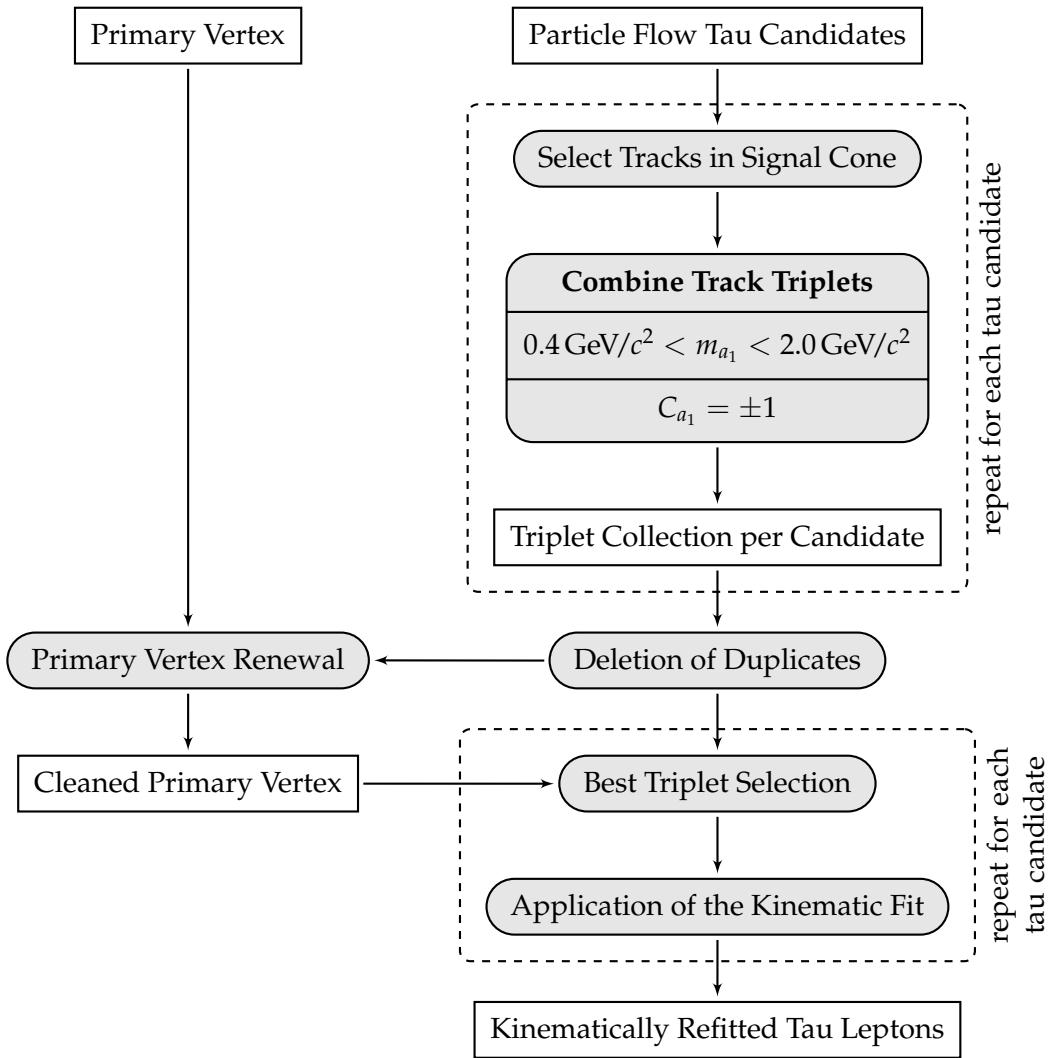


Figure 2.12.: Flowchart describing the workflow of the kinematic fit.

#### 2.4.2.1. Selection of Tracks within Tau Signal Cones

The particle-flow algorithm provides tau candidates, which are divided into three classes, as described in Section 2.4.1. As the kinematic fit algorithm will allow for applying new discriminants against background contribution, it is reasonable to choose the collection with

the highest initial frequency of tau candidates: the already discussed `FixedConeHigh-Efficiency` collection. Especially the tau decay into three prongs benefits from the larger signal-cone size, which is used for the creation of this tau-candidate collection.

Depending on the specific physics-analysis requirements, the accompanying particle-flow discriminators can further distinguish these candidates. To be independent of the impact of these requirements and to analyze the efficiency of this kinematic fit itself, none of these discriminators are applied in the following discussions.

Once the candidates are obtained, the tracks located within their signal cone can be accessed and are a priori considered to derive from the tau-lepton decay. As their number is not fixed and at least three reconstructed tracks are needed for the kinematic refit of 3-prong tau decays, the algorithm accordingly requires to find at least three tracks within this annulus.

#### 2.4.2.2. Track-Triplet Combination

If more than three tracks are associated with one particle-flow candidate, all possible track-triplet combinations are formed. Apparently, this step can lead to a multitude of track triplets. As there is only one correct track triplet, it is convenient to get rid of most of the wrong combinations as soon as possible. This is achieved by applying two constraints on the track triplets. The first one is a check of the cumulated charge of the track triplet  $C_{a_1}$ . This value is expected to be equal to  $\pm 1$  since the triplet represents the  $a_1$  resonance. The invariant mass of the triplet should loosely reside inside the kinematically allowed region of the 3-prong tau decay. According to Figure 2.1(a) in Section 2.2.1.1, the bounds on the invariant mass are chosen to  $0.4 \text{ GeV}/c^2 < m_{a_1} < 2.0 \text{ GeV}/c^2$ . Triplets, which do not fulfill both requirements, are deleted, keeping a number of track triplets per particle-flow tau candidate greater or equal zero for further processing.

#### 2.4.2.3. Deletion of Triplet Duplicates

In this step the selected track triplets are considered within the over-all scope of the proton-proton collision. Without loss of generality there can be more than one particle-flow tau candidate per event. For each selected tau candidate there might also be more than one track triplet, which survived the former processing step. It is possible that the tracks in the different triplets are not disjoint. If there are identical triplets for some reason, the duplicates get removed. This action is necessary to avoid double counting of tau-lepton candidates.

#### 2.4.2.4. Primary Vertex Renewal

The standard reconstruction of the primary vertex considers all tracks available in one event. As tracks produced by tau-lepton decays originate from secondary vertices, the primary vertex resolution should increase by ignoring those tracks in the primary vertex reconstruction. A new primary vertex is achieved by subtracting all tracks of the track-triplet collection from the complete set of reconstructed tracks within the event. The reduced track collection is used as input for the standard primary vertex reconstruction. In the rare case of a failing vertex reconstruction, the entire event is not processed any further. If the primary vertex could be built, it is reasonable to check its quality. Therefore, the algorithm requires at least three tracks associated to the new primary vertex. Otherwise, a three-dimensional vertex position is not achievable without additional constraints. There is a user-definable constraint on the quality of the primary vertex reconstruction via the  $\chi^2$  value of the fit.

### 2.4.2.5. Selection of the Best Track Triplet

Up to this state of the procedure there is only an ambiguous association of tau-lepton candidates and track triplets. This is due to the possibility of assigning several track triplets to one particle-flow tau candidate. The following procedure is applied individually to each of the track-triplet collections. At first, an adaptive vertex fit [67] is executed on the track triplets. If the vertex fit fails, it is quite obvious that these three tracks do not belong to a single tau-lepton decay and have been randomly selected by the triplet-combination mechanism. Thus this triplet is deleted. As the adaptive vertex fit may disregard tracks while performing the actual fit, it is convenient to look at the number of tracks, which are associated with the final vertex. If this number is smaller than three, this track triplet is also excluded from further processing.

The next step in the process depends on the number of track triplets, which passed the secondary vertex fit. If there is only one triplet left, these three tracks are identified as the three charged pions from the 3-prong tau decay. The association of the initial particle-flow tau candidate and this track triplet is unambiguous. Accomplishing the same uniqueness for the case of several track triplets, which pass the secondary vertex fit, one needs to perform an additional step.

In Figure 2.3 (Section 2.2.2.1) the maximally allowed Gottfried-Jackson angle  $\theta_{GJ}^{\max}$  was introduced. It is necessary that the measured Gottfried-Jackson angle  $\theta_{GJ}$  is equal or smaller than  $\theta_{GJ}^{\max}$ . Otherwise, the constraints, which are used in the kinematic fit, are not satisfiable. The idea is to sort the remaining track triplets according to the significant movement of their primary vertex, which is needed to maintain a tau-decay topology of  $\theta_{GJ} = \theta_{GJ}^{\max}$ . Unfortunately,  $\theta_{GJ}$  is very sensitive to mismeasurements of the primary and the secondary

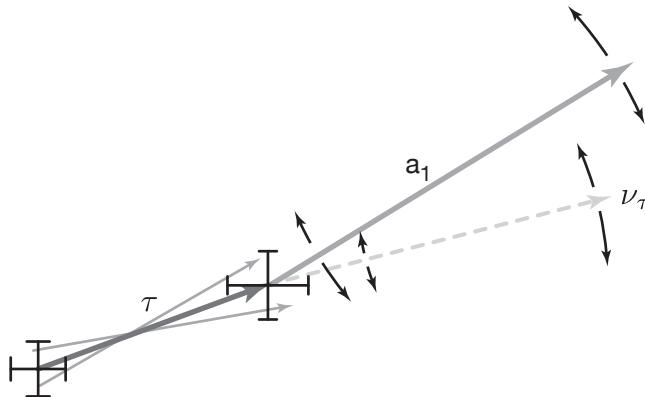


Figure 2.13.: Positional uncertainties of the primary and the secondary vertex, and mis-measurements of the track reconstruction. The Gottfried-Jackson angle  $\theta_{GJ}$  is highly sensitive to variations of the tau flight direction within these uncertainties.

vertex position, thus it is not unusual that the Gottfried-Jackson angle exceeds its kinematically allowed region. In this case the tau candidate would normally be dropped. But, as outlined in Figure 2.13, the tau flight direction can be varied within the positional uncertainties of the vertices. This leads to a change of the Gottfried-Jackson angle and can in principle once more establish a physical situation of  $\theta_{GJ} \leq \theta_{GJ}^{\max}$ . As long as this variation is done within the uncertainties of both vertices, there is no reasonable objection against this technique.

For technical reasons it is convenient to vary the tau flight direction by a rotation of the primary vertex around the secondary vertex, as shown in Figure 2.14. The rotational axis is perpendicular to the plane formed by the  $a_1$  flight direction and the link between the primary and the secondary vertex. The rotation is done for all tau-lepton candidates until  $\theta_{GJ} = \theta_{GJ}^{\max}$ . Even for track triplets, which form a tau-lepton decay with an initial  $\theta_{GJ} < \theta_{GJ}^{\max}$ , the primary vertex is rotated as well. Afterwards, the significance of the distance between the initial and

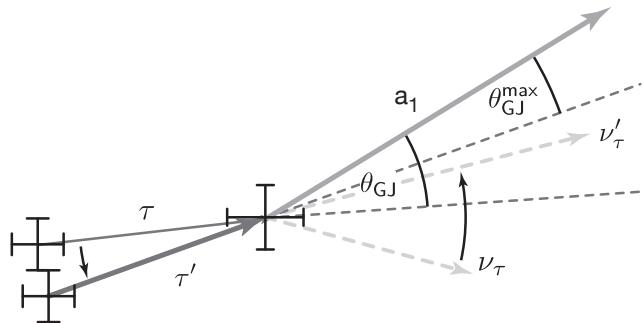


Figure 2.14.: Variation of the Gottfried-Jackson angle  $\theta_{GJ}$  due to the rotation of the primary vertex around the secondary vertex. The rotational axis is perpendicular to the plane formed by the flight direction of the  $a_1$  resonance and the vertex link.

the rotated primary vertex is calculated, as described in Appendix A.2. The track triplets are sorted with respect to this value, and in conclusion the triplet with the smallest significant movement is selected.

On an optional basis it is also possible to select the best track triplet via the goodness of the already applied secondary vertex fit instead of the selection via the primary vertex rotation.

#### 2.4.2.6. Application of the Kinematic Fit

The previous step ensures that exactly one track triplet is associated with each of the remaining particle-flow tau candidates. Now that all requirements are complied, the kinematic fit is executed for all present track triplets. The actual fitting algorithm is based on a least mean squares minimization with Lagrange multipliers, as described in Section 2.3. For this purpose it is stringently necessary to create a physically meaningful start situation. The momentum vector of the tau neutrino is calculated by following the instructions in Section 2.2.2. As we take the already constructed rotated primary vertex as an input for this calculation, the ambiguity for the neutrino momentum vanishes and there is just one solution left. In principle, the hypothesis that all 3-prong tau decays realize a maximal Gottfried-Jackson angle  $\theta_{GJ}$  is wrong. But, as illustrated in Figure 2.6 in Section 2.2.2.1, tau-decay topologies with  $\theta_{GJ} \approx \theta_{GJ}^{\max}$  are highly favored. That is why the former assumption is justified to a large extent.

It is assumed that the tau neutrino and the pion tracks of the corresponding triplet have a common origin. The vertex, which was calculated within the selection of the best track triplet, is chosen for this purpose. The positional uncertainties of the vertex are assigned to those of the four particles. The momentum and the uncertainties of the three tracks originate from the standard track reconstruction. As the neutrino is implemented as measured particle within the kinematic fit, the Huge Error Method is applied, assuming large uncertainties with respect to the calculated neutrino momentum. The masses of the three pions

and the neutrino are set to their values published by the Particle Data Group in [85]. The mass uncertainties derive from the same source. In case of the neutrino, we end up with a block diagonal covariance matrix. The three resulting blocks correspond to the covariance matrices of the vertex, the neutrino momentum and its mass. Correlations between these blocks are not taken into account, thus off-diagonal blocks remain zero.

At this state of the workflow the start situation is completely determined, and the kinematic fit is now applied to all available tau-lepton candidates. The constraints, which are used within the fit, are described in Section 2.2.3. Based on its linearized minimization technique (see Section 2.3) the fit can pass through several iterations. Due to the underlying fit algorithm, each iteration can modify the kinematic parameters and the associated covariance matrices of each of the involved particles. As shown in Figure 2.15, the average number of iterations is small. There are two main reasons for the fast convergence of the fit. First of all,

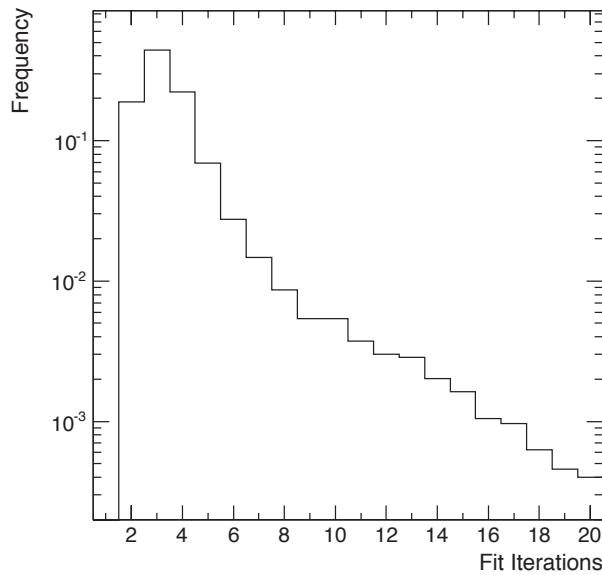


Figure 2.15.: Number of iterations per decay used by the kinematic fit in case of a pure  $\tau \rightarrow 3\pi^\pm + \nu_\tau$  sample.

the linearization of the applied constraints is a valid approximation in our case. Second, the educated guess of the neutrino parameters models a kinematically allowed start situation for the fit, which is already located in the region of the realized neutrino parameters.

The three pions and the tau neutrino are equally treated by the fit. This leads to a stepwise modification of the particle momenta and masses, and to the deployment of a common vertex of these four particles. The algorithm stops if either the maximally allowed number of iterations or the targeted constraint sum is reached. The constraint sum, which is derived in (2.19), is an overall measure of the fulfillment of the different constraints, thus characterizes the goodness of the kinematic fit. Both convergence criteria are user-definable and can be utilized to control the behavior of the kinematic fit.

#### 2.4.2.7. Kinematically Refitted Tau Leptons

If the kinematic fit converges, it delivers a kinematically refitted tau-decay tree. This tree consists of the tau lepton and its decay products. As already mentioned before, the tau

daughters are refitted, too. The kinematic tau-lepton properties are calculated by summing the refitted four-momenta of the three pions and the tau neutrino. The fitted vertex, which is shared by the four daughter particles, is assigned to the tau lepton. It is important to note that this vertex is the decay vertex of the tau lepton and not its production vertex as in the daughter particle's case. Instead, the rotated primary vertex is associated with the tau-production vertex. All in all, we obtain a fully determined tau-lepton decay.

The user has direct access to a complete set of kinematic tau-lepton parameters. The same is true for the refitted  $a_1$ -resonance properties by summing the four-momenta of the three pions. The initially unmeasured tau-neutrino attributes are available. In addition to this, the tau-decay vertex is measured to a very high precision. According to (2.18) in Section 2.3.1, the kinematic fit delivers a  $\chi^2$  value, which constitutes additional information on the quality of the refitted parameters. There are more sophisticated variables available, which help to discriminate 3-prong tau decays. The derivation of these quality criteria is described in the following Section.

### 2.4.3. Quality Criteria

The kinematic fit workflow itself provides a basic discrimination against events, which are not containing 3-prong tau decays. The main discrimination comes from the preselection when creating the consistent start scenario. The convergence of the kinematic fit has only a smaller impact (see Section 2.4.5 for detailed numbers).

In addition to this intrinsic 3-prong-decay selection, the algorithm gives access to some new observables, which allow for a more advanced discrimination. These additional criteria are intended to be used as a universal tau discriminator. It should be independent of the tau-production mechanism. Therefore, only rather loose criteria are required.

To cover a broad range of tau-production processes, the following signal-like tau sources are tested:  $W^\pm \rightarrow \tau\nu_\tau$ ,  $Z^0 \rightarrow \tau\tau$ ,  $H^0_{\text{GF}} \rightarrow \tau\tau$  and  $H^0_{\text{VBF}} \rightarrow \tau\tau$ . The indices GF and VBF denote Higgs bosons from gluon fusion and vector-boson fusion. The Higgs mass in these samples is chosen to be  $145 \text{ GeV}/c^2$ . The behavior of the following distributions is stable for a tested mass range of  $115 \text{ GeV}/c^2 < M_{H^0} < 145 \text{ GeV}/c^2$ .  $W^\pm$  and  $Z^0$  represent the Standard Model gauge bosons. In these event samples all tau leptons are forced to decay into three charged pions and neutrino and are compared to jets from QCD events imitating the signal. The QCD sample contains  $2 \rightarrow 2$  processes with  $30 \text{ GeV}/c < \hat{p}_T < 80 \text{ GeV}/c$ . These events are enriched by electrons, photons, charged pions, and charged kaons, which are above a transverse energy threshold of  $E_T > 20 \text{ GeV}$ .

To estimate efficiencies we start from the official CMS tau trigger. To obtain a conservative estimate only the following basic trigger requirements are applied: The event has to contain either a single tau with a transverse momentum of more than  $20 \text{ GeV}/c$  or two tau leptons with at least  $15 \text{ GeV}/c$  each. A detailed description of these high-level triggers can be obtained from [97]. A different choice of triggers will likely change the resulting tau leptons, and the quality requirements will have to be adapted. The QCD-background events are skimmed with respect to the tau-lepton triggers. They pass either the single-tau and/or the double-tau trigger.

The final 3-prong tau discriminator uses six discriminants depicted in the following figures. They are chosen to be as independent from each other as possible. The selection consists of observables from the fit itself, the decay kinematics, and the resolved secondary vertex of the

tau decay. The optimal values of the requirements are chosen by maximizing the product of signal efficiency  $\epsilon$  and purity  $\mathcal{P} = S/(S+B)$ , where  $S$  and  $B$  denote the frequency of signal-like and background-like decays.

For every observable the related distribution is shown with and without already requiring the other five discriminants. The former will be referred to as stand-alone, the latter as  $n-1$  representation. This shows the impact of the other criteria on the given observable. In each figure the distributions are normalized to an integral of one over the whole, not only the visible range.

The first discriminating observable is the  $\chi^2$ -probability of the kinematic fit. It is calculated from the minimized  $\chi^2$  (equation 2.18 in Section 2.3.1) and the degrees of freedom (Table 2.1). It gives an estimate, how well the given decay is described by the assumed model. Quantitatively, it denotes the probability that an observed  $\chi^2$  exceeds the minimized value by chance, even for a correct model. Therefore, entries in the first bins of Figure 2.16 correspond to tau decays with a poor quality as they have a too large minimized  $\chi^2$ . If the probability does not exceed 0.03, the corresponding tau candidates are discarded. In the left plot of

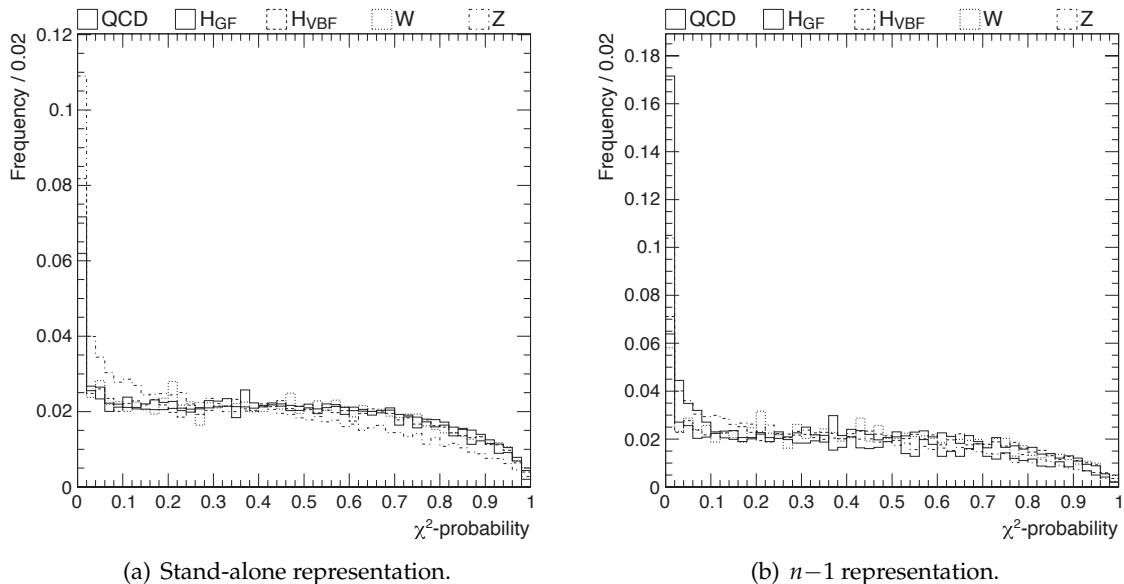


Figure 2.16.: Quality criterion:  $\chi^2$ -probability of the kinematic tau fit. Tau-lepton candidates need a  $\chi^2$ -probability  $\geq 0.03$  to pass this requirement.

Figure 2.16 it is remarkable, that before applying any other criteria the vetoed region is dominated by signal-like tau decays. Therefore, this stand-alone discriminant cannot enrich the signal-over-background ratio, but ensures a minimal fit quality. In the  $n-1$  representation the situation is inverted (Figure 2.16(b)). Now, mostly QCD decays are rejected. The poorly fitted signal-like decays, already discarded, fail other criteria.

The invariant mass of the hadronic part of the tau decay, the  $a_1$  mass, provides a kinematic criterion. It is calculated from the measured tracks (pion masses are assumed). It has a resolution of about  $\mathcal{O}(10 \text{ MeV}/c^2)$  (see Figure 2.28). The shape of the  $a_1$  mass distribution is similar for all true tau decays and is shifted to lower values in QCD decays (Figure 2.17(a)). This behavior is not changed by the other criteria. Decays with masses smaller than  $0.8 \text{ GeV}/c^2$

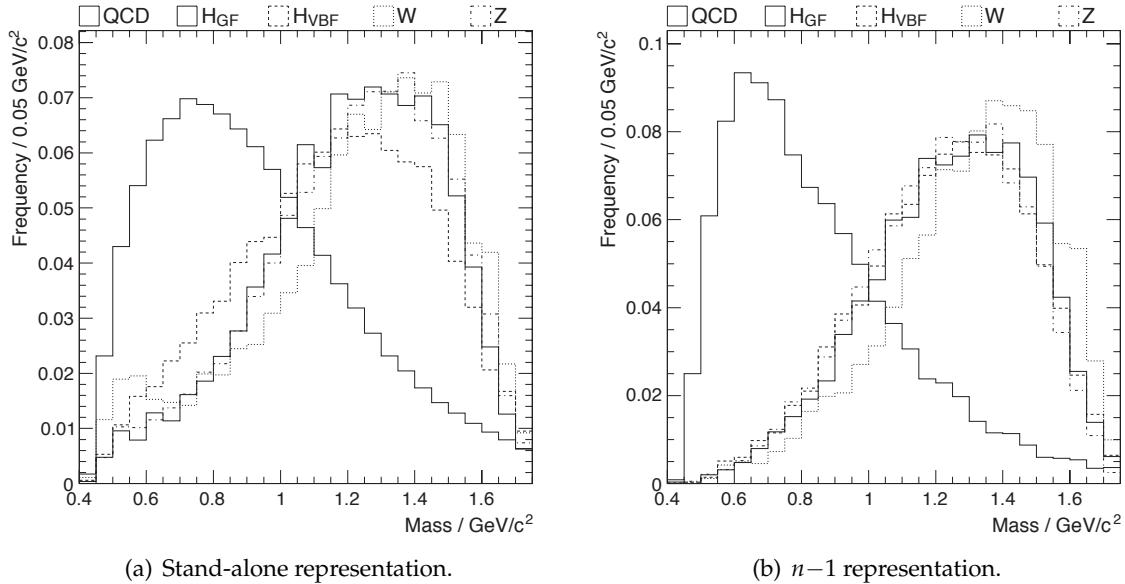


Figure 2.17.: Quality criterion: The invariant mass of the hadronic part of the tau decay  $m_{a_1}$ . Tau-lepton candidates need an  $a_1$  mass  $\geq 0.8 \text{ GeV}/c^2$  to pass this requirement.

are discarded. In addition to its strong discrimination, this requirement also improves the momentum resolution of the final kinematic tau. As discussed in Section 2.4.6.1, the resolution suffers from light  $a_1$  masses. Depending on the aimed efficiency of a final analysis, this criterion might be tightened to improve the purity.

An additional kinematic criterion can be derived from the ratio between the initial transverse energy of the tau candidate measured by particle flow  $E_T^{\text{PF}}$  and the final transverse energy  $E_T^{\text{kin}}$  obtained from the kinematic fit. As the neutrino is not included in  $E_T^{\text{PF}}$  but added during the kinematic fit, this ratio should always be smaller than one. This behavior of  $E_T^{\text{PF}}/E_T^{\text{kin}}$  is observed in the signal-like sources but not in all QCD decays, where this mass bound is often exceeded (Figure 2.18). This results in a strong QCD suppression.

The kinematic fit algorithm starts from at least three tracks within the signal cone of the corresponding particle-flow tau candidate (Section 2.4.2.1). The initial amount of these signal tracks differs between 3-prong tau leptons and QCD jets. This is due to the fact, that a QCD jet is expected to produce more activity in its vicinity than a tau decay. As shown in Figure 2.19, requiring exactly three tracks reduces QCD decays.

During the kinematic tau reconstruction the initial primary vertex of the current event may be modified by a rotation around the tau-decay vertex (see Section 2.4.2.5). The magnitude of this modification should be compatible with the uncertainty of the primary vertex. Therefore, one can calculate the significance of the separation between both versions of the vertex, as described in Appendix A.2. The resulting significance between the rotated and the initial primary vertex may not exceed two standard deviations. Both representations in Figure 2.20 demonstrate that for QCD jets the fit sometimes exceeds the allowed vertex rotation. These QCD decays can be discarded.

An additional discriminant concerning vertices is the separation between the primary and

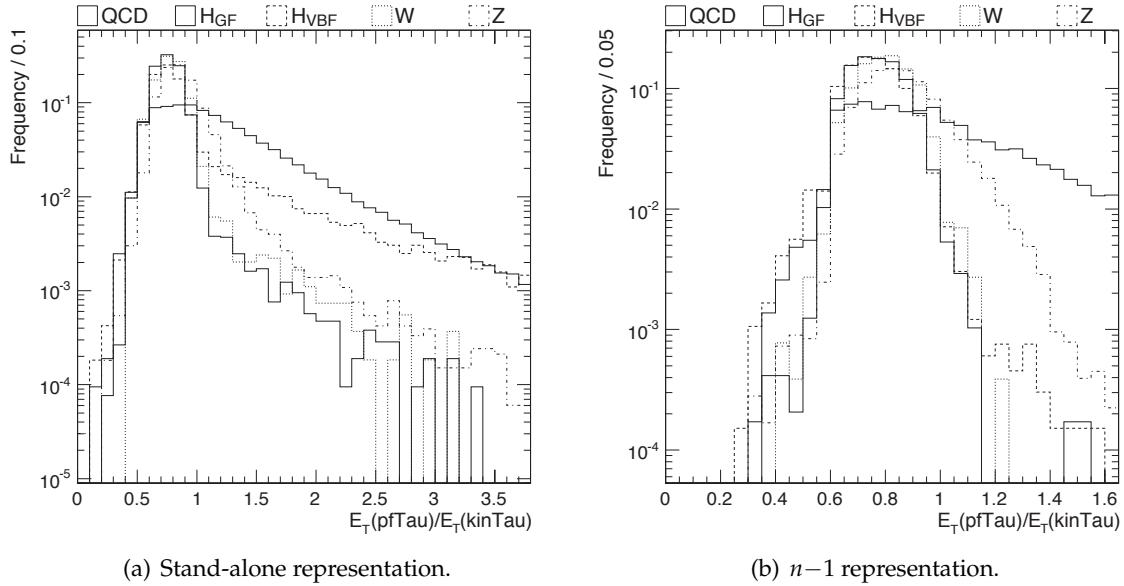


Figure 2.18.: Quality criterion: Transverse-energy fraction of tau leptons reconstructed with the particle-flow algorithm and the kinematic fit  $E_T^{\text{PF}}/E_T^{\text{kin}}$ . Tau-lepton candidates need to fulfill  $E_T^{\text{PF}}/E_T^{\text{kin}} \leq 1$  to pass this requirement.

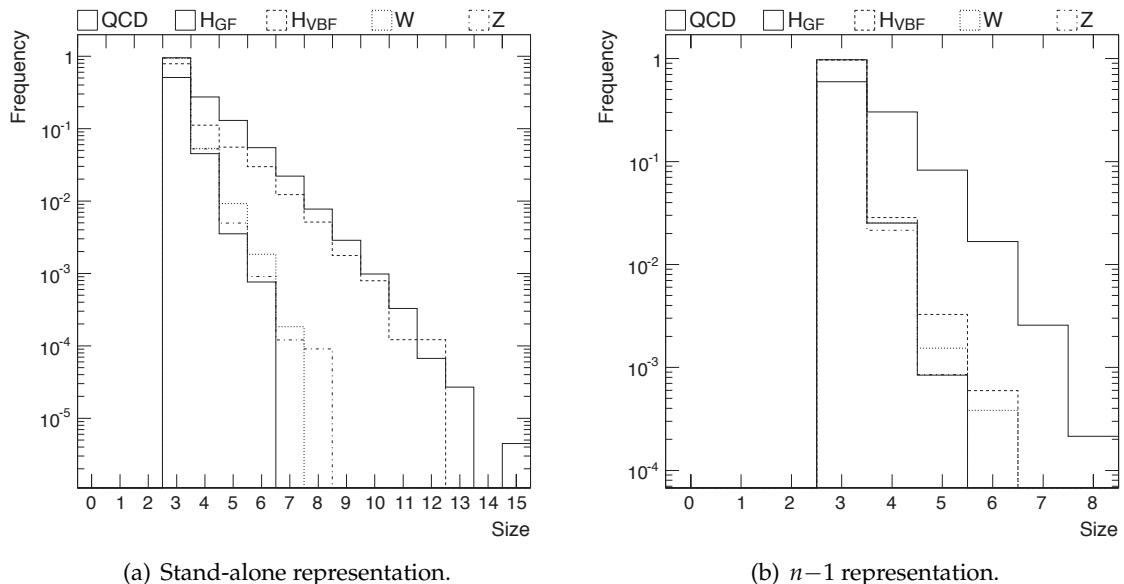


Figure 2.19.: Quality criterion: Number of tracks within the signal cone of the corresponding particle-flow tau candidate. The track count  $n_{\text{signal cone}}$  must be equal to three in order to pass this requirement.

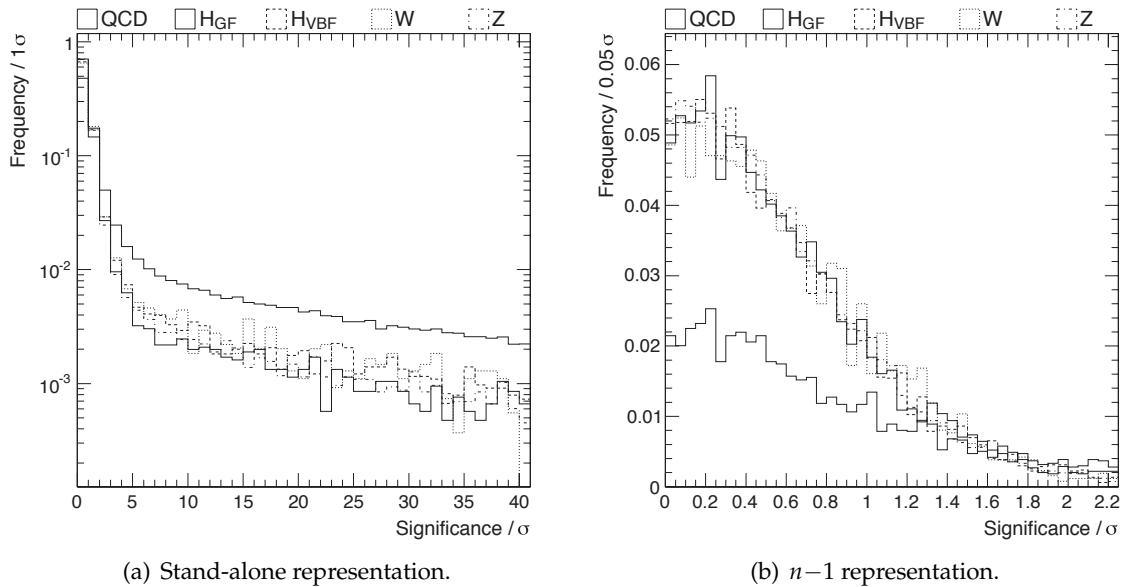


Figure 2.20.: Quality criterion: Separation significance  $\sigma$  between the initial and the rotated primary vertex. Tau-lepton candidates pass this requirement if  $\sigma \leq 2$ .

the secondary vertex. As discussed in Section 2.2.1, tau leptons have a significant lifetime. Therefore, signal-like tau candidates are expected to have a longer flight length than typical QCD jets. The tau-decay vertex is separated from its production point, whereas secondary vertices from soft jets may have a larger overlap with the primary vertex. For a clean distinction a minimal significance of two standard deviations is required. As expected, the selected region in Figure 2.21 is clearly dominated by signal-like tau decays. There is no large impact when applying the other requirements.

Combining these six criteria, the resulting discriminator should not interfere with those already available within the particle-flow framework (Section 2.4.1). Thus, this new set of discriminants is determined without any preselection obtained from the existing discriminators. The remaining overlap is depicted in Figure 2.22 for one representative tau source. The signal and the background case is shown. A translation from the tested discriminator ID's into the official name is integrated within the left plot. In both cases the efficiency of a broad range of common discriminators is hardly affected by applying the newly proposed 3-prong tau selection.

The defined quality criteria provide a loose selection of 3-prong tau decays, which can be implemented into the CMS tau framework as a stand-alone discriminator (see Section 2.5.3). The efficiencies on several decay types are summarized in Section 2.4.5. For a distinct analysis concerning tau leptons from a particular production process, the discrimination can be increased further by ignoring the source independence of the current set. Therefore, the discussed requirements can be tightened and additional quantities are available, such as kinematic observables and isolation criteria. Especially the transverse-momentum distribution of the tau lepton is a strong discriminator for separating common signal-like tau sources from QCD. In addition, it is only weakly correlated to the currently defined selection. But, this distribution strongly depends on the production mechanism of the tau lepton. It is therefore

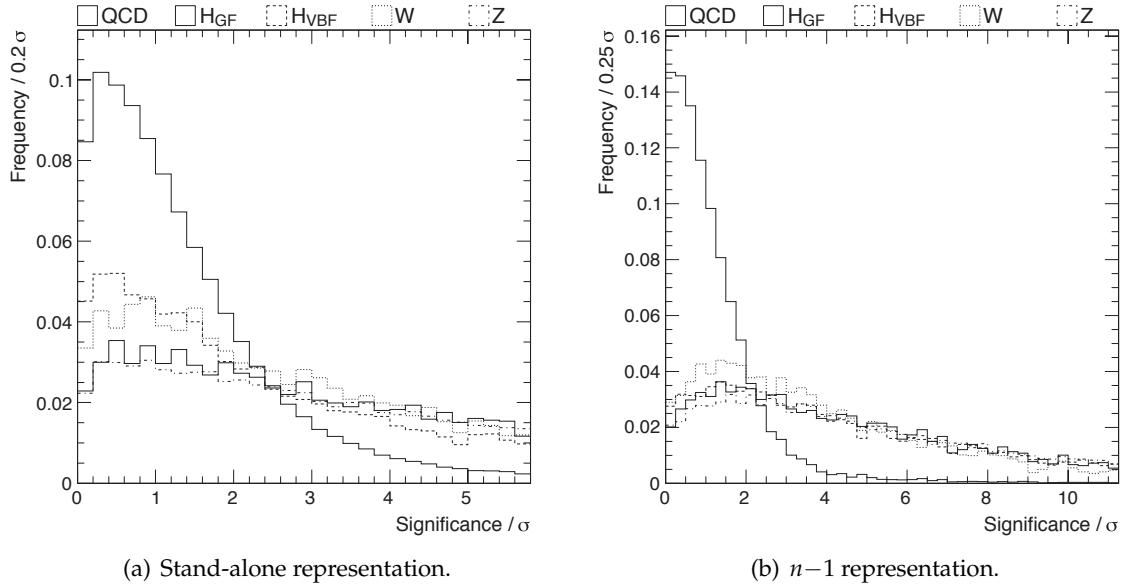


Figure 2.21.: Quality criterion: Separation significance  $\sigma$  between the primary and the secondary vertex. Tau-lepton candidates pass this requirement if  $\sigma \geq 2$ .

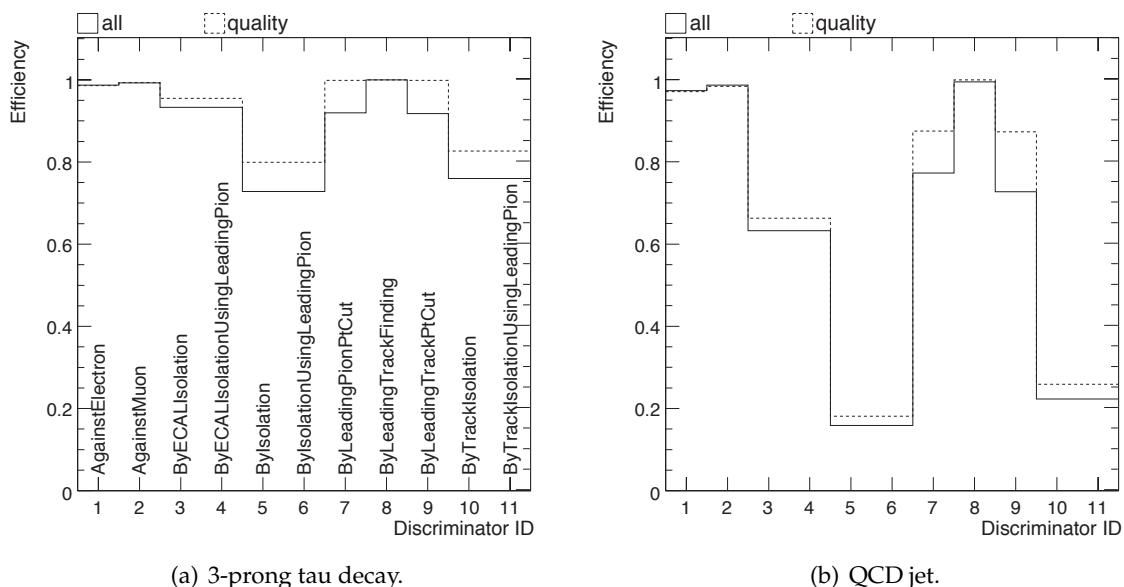


Figure 2.22.: Overlap between the existing particle-flow discriminators and the newly proposed 3-prong tau discrimination.

fore unsuitable for analyses covering less energetic tau decays (e.g. SUSY cascades). Such sample-dependent observables are not utilized in this universal discriminator.

### **2.4.3.1. Criteria Interrelation**

In the process of finding an appropriate set of quality requirements, it is crucial to know the level of interrelation between different criteria. In this regard the interrelation  $\kappa_{ij}$  of two quality criteria  $i$  and  $j$  is a useful measure. We define it as

$$\kappa_{ij} := \frac{\sum_{k=1}^C \delta_k(i, j)}{C} \quad , \text{with } \delta_k(i, j) := \begin{cases} 1, & \text{if } \mathcal{P}_i(k) \wedge \mathcal{P}_j(k), \\ -1, & \text{else.} \end{cases}$$

The normalization factor  $C$  is equal to the number of processed tau decays, which pass at least one of both requirements. For all of these decays the sum of the coefficients  $\delta_k(i, j)$  is calculated.  $\mathcal{P}_m(k)$  is a boolean expression, which is true if a tau lepton  $k$  passes the corresponding criterion  $m$ , otherwise it is false. Therefore, if two requirements provide the same result for all tau leptons their interrelation is maximal, and  $\kappa_{ij} = 1$ . In contrast to this, if both criteria never show the same result they are not interrelated, thus  $\kappa_{ij} = -1$ . Obviously,  $\kappa_{ii}$  is always one. In case of real tau decays, the coincidences should be as high as possible, whereas we aim to achieve low values for the QCD background case.

In Figure 2.23 the interrelation matrix is shown for the set of quality requirements, which are described in Section 2.4.3. For the signal case the interrelations are high and they contrast strongly with the background interrelation matrix.

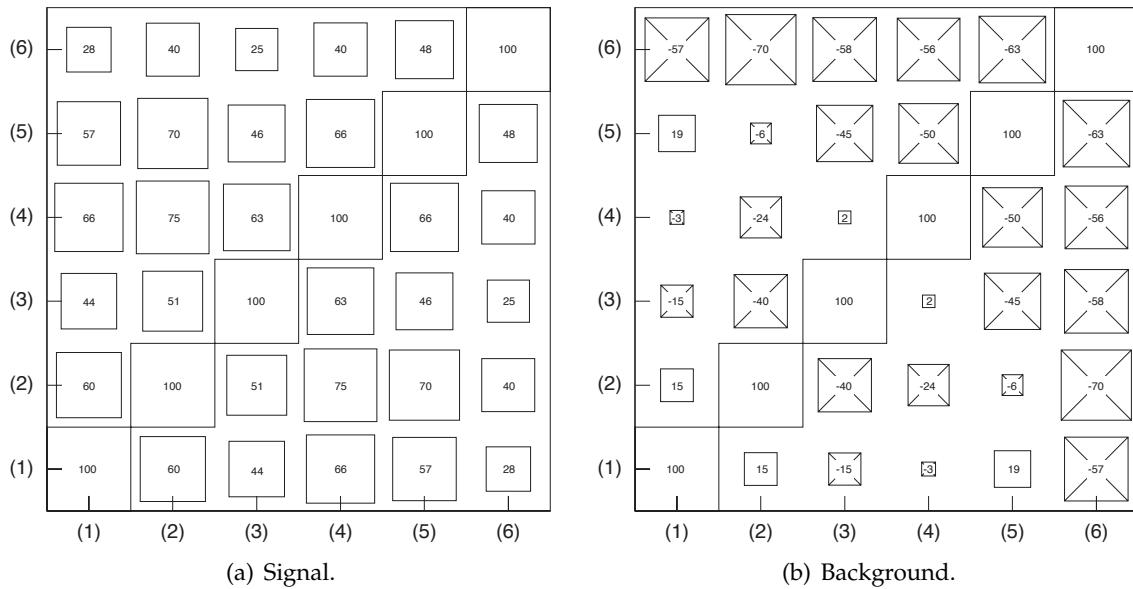


Figure 2.23.: Interrelation matrix for the set of quality requirements for signal-like 3-prong tau decays and background from QCD jets. The numerical values in the matrices are given in percent. The numeration of the quality criteria corresponds to the scheme used in Table 2.3 in Section 2.4.5.

### 2.4.4. Monto Carlo Truth Matching

The six discriminators, discussed in Section 2.4.3, are chosen to distinguish QCD-like jets from 3-prong tau decays. In addition, they enhance the matching purity for signal-like decays. To determine how many reconstructed tau candidates can be associated with a true tau lepton at the Monte Carlo generator level, the following definition is used: A whole tau decay is declared as matched if all three charged daughters can be assigned to their corresponding generator particles within an annulus of  $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} = 0.01$ . Mismatches in charge or transverse momentum ( $\Delta p_T / p_T = 10\%$ ) within this single-track assignment are vetoed.

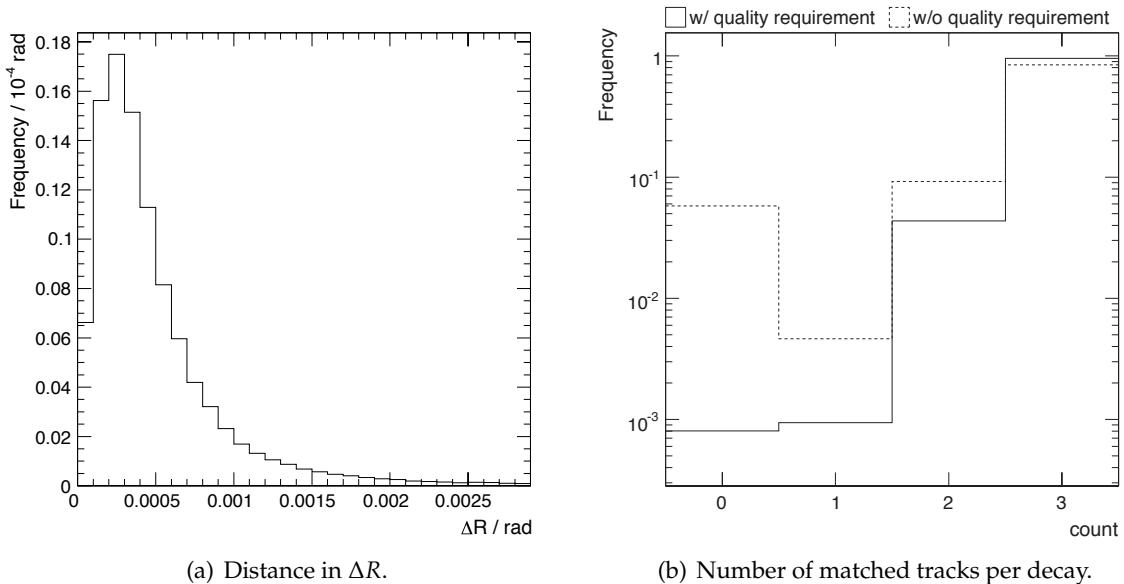


Figure 2.24.: Monte Carlo truth matching based on  $\Delta R$  assignment.

Figure 2.24(a) justifies the chosen size of the  $\Delta R$  cone. It depicts the distance between the measured tracks and their assigned generator counterparts. Whenever a matching is found, the distance between the related objects is much smaller than the minimally required value.

Applying the quality criteria, as proposed in the last Section, increases the matching purity. Figure 2.24(b) counts the number of tracks per tau decay, which can be assigned to a generator pion. Without using the quality discriminator there is a large contribution in the very first bin where none of the tau daughters can be matched. These decays are suppressed by the quality criteria. This results in a final matching purity for signal-like tau decays between 80 % and 90 % percent.

This purity gain can also be demonstrated by Figure 2.25. As an example, it shows two of the six observables, which are used to define the quality criteria. The solid line contains all tau decays, whereas the dashed one corresponds to the matched decays. Applying one particular requirement discards regions, which are dominated by decays not assigned to a generator particle.

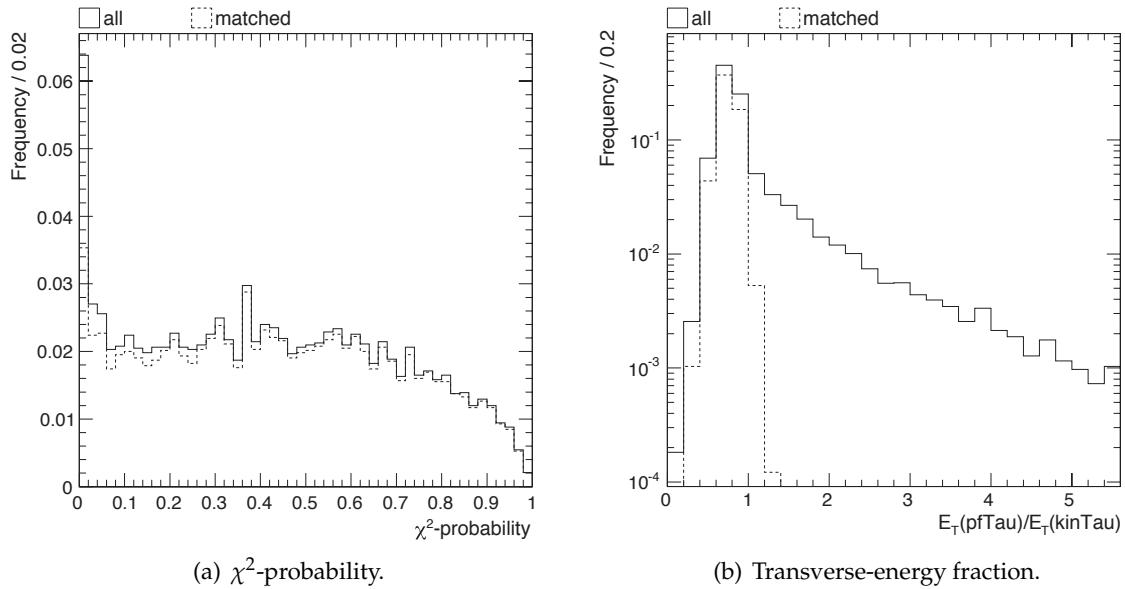


Figure 2.25.: Purity gain from the application of tau quality requirements.

#### 2.4.5. Efficiency of the Kinematic Tau Lepton Identification

The efficiency of the kinematic fit as well as the purity of its tau-lepton identification are crucial parameters, which determine the performance of the whole method. It is important to know the behavior of the fit for different tau-production processes and background samples. In Table 2.2 the efficiencies are listed for several tau-production processes and fake tau leptons from QCD jets. The samples are equal to the ones used in Section 2.4.3. The efficiencies in each row are not cumulative. The final row gives the cumulated number. For each requirement further information can be found in the description of the corresponding module in Section 2.5.

Table 2.2.: Event-based efficiencies of the kinematic fit for several tau-production processes and fake tau leptons from the enriched QCD dataset. The cumulated efficiencies are given in the last row of the table. The requirements are described in the associated workflow modules in Section 2.5.

requirement	event efficiency in %				
	$W^\pm$	$Z^0$	$H^0_{\text{GF}}$	$H^0_{\text{VBF}}$	QCD
(a) tau HLT	9.0	31.0	76.9	96.1	46.7
(b) valid primary vertex	96.8	93.8	99.4	98.4	100.0
(c) $n_{\text{signal cone}} \geq 3$	62.4	79.0	89.5	93.5	64.7
(d) valid track triplet	96.7	97.0	99.1	96.7	95.7
(e) converged fit	96.8	98.4	98.8	98.4	95.2
(f) $n_{\text{tau}}^{\text{quality}} \geq 1$	49.8	57.5	68.8	62.8	1.7
cumulated	2.5	12.4	44.8	51.6	0.4
expected cross section in pb	18.9	0.8	$1.5 \cdot 10^{-3}$	$2.0 \cdot 10^{-4}$	$1.5 \cdot 10^4$

The first row in Table 2.2 covers the high-level trigger with the tau trigger paths described

in Section 2.4.3. The di-tau samples are passing this trigger with a high rate of up to 96 %, whereas single tau events from  $W^\pm$  bosons suffer from an efficiency below 10 %. The QCD dataset is already preselected to achieve a reasonable efficiency passing the HLT. Another fundamental requirement to reconstruct a tau event with CMS is a valid primary vertex. Its quality depends on the  $\chi^2$  of the vertex fit and the number of assigned tracks. Row (b) requires at least three tracks and  $\chi^2/\text{ndf} < 10$ . It has only a small impact on the inspected samples. The next row corresponds to the module `InputTrackSelector` (Section 2.5.1), which tests the initial particle-flow tau candidate to contain at least three tracks in its signal cone. This relies on the successful reconstruction of three narrow tracks. This can fail even in a clean single-tau environment like the  $W^\pm$  source, if the tau leptons are too low in energy. The efficiency varies from 62 % to 94 %. In row (d) the `ThreeProngInputSelector` applies some basic tests on the selected tracks and defines a triplet, which fits best to a loose 3-prong assumption. After this preselection, valid tau candidates enter the kinematic fit, and row (e) ensures its convergence. The fit can only fail due to very inconsistent start scenarios, which is quite rare after passing all the previous modules. Therefore, one obtains a large efficiency of more than 95 % in all samples. Finally, the tau leptons have to pass the quality criteria defined in Section 2.4.3. Row (f) requires at least one tau lepton to pass this selection, which leads to rather large efficiencies in di-tau sources like  $Z^0$  or  $H^0_{\text{GF}}$ , medium rates of about 50 % for single-tau sources, like  $W^\pm$ , and low frequencies of less than 2 % for QCD jets faking a tau candidate.

Combining these numbers one obtains the cumulated efficiency of up to 50 % in the signal case. The inspected QCD sample can be suppressed by nearly 3 orders of magnitude. This translates to an expected cross section for the given tau sources. The  $W^\pm \rightarrow \tau\nu_\tau$  decay has a quite large value of about 19 pb compared to  $Z^0 \rightarrow \tau\tau$  with less than 1 pb. This can be explained by the second tau lepton involved in the latter decay. Here, the branching fraction of  $\tau \rightarrow 3\pi^\pm + \nu_\tau$  has to be taken into account twice. The same holds for both Higgs boson decays. Their cross section is still 7 orders of magnitude smaller than the expected QCD background.

Table 2.3 gives a detailed overview of the selection efficiencies of the quality criteria for signal and background samples. In contrast to the previous event-based Table, single-tau decays are counted here. The actual value of the specific discriminant is given in the first column of this Table. The cumulated selection efficiency of the tau decays slightly depends on the tau production mechanism. In case of a  $H^0_{\text{VBF}}$  source, nearly 39 % of the 3-prong tau decays pass these criteria, whereas the selection efficiency for the  $H^0_{\text{GF}}$  source rises up to 53 %.

Table 2.3.: Decay-based efficiencies of the quality criteria for several tau-production processes and fake tau leptons from QCD jets.

requirement	decay efficiency in %					
	$W^\pm$	$Z^0$	$H^0_{\text{GF}}$	$H^0_{\text{VBF}}$	non 3-pion	QCD
(1) $\chi^2$ -probability $\geq 0.03$	92.8	87.0	91.5	90.6	92.3	92.5
(2) $m_{a_1} \geq 0.8 \text{ GeV}/c^2$	88.5	91.3	91.9	87.1	53.3	61.1
(3) $E_T^{\text{PF}}/E_T^{\text{kin}} \leq 1$	95.1	80.2	96.8	77.7	81.0	44.2
(4) $n_{\text{signal cone}} = 3$	93.6	94.2	95.1	78.3	87.2	50.8
(5) $\sigma(\text{PV rotation}) \leq 2.0$	83.3	87.8	88.3	83.6	64.9	62.4
(6) $\sigma(\text{PV-SV}) \geq 2.0$	60.6	72.3	69.9	57.0	51.9	24.1
cumulated	46.6	45.0	54.8	39.5	18.5	1.2

The column labeled with "non 3-pion" refers to tau leptons from  $Z^0 \rightarrow \tau\tau$  decaying according to their Standard Model branching ratio. Here, the signal-like mode  $\tau \rightarrow 3\pi^\pm + \nu_\tau$  is vetoed to estimate the effect of a wrong decay assumption. Within the remaining sample 18.5 % of the tau leptons survive the quality selection. These are dominated by other 3-prong modes without pions. Therefore, the only error the kinematic fit commits by its 3-pion assumption is applying a wrong pion mass. In case of additional neutrals, this results also in a non-vanishing neutrino mass. These two effects lead to kinematic tau leptons, which are very similar to those decaying into three pions and therefore pass the quality criteria. If these "3-prong+X"-modes including arbitrary numbers of neutrals are also ignored, the resulting efficiency drops to approximately 1 %.

Finally, the last column of Table 2.3 covers fake tau leptons from quark and gluon jets. These are highly suppressed and just 1.2 % pass the quality requirements. All in all, the defined quality criteria have a reasonable effect on the selection of 3-prong tau leptons while suppressing QCD background to a high extent.

#### 2.4.6. Resolution

The resolution of the different tau-lepton parameters  $\alpha_i$  is shown with respect to the generated truth and is always calculated as  $\Delta\alpha_i = \alpha_i^{\text{fit}} - \alpha_i^{\text{generator}}$ . The generator counterpart is selected according to the assignment rules of Section 2.4.4. Figure 2.26 illustrates the angular resolution of the kinematically fitted tau-lepton momentum. The azimuthal angle  $\phi$  and the pseudorapidity  $\eta$  can be reconstructed very well. Both corresponding distributions peak sharply at the generated value and show a symmetric shape.

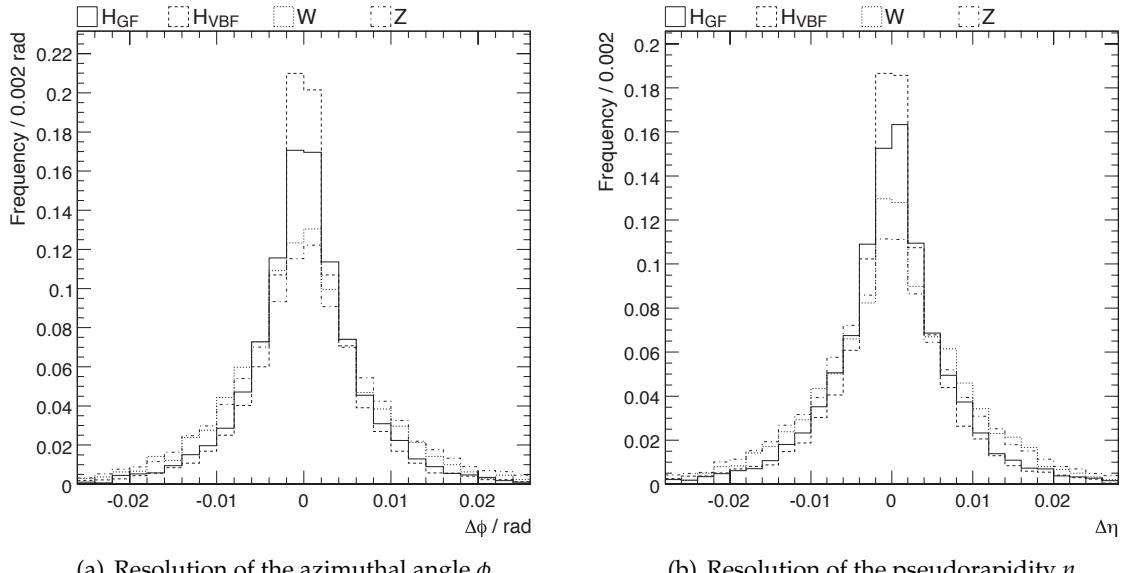


Figure 2.26.: Angular resolution of the kinematically fitted tau-lepton momentum. The quality criteria of the kinematic fit are applied in these plots.

The transverse-momentum resolution of the fitted tau lepton is depicted in Figure 2.27. The distribution for the z-component is independent of the tau source, whereas the transverse

part differs between  $W^\pm$  and  $Z^0$  sources and tau leptons from Higgs decays. The latter yields a symmetric distribution around zero but with a broader shape than the distribution for  $W^\pm$  and  $Z^0$ . The systematical shift towards higher values in Figure 2.27(a) can be explained by intrinsic effects, which are described in Section 2.4.6.1.

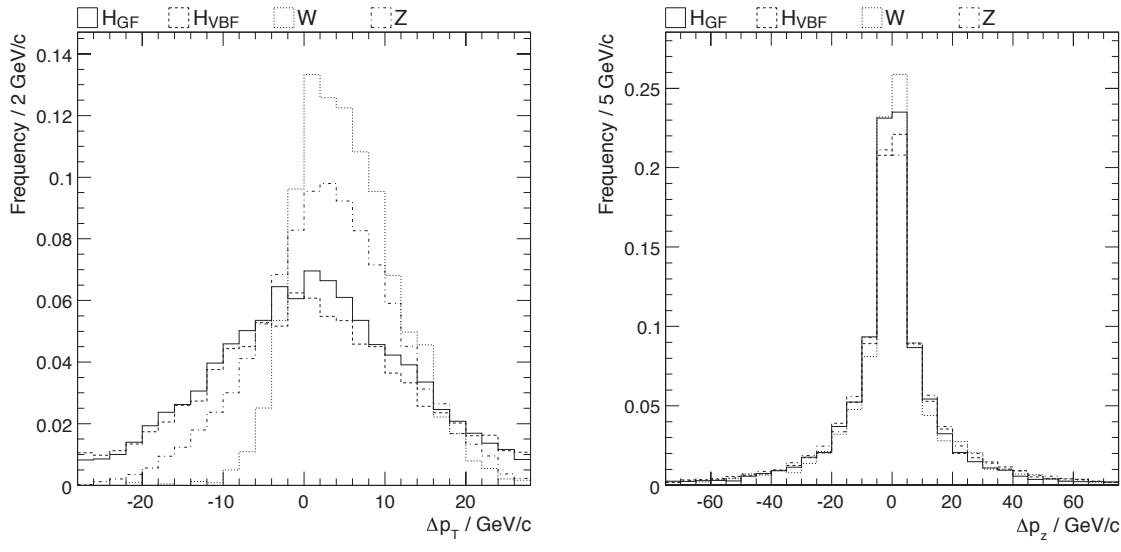


Figure 2.27.: Momentum resolution of the kinematically fitted tau lepton. The quality criteria of the kinematic fit are applied in these plots.

In addition to the full tau-lepton momentum, the resolution plots for the invariant mass  $M$  and the transverse momentum  $p_T$  of the  $a_1$  resonance are shown in Figure 2.28. The narrow Gaussian shape with a mean value of zero in both cases implies a very good agreement between the generated parameter and the outcome of the kinematic fit. As the  $a_1$  resonance corresponds to the hadronic energy of the tau lepton, it can be directly measured by the CMS tracker. One can estimate the impact of the initially unmeasured tau neutrino on the resolution of the transverse tau momentum by comparing Figure 2.27(a) and 2.28(b).

Besides the reconstruction of the momentum components of the particles involved, the kinematic fit also determines the position of the tau-decay vertex. In Figure 2.29 the spatial resolution of the fitted tau-decay vertex is illustrated in Cartesian coordinates. One should keep in mind that this coordinate system is the one of the CMS detector, thus it is fixed and not relative to the tau-lepton momentum. Independent of the three space coordinates, the obtained vertex resolution is represented by a peak without any systematic shift. However, the resolution of the  $z$  coordinate is worse than for the transverse coordinates  $x$  and  $y$ . Nevertheless, the fit is able to reconstruct the position of the tau-decay vertex to high precision.

#### 2.4.6.1. Intrinsic Resolution

The resolution of the final tau parameters is the combined effect of the track reconstruction and the application of the kinematic fit adding an unmeasured particle. The result is limited by the secondary vertex and the derived Gottfried-Jackson angle  $\theta_{GJ}$ . The current algorithm

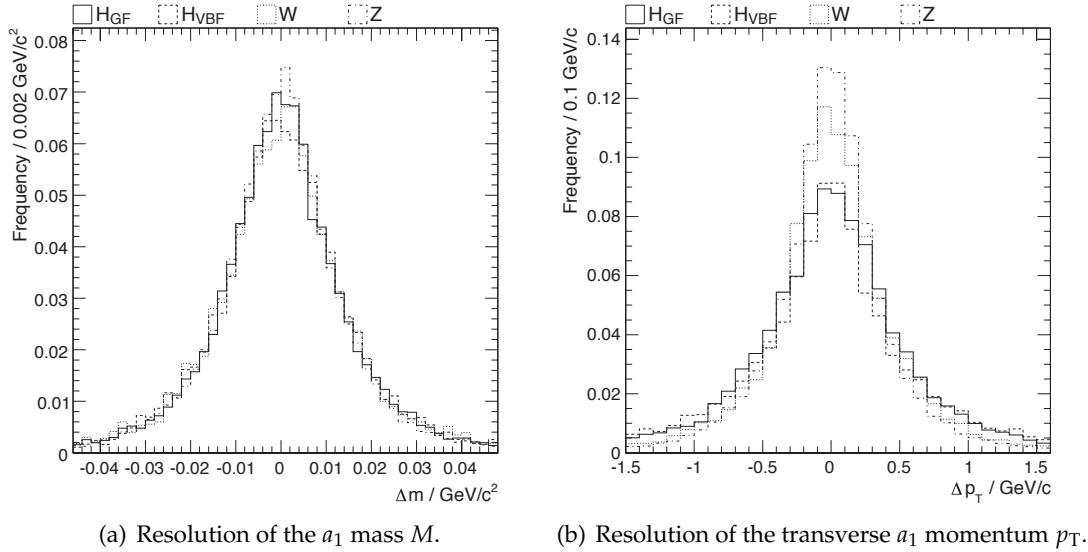


Figure 2.28.: Mass and transverse-momentum resolution of the  $a_1$  resonance. The quality criteria of the kinematic fit are not applied in these plots.

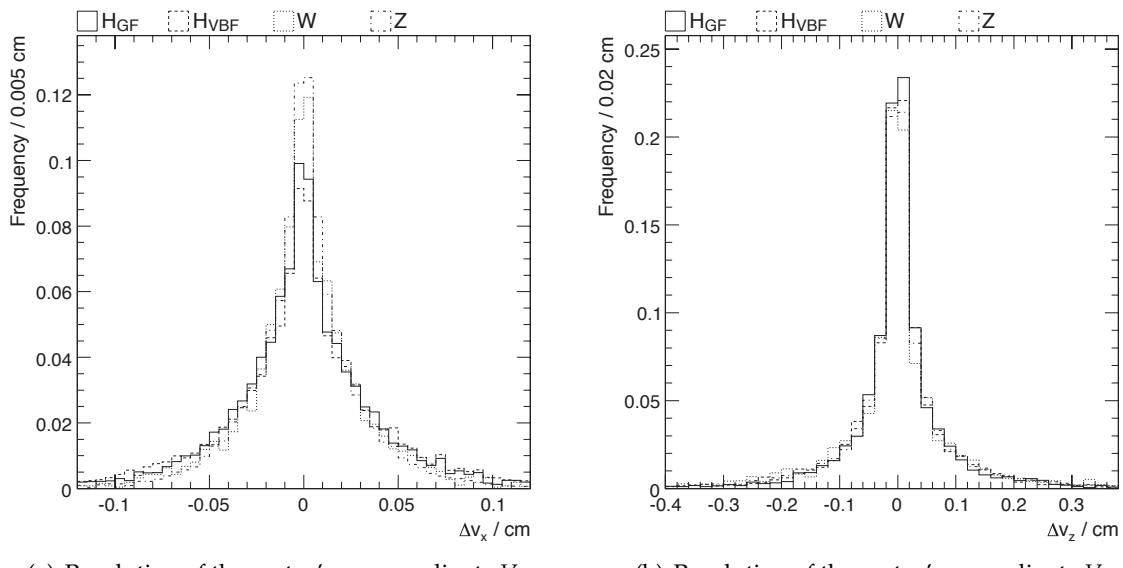


Figure 2.29.: Spatial resolution of the kinematically fitted tau-decay vertex in Cartesian coordinates. The plot in (a) is valid for both the  $x$  and the  $y$  coordinate, as their resolution is identical within numerical fluctuations. The quality criteria of the kinematic fit are applied in these plots.

starts from an approximated value  $\theta_{\text{GJ}}^{\max}$  (see Section 2.2.2). The initial value of the tau momentum is biased according to (2.1). This offset can be calculated as the difference between the tau momentum obtained from the maximum  $\theta_{\text{GJ}}^{\max}$  and the true angle  $\theta_{\text{GJ}}$ :

$$\Delta\tau(\theta_{\text{GJ}}, m_{a_1}, p_{a_1}) = |\vec{p}_\tau|(\theta_{\text{GJ}}^{\max}, m_{a_1}, |\vec{p}_{a_1}|) - |\vec{p}_\tau|(\theta_{\text{GJ}}, m_{a_1}, |\vec{p}_{a_1}|) . \quad (2.20)$$

Replacing  $\theta_{\text{GJ}}^{\max}$  by using (2.4) in Section 2.2.2.1 and transforming  $\theta_{\text{GJ}}$  into the tau rest frame yields an analytic equation  $\Delta\tau(\cos\theta_{\text{GJ}}^*, m_{a_1}, p_{a_1})$ . This expression depends strongly on the particular value of  $\cos\theta_{\text{GJ}}^*$ , as shown in Figure 2.30. For typical  $a_1$  parameters it varies over a broad range of about  $40\text{ GeV}/c$ . The Lorentz transformation of  $\cos\theta_{\text{GJ}}$  results in an asymmetry in the plotted distribution between the region corresponding to the positive square root in (2.1) (labeled  $\tau^{(+)}$ ) and the negative one (labeled  $\tau^{(-)}$ ).

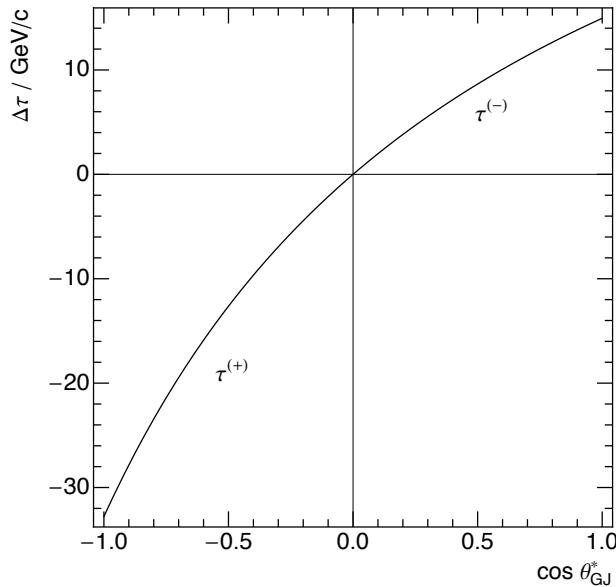


Figure 2.30.: Dependency of the calculated momentum resolution of the tau lepton  $\Delta\tau(\cos\theta_{\text{GJ}}^*, m_{a_1}, p_{a_1})$  w.r.t.  $\cos\theta_{\text{GJ}}^*$  assuming  $p_{a_1} = 40\text{ GeV}/c$  and  $m_{a_1} = 1.2\text{ GeV}/c^2$ .

The intrinsic error of the algorithm is estimated by always assuming  $\theta_{\text{GJ}} = \theta_{\text{GJ}}^{\max}$ . One can calculate the expectation value of (2.20) by

$$\mathbb{E}(m_{a_1}, p_{a_1}) \equiv \langle \Delta\tau(\cos\theta_{\text{GJ}}^*, m_{a_1}, p_{a_1}) \rangle = \int_{-1}^1 d\cos\theta_{\text{GJ}}^* \Delta\tau(\cos\theta_{\text{GJ}}^*, m_{a_1}, p_{a_1}) \mathbb{P}(\cos\theta_{\text{GJ}}^*) . \quad (2.21)$$

Here,  $\mathbb{P}(\cos\theta_{\text{GJ}}^*)$  denotes the probability density of the Gottfied-Jackson angle distribution in the tau rest frame. The integral is taken over the domain of  $\cos\theta_{\text{GJ}}^*$ . The resulting function  $\mathbb{E}(m_{a_1}, p_{a_1})$  only depends on the mass of the  $a_1$  and its momentum  $p_{a_1}$ , and can therefore be measured to a high precision for a given density. Assuming a uniform distribution for  $\mathbb{P}(\cos\theta_{\text{GJ}}^*)$ , one obtains the behavior depicted in Figure 2.31. The mean error of the algorithm vanishes with heavy  $a_1$  masses and increases with growing  $p_{a_1}$ . The same holds for the evolution of the one sigma region. The center value of both curves is consistently shifted to one direction depending on the shape of  $\mathbb{P}(\cos\theta_{\text{GJ}}^*)$ .

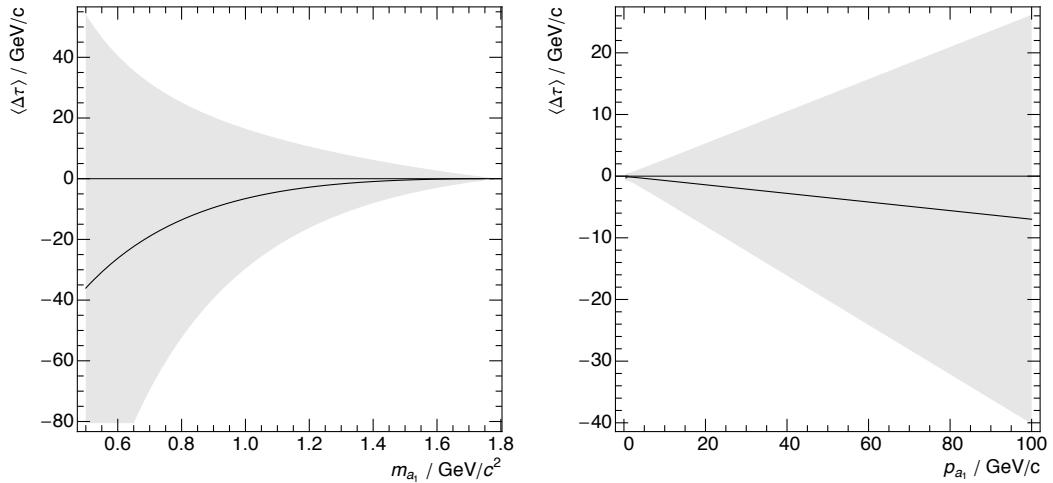


Figure 2.31.: Dependencies of the expectation value of the tau-momentum resolution for uniformly distributed  $\cos \theta_{GJ}^*$ . The filled region represents the one sigma belt. Left: assuming  $p_{a_1} = 40 \text{ GeV}/c$ , Right: assuming  $m_{a_1} = 1.2 \text{ GeV}/c^2$ .

Considering a more suitable distributed  $\mathbb{P}(\cos \theta_{GJ}^*)$ , one can explain the conspicuous offset of the transverse momentum in gauge-boson sources (Figure 2.27(a)). The calculated dependencies in Figure 2.31 can be qualitatively reproduced in the generated tau decays. Even a correction of this offset can be derived, especially for the linear increase w.r.t. the  $a_1$  momentum. This does not only result in a corrected mean of the transverse-momentum resolution but also affects its shape. After applying a roughly estimated correction function, one can obtain a narrow Gaussian shape. The precise equation of this correction has to be adapted to the particular tau-production mechanism. Also different event selections, e.g. by the high-level trigger, modify the probability density of the Gottfied-Jackson angle distribution in the tau rest frame and therefore change the dependencies.

## 2.5. Technical Implementation in the CMS Software Framework

The implementation of the kinematic tau fit, as described in Section 2.4.2, is distributed over several components. The package consists of the `InputTrackSelector`, the `ThreeProngInputSelector`, the `KinematicTauCreator`, the `ThreeProngTauCreator` and the `KinematicTauProducer`. These modules are located in `CMSSW/RecoTauTag/KinematicTau`. As illustrated in Figure 2.32, they need to be executed in a specific order. One should keep in mind that this workflow is only valid for the kinematic reconstruction of 3-prong tau decays. Additional decay modes can be added by developing appropriate selector and/or creator modules. The implementation of these additional modules within the fit can be done in the same way as for the 3-prong tau decays. The structure of the `KinematicTau` package is designed to be very flexible, allowing the user for example to specify the collection of particle candidates or tracks, which shall be used as input for the fit.

The workflow is shown for 3-prong tau decays. It uses particle-flow tau candidates as the input. While the particle-flow algorithm is executed via the `PFTauProducer` during the CMS event reconstruction, the `KinematicTau` package needs to be subsequently executed by the user. In the following Sections the different modules are discussed in detail.

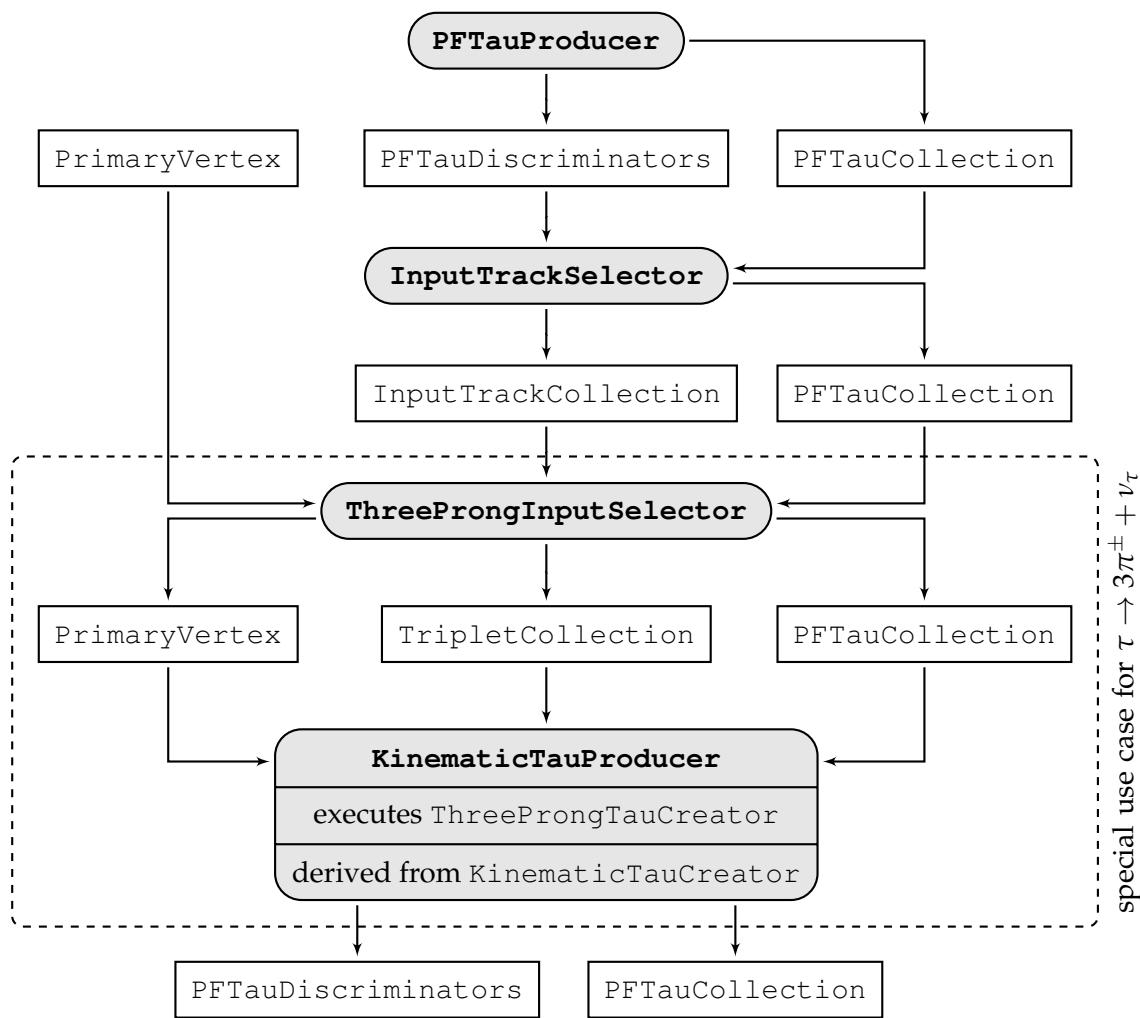


Figure 2.32.: Technical workflow of the kinematic fit. The specialized reconstruction of 3-prong tau decays is shown here.

### 2.5.1. The `InputTrackSelector` and the `ThreeProngInputSelector`

The `InputTrackSelector` is used to select particle-flow tau candidates for processing. This corresponds to the steps described in Section 2.4.2.1. There are several user-definable switches. First of all, one can set the type of the `PFTauProducer`. The default for this option is the fixed cone method. Second, the user can request a minimal number of tau candidates, which should be selected by this module. If this number is not reached by an event, the return value of the `InputTrackSelector` is set to false, which results in the abortion of the rest of the workflow for this particular event. In addition, the minimal number of tracks within the signal cone of the particle-flow tau candidate can be specified. One-prong and multi-prong decay modes can be efficiently disentangled by this value. Optionally, the user can choose to filter particle-flow tau candidates via a set of existing `PFTauDiscriminators`. The output of the `InputTrackSelector` is a set of track collections and a `PFTauRef` collection storing all particle-flow tau candidates, which are selected by the module.

In contrast to the `InputTrackSelector`, the `ThreeProngInputSelector` is not independent of the tau-decay mode, but specialized in 3-prong tau decays. This module implements the steps and functionalities discussed in Sections 2.4.2.2 to 2.4.2.5. Therefore, it receives the output of the `InputTrackSelector`, creates track-triplet combinations out of the track collections, and deletes duplicates. In order to determine the best track triplet for a particle-flow tau candidate, a primary vertex collection needs to be provided to this module. Within this module the primary vertex is also renewed (see Section 2.4.2.4). The user can choose to filter the new vertex by the minimal number of tracks, which are used for the vertex fit and by an upper bound on the normalized  $\chi^2$  value of the applied vertex fit. If the new vertex fulfills these criteria, the `ThreeProngInputSelector` releases it to the event stream. As for the `InputTrackSelector` there is a user-definable minimal number of tau-lepton candidates, which should survive this module. All in all, the output of the `ThreeProngInputSelector` consists of a primary vertex, a collection of track triplets and a `PFTauRef` collection. The last one is stored to enable an unambiguous reference between the initial particle-flow tau candidate and the selected track triplet.

### 2.5.2. The `KinematicTauCreator` and the `ThreeProngTauCreator`

The `ThreeProngTauCreator` derives from the `KinematicTauCreator` class, which is implemented as a purely abstract base class. The `KinematicTauCreator` provides basic functionalities to interact with the underlying core of the kinematic fit. Furthermore, the generic retrieval of the refitted tracks and the refitted tau lepton are realized by this class. As these features are completely independent of a specific tau-decay mode, various decay-specific refit strategies can be easily deployed within the `KinematicTauProducer`. In principle, the execution of the kinematic fit by this class can be done at an arbitrary step of a physics analysis and does not necessarily need to be performed by a special framework module.

The `ThreeProngTauCreator` is a realization of the `KinematicTauCreator`, which focusses on the kinematic reconstruction of 3-prong tau decays. With the help of this class, the physically correct start situation, which is needed by the kinematic fit, is established. Therefore, as described in Section 2.4.2.6, a primary vertex rotation is applied to achieve a tau-decay topology of  $\theta_{GJ} = \theta_{GJ}^{\max}$ . The tau neutrino is built as a virtual particle within this process. Its momentum is calculated and initially the Huge Error Method is applied.

Subsequently, all four tau-decay products are handed over to the core of the kinematic fit. The kinematic fit is performed using a specific set of constraints, which is described in Section 2.2.3, and finally delivers the refitted 3-prong tau decay.

### 2.5.3. The KinematicTauProducer

The `KinematicTauProducer` module is a convenient way to control the execution of the `ThreeProngTauCreator` or any other realization of the `KinematicTauCreator` class within the CMS software framework. The producer needs the output of the `ThreeProngInputSelector` module as input data. It requires a primary vertex, a track triplet for each potential tau-lepton candidate, and the reference to the initial particle-flow tau collection. Optionally, the user may specify parameters to control the convergence of the kinematic fit (see Section 2.4.2.6).

For each event the `KinematicTauProducer` returns a `PFTauCollection` and two `PFTauDiscriminators`. The output of the entire workflow is a collection of refitted tau leptons compatible with the initial particle-flow tau collection. However, one should keep in mind that not every particle-flow tau candidate, which enters this workflow, actually results in a refitted tau lepton. The `PFTauCollection`, which is produced by the `KinematicTauProducer`, consists of the same entries as the initial collection. Only if the kinematic fit succeeds in creating a refitted tau lepton, a particle-flow tau is created using the four-vector of the refitted  $a_1$  resonance. Accordingly, the vertex of the particle-flow tau is set to the refitted secondary vertex. The cause for taking the momentum of the  $a_1$  resonance instead of that of the tau lepton is to achieve a comparability between the particle-flow algorithm and the kinematic refit. As particle flow accounts for the visible tau energy (see Section 2.4.1), it seems reasonable to choose the  $a_1$  resonance, as it exactly represents this state. The parameters corresponding to the complete tau decay including the fitted neutrino information is stored as an alternative four-vector.

The updated tau leptons can be selected from the `PFTauCollection` by a special `PFTauDiscriminator`. The discriminator is true for a successfully refitted tau lepton and false for all other cases. In addition to this, the user may discriminate the refitted tau leptons on the basis of the quality of the kinematic fit. The return value of the according `PFTauDiscriminator` determines if all of the quality criteria are passed by the tau candidate. A detailed description of the applied quality criteria can be found in Section 2.4.3. It should be noted that this `PFTauDiscriminator` corresponds to the set of loose quality criteria only. In this way it can be efficiently used for a large variety of tau-production processes and thus a wide range of tau-lepton energies.



## Chapter 3

# Higgs Boson Reconstruction from Hadronic Tau Pair Decays

There are several analyses within the CMS collaboration searching for the Higgs boson in the Standard Model and its extensions like the MSSM. The various decay channels are covered independently. Analyses using tau leptons mostly rely on at least one leptonic tau decay. The identification and selection of the isolated lepton in the final-state is a key signature of the events. The analysis introduced in this Chapter complements these channels by studying the SM Higgs decay into two tau leptons both decaying hadronically according to  $\tau \rightarrow 3\pi^\pm + \nu_\tau$ . Their reconstruction utilizes the kinematic approach described in Chapter 2. The prospects for a Higgs search in this channel are studied by a Monte Carlo analysis.

The experimental setup and the analyzed datasets are described in Section 3.1. After discussing the event topology in Section 3.2, the selection strategy is introduced in Section 3.3. The behavior of the HLT is inspected in Section 3.4. The analysis objects are selected in Section 3.5, from which the event selection is derived in Section 3.6. The final variable, the invariant mass of the reconstructed Higgs boson, is analyzed in Section 3.7. The influence from background contamination and other systematics is estimated in Section 3.8. Finally, Section 3.9 gives the sensitivity of the analysis.

### 3.1. Datasets and Expected Event Rates

The Monte Carlo analysis is carried out for a center-of-mass energy of  $\sqrt{s} = 7\text{ TeV}$ . The datasets are generated with PYTHIA (version 6.4, using CTEQ6 [98]) and the event simulation and reconstruction is done with CMSSW in version 3.6.2 (see Section 1.4.). The simulation uses a magnetic field of 3.8 T. The detector conditions are chosen according to a scenario expected during the startup of the LHC operation, START36\_v9. A realistic displacement and smearing of the interaction point is considered.

#### 3.1.1. Signal-like Processes

The two leading production mechanisms of SM Higgs bosons at hadron colliders are considered as signal processes. Separate datasets are generated for Higgs decays into tau pairs from gluon fusion  $H^0_{\text{GF}} \rightarrow \tau\tau$  (referred to as GF) and from vector-boson fusion  $H^0_{\text{VBF}} \rightarrow \tau\tau$  (referred to as VBF) in a private production. In both processes the branching fraction of the desired decays is set to 100 %. The Higgs boson is forced to decay into a tau pair. Both tau-leptons decay into three charged pions and a neutrino  $\tau \rightarrow 3\pi^\pm + \nu_\tau$ .

Thus, every event contains the decay chain

$$\begin{aligned} pp \rightarrow H^0 &\rightarrow \tau^- \tau^+ \\ \tau^- &\rightarrow \pi^- \pi^+ \pi^- + \nu_\tau \\ \tau^+ &\rightarrow \pi^+ \pi^- \pi^+ + \bar{\nu}_\tau . \end{aligned} \quad (3.1)$$

Four different Higgs masses are chosen between  $115 \text{ GeV}/c^2$  and  $145 \text{ GeV}/c^2$ . The range regards the lower experimental exclusion limit and the onset of the Higgs decay into  $W^\pm$  pairs. The width of the Higgs boson is neglected. In the Standard Model it varies from about 3 MeV to 11 MeV within the selected mass range [20].

Table 3.1 summarizes the datasets and the number of generated events. The last column contains the expected cross sections of the whole decay chain  $\sigma_{\text{sample}}$  according to

$$\sigma_{\text{sample}} = \sigma(pp \rightarrow H^0) \cdot \mathcal{B}(H^0 \rightarrow \tau\tau) \cdot \mathcal{B}(\tau \rightarrow 3\pi^\pm + \nu_\tau)^2 . \quad (3.2)$$

The production and decay of the Higgs boson are discussed in Section 1.1.2. The values for the Higgs properties at each mass point are calculated in [20]. For the decay  $\tau \rightarrow 3\pi^\pm + \nu_\tau$  the branching ratio is taken from [85].

Table 3.1.: Number of generated events and the expected cross sections of the analyzed datasets. The Higgs cross sections and branching ratios are derived from [20].

process	$m_{H^0} / \text{GeV}/c^2$	generated events	cross section / pb
GF	115	100 000	$1.200 \cdot 10^{-2}$
GF	125	100 000	$8.43 \cdot 10^{-3}$
GF	135	100 000	$5.11 \cdot 10^{-3}$
GF	145	100 000	$2.55 \cdot 10^{-3}$
VBF	115	100 000	$8.8 \cdot 10^{-4}$
VBF	125	100 000	$6.7 \cdot 10^{-4}$
VBF	135	100 000	$4.3 \cdot 10^{-4}$
VBF	145	100 000	$2.3 \cdot 10^{-4}$
Z3pr	-	1 000 000	6.3
QCD	-	$7.18 \cdot 10^8$	$3.5 \cdot 10^6$

There are no acceptance requirements applied during the generation. No pile-up or underlying events are simulated for the signal.

The polarization of the tau pair may be taken into account with the TAUOLA package. To estimate the effects of the spin correlation, a small dataset is generated separately using TAUOLA. A comparison to the PYTHIA samples shows no significant differences in the important distributions and efficiencies. To simplify the generation process, the TAUOLA package is not used in the final samples.

### 3.1.2. Background Processes

Tau pairs from  $Z^0$ -boson decays and hadronic jets from QCD multi-jet events are considered as main background sources in the analysis.

The first background contains two signal-like tau leptons. Their invariant mass is close to the expected Higgs mass. Off-shell  $Z^0$  bosons are indistinguishable from a Higgs boson produced by gluon fusion and have a comparable rate to the signal.

Electroweak processes that result in a  $Z^0$  boson accompanied by additional jets are not considered due to their small cross sections in the signal region. The  $Z^0$  production via vector-boson fusion provides the same final state as the VBF Higgs decay. The cross section is in the same order as the Higgs signal [99]. The fraction of off-shell  $Z^0$  bosons from this production type is small compared to the signal.

QCD jets fake real tau signatures. There is a huge cross section for this process. It exceeds other electroweak backgrounds such as  $t\bar{t}$  processes by a large factor (compare Figure 1.9).

In a private production the background from  $Z^0$  bosons (referred to as Z3pr) is generated in the same way as the signal samples. Both tau leptons decay hadronically as  $\tau \rightarrow 3\pi^\pm + \nu_\tau$ . The inclusive decay modes can be ignored as they pass the applied reconstruction with little efficiency. The cross section is determined from PYTHIA by generating 10000 events with Standard Model cross sections. The branching ratio into tau pairs is obtained from [85].

The QCD sample was introduced in Section 2.4.3. It is obtained from an official production, Summer10, and is labeled as QCD\_EMEnriched\_Pt30to80 there. The dataset is pre-filtered at generator level to enhance the trigger efficiency. It is enriched by electrons, photons, charged pions, and charged kaons, which are above a transverse energy threshold of  $E_T > 20\text{ GeV}$ . The filter efficiency is 5.9 %. Reconstructed jets are selected in the range  $30\text{ GeV}/c < \hat{p}_T < 80\text{ GeV}/c$ . Lower energies are expected to fail the selection criteria during the analysis. Higher energies have a cross section too small to contribute. The dataset is referred to as QCD in the following.

The number of generated background events and the expected cross section of each sample are given in Table 3.1. For the Z3pr sample the cross section of the decay chain is calculated replacing  $H^0$  with  $Z^0$  in (3.2). The numbers for the QCD sample include the filter efficiency in the generator.

## 3.2. Event Topology

Selection criteria are derived from the topology of the analyzed events. The features of the signal-like decays originate from the Higgs properties described in Section 1.1.2. All distri-

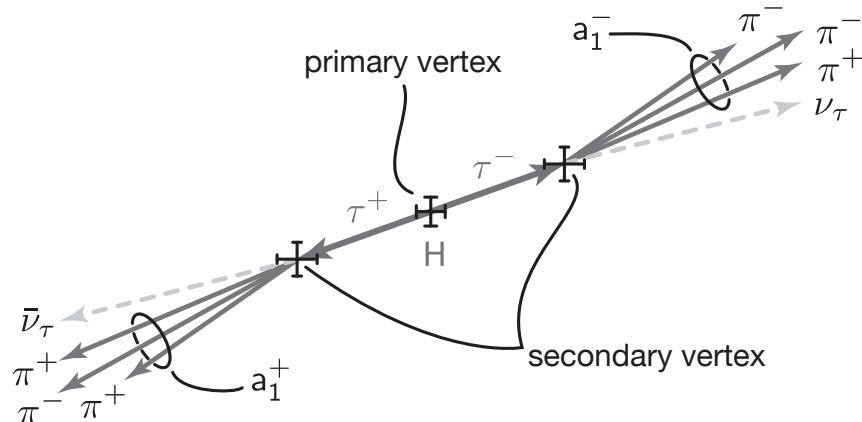


Figure 3.1.: Illustration of the Higgs decay into a tau pair. Both tau leptons are well separated and decay into three charged pions.

butions shown in this Section are obtained at generator level. They are scaled to the same integral to emphasize their shape, not their absolute rate. Higgs decays from gluon fusion and from vector-boson fusion show significant differences and have to be treated independently.

Figure 3.1 shows an illustration of the signal decay. The short-lived Higgs bosons decay instantaneously at their production point into pairs of tau leptons. Their average lifetime is large enough to provide distinguishable decay vertices. The tau leptons cover a large solid angle in the laboratory frame, which decreases for highly boosted Higgs bosons. The decay products of both tau leptons are collinear in the  $r\phi$ -plane and well separated within the detector.

### 3.2.1. Higgs Kinematics

The Higgs bosons from both production processes are uniformly distributed in the azimuthal angle  $\phi$  and differ in the polar angle  $\theta$ . Figure 3.2 compares the pseudorapidity distributions from GF and VBF. In both histograms there are no significant differences among the four curves, which represent different Higgs masses. Higgs bosons from GF are produced with large polar angles. The pseudorapidity reaches values up to  $|\eta| = 8$  and has two separated peaks at  $\eta \approx \pm 5$ . The VBF process results in lower values  $|\eta| < 4$ . The maximum is between  $-2 < |\eta| < 2$ .

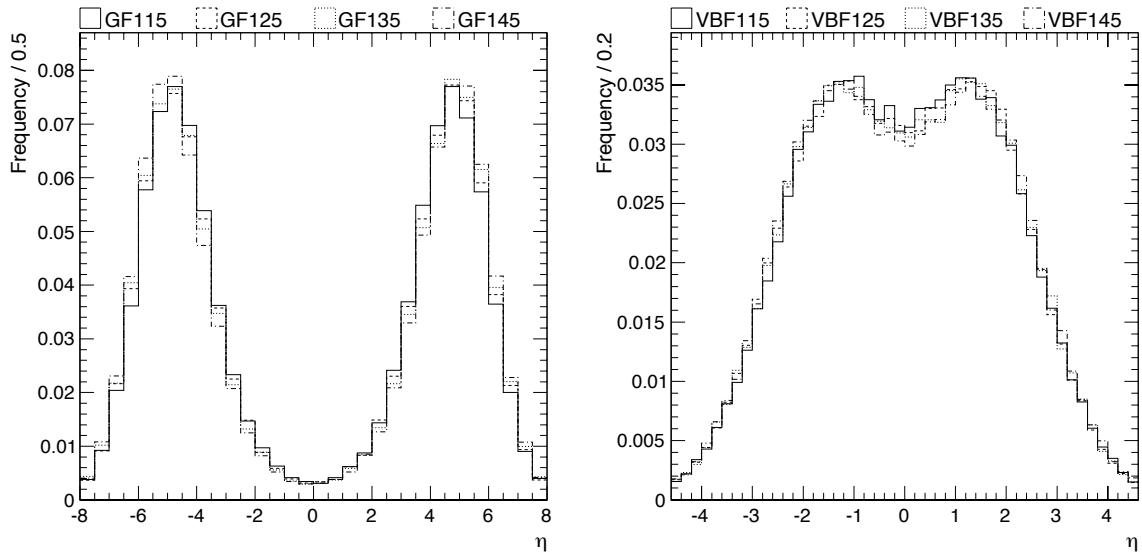


Figure 3.2.: Pseudorapidity distributions for Higgs bosons from gluon fusion (left) and vector-boson fusion (right) for four different Higgs masses.

Both processes provide a similar distribution of the momentum magnitude (Figure 3.3) with a decreasing tail beyond  $|\vec{p}| \approx 1$  TeV. The GF distributions peak at small values, the VBF curves at about  $|\vec{p}| \approx 150$  GeV/c. Combining the distributions of  $\eta$  and  $|\vec{p}|$  yields the transversal momentum component  $p_T$  as shown in Figure 3.4. Higgs bosons from GF are produced almost at rest w.r.t. to the transversal axis. At leading order they acquire no transversal momentum. The residual  $p_T$  is caused by higher order corrections. The VBF process provides a significant transversal momentum of up to a few hundred GeV. The distributions peak at  $p_T \approx 50$  GeV/c.

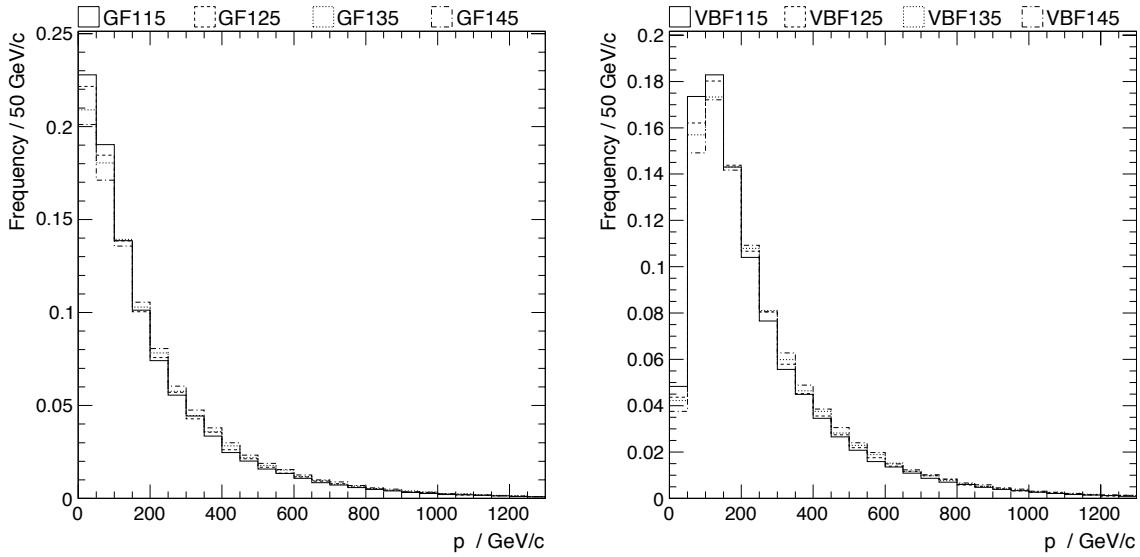


Figure 3.3.: Magnitude of the momentum vector  $p \equiv |\vec{p}|$  for Higgs bosons from gluon fusion (left) and vector-boson fusion (right) for four different Higgs masses.

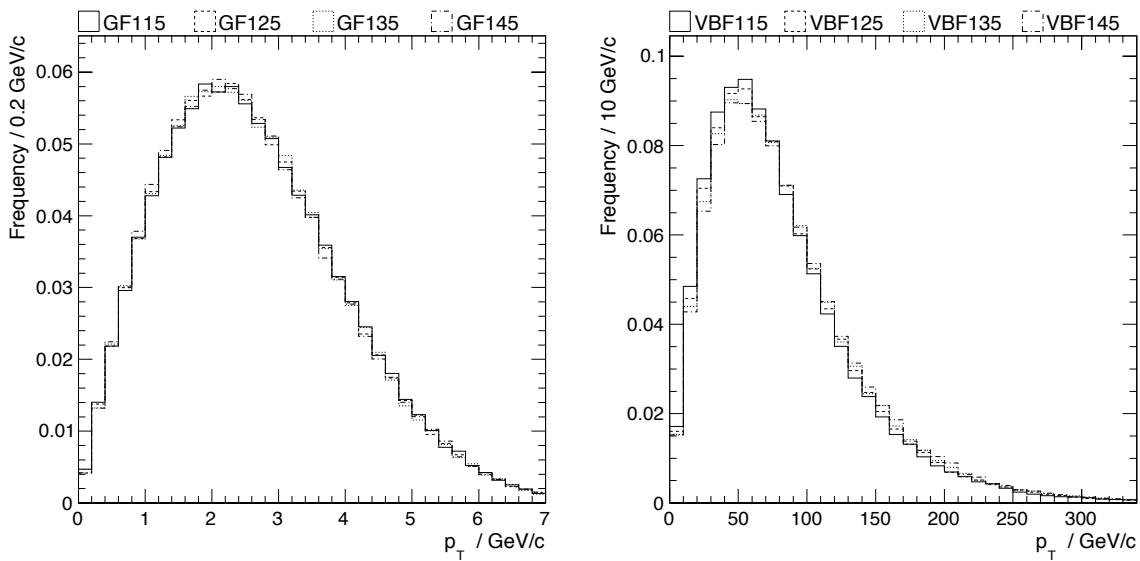


Figure 3.4.: Transversal momentum  $p_T$  of Higgs bosons from gluon fusion (left) and vector-boson fusion (right) for four different Higgs masses.

### 3.2.2. Tau Pair Kinematics

The properties of the Higgs production determine the distributions of the tau pairs. The opening angle between both tau leptons in the detector frame depends on the Lorentz boost of the Higgs boson. Figure 3.5 depicts the transversal angle  $\phi$  enclosed between both tau leptons. Tau pairs from the GF process are almost back-to-back due to their low transversal

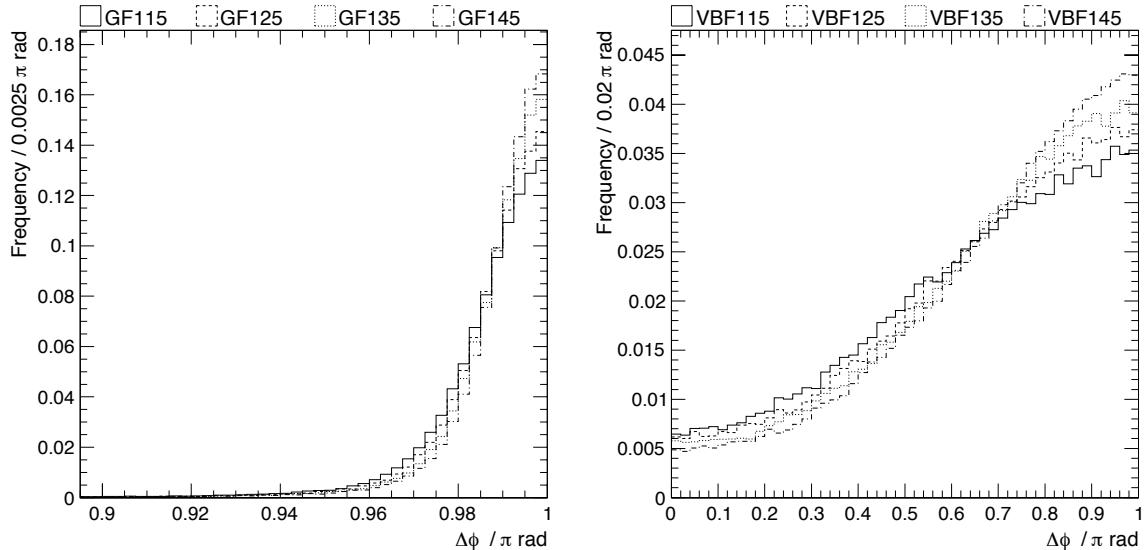


Figure 3.5.: Angular difference between the two tau leptons from Higgs decays in the transversal plane. Four different Higgs masses are shown for gluon fusion (left) and vector-boson fusion (right).

boost. There are only few tau pairs with an angular difference below  $\Delta\phi = 3$  rad. In the VBF process the distribution peaks at large angles. It has a tail towards  $\Delta\phi = 0$  rad caused by heavily boosted tau pairs. There is a small mass dependency for both production processes. The angles between tau leptons from heavier Higgs bosons are less affected by the transversal boost and exceed the contribution of light Higgs bosons in the region near  $\Delta\phi = \pi$ .

The angular difference in  $\eta$  has a maximum at zero for both processes and a falling slope up to  $\eta \approx 3$  for tau leptons from GF and up to  $\eta \approx 5$  for tau leptons from VBF.

The combined distribution,  $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ , is shown in Figure 3.6. Tau pairs from GF result in a narrow distribution with a sharp edge at  $\Delta R = \pi$  and a tail up to larger values. The VBF distribution is much wider. It peaks below  $\Delta R = \pi$  and the mean is shifted towards lower values. Both processes show a slight mass dependence. The distributions for lighter Higgs masses have a small shift towards lower values.

The different transversal Higgs momenta of the two production mechanisms cause a large difference in the transversal momentum balance  $\Delta p_T$ . As shown in Figure 3.7, tau pairs from GF are almost balanced. The distribution is centered at zero and has a width of about 2 GeV. The curves for VBF have large tails up to several 100 GeV.

### 3.2.3. Kinematics of Single Tau Leptons

The distributions for single tau leptons were introduced in Chapter 2. They cover the features of the decay  $\tau \rightarrow 3\pi^\pm + \nu_\tau$  independent of the production process. Here, additional variables are discussed that distinguish between production types.

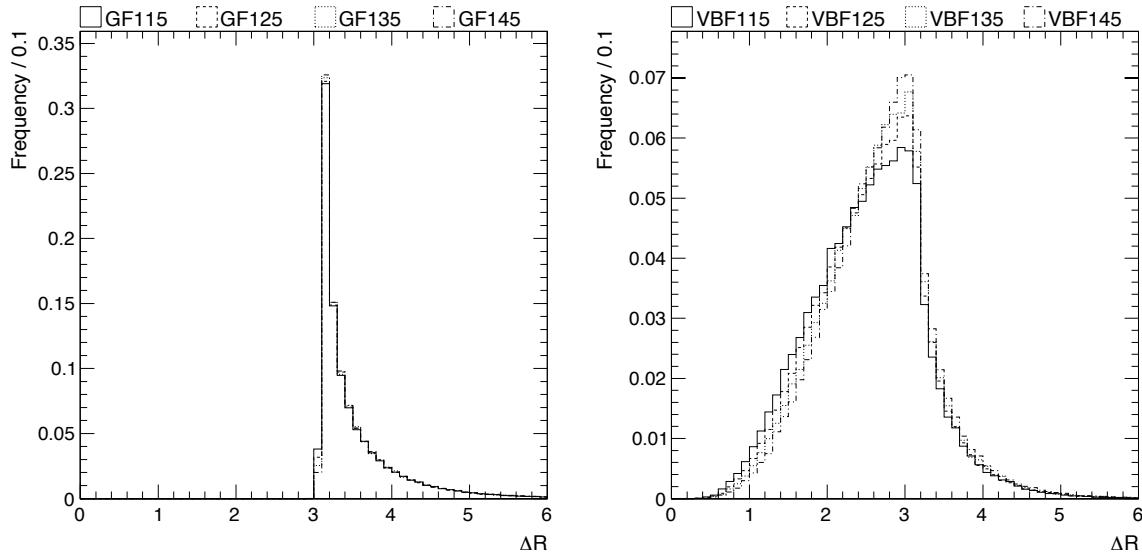


Figure 3.6.: Difference in  $\Delta R$  between the two tau leptons from Higgs decays. Four different Higgs masses are shown for gluon fusion (left) and vector-boson fusion (right).

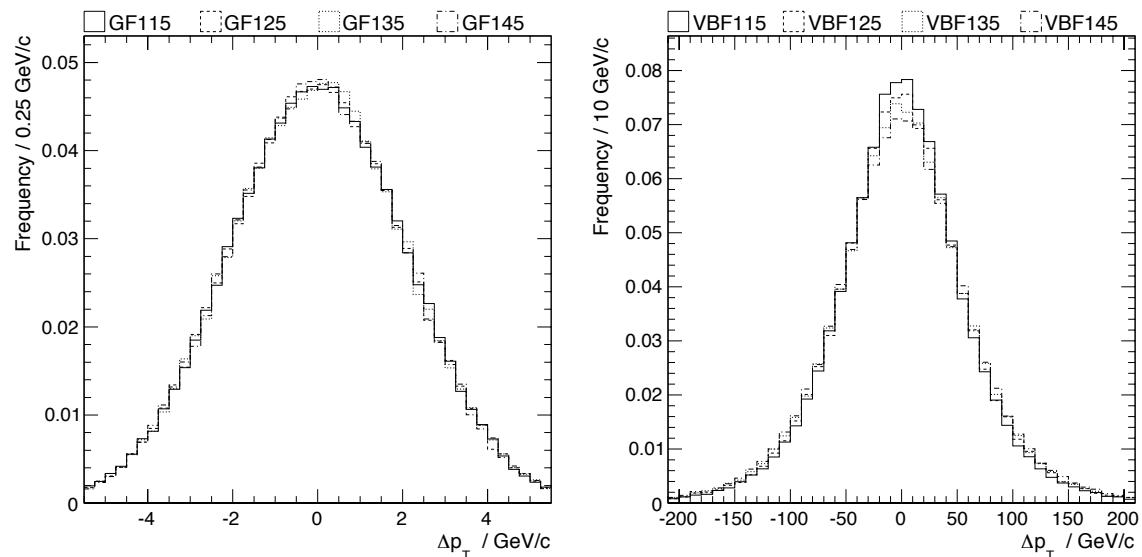


Figure 3.7.: Balance of the transversal momentum between the two tau leptons from Higgs decays. Four different Higgs masses are shown for gluon fusion (left) and vector-boson fusion (right).

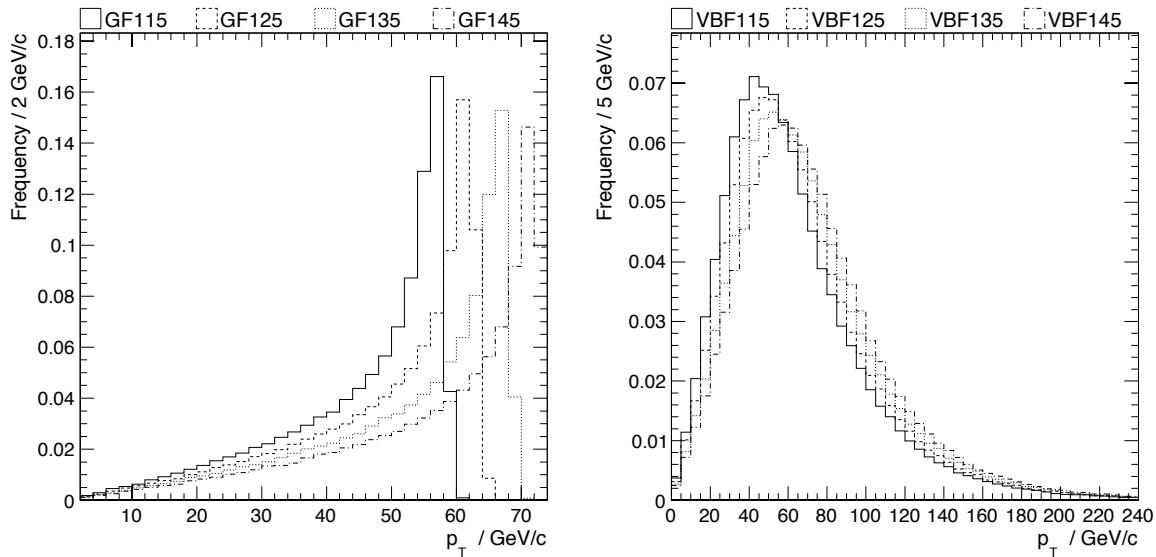


Figure 3.8.: Transversal momentum of tau leptons from Higgs decays. Four different Higgs masses are shown for gluon fusion (left) and vector-boson fusion (right).

The transversal boosts of the Higgs bosons result in the distributions of the tau  $p_T$  shown in Figure 3.8. For the GF process one obtains the typical shape of almost unboosted two-body decays. The upper edge is determined by half the mass of the decaying particle. It varies from  $p_T \approx 55 \text{ GeV}$  to  $p_T \approx 75 \text{ GeV}$ . This behavior is not found in case of the VBF. The large transversal boost exceeds the impact of the mass, which only slightly modifies the position of the peak. Independent of the Higgs mass, the distributions have large tails to higher momenta.

In both processes tau leptons are produced in the central region of the polar angle. As depicted in Figure B.3 of Appendix B.2, the  $\eta$  distribution is almost gaussian shaped and centered around zero. Its width is below 1.5 for both processes and all generated masses.

The opening angle of the 3-pion system decreases with larger tau boosts. The pion tracks become more and more collinear. This can lead to reconstruction inefficiencies and especially affect the higher boosted pions from VBF.

### 3.2.4. Kinematics of Quarks from VBF

In case of the vector-boson fusion, the Higgs boson is accompanied by the two initial quarks, which emitted the vector bosons (see the second diagram of Figure 1.3). Typically, they are only deflected slightly by the initial scattering and are detected as jets in the forward region. Jet activity in the central part is suppressed as there is no color exchange between the initial quarks.

The left part of Figure 3.9 shows the  $\eta$  distribution of the quarks, which peaks at  $\eta \approx \pm 2.5$  and has a minimum at zero. One expects only little overlap between the quark jets and the tau leptons from the Higgs decay. The quarks are uniformly distributed in  $\phi$ . As the initial quark momenta have opposite signs in  $p_z$ , the quarks are well separated after the collision. The right part of Figure 3.9 depicts this separation in terms of  $\Delta R$ . There are no entries below  $\Delta R \approx 1$ .

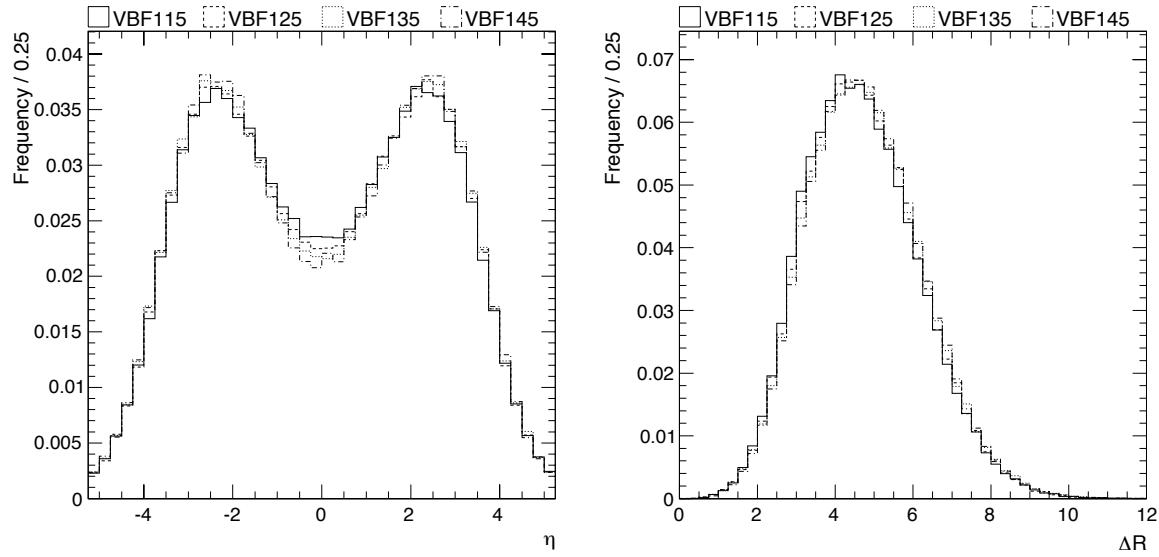


Figure 3.9.: Kinematics of the quarks in vector-boson fusion. The pseudorapidity (left) and the separation in terms of  $\Delta R$  (right) are shown.

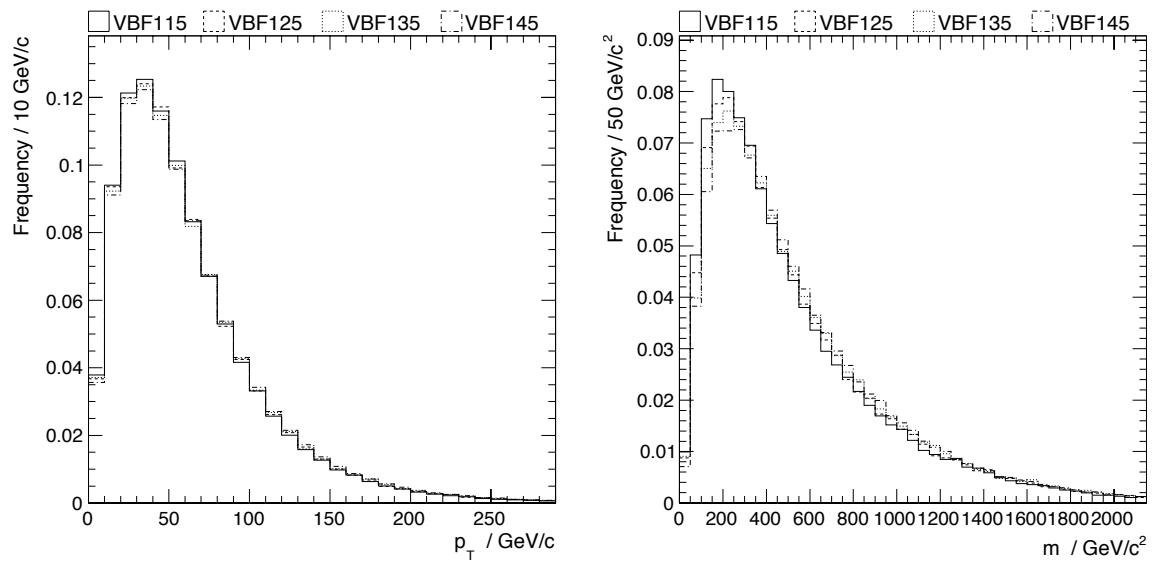


Figure 3.10.: Kinematics of the quarks in vector-boson fusion. The transversal momentum (left) and the invariant mass of the quark pair (right) are shown.

The production process results in quarks with significant transversal momentum, which is shown in the left part of Figure 3.10. The right part depicts the invariant mass of the quark pair. The distribution peaks at 200 GeV and extends to masses beyond 1 TeV.

### 3.2.5. Background Kinematics

The topology of the decay  $Z^0 \rightarrow \tau\tau$  is very similar to the GF process. The  $Z^0$  boson production leads to a comparable distribution of the momentum components, especially  $p_T$ . The pseudorapidity is slightly shifted to larger absolute values. It shows two maxima at  $\eta \approx \pm 5.5$ . The distributions of the transversal momentum and the pseudorapidity can be found in Figure B.4 of Appendix B.2.

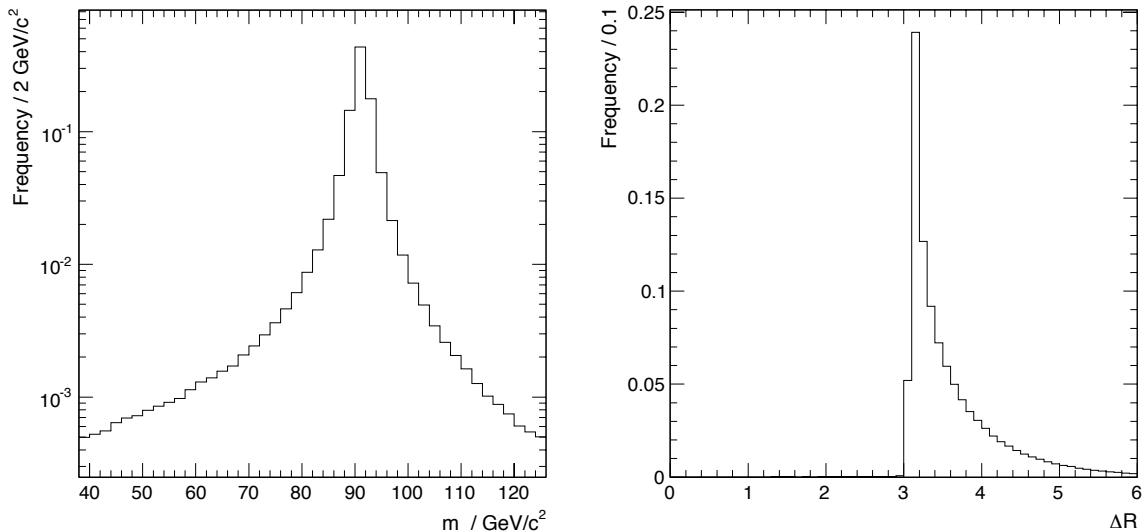


Figure 3.11.: Event topology of the  $Z^0$ -boson decay into tau pairs. The left figure shows the  $Z^0$ -boson mass. The right plot depicts the angle between both tau leptons expressed in  $\Delta R$ .

The major observable to distinguish between tau pairs from  $Z^0$  boson decays and from GF is the boson mass. The left plot in Figure 3.11 shows the mass distribution of the  $Z^0$  boson. Ignoring the experimental resolution, the maximum is significantly below the expected mass region for Higgs bosons. Small tails with rates below one per mille reach this region.

The kinematic distributions of the tau pairs do not provide much handle to separate the GF process. The angular differences in  $\phi$  have a minimal offset towards lower values due to the lighter invariant mass of the di-tau system. This affects the  $\Delta R$  distribution in the right plot of Figure 3.11. The left edge is even sharper as in the GF process and is shifted to  $\Delta R \approx 3$ .

The absence of a significant transversal boost results in a  $p_T$  distribution of both tau leptons, which is limited by  $m_{Z^0}/2$ . The upper edge lies at  $p_T \approx 46 \text{ GeV}/c$ , which is below the peaks of the GF distributions. This clean separation from the signal is reduced by the limited resolution of the reconstruction.

The backgrounds from QCD are highly affected by the event reconstruction. The sample contains jets that imitate the signal. There is no meaningful way to inspect the events at generator level.

### 3.3. Selection Strategy

The initial signal over background ratio  $S/B$  between the Higgs samples and the backgrounds listed in Table 3.1 is about  $10^{-10}$ . An efficient selection has to be developed to enhance this ratio. The selection is divided into several steps, which are summarized in the flowchart in Figure 3.12.

Two independent but mostly parallel selections are defined, one for the GF and one for the VBF process. Both selections start from a common trigger path. First, the required objects are reconstructed. These are tau candidates and, in case of the VBF selection, an additional jet pair.

Both selections rely on the kinematic tau reconstruction described in Chapter 2 and apply the quality criteria defined in Section 2.4.3. The selected tau decays are referred to as quality tau candidates. They have to pass further requirements in the common step labeled as custom tau selection (Section 3.5.1.1). The additional selection may be bypassed by fake jets from QCD to increase the statistics of the background sample. This is valid under assumptions that will be discussed in Section 3.5.2. The remaining tau decays are combined to pairs in Section 3.5.1.2. If more than one pair is reconstructed, exactly one is selected for further processing.

A separate jet selection collects jet candidates (Section 3.5.3) and delivers the best pair for the VBF analysis. These jets are vetoed against the selected tau candidates. At this stage exactly one pair of tau candidates and for the VBF one unambiguous jet pair is required. This selection layer may be bypassed by QCD and  $Z^0$  events.

The next step is different for the selection of Higgs decays from GF and VBF production. In Section 3.6 two specialized selections are described that combine the collected objects and apply further discrimination.

### 3.4. Event Trigger

The analysis relies on the trigger paths for events containing tau candidates. The trigger requirements depend on the instantaneous luminosity. Here, trigger parameters intended for  $\mathcal{L} = 8 \cdot 10^{29} \text{ cm}^{-2} \text{ s}^{-1}$  are used. This luminosity corresponds to the scenario of early data taking considered in the generated datasets.

Table 3.2.: Trigger requirements for single-tau and double-tau events.

trigger label	HLT requirements	L1 seeds
HLT_SingleLooseIsoTau20	$E_T > 20 \text{ GeV}$ , $a = 5 \text{ GeV}$ , $b = 0.025$ , $c = 0.0075$	L1_SingleTauJet20U or L1_SingleJet30U
HLT_DoubleLooseIsoTau15	$E_T > 15 \text{ GeV}$ , $a = 5 \text{ GeV}$ , $b = 0.025$ , $c = 0.0075$	L1_DoubleTauJet14U or L1_DoubleJet30U

The HLT menu (see Section 1.3.4) provides a trigger for single-tau and for double-tau events, `HLT_SingleLooseIsoTau20` and `HLT_DoubleLooseIsoTau15`. Both dedicated tau triggers require isolated tau jets with a minimum of transversal energy  $E_T$ . The criteria and the corresponding L1 seeds are given in Table 3.2. The parameters  $a$ ,  $b$ , and  $c$  determine the electromagnetic isolation. It is defined as  $I_{\text{EM}} < a + b \cdot E_T + c \cdot E_T^2$ , where  $I_{\text{EM}}$  is the energy deposition in the ECAL collected in an annulus of  $0.15 < \Delta R < 0.5$  around the jet axis. There

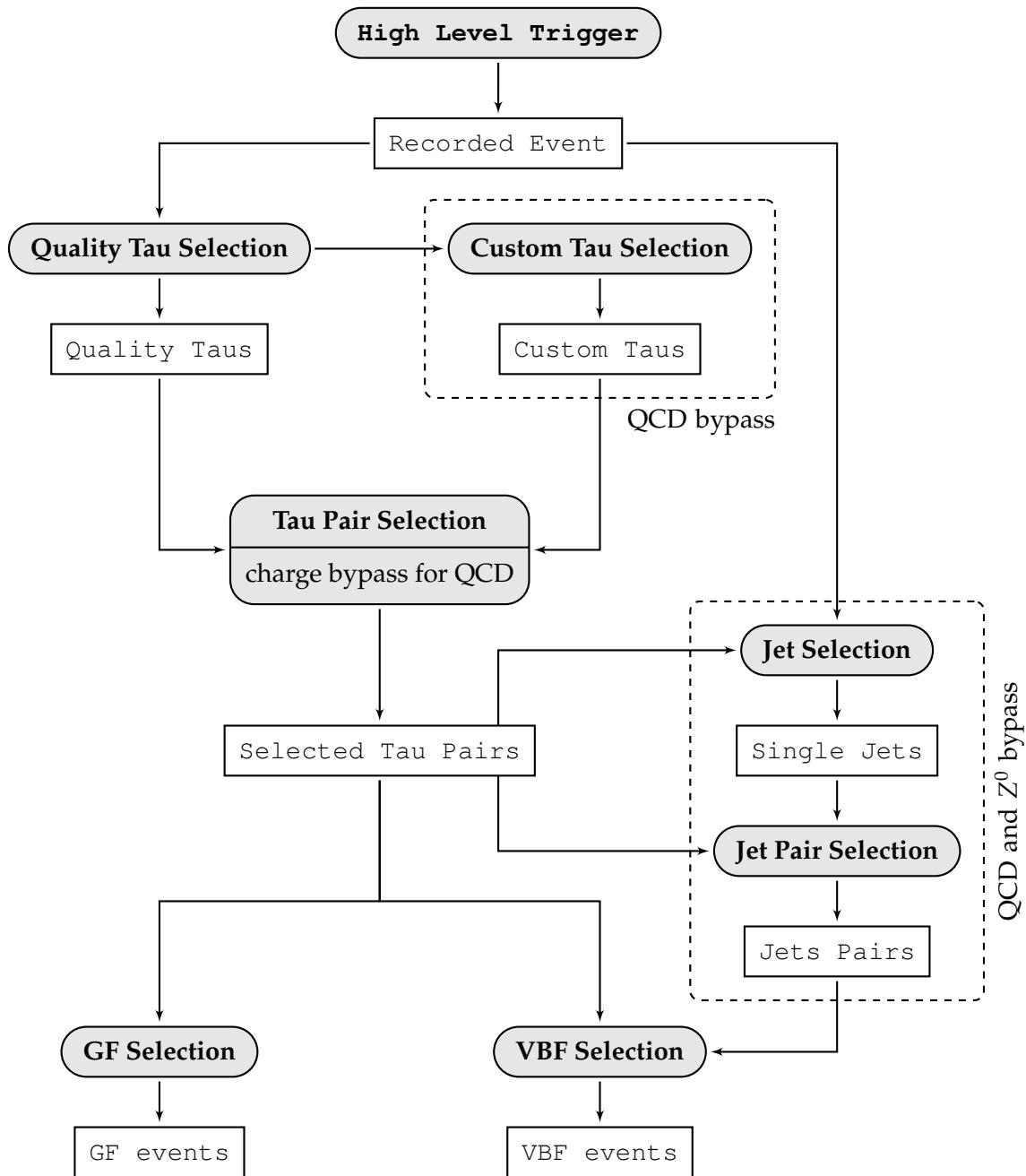


Figure 3.12.: Sketch of the selection strategy. See text for description.

is no pre-scale factor applied in both tau triggers. Every event meeting the requirements is accepted. A detailed description of the applied trigger algorithms at L1 and HLT and their expected performance can be found in [72].

Table 3.3.: Trigger efficiencies of the analyzed processes in percent and the recorded cross section. The value for the QCD sample marked with an asterisk considers the pre-filter at generator level.

$m_{H^0} / \text{GeV}/c^2$	$H^0_{\text{GF}}$				$H^0_{\text{VBF}}$				Z3pr	QCD
	115	125	135	145	115	125	135	145	-	-
HLT efficiency	60.5	67.1	72.3	76.6	93.2	94.6	95.4	96.1	28.7	46.7
cross section / fb	7.3	5.7	3.7	2.0	0.8	0.6	0.4	0.2	$1.8 \cdot 10^3$	$*9.6 \cdot 10^7$

The analysis accepts events that pass at least one of the two trigger paths. The combined efficiencies for the inspected processes are summarized in Table 3.3. Small deviations from the rates obtained in Table 2.2 originate from a larger statistics. The highly boosted tau leptons from VBF pass the trigger at a rate of more than 93 %. The trigger efficiency of the GF process varies between 60 % and 77 % depending on the generated Higgs mass. The additional energy in the di-tau system from larger Higgs masses increases the average tau momentum. The mass dependency has a higher impact on the less boosted tau leptons from GF. The backgrounds are triggered at rates below 50 %.

## 3.5. Object Selection

Events passing the HLT paths are available to the offline reconstruction (see Section 1.4.3), which creates higher-level objects and particle candidates. The analysis defines a selection of objects that are needed to reconstruct the signal-like events. For the Higgs reconstruction two tau leptons and for VBF two additional jet candidates are required. These objects have to be identified among other candidates in the event. For both kinds of objects, the stand-alone candidates are selected separately and are afterwards combined to pairs. Further quality requirements aim to enhance the background rejection. Histograms in this section are each normalized to an integral of one.

### 3.5.1. Tau Selection

The tau selection is divided into two parts. First an arbitrary number of single tau candidates is identified within the events. Out of these candidates tau pairs with opposite charge are combined. If more than one combination is possible, a criterion is defined to select one unambiguous pair.

#### 3.5.1.1. Single Tau Selection

The single tau selection regards only one distinct tau decay. Global event observables are not considered. The analysis is based on the tau leptons reconstructed by the kinematic fit described in Chapter 2. Tau candidates have to pass the quality criteria defined in Section 2.4.3 and are referred to as quality tau candidates. Events not containing at least one quality candidate are rejected. The efficiency of the analyzed samples is given in the first row of Table 3.5 and are already discussed in Section 2.4.5.

The remaining quality tau candidates are subjected to a tighter selection, referred to as custom tau selection. Again, the tau candidates are inspected, not the events. Further dis-

crimination is achieved by additional requirements. The following plots are based on tested decays, not events. Any number of quality decays per event is accepted. If the discussed behavior of the signal processes is independent of the generated Higgs mass, only one mass is shown in the plots.

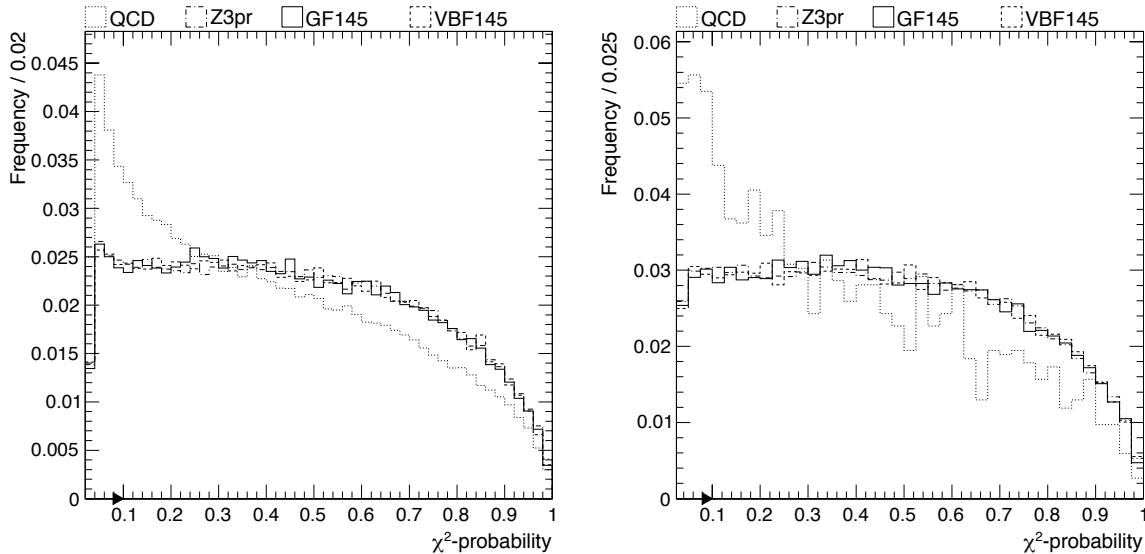


Figure 3.13.: Selection criterion:  $\chi^2$ -probability of the kinematic tau fit as stand-alone (left) and  $n-1$  representation (right). Tau candidates must have a  $\chi^2$ -probability  $\geq 0.1$  to pass this requirement.

Some of the quality criteria are tightened. The  $\chi^2$ -probability of the kinematic fit is required to exceed an increased value of 0.1. Figure 3.13 shows the corresponding distribution. The left plot shows tau candidates that pass the quality criteria. It differs from Figure 2.16(b) due to larger samples. The right plot contains decays that fulfilled all other requirements of the custom tau selection. The depicted criterion is not applied ( $n-1$  representation). The distributions from real tau leptons are almost flat while the curve for the QCD sample shows a nearly linear decrease with increasing  $\chi^2$ -probability.

The transversal energy fraction shown in Figure 3.14 provides further suppression of fake jets from QCD. The definition of this value was discussed in Section 2.4.3. The stand-alone representation in the left plot corresponds to Figure 2.18(b). The signal-like tau decays peak at about  $E_T^{\text{PF}}/E_T^{\text{kin}} \approx 0.75$ . QCD jets have a broad distribution up to  $E_T^{\text{PF}}/E_T^{\text{kin}} = 1$ , which is the maximal allowed value. The initial requirement is tightened to  $E_T^{\text{PF}}/E_T^{\text{kin}} = 0.95$ .

The other criteria of the quality discriminator are untouched.

The transversal momentum provides an effective rejection of QCD jets. Figure 3.15 depicts the distribution for the visible part of the tau momentum. Due to a better resolution, the  $a_1$  provides an improved discrimination compared to the tau lepton. Its transversal momentum distribution is depicted in Figure B.2 of Appendix B.2. The QCD curve shows a rapidly falling slope. Real tau decays exhibit distinct maxima at larger values. Their position depends on the boost and the mass of the tau mother. The larger the available energy the broader the curve. In case of the  $Z^0$  and the Higgs from gluon fusion, the boost is negligible. In average, the  $a_1$  momentum corresponds to half the mass of the tau mother diminished by the neutrino energy. The curve peaks at  $p_T \approx 35 \text{ GeV}$  for the  $Z^0$  and  $p_T > 50 \text{ GeV}$  for the GF

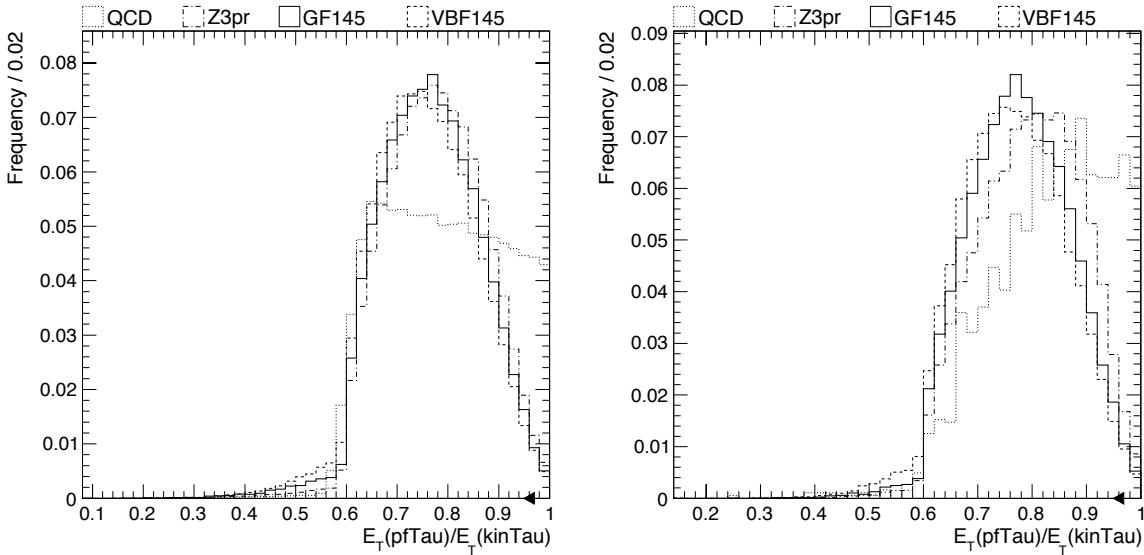


Figure 3.14.: Selection criterion: Transverse-energy fraction of tau leptons reconstructed with the particle-flow algorithm and the kinematic fit  $E_T^{\text{PF}}/E_T^{\text{kin}}$  as stand-alone (left) and  $n-1$  representation (right). Tau candidates need to fulfill  $E_T^{\text{PF}}/E_T^{\text{kin}} \leq 0.95$  to pass this requirement.

with a Higgs mass of  $m_{H^0} = 145 \text{ GeV}/c^2$ . The same generated mass results in a slightly lower maximum but large tails to higher values for the boosted tau leptons from VBF.

The mean of the curves shifts towards lower momenta with decreasing masses. This will cause more overlap to the distribution from Z3pr and weakens the separation. Figure 3.16 compares all four generated masses for both production mechanisms. The impact on the peak position is larger for the gluon fusion than in case of the large transversal boost of the VBF process.

The custom tau selection requires a transversal momentum of  $p_T(a_1) > 35 \text{ GeV}$ . This momentum region is chosen to contain all signal peaks but to exclude a large fraction of the QCD jets. The maximum of Z3pr is included. A tighter requirement is defined at a later stage of the dedicated gluon-fusion selection.

The kinematic fit modifies the initial direction of the tau candidate by taking the neutrino into account. This affects the orientation of the veto cone, which is defined on the particle-flow candidate. The distribution in Figure 3.17 counts the multiplicity of charged tracks  $n_{\text{iso}}$  within an annulus of  $\Delta R = 0.5$  around the new tau direction ignoring the charged tau daughters. If an additional track is found, the tau decay is rejected. True tau decays pass this requirement with a rate larger than 70 %. The curve for QCD jets is broad. It contributes with less than 15 % to the selected bin in both curves. The value is not affected by applying the other requirements of the custom tau selection.

The efficiencies of the single requirements are summarized in Table 3.4. True tau decays from the Higgs samples and Z3pr show similar values for the single requirements except for the transversal momentum. This criterion is mass dependent and distinguishes the different production types of the tau decays. Fake jets from QCD are mainly rejected by the momentum and isolation requirements. The cumulated efficiency is  $\epsilon_{\text{custom}}(\text{QCD}) < 1 \%$ . The Z3pr passes at  $\epsilon_{\text{custom}}(\text{Z3pr}) \approx 30 \%$  and the signal samples at  $\epsilon_{\text{custom}}(H^0) \approx 50 \%$ .

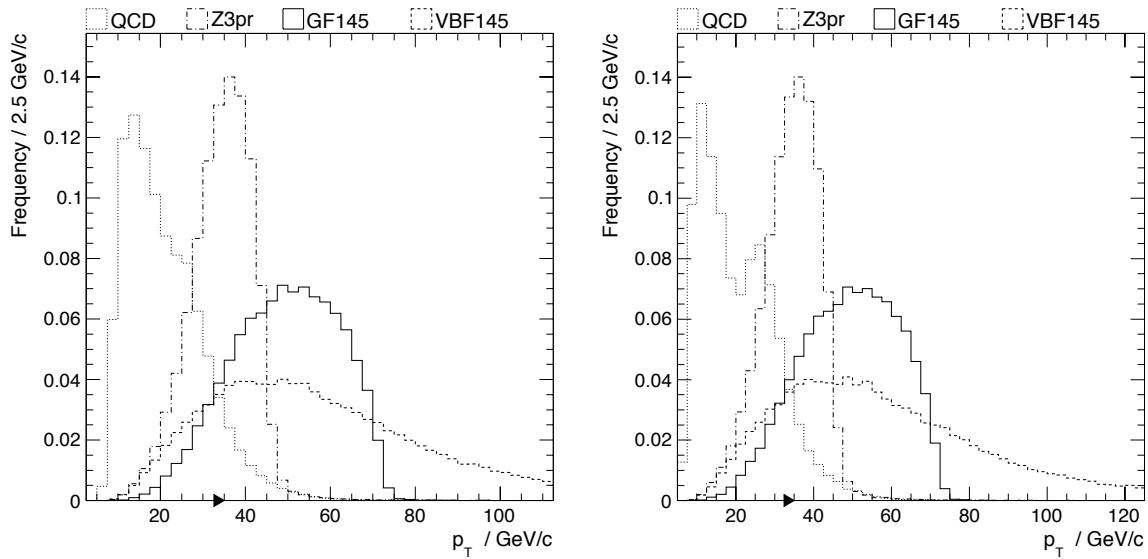


Figure 3.15.: Selection criterion: Transversal momentum of the  $a_1$  as stand-alone (left) and  $n-1$  representation (right). Tau decays have to exceed  $p_T > 35$  GeV to pass this requirement.

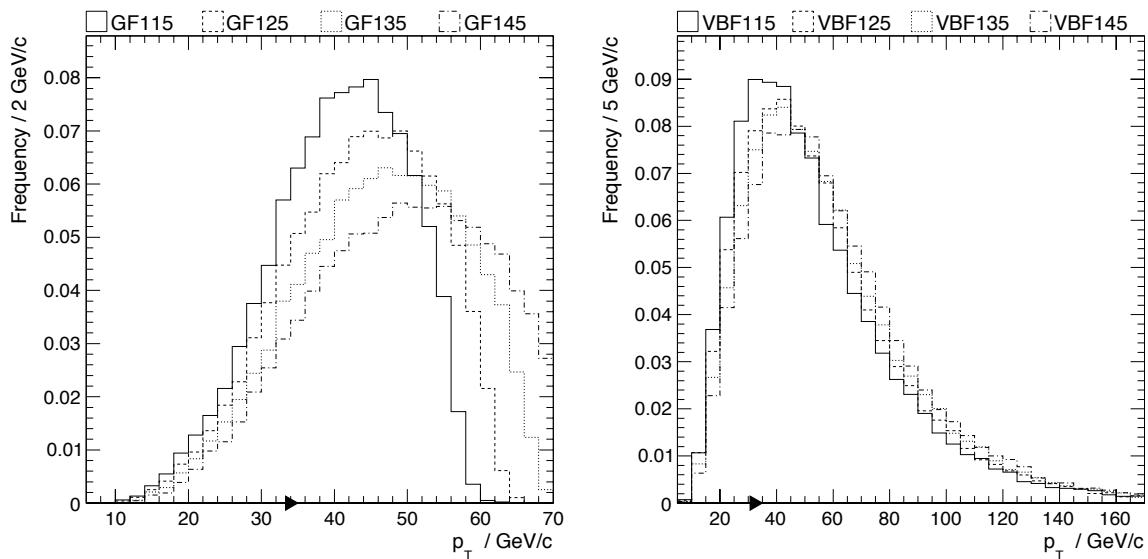


Figure 3.16.: Mass dependence of the transversal  $a_1$  momentum. The four generated Higgs masses are shown as  $n-1$  representation for Higgs production from gluon-fusion (left) and from vector-boson fusion (right).

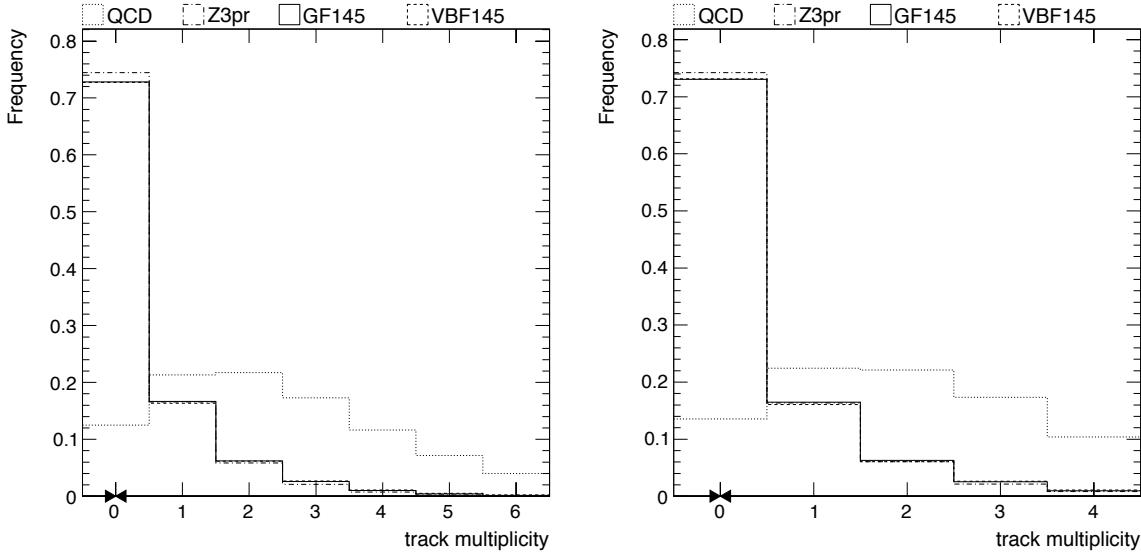


Figure 3.17.: Selection criterion: Track multiplicity within  $\Delta R < 0.5$  around the direction of the kinematic tau as stand-alone (left) and  $n-1$  representation (right). Both plots are zoomed to contain 98 % of the tested decays. The three tau daughters are excluded. Tau candidates are rejected if additional tracks exist in the annulus.

Table 3.4.: Efficiencies of the single tau selection. The numbers are given in percent of the tested tau decays (not events). Events contain at least one tau candidate that passed the quality discriminator of Section 2.4.3. Decays that pass all requirements are referred to as custom tau candidates.

$m_{H^0} / \text{GeV}/c^2$	$H^0_{\text{GF}}$				$H^0_{\text{VBF}}$				Z3pr	QCD
	115	125	135	145	115	125	135	145	-	-
$\chi^2$ -probability $\geq 0.1$	91.1	91.1	91.0	91.1	91.0	90.8	90.9	91.1	91.0	85.9
$E_T^{\text{PF}}/E_T^{\text{kin}} \leq 0.95$	97.6	97.8	97.8	97.8	98.0	97.9	98.1	98.0	97.3	89.1
$p_T(a_1) > 35 \text{ GeV}/c$	73.7	78.7	82.8	86.0	72.3	75.5	78.2	80.7	50.1	8.0
$n_{\text{iso}}(\Delta R < 0.5) = 0$	73.3	73.0	72.9	72.9	73.0	73.5	72.9	72.8	74.4	12.5
cumulated	47.9	51.1	53.7	55.9	47.1	49.6	51.1	52.6	32.5	0.8

The selection provides clean tau decays with high matching purity. The matching of tau candidates to the generated decays is defined in Section 2.4.4. In all tested samples that contain true tau leptons, the selected candidates can be matched to the generated tau lepton at about 97 % probability.

Table 3.5.: Efficiencies of the single tau selection. The numbers are given in percent of the tested events that contain at least one kinematically reconstructed tau candidate. The custom tau selection,  $n_{\text{tau}}^{\text{custom}} \geq 1$ , is normalized to events that pass the quality tau selection,  $n_{\text{tau}}^{\text{quality}} \geq 1$ .

$m_{H^0} / \text{GeV}/c^2$	$H^0_{\text{GF}}$				$H^0_{\text{VBF}}$				Z3pr	QCD
	115	125	135	145	115	125	135	145	-	-
$n_{\text{tau}}^{\text{quality}} \geq 1$	66.9	67.6	68.4	68.6	60.4	61.4	62.2	62.3	62.5	1.6
$n_{\text{tau}}^{\text{custom}} \geq 1$	53.4	56.6	59.4	61.4	53.0	55.4	56.9	58.4	36.7	0.8
cumulated	35.7	38.3	40.6	42.1	32.0	34.0	35.4	36.4	23.0	0.012

The decay-based numbers translate into the event-based values of Table 3.5 by requiring at least one remaining tau candidate per event. The first row repeats the efficiencies from the quality tau discriminator of Table 2.2 and adds the missing Higgs masses. Small deviations arise from larger sample statistics. The second line summarizes the event efficiencies for the custom tau selection. The cumulated value still excludes the trigger and reconstruction inefficiencies. Signal-like tau decays pass the custom tau selection with a rate between 32 % and 42 %. Jets from QCD are rejected with an efficiency of  $1.2 \cdot 10^{-4}$ .

### 3.5.1.2. Tau Pair Selection

The pair selection can either start from tau candidates that pass the quality criteria defined in Section 2.4.3, the quality tau candidates, or that survive the tightened requirements of Section 3.5.1.1, the custom tau candidates. In both cases the remaining tau candidates are combined into pairs if their total charge is zero.

Figure 3.18 counts the number of tau pairs at the two different stages of the single tau selection. For all processes only a negligible amount of events contains more than one tau pair. In the rare case of an ambiguity, the pair that consists of the tau candidates with the smallest normalized  $\chi^2$  of the tau fit is selected. This criterion is found to have little impact on the further event selection.

The efficiency of the custom tau pair selection is summarized in Table 3.6 starting from quality tau candidates. The charge control, which tests for a neutral pair, is denoted explicitly. The samples containing real tau leptons pass this charge test with more than 98 % for the quality tau selection. With increasing purity of the custom tau selection the rate almost reaches 100 %. The charge of fake jets from QCD is randomly distributed. The efficiency of the charge test is close to 50 %. The cumulated value and the expected cross section consider all efficiencies of the passed selection chain. The selection efficiency of custom tau pairs is at a few percent for the Higgs samples and below 1 % for Z3pr. The background from QCD is suppressed by about  $10^{-11}$ . This exceeds the sample statistics, no generated events are left. The numbers marked with an asterisk are conservative estimates that hold under certain assumptions. They are discussed in the following section. The ratio between the smallest signal and the largest background is in the order of  $10^{-4}$ . The QCD cross section is in the same order of magnitude as the signal expectation.

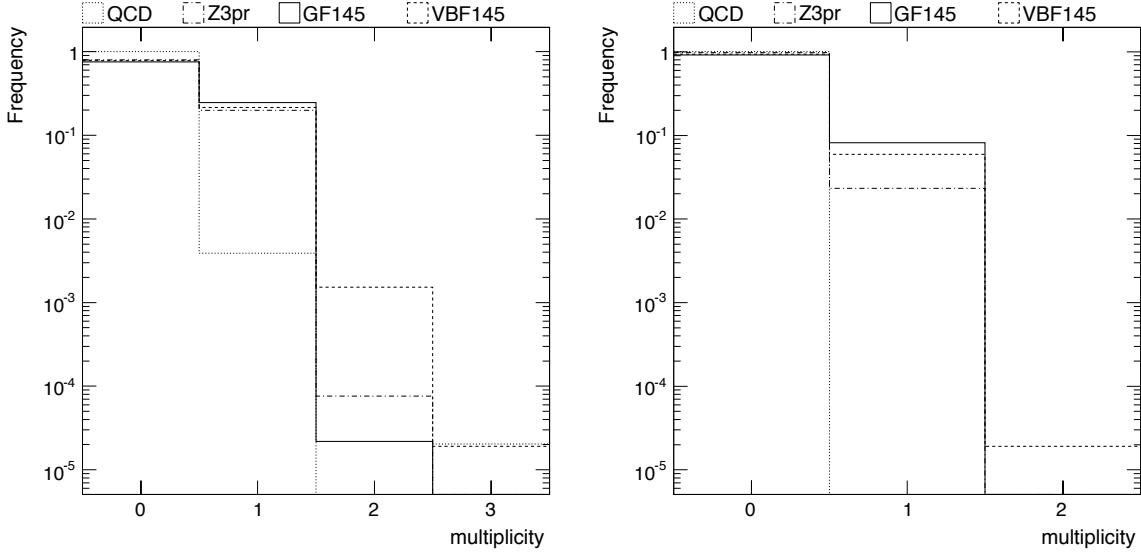


Figure 3.18.: Number of tau pairs per event that pass the quality selection (left) and the custom selection (right). The events are pre-filtered to contain at least one quality tau candidate.

Table 3.6.: Efficiency of the pair selection for quality tau candidates (first two rows) and custom tau candidates (rest). The numbers are given in percent of the tested events. The first row is normalized to events containing at least one tau candidate that passed the quality discriminator of Section 2.4.3. The cumulated value and the expected cross section  $\sigma_{\text{exp}}$  consider all efficiencies of the passed selection chain. The values for the QCD sample marked with an asterisk are valid under the assumptions discussed in Section 3.5.2.

$m_{H^0} / \text{GeV}/c^2$	$H^0_{\text{GF}}$				$H^0_{\text{VBF}}$				Z3pr	QCD
	115	125	135	145	115	125	135	145	-	-
$n_{\text{tau}}^{\text{quality}} \geq 2$	24.0	24.1	24.8	24.7	20.7	21.2	22.0	22.1	20.0	0.4
total charge = 0	99.6	99.6	99.6	99.8	98.1	98.4	98.4	98.2	99.6	57.7
$n_{\text{tau}}^{\text{custom}} \geq 2$	24.7	28.3	30.9	33.2	19.2	22.9	25.2	27.3	11.7	$*6 \cdot 10^{-3}$
total charge = 0	100	100	100	100	99.7	99.8	99.8	99.9	100	$*60$
cumulated	2.1	2.7	3.3	3.7	1.9	2.4	2.9	3.1	0.3	$*2 \cdot 10^{-9}$
$\sigma_{\text{exp}} / \text{fb}$	0.25	0.23	0.17	0.10	0.017	0.016	0.012	0.007	19.2	$*0.08$

Events that contain at least one custom tau pair pass this stage of the selection. Quality pairs are only used to estimate further selection efficiencies of the QCD sample.

### 3.5.2. Assumptions on Backgrounds

The size of the generated QCD sample is not large enough to investigate a selection that provides the required background rejection. To estimate the suppression, a few crucial assumptions are made. The reconstructed tau candidates in the QCD sample are jets that fake the signal. They do not originate from the decay of a heavy particle but are randomly combined to pairs. One expects little correlation between the two hemispheres. They can be regarded as almost independent.

The first assumption is the absence of charge correlation between both tau candidates. In the signal the tau pairs add up to a charge of zero. Fake tau leptons from QCD are each uniformly distributed in charge. There is no correlation between the tau candidates in the event. The charge test of quality pairs in Table 3.6 rejects almost half of the QCD events. This charge veto can be bypassed to increase the statistics of the remaining sample by a factor of about two.

Tau candidates emerge from independent kinematic fits and have to pass the quality requirements of Section 2.4.3. They only consider observables of the tau decay, not the whole event. The efficiency to pass the the quality requirements is given in the first rows of Table 3.5 and Table 3.6. The efficiency of the second quality tau is about half of the first one. If only the single tau rate  $\epsilon(n_{\text{tau}} \geq 1)$  is measured, one can estimate the cumulated di-tau rate as

$$\epsilon(n_{\text{tau}} \geq 2) = \epsilon^2(n_{\text{tau}} \geq 1) . \quad (3.3)$$

The custom tau selection also considers independent tau decays. Little correlation to the rest of the event is expected. It rejects QCD jets mostly by the isolation and momentum requirement. Figure 3.19 shows the momentum correlation between two tau decays that pass the quality criteria. The visible  $a_1$  momenta within the events are compared. The signal events from GF tend to be balanced. Fake jets from QCD exhibit no dependency between the two momenta.

There are no remaining QCD events containing two custom tau candidates. The custom tau selection of Section 3.5.1.1 is regarded as independent for both tau candidates. The rate of the second hemisphere is conservatively estimated with (3.3). To combine to a neutral pair, an additional suppression of 0.6 is considered, which is still a modest estimation.

One additional assumption only affects the selection of the VBF process. Two quark jets derive from the production process. Their selection will be described in Section 3.5.3. It relies on the rejection of jets that are to close to tau candidates. This induces a small dependency between the tau and the jet selection. The impact can be determined by comparing the efficiencies of the different jet selections in Table 3.9 and Table 3.10. The numbers of the signal from VBF are compatible. The low rates of fake quark jets in GF and Z3pr are varied slightly. As no strong dependencies are present, the efficiency of the jet selection that vetoes custom tau candidates in the QCD sample is estimated by the jet selection concerning quality tau candidates. The jet pair discrimination is assumed to be independent of the other requirements of the VBF selection. To estimate the QCD and  $Z^0$  efficiencies of this selection layer, the jet pair selection is bypassed.

The rejection factors of the dedicated event selections for GF are determined from quality tau

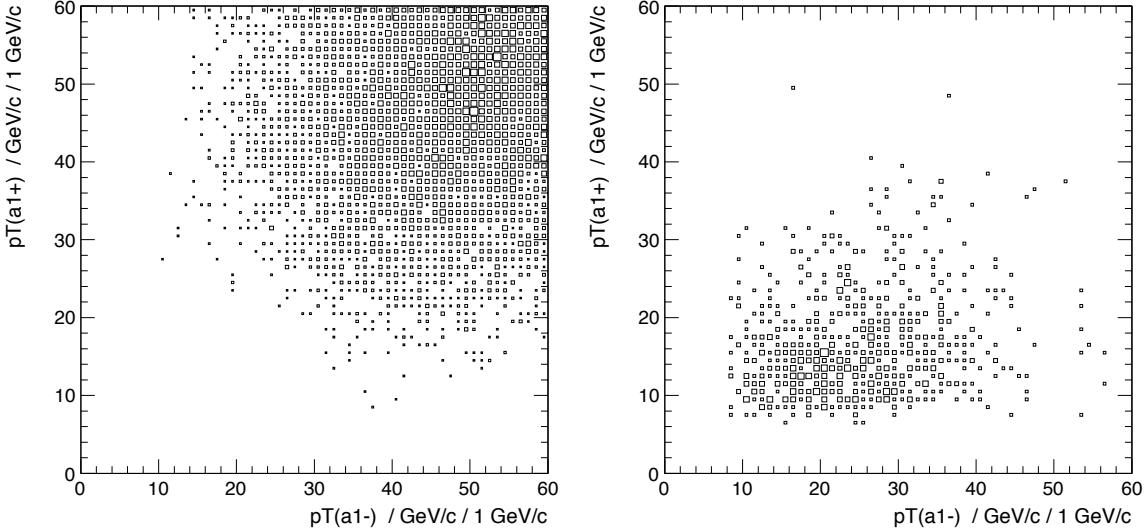


Figure 3.19.: Kinematic correlation between the  $a_1$  momenta from both tau hemispheres. The tau candidates passed the quality discrimination and originate from gluon fusion (left) or QCD (right). The left plot is zoomed to match the right one.

decays in case of the QCD sample. The selection applies a tightened version of the  $a_1$  momentum requirement, which is directly correlated to the custom tau selection. It is intended to reject tau pairs from  $Z^0$  decays and is bypassed in the determination of the QCD efficiency. The resulting rate can be used to estimate the efficiency starting from custom tau pairs.

### 3.5.3. Jet Selection

The VBF selection requires two jet candidates in addition to the tau pair. Considering the kinematics of the quarks involved in the Higgs production (Section 3.2.4), one expects two jets with significant transversal momenta and a large invariant mass. The pair should be separated in  $\eta$  and have little overlap with the tau pair in the central region. This behavior is independent of the Higgs mass. Thus, only one curve is shown in the following distributions.

The analysis starts from particle-flow jets [95]. An iterative-cone algorithm clusters all particles derived from the particle-flow algorithm within an annulus of  $\Delta R = 0.5$ . This approach improves the jet efficiency and purity of basic jet algorithms based on calorimetry information (referred to as calo jets).

A comparison of the transversal momentum resolution between both algorithms is depicted in Figure 3.20. Both plots are obtained from the same dataset ( $m_{H^0} = 145 \text{ GeV}/c^2$ ) and show the jets that are matched to the quarks from the VBF process within a cone of  $\Delta R = 0.1$ . The jet matching is defined in Appendix B.1. The distribution for calo jets peaks at about  $\Delta p_T = -10 \text{ GeV}/c$  and has a tail below  $\Delta p_T = -50 \text{ GeV}/c$ . As the energy deposition in the calorimeter is exclusively considered, the reconstructed momenta are shifted towards lower values. The right curve from particle-flow jets has a mean close to zero and a smaller width of about  $7.5 \text{ GeV}/c$ . A detailed comparison between both jet types can be found in [95].

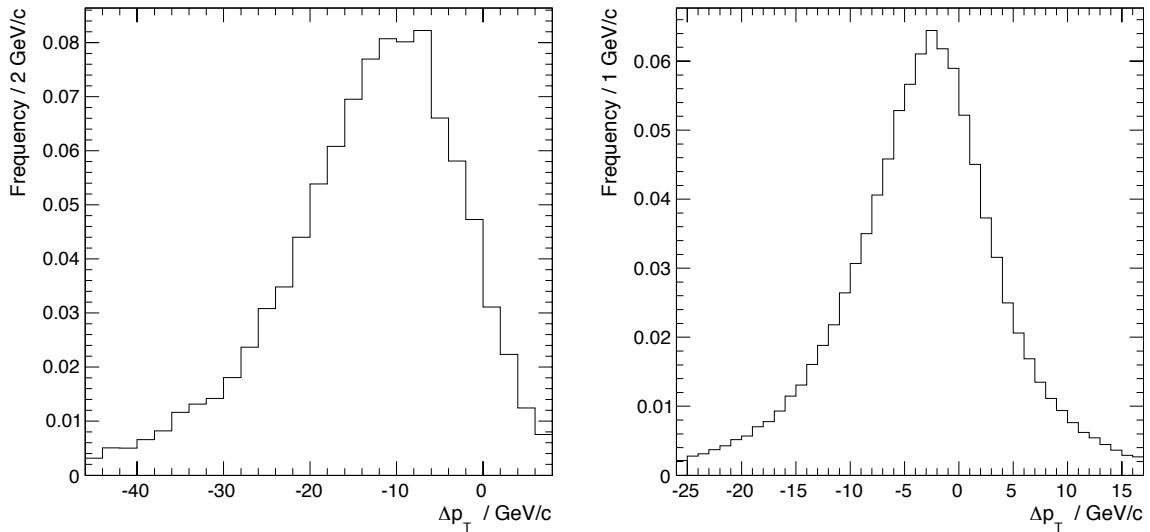


Figure 3.20.: Transversal momentum resolution  $\Delta p_T = p_T^{\text{rec}} - p_T^{\text{gen}}$  for jets in vector-boson fusion reconstructed with calorimetry information (left) and with the particle-flow algorithm (right). The jets passed a soft pre-filter of  $p_T(\text{jet}) > 5 \text{ GeV}/c$ .

### 3.5.3.1. Single Jet Selection

The jet selection is executed after a final pair of tau candidates is chosen. It is possible to reject jets that do not originate from the signal quarks. The main criterion is the overlap with the selected tau candidates. In addition to this tau veto, the jet candidates have to pass a rather soft  $p_T$  threshold.

For the signal one obtains the transversal momentum distribution of reconstructed jets depicted in Figure 3.21. The plot contains events that pass the HLT and the custom tau pair selection of Section 3.5.1. Jets with a minimal transversal momentum of  $p_T(\text{jet}) > 5 \text{ GeV}/c$  and no further discrimination enter the left histogram. As tau leptons contribute to the jet candidates, one expects four entries per event. The highest bin close to the threshold contains almost no matched jets. The solid curve has a significant contribution from signal jets beyond  $p_T \approx 20 \text{ GeV}/c$ . A second maximum at about  $p_T \approx 60 \text{ GeV}/c$  is caused by tau jets. Its position matches the peak of the momentum distribution of identified tau candidates (compare Figure B.2 of Appendix B.2). After applying a tau veto, this second maximum vanishes. This is shown in the right part of the figure. The dashed curve for the matched jets is marginally affected. To reject the fake jets with low momenta, a transversal momentum of at least  $p_T = 20 \text{ GeV}/c$  is required.

The tau veto requires a separation of the jet candidates from the selected tau pair. The minimal distance of the jets to both tau candidates is shown in Figure 3.22. The two plots depict jets with (right) and without (left) applying the momentum requirement. Both distributions show a large peak in the first bin, which contains tau jets instead of quarks. As the jet and the tau algorithm define an annulus of  $\Delta R = 0.5$ , matched jets start to contribute at this minimal distance. The broad part of the distribution contains the quark jets from the production process, which reach separation angles beyond  $\Delta R = 4$ . The left plot includes low energetic fake jets randomly spread over the distribution. They fail the momentum requirement and do not enter the right plot. Here, the distribution clearly divides into jet candidates from identified

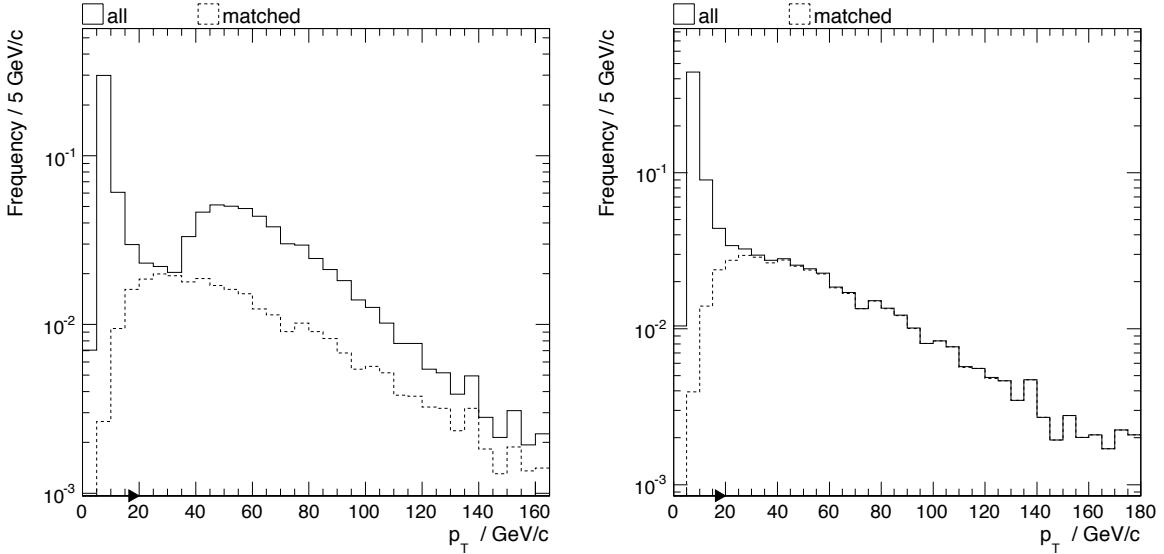


Figure 3.21.: Transversal momentum of reconstructed jets from vector-boson fusion above a threshold of  $p_T(\text{jet}) > 5 \text{ GeV}/c$ . The events contain a custom tau pair. The dashed curve corresponds to jets matched to the generated quarks. The solid line contains all jet candidates. Left: without applying further discrimination, right:  $n-1$  representation.

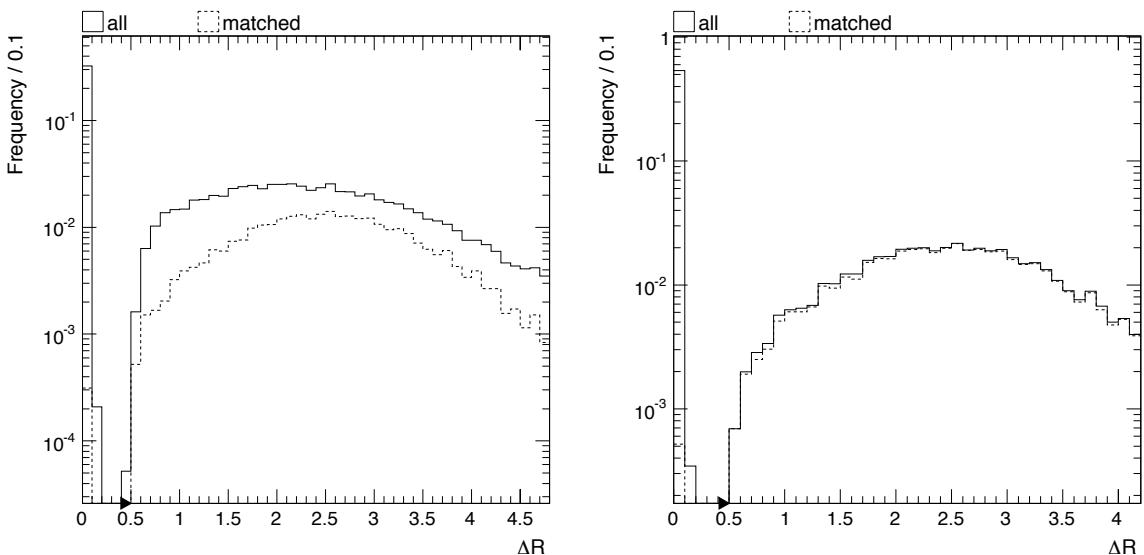


Figure 3.22.: Separation in  $\Delta R$  between tau and quark jets obtained from vector-boson fusion. The events contain a custom tau pair. The dashed curve corresponds to jets matched to the generated quarks. The solid line contains all jet candidates. Left: without applying further discrimination, right:  $n-1$  representation.

tau leptons and from signal quarks. About 50 % of all reconstructed jets are contained in the first bin, the other half in the rest of the distribution. Jets with a separation above  $\Delta R = 0.5$  are selected.

Single jets that pass the requirements have a large purity. About 97% of the selected jet candidates can be matched to the generated quarks.

The combination of tau veto and momentum criterion identifies the two quark jets in signal events. The impact of these requirements on background events is analyzed in the following. The signal samples and the Z3pr contain enough custom tau candidates to directly obtain the efficiency of the jet selection, referred to as custom jet selection. For the QCD sample, one can only determine the efficiency of a jet selection starting from quality tau candidates. This quality jet selection can be used to estimate the custom jet selection using the assumptions of Section 3.5.2. Thus, the following distributions consider events that contain tau pairs passing the quality criteria.

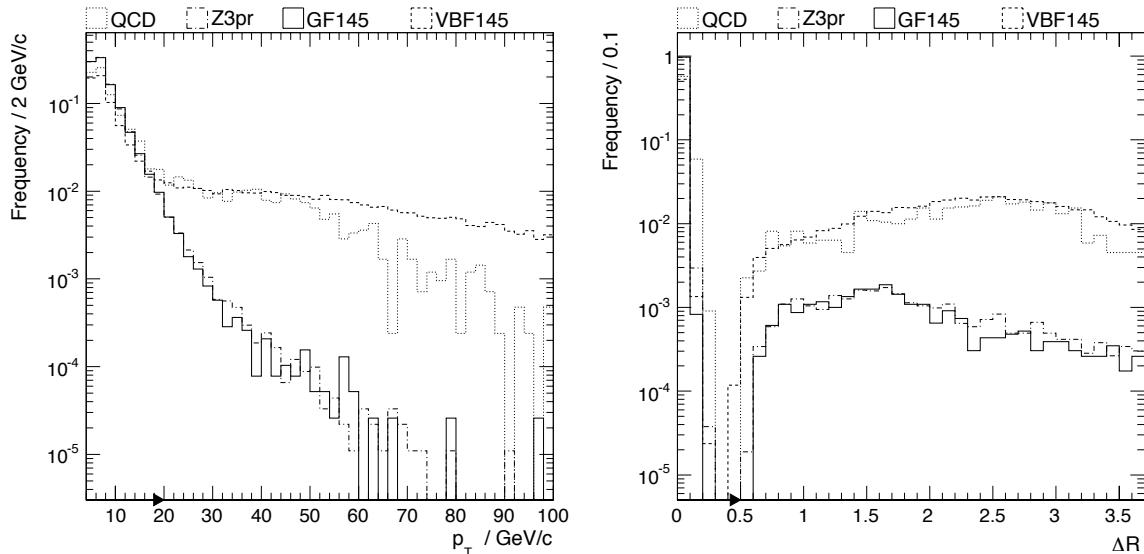


Figure 3.23.: Single jet selection in vector-boson fusion: Transversal momentum (left) of jet candidates and tau veto in  $\Delta R$  (right). The events contain at least one pair of tau candidates that pass the quality discriminator of Section 3.5.1. Both plots give a  $n-1$  representation, each including the other requirement.

Figure 3.23 compares the signal from VBF to backgrounds from QCD and Z3pr. The fourth curve shows the gluon-fusion process. It can be regarded as background in the jet selection as it contains no quark jets. The left plot shows the transversal momentum distribution including the tau veto ( $n-1$  representation). The signal curve exhibits no relevant deviations from the one obtained with the custom tau requirements (Figure 3.21). After tau jets are vetoed, the distributions for GF and Z3pr include no further jets with significant momentum. Both curves decrease exponentially. A fraction of about 1/1000 reaches momenta beyond  $p_T = 20 \text{ GeV}/c$ . Jets from the QCD sample result in a signal-like curve up to  $p_T \approx 50 \text{ GeV}/c$ . For larger momenta the rate drops and a distinction from the signal is possible. The applied momentum requirement of  $p_T > 20 \text{ GeV}/c$  could be tightened if further QCD suppression is needed.

The right plot in Figure 3.23 depicts the distance from tau candidates for jets that pass the

momentum criterion. Events from GF and Z3pr contain only the two tau jets in average and contribute almost exclusively to the first bin. Beyond the required separation of  $\Delta R = 0.5$ , the curves for QCD and the signal are similar in shape and rate.

Table 3.7.: Efficiency of the single jet selection. The numbers are given in percent of the tested jets (not events). Background events from QCD contain at least one pair of tau candidates that passed the quality discriminator of Section 2.4.3. For the other samples the tighter requirements of Section 3.5.1 are applied.

$m_{H^0} / \text{GeV}/c^2$	$H^0_{\text{GF}}$				$H^0_{\text{VBF}}$				Z3pr	QCD
	115	125	135	145	115	125	135	145	-	-
$p_T(\text{jet}) \geq 20 \text{ GeV}/c$	44.9	44.9	45.6	45.6	61.2	61.0	60.5	60.4	44.4	38.6
$\Delta R(\tau, \text{jet}) \geq 0.5$	55.9	55.9	55.3	55.2	67.5	67.4	67.3	67.5	56.6	73.1
cumulated	0.8	0.7	0.8	0.8	28.7	28.4	27.8	27.9	1.0	14.5

The efficiency of the single jet selection is summarized in Table 3.7. It considers events containing quality tau pairs for the QCD sample and otherwise custom tau pairs. The signal from VBF and the fake jets from the other samples pass the stand-alone requirements with compatible rates. In case of the backgrounds, the two criteria select almost disjunct sets of jet candidates. Jets passing the first line are mostly tau candidates, which are vetoed in the second line. Thus, the cumulated efficiencies drop to 1 % for samples not containing quark jets. QCD events have a larger jet multiplicity, which results in a weaker separation of both criteria. The cumulated efficiency is significantly larger. Jets from the VBF process pass the single jet selection with about 28 %. The rate is independent of the generated Higgs mass.

The event-based efficiencies are discussed in the pair selection of the following section. The values of Table 3.9 start from quality tau candidates and Table 3.10 from custom tau candidates.

### 3.5.3.2. Jet Pair Selection

The single jet selection suppresses jets in the processes Z3pr and GF to a large extent. To achieve a further rejection of QCD, the remaining jets are combined to pairs that have to fulfill additional criteria. The pair selection is tested on events that either contain quality or custom tau pairs. To include a sufficient statistics in the QCD curves, the following distributions consider the quality tau selection. The GF curves are not drawn in the distributions in favor of a second VBF curve. The lightest and the heaviest generated Higgs masses are shown. This demonstrates the mass independence of the jet selection. More than one jet pair per event can enter the distributions.

One efficient discriminator against QCD jets is the invariant mass of the di-jet system  $m_{\text{di-jet}}$ . As discussed in Section 3.2.4, the quark jets from VBF are expected to provide large masses while small values are expected from fake jets. This behavior is not much affected by the reconstruction and can be found in Figure 3.24. The left plot gives the stand-alone, the right the  $n-1$  representation w.r.t. the further selection of jet pairs. The shape of the signal and the position of the maximum at  $m_{\text{di-jet}} = 200 \text{ GeV}/c^2$  are close to the expectation in the right part of Figure 3.10. Jet pairs from the heavier Higgs sample are slightly shifted towards larger values. Jets that do not originate from a signal-like quark exhibit a peak at low masses with a rapidly decreasing tail. The curves for Z3pr and QCD are similar. The jet selection requires a minimal di-jet mass of  $m_{\text{di-jet}} > 150 \text{ GeV}$ . This is a rather soft criterion but provides sufficient rejection of background processes.

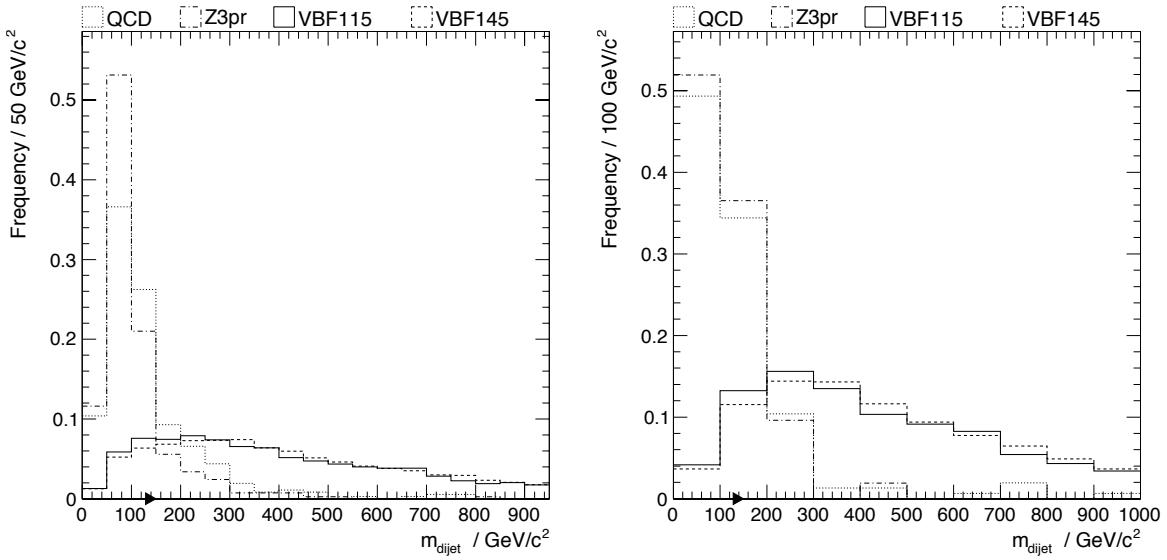


Figure 3.24.: Selection criterion: The invariant masses of jet pairs have to exceed  $m_{\text{di-jet}} > 150 \text{ GeV}$ . Jets that pass the quality selection of Section 3.5.3.1 are considered. Left: without applying further discrimination, right:  $n-1$  representation.

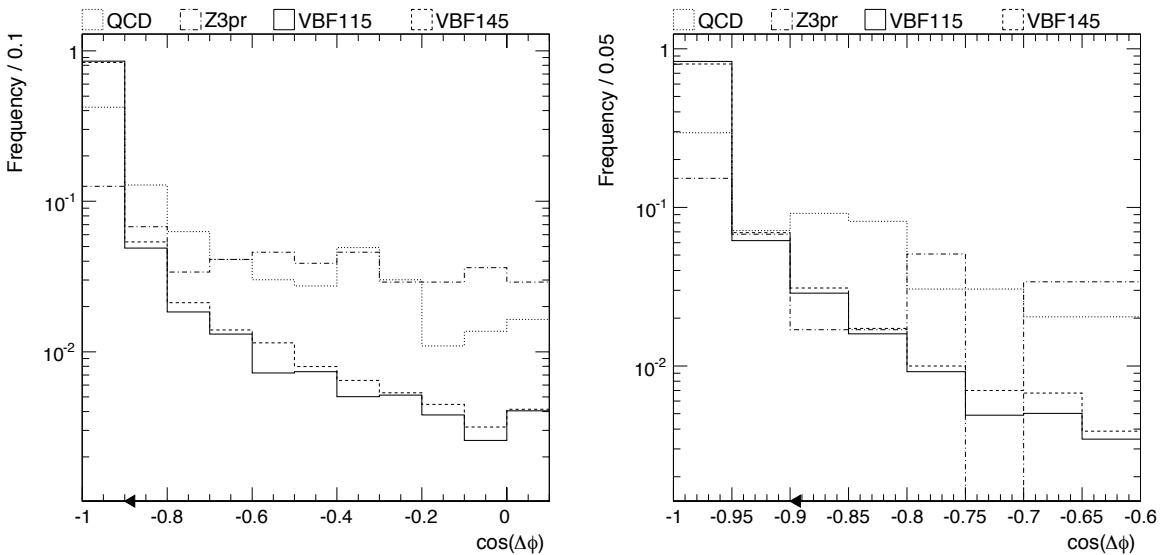


Figure 3.25.: Selection criterion: Acoplanarity between the di-jet and the di-tau systems  $\cos(\Delta\phi) = \cos(|\phi_{\text{di-tau}} - \phi_{\text{di-jet}}|)$ . Jets that pass the quality selection of Section 3.5.3.1 are considered. Left: without applying further discrimination, right:  $n-1$  representation. The selection requires  $\cos(\Delta\phi) < -0.9$ .

Single jets have no preferred orientation in  $\phi$  and the angle enclosed by the jet pair in the transversal plane is nearly uniformly distributed. But the sum of their momenta has to be aligned with the tau pair. One can define the acoplanarity between the di-jet and the di-tau systems as

$$\cos(\Delta\phi) \equiv \cos(|\phi_{\text{di-tau}} - \phi_{\text{di-jet}}|) . \quad (3.4)$$

The distribution of this observable is shown in Figure 3.25. Jets from quarks of the VBF process peak in the very first bin of the left histogram, which corresponds to an enclosed angle between the two systems of  $\Delta\phi > 2.7$  rad. In both representations the jets from QCD and Z3pr backgrounds show flat curves that only slightly prefer values near  $\cos(\Delta\phi) = -1$ . The  $n-1$  representation on the right contains only very little events of the generated backgrounds. This causes large fluctuations, especially for the Z3pr sample. Jet pairs are accepted by the selection if the acoplanarity holds  $\cos(\Delta\phi) < -0.9$ .

Table 3.8.: Efficiencies of the jet pair selection. The numbers are given in percent of the tested jet pairs (not events). The selection is tested on events containing at least one custom tau pair. The efficiencies for the QCD samples are estimated by considering the quality tau selection.

$m_{H^0} / \text{GeV}/c^2$	$H^0_{\text{GF}}$				$H^0_{\text{VBF}}$				$Z3\text{pr}$	$\text{QCD}$
	115	125	135	145	115	125	135	145	-	-
$m_{\text{di-jet}} > 150 \text{ GeV}/c^2$	6.3	11.1	20.0	19.4	87.6	89.5	89.2	89.4	19.1	26.8
$\cos(\Delta\phi) < -0.9$	12.5	7.4	20.0	19.4	87.6	86.6	84.8	84.8	14.3	54.9
cumulated	< 0.1	< 0.1	< 0.1	3.2	80.0	80.5	78.0	78.5	4.8	14.5

The efficiencies of the jet pair selection are summarized in Table 3.8. Except for the QCD sample, the custom tau selection is considered. The estimated values for fake jet pairs from QCD are only valid under the assumptions discussed in Section 3.5.2. The signal jets pass the selection at about 80 %. There is only a small variation for different Higgs masses. Jet pairs from events with lighter Higgs bosons exceed the minimal di-jet mass with a smaller efficiency but tend to be more balanced with the tau system. These two effects almost cancel each other. The Z3pr and the GF process are suppressed to a large extent. There is only a negligible amount of generated events left. The QCD events contain more jet pairs, which may pass the jet pair selection. The obtained efficiency is conservatively estimated to about 15 %.

The jet pair selection can provide an arbitrary number of jet pairs per event. If more than two jets pass the single jet selection, all combinations are tested. There is no charge requirement to reduce the number of permutations as in the tau pair combination. The probability that more than one jet pair fulfills the acoplanarity requirement is rather small. The remaining number of jets per event is summarized in Figure 3.26 for all generated VBF samples. The left part considers quality tau pairs, the right part custom tau pairs. In both cases the efficiency to find more than one jet pair passing the selection is at percent level. Starting from custom tau pairs, the amount of events with ambiguous jet pairs decreases with heavier Higgs masses.

This ambiguity has to be resolved. As discussed in Section 3.2.4, the quarks from the vector-boson fusion tend to be separated in the detector. In the rare case of more than one remaining pair, the one that has the largest separation in pseudorapidity between the two jets is selected.

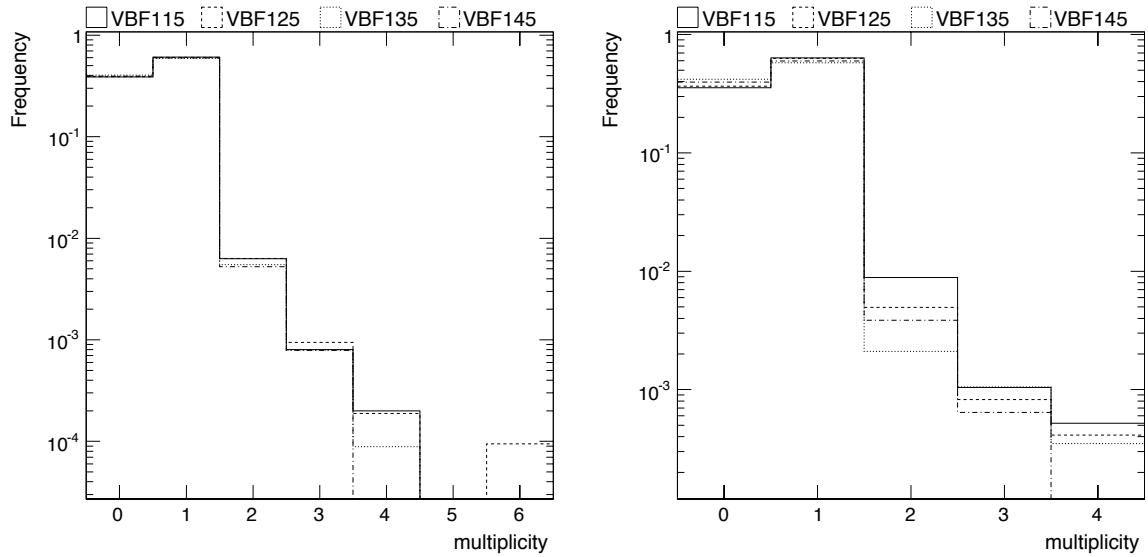


Figure 3.26.: Number of jet pairs per event passing the jet pair selection. The events contain at least one quality tau pair (left) or one custom tau pair (right).

Table 3.9.: Efficiency of the jet pair selection that vetoes quality tau decays. The numbers are given in percent of the tested events. The first row is normalized to events containing at least one tau pair passing the quality discriminator of Section 2.4.3.

$m_{H^0} / \text{GeV}/c^2$	$H^0_{\text{GF}}$				$H^0_{\text{VBF}}$				Z3pr	QCD
	115	125	135	145	115	125	135	145	-	-
$n_{\text{jet}}^{\text{quality}} \geq 2$	1.3	1.2	1.3	1.1	74.0	73.7	72.8	73.2	1.3	27.2
$n_{\text{jet pair}}^{\text{quality}} \geq 1$	0.9	< 0.1	2.3	7.3	82.9	82.9	82.0	82.6	4.3	22.0
cumulated	0.01	< 10⁻²	0.03	0.08	61.3	61.1	59.7	60.5	0.06	6.0

Table 3.10.: Efficiency of the jet pair selection that vetoes custom tau decays. The numbers are given in percent of the tested events. The first row is normalized to events containing at least one tau pair passing the custom requirements of Section 3.5.1. There are no remaining events from the QCD sample. The efficiencies are estimated in Section 3.5.2.

$m_{H^0} / \text{GeV}/c^2$	$H^0_{\text{GF}}$				$H^0_{\text{VBF}}$				Z3pr	QCD
	115	125	135	145	115	125	135	145	-	-
$n_{\text{jet}}^{\text{custom}} \geq 2$	0.8	0.9	1.0	0.8	75.1	74.5	71.0	72.1	1.1	—
$n_{\text{jet pair}}^{\text{custom}} \geq 1$	< 0.1	< 0.1	< 0.1	3.4	85.9	85.1	81.8	83.8	5.7	—
cumulated	< 10⁻²	< 10⁻²	< 10⁻²	0.03	64.5	63.4	58.1	60.4	0.07	—

Events with a remaining jet pair are selected for the VBF analysis. The event-based efficiencies are summarized in Table 3.9 and Table 3.10 considering quality or custom tau pairs. The two selections show compatible numbers, which allows to estimate the rate of QCD jets as discussed in Section 3.5.2. In each Table the first row counts the events with more than two jets, the second requires a remaining pair. The signal samples pass the selection with efficiencies of about 60 %. The efficiency of QCD events is estimated with 6 % while backgrounds from Z3pr are suppressed by 3 orders of magnitude. This additional background rejection of the vector-boson fusion analysis is the major benefit compared to the gluon fusion.

In case of the VBF process, the jet pair selection provides a clean sample of signal-like jets. Both jets of a pair passing the selection can be matched to the generated quarks with an efficiency of 94.7 %. This value corresponds to the squared single-jet purity including a small improvement by the pair selection.

## 3.6. Event Selection

All objects that are needed for the reconstruction of events from GF and VBF are selected by the criteria described in the last section. They are now combined and further discrimination of background processes is applied.

The dedicated selections for the two types of Higgs production are analyzed separately to exploit the different topologies. There is only a negligible amount of signal events that pass both selections. Thus, the distributions consider only one production type. In each selection the samples with the lightest and the heaviest generated Higgs mass are plotted to reveal mass effects.

As discussed in Section 3.5.2, the following QCD distributions are estimated without applying the custom tau selection. The generated Z3pr sample is large enough to consider the custom tau selection. To analyze both background processes as accurate as possible, the following distributions are divided. The left part of each figure is based on the quality selection and includes a curve for the QCD events, the right part applies the custom tau selection. This allows for a direct comparison of the shapes in both histograms to estimate the impact of the tightened tau selection. All distributions are normalized to an integral of one.

As far as the sample statistics are sufficient, the  $n-1$  representations of the applied requirements have been studied to reduce correlations between the observables. In case of the QCD process, the correlation can only be tested within the factorized selection layers. Little correlation among these layers is assumed.

### 3.6.1. Selection of Vector Boson Fusion

The VBF selection benefits from the additional jet signature and the significant transversal momentum of the Higgs boson. Both features provide strong rejections against backgrounds from QCD and Z3pr. The jet selection is assumed to be independent of further criteria that focus on the tau candidates. To enlarge the statistics of the background samples, the events in the following distributions are not required to contain a selected jet pair.

The kinematic tau reconstruction modifies the primary vertex of the event. As described in Chapter 2, the tracks assigned to the tau decay are eliminated from the primary vertex. Due to the kinematic constraints, the production point of the tau lepton is changed. These modifications are applied on both tau decays independently. Thus, the event contains two

tau production points, which possibly differ from each other. The significance of the distance between both vertices  $\sigma(PV^-, PV^+)$  is calculated according to Appendix A.2. The vertices are denoted as  $PV^-$  and  $PV^+$  depending on the charge of the tau candidate.

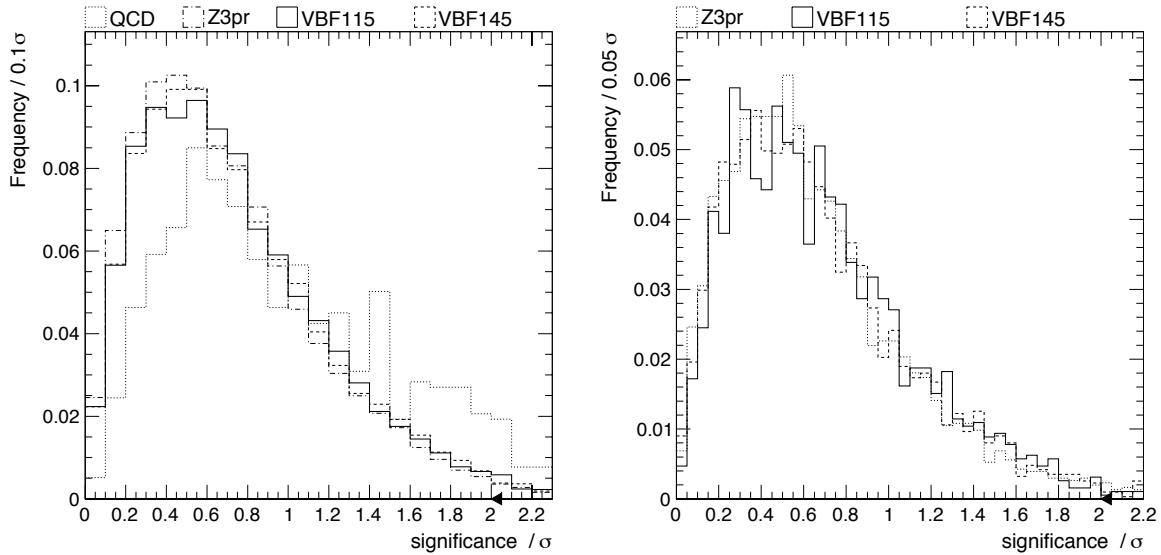


Figure 3.27.: Event selection for vector-boson fusion: Vertex significance between the rotated primary vertices from both hemispheres considering the quality (left) and custom (right) tau selection. The significance has to fulfill  $\sigma(PV^-, PV^+) < 2$ .

The corresponding distribution is shown in Figure 3.27. Events that contain true tau decays result in insignificant offsets between the vertices. This confirms the assumption that both tau candidates originate from the same mother. The significance has a maximum at about  $\sigma(PV^-, PV^+) = 0.4$ . Jet pairs from QCD are slightly shifted to higher values. To obtain unambiguous events with both tau candidates emerging from a common vertex, the significance has to fulfill  $\sigma(PV^-, PV^+) < 2$ . This criterion provides no large suppression of the QCD background but ensures the consistency of the reconstruction.

In case of the VBF process, the tau leptons from a Higgs decay are expected to enclose rather small azimuth angles. As discussed in Section 3.2.2, this results in a characteristic separation in the  $(\eta, \phi)$ -plane. The distributions in Figure 3.28 express this separation in terms of  $\Delta R(\tau, \tau)$ . Fake tau candidates from QCD jets result in a broad curve with a maximum close to  $\Delta R = 3$ . The little boosted  $Z^0$  boson decays into tau pairs that are almost perfectly aligned in the transversal plane. The distributions with and without the custom tau requirements exhibit a sharp peak near  $\Delta R = 3$ . The signal curves are slightly affected by the tightened tau selection. Especially the lightest generated Higgs mass is shifted towards lower values. In both cases there is a negligible amount of signal-like events with a separation beyond  $\Delta R = 3.2$ . This defines the selected region. Events are accepted if the separation holds  $\Delta R(\tau, \tau) < 3.2$ .

Higgs bosons from vector-boson fusion are emitted into the central region while the  $Z^0$  bosons are produced with large pseudorapidities. This behavior is retained in the reconstructed di-tau system. Figure 3.29 depicts the pseudorapidity  $\eta$  of the tau pair. The signal has a maximum at zero and small tails beyond  $\eta = \pm 4$ . Tau pairs from  $Z^0$  bosons show a minimum in the central region and peak at larger values. The custom tau selection has no

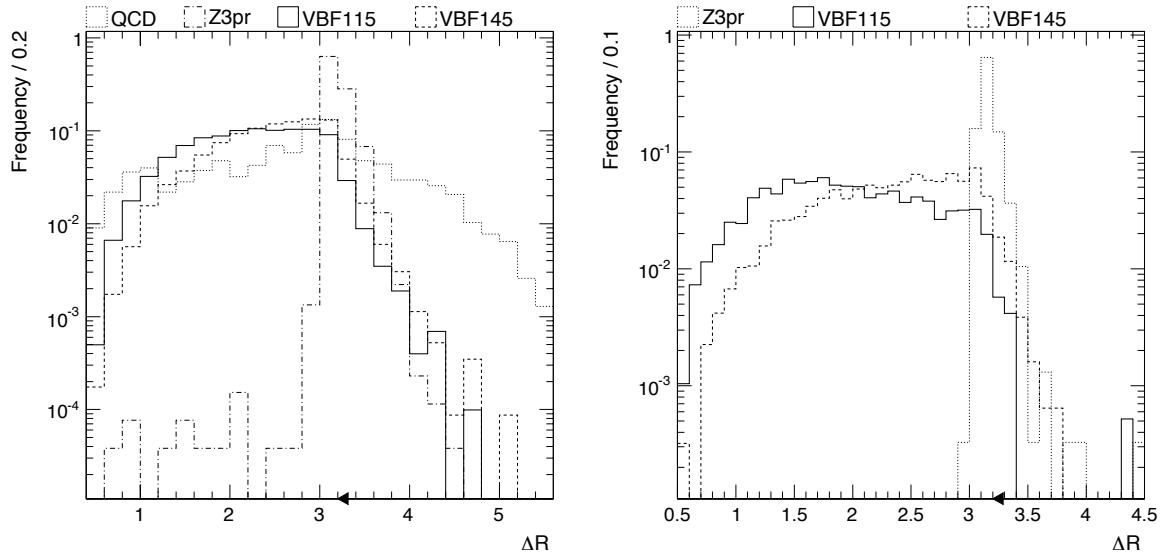


Figure 3.28.: Event selection for vector-boson fusion: Separation of the tau pair in the  $(\eta, \phi)$ -plane considering the quality (left) and custom (right) tau selection. The separation has to fulfill  $\Delta R(\tau, \tau) < 3.2$ .

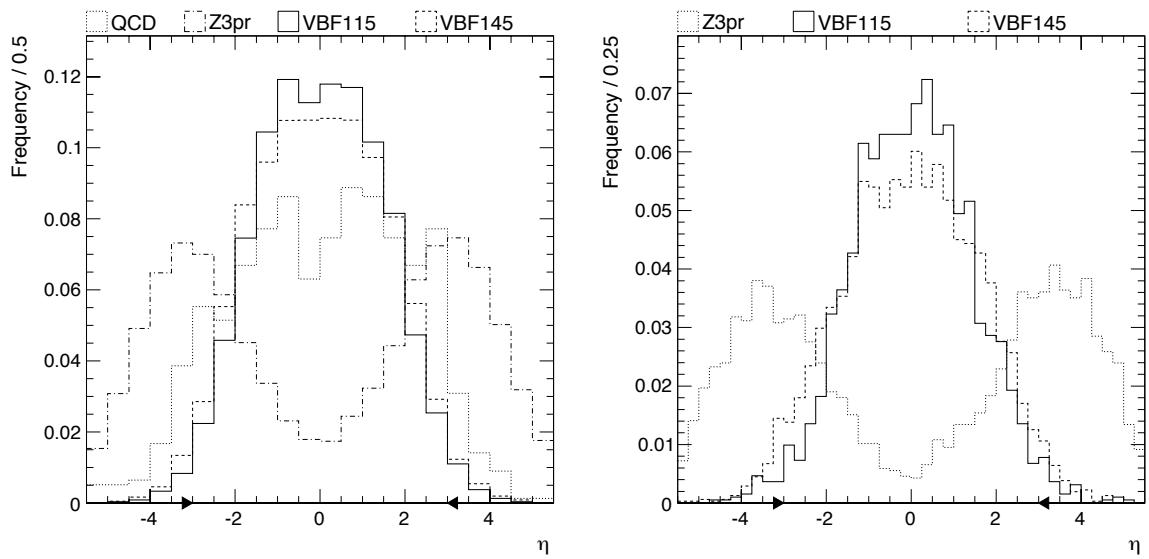


Figure 3.29.: Event selection for vector-boson fusion: Pseudorapidity  $\eta$  of the di-tau system considering the quality (left) and custom (right) tau selection. The region defined by  $-3 < \eta < 3$  is selected.

impact on this behavior. Events from QCD provide a broad curve around zero. The selection rejects events that are not contained in the region of  $-3 < \eta < 3$ .

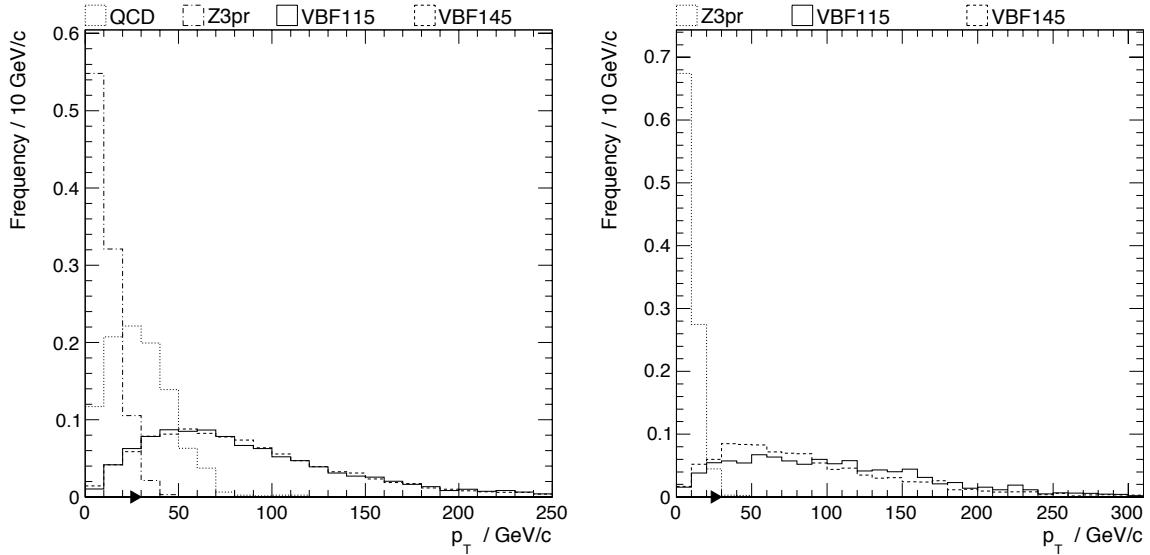


Figure 3.30.: Event selection for vector-boson fusion: Transversal momentum of the di-tau system considering the quality (left) and custom (right) tau selection. The selection requires  $p_T^{\tau\tau} \geq 30 \text{ GeV}/c$ .

The significant transversal boost of Higgs bosons from VBF provides further discrimination against QCD jets and tau pairs from  $Z^0$  decays. The transversal momentum of the di-tau system  $p_T^{\tau\tau}$  is shown in Figure 3.30. The signal curves have a maximum at about  $p_T^{\tau\tau} \approx 50 \text{ GeV}/c$  and large tails to higher values. The shape and the peak position are almost independent of the generated Higgs mass. Tau pairs from Z3pr have almost no transversal momentum. The curves from the two different tau selections exhibit no significant differences. Selected pairs from QCD are not as balanced as the di-tau system from  $Z^0$  decays and peak at about  $p_T^{\tau\tau} \approx 20 \text{ GeV}/c$ . The distribution rapidly decreases towards higher momenta. A major part of the backgrounds is rejected by requiring a minimal transversal momentum of  $p_T^{\tau\tau} \geq 30 \text{ GeV}/c$ .

The major observable of the VBF selection is the invariant mass of the tau pair  $m_{\tau\tau}$ , which is depicted in Figure 3.31. The sharp peaks of the resonances are affected by the limited resolution of the reconstruction. The maxima become broader and the distributions overlap. The mass range that is dominated by the signals from true di-tau events is defined as the signal region. Invariant masses between  $80 \text{ GeV}/c^2 < m_{\tau\tau} < 200 \text{ GeV}/c^2$  are accepted by the selection. This range contains most of the signal events from the lightest to the heaviest generated Higgs mass. It also includes the maximum from the  $Z^0$  decay, which will be suppressed by other requirements. The curve from the QCD background has a falling slope in the signal region. A large fraction of this background is rejected. The details of the distributions will be discussed in Section 3.7.

Table 3.11 gives an overview of the VBF selection. Except for the QCD sample, the efficiencies consider the custom tau discrimination. The rates of the QCD background are estimated under the assumptions of Section 3.5.2 and derive from the quality tau discrimination. The jet-pair selection is bypassed for the backgrounds to enhance the statistics.

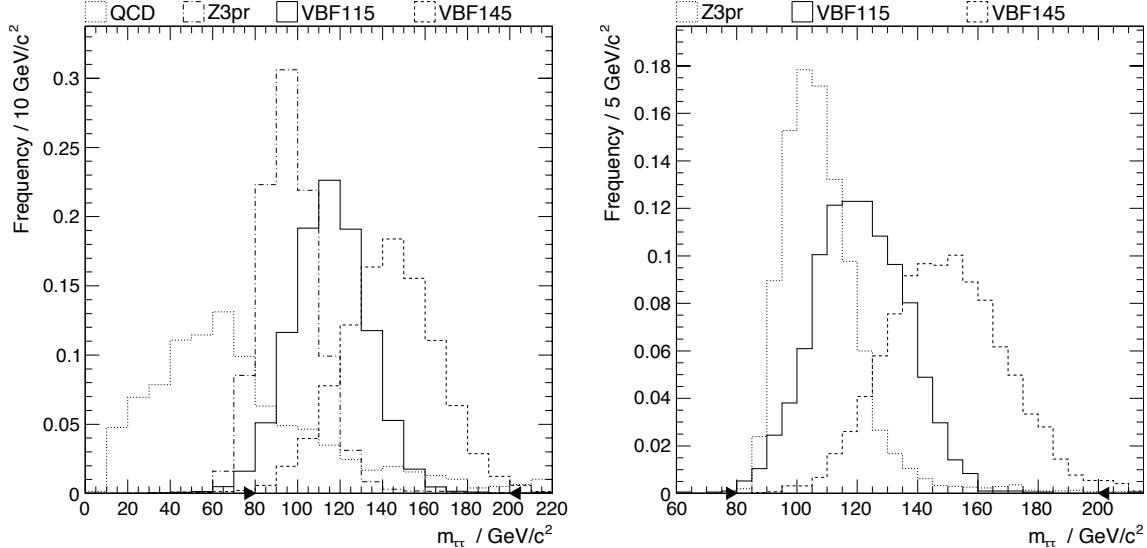


Figure 3.31.: Event selection for vector-boson fusion: Invariant mass of the di-tau system considering the quality (left) and custom (right) tau selection. Invariant masses between  $80 \text{ GeV}/c^2 < m_{\tau\tau} < 200 \text{ GeV}/c^2$  are selected.

Table 3.11.: Efficiency of the VBF selection. The numbers are given in percent of the tested events. Background events from QCD contain at least one pair of tau candidates that pass the quality discriminator. For the other samples the tighter requirements of Section 3.5.1 are applied. The jet pair selection is only considered for the signals from VBF.

$m_{H^0} / \text{GeV}/c^2$	$H^0_{\text{VBF}}$				Z3pr	QCD
	115	125	135	145	-	-
$\sigma(\text{PV}^-, \text{PV}^+) < 2$	98.8	98.5	98.1	98.5	98.7	88.6
$\Delta R(\tau, \tau) < 3.2$	99.8	99.4	99.2	98.2	80.2	69.4
$-3 < \eta < 3$	99.2	98.2	98.1	97.6	41.7	86.9
$p_T^{\tau\tau} \geq 30 \text{ GeV}/c$	96.3	95.0	94.3	95.3	0.6	45.4
$80 \text{ GeV}/c^2 < m_{\tau\tau} < 200 \text{ GeV}/c^2$	99.8	99.9	99.9	98.6	36.6	30.1
cumulated	94.8	92.6	91.5	90.4	1.1	5.7

The signal from VBF passes the criteria with an efficiency of more than 90 %. The rate is slightly mass dependent. A consistent primary vertex is found in almost 99 % of the events that contain true tau pairs. The  $Z^0$  decays are mostly rejected by an insufficient transversal momentum of the di-tau system and too large pseudorapidities. The other requirements are intended to diminish the backgrounds from QCD, which is suppressed to about 6 %. The cumulated efficiency for the Z3pr process is one percent. The efficiency of the entire selection chain, including the tau-pair and jet-pair selections, will be discussed in Section 3.6.3.

### 3.6.2. Selection of Gluon Fusion

The gluon-fusion process generates Higgs bosons in the central region of the detector that result in a balanced di-tau system. There are no additional particles in the event. Only the tau pair can be analyzed to discriminate against background processes.

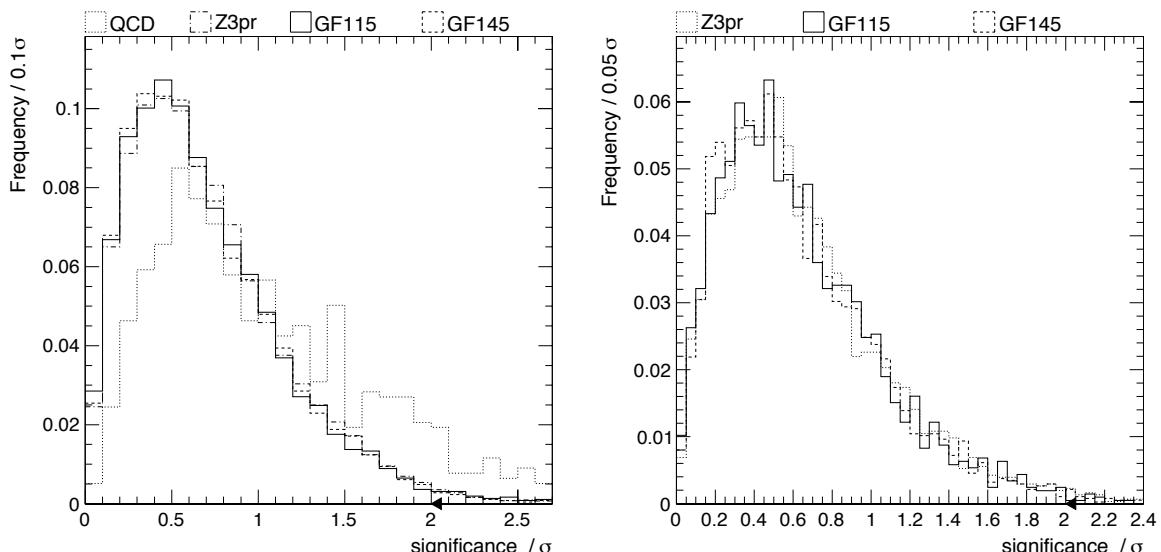


Figure 3.32.: Event selection for the gluon-fusion process: Vertex significance between the rotated primary vertices from both hemispheres considering the quality (left) and custom (right) tau selection. The significance has to fulfill  $\sigma(\text{PV}^-, \text{PV}^+) < 2$ .

The first selection criterion ensures the consistency of the event. The same observable as in the VBF selection is used. The separation of the primary vertices deriving from the kinematic reconstruction of both tau decays must not exceed a significance of  $\sigma(\text{PV}^-, \text{PV}^+) < 2$ . The corresponding distributions are depicted in Figure 3.32. The signal events for the GF contain true tau decays and show the same behavior as observed in Figure 3.27 for the VBF selection. The background events pass the same selection as for the VBF selection. The curves are identical.

The two selections share another common observable. The separation of the tau pair in the  $(\eta, \phi)$ -plane provides a discrimination from backgrounds also for the gluon-fusion process. In contrast to the vector-boson fusion, the tau pairs are barely boosted and almost perfectly back to back in the transversal plane. Combined with the large pseudorapidities of the Higgs boson (compare Section 3.2), this results in the  $\Delta R$  distribution of Figure 3.33. The signal processes show similar curves as the  $Z^0$  decay. There is a sharp peak near  $\Delta R(\tau, \tau) = 3$  with a rapidly falling slope towards larger values. To reject jets from QCD, a region defined

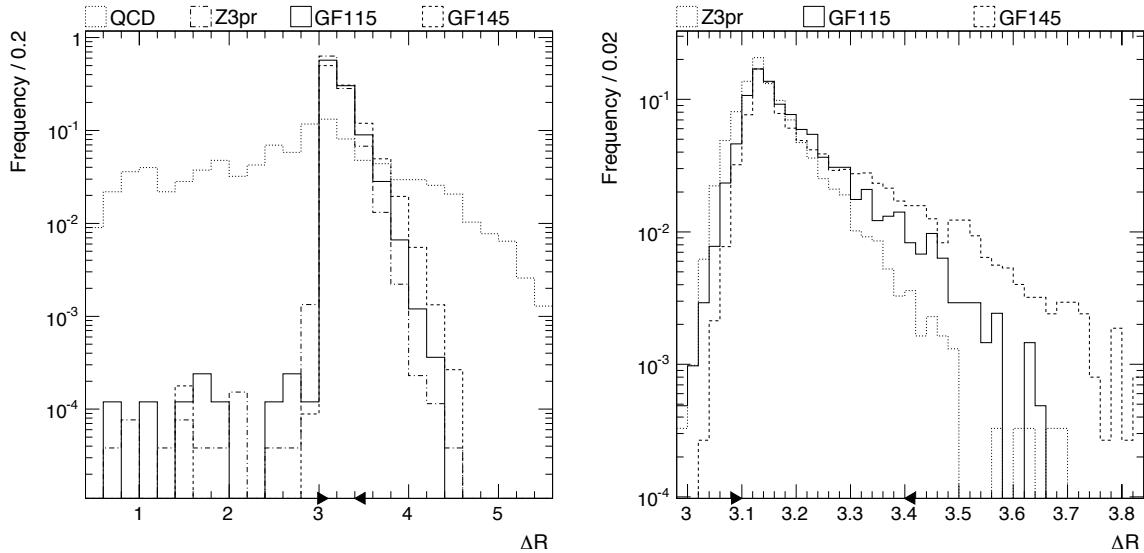


Figure 3.33.: Event selection for the gluon-fusion process: Separation of the tau pair in the  $(\eta, \phi)$ -plane considering the quality (left) and custom (right) tau selection. The separation has to fulfill  $3.1 < \Delta R(\tau, \tau) < 3.4$ .

by  $3.1 < \Delta R(\tau, \tau) < 3.4$  is selected. The zoomed view of the right histogram reveals little differences between the signal and the Z3pr curves. The shape broadens with increasing mass of the di-tau system. The  $Z^0$  peak has a steeper rise of the left edge. This provides a small suppression of  $Z^0$  decays.

Additional information can be extracted from the angle between the two tau leptons. Figure 3.34 depicts the cosine of the angle  $\alpha$  enclosed by the tau pair. The full codomain is populated. Tau pairs from GF exhibit a maximum near  $\cos(\alpha(\tau, \tau)) = -1$ , which corresponds to a balanced di-tau system in the detector frame. Tau pairs from background processes tend to enclose larger angles. Both curves from QCD and Z3pr have a slight ascent to positive values. These features are retained by the custom tau selection. A modest suppression of background events is achieved by requiring  $\cos(\alpha(\tau, \tau)) < 0.4$ .

A powerful discrimination is provided by the transversal momenta of the tau candidates. To exploit the improved resolution, the momenta of the  $a_1$  are considered. Sorting the pair with respect to the  $p_T$  of the  $a_1$  increases the discrimination. Figure 3.35 shows the distribution for the  $a_1$  with the larger momentum. With increasing Higgs mass the signal is more and more separated from the background. The peak for the lightest generated Higgs mass is close to the  $Z^0$  peak. Requiring  $p_T^{\max}(a_1) \geq 48 \text{ GeV}/c$  rejects the bulk of the  $Z^0$  peak but accepts a significant fraction of signal events.

The QCD distribution has a broad maximum at low momenta and falling tails into the region dominated by the signal. The custom tau discrimination already applied a soft momentum requirement of  $p_T(a_1) > 35 \text{ GeV}$ . As the custom tau selection considers single tau decays, this affects the  $p_T$  trailing  $a_1$  and determines the left bound of the right plot. The discussion in Section 3.5.2 reveals that there is little momentum correlation between the two hemispheres in case of the QCD background. But to avoid a possible double counting in the estimation of the selection efficiencies for QCD processes, the momentum requirement is bypassed here. This is a conservative estimate accounting for the large uncertainties of this process.

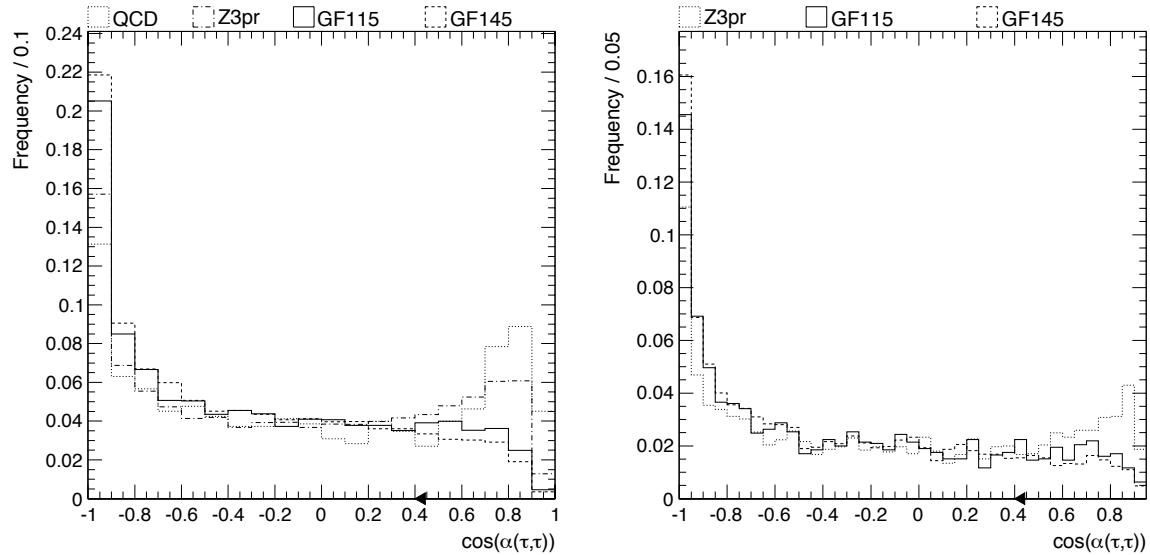


Figure 3.34.: Event selection for the gluon-fusion process: The angle  $\alpha$  enclosed by the di-tau system based on the quality (left) and custom (right) tau selection. The selection requires  $\cos(\alpha(\tau, \tau)) < 0.4$ .

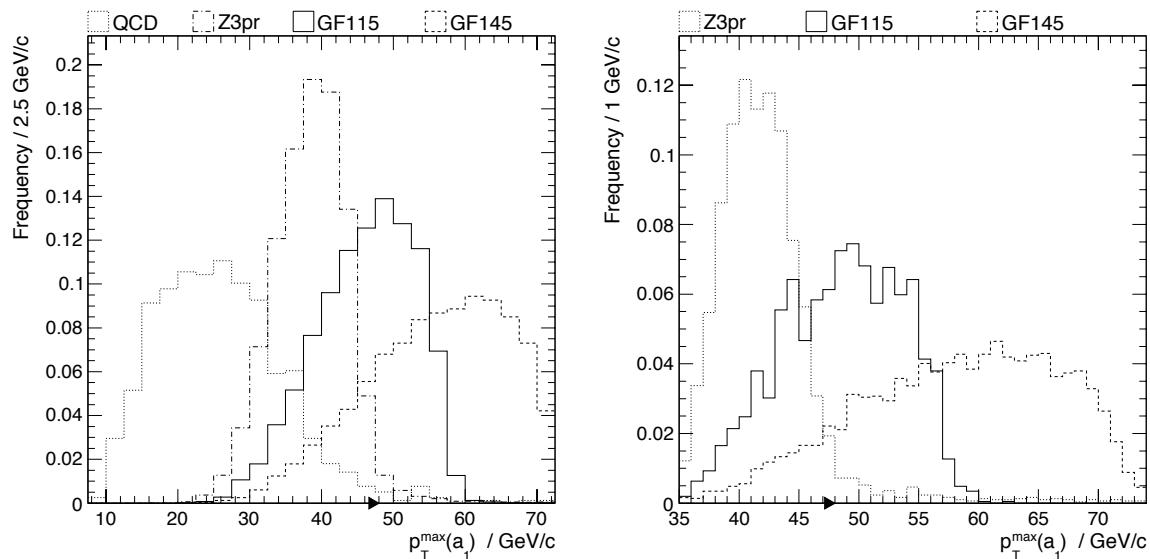


Figure 3.35.: Event selection for the gluon-fusion process: Transversal momentum of the  $p_T$ -leading  $a_1$  considering the quality (left) and custom (right) tau selection. A momentum of  $p_T^{\max}(a_1) \geq 48 \text{ GeV}/c$  is required.

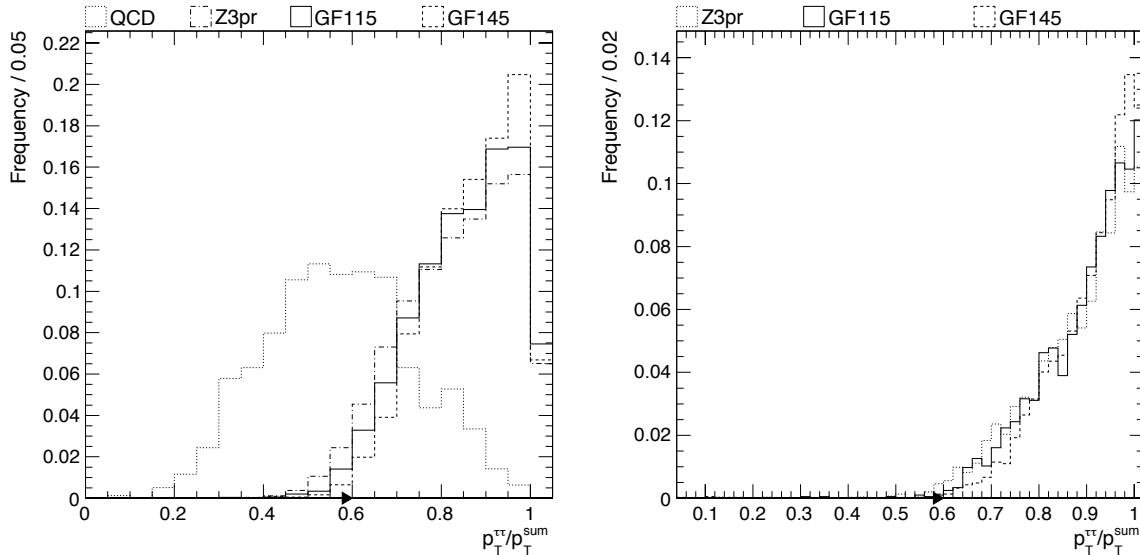


Figure 3.36.: Event selection for the gluon-fusion process: Weighted transversal momentum of the di-tau system  $p_T^{\tau\tau}/p_T^{\text{sum}}$  considering the quality (left) and custom (right) tau selection. A ratio of  $p_T^{\tau\tau}/p_T^{\text{sum}} > 0.6$  is required.

Further discrimination against events from QCD can be derived from the weighted transversal momentum of the di-tau system  $p_T^{\tau\tau}/p_T^{\text{sum}}$ . The numerator corresponds to the transversal momentum of the reconstructed Higgs boson, which includes the neutrino momenta. The denominator sums all transversal track momenta of the event including the visible part of the tau decays. To avoid correlation to the isolation requirement of the custom tau selection (see Figure 3.17 of Section 3.5.1.1), the sum ignores additional tracks within a cone of  $\Delta R(\tau, \text{track}) < 0.5$  for both tau candidates. The distribution of this observable is depicted in Figure 3.36. True tau decays contain the major part of the transversal momentum sum. The signal events and the  $Z^0$  decay provide a maximum close to one. Values larger than one are possible due to the kinematic tau reconstruction. The custom tau criteria have no impact on these distributions. The curve from fake tau candidates has a broad shape and peaks at a ratio of about 50 %. The GF selection applies a soft requirement of  $p_T^{\tau\tau}/p_T^{\text{sum}} > 0.6$ . This rejects more than half of the QCD background and marginally affects the other processes.

As in the VBF selection, the final observable is the invariant mass of the di-tau system depicted in Figure 3.37. Again a signal region of  $100 \text{ GeV}/c^2 < m_{\tau\tau} < 200 \text{ GeV}/c^2$  is defined. The shapes of the signal curves exhibit no significant deviations from those obtained in the VBF selection. A comparison of the invariant di-tau mass of all generated Higgs samples will be discussed in Section 3.7.

A summary of the gluon-fusion selection is given in Table 3.12. As in the VBF case, the efficiencies of the QCD process are estimated by events passing the quality tau selection. The other Monte Carlo samples are large enough to apply the entire selection including the custom tau discrimination. It is crucial to note that the momentum criterion labeled with  $p_T^{\max}(a_1)$  is bypassed and not included in the cumulated value. This avoids overlap to the custom tau selection (compare the assumptions made in Section 3.5.2).

With this conservative estimate, the QCD process can be suppressed below one percent. Background from the  $Z^0$  decay can be reduced to about 3 %. Here, the main rejection is

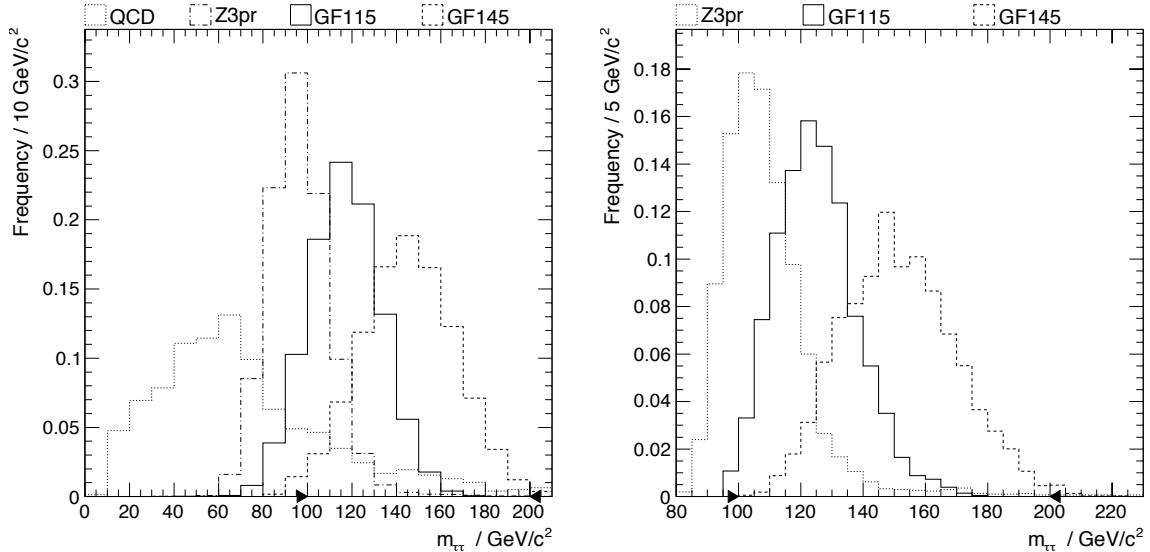


Figure 3.37.: Event selection for the gluon-fusion process: Invariant mass of the di-tau system considering the quality (left) and custom (right) tau selection. The mass has to fulfill  $100 \text{ GeV}/c^2 < m_{\tau\tau} < 200 \text{ GeV}/c^2$ .

Table 3.12.: Efficiency of the GF selection. The numbers are given in percent of the tested events. Background events from QCD contain at least one pair of tau candidates that pass the quality discriminator. For the other samples the entire selection chain including the custom tau selection is considered. The requirement labeled with  $p_T^{\max}(a_1)$  is bypassed for the QCD background.

$m_{H^0} / \text{GeV}/c^2$	$H^0_{\text{GF}}$				$Z3\text{pr}$	$\text{QCD}$
	115	125	135	145	-	-
$\sigma(\text{PV}^-, \text{PV}^+) < 2$	98.9	99.0	99.1	98.9	98.7	88.6
$3.1 < \Delta R(\tau, \tau) < 3.4$	87.3	85.6	83.8	82.5	82.8	13.9
$\cos(\alpha(\tau, \tau)) < 0.4$	82.0	85.3	84.6	85.5	72.2	67.4
$p_T^{\max}(a_1) \geq 48 \text{ GeV}/c$	56.6	72.6	80.0	87.5	6.3	—
$p_T^{\tau\tau} / p_T^{\text{sum}} > 0.6$	99.6	99.6	99.9	99.6	99.0	42.9
$100 \text{ GeV}/c^2 < m_{\tau\tau} < 200 \text{ GeV}/c^2$	98.9	99.9	99.9	99.3	72.3	18.9
cumulated	41.0	55.9	59.8	62.7	2.9	0.5

achieved by the momentum requirement. The signal events from GF pass the selection with rates between 41 % and 63 %. The different generated Higgs masses directly affect the  $a_1$  momenta and dominate this mass dependence of the efficiencies.

### 3.6.3. Selection Summary

Two independent but mostly parallel selections have been introduced to reconstruct Higgs events from either GF or VBF production. The individual efficiencies of the selection steps are combined to a cumulated value in the following Tables. The first row considers the entire selection chain starting from the HLT trigger. The second line depicts the corresponding cross section  $\sigma_{\text{exp}}$ , which combines the selection efficiencies and the initial cross sections from Table 3.1. The last line gives the expected number of events  $N_{\text{exp}}$  for each analyzed process. It is calculated using (1.17) as

$$N_{\text{exp}} = \sigma_{\text{exp}} \cdot \mathcal{L} \quad . \quad (3.5)$$

An integrated luminosity of  $\mathcal{L} = 30 \text{ fb}^{-1}$  is assumed, which can be expected for the 2011/12 run of the LHC.

Table 3.13.: Summary of the entire selection chain for VBF Higgs decays. The numbers are given in percent of the tested events. The values for the QCD background are valid under the assumptions of Section 3.5.2.

$m_{H^0} / \text{GeV}/c^2$	$H^0_{\text{VBF}}$				Z3pr	QCD
	115	125	135	145	-	-
cumulated efficiency	1.2	1.4	1.5	1.7	$7 \cdot 10^{-7}$	$6 \cdot 10^{-12}$
cross section / ab	10.3	9.5	6.5	3.9	0.05	0.21
exp. events @ $30 \text{ fb}^{-1}$	0.3	0.3	0.2	0.1	0.001	0.006

The selection of Higgs decays from VBF is summarized in Table 3.13. About 10 ab of the lightest generated Higgs mass pass the event selection. This corresponds to a cumulated efficiency of more than 1 %. The heavier Higgs masses can be selected with up to 1.7 %. The increasing reconstruction and selection efficiencies for heavier Higgs bosons is not sufficient to compensate the smaller initial cross section. The backgrounds are suppressed below one attobarn. The leading one is the QCD process, which is six times larger than the contribution from  $Z^0$  decays. The di-tau signature and the additional jet pair provide a clean remaining sample. But the VBF suffers from a very small production cross section. With an integrated luminosity of  $30 \text{ fb}^{-1}$ , the probability is about 30 % to observe one event.

Table 3.14.: Summary of the selection chain for GF Higgs decays. The numbers are given in percent of the tested events. The values for the QCD background are valid under the assumptions of Section 3.5.2.

$m_{H^0} / \text{GeV}/c^2$	$H^0_{\text{GF}}$				Z3pr	QCD
	115	125	135	145	-	-
cumulated efficiency	0.8	1.5	2.0	2.3	$9 \cdot 10^{-3}$	$2 \cdot 10^{-11}$
cross section / fb	0.10	0.13	0.10	0.06	0.55	$8 \cdot 10^{-4}$
exp. events @ $30 \text{ fb}^{-1}$	3	4	3	2	17	0.02

The situation is different for the GF selection in Table 3.14. The signal events pass the selection with efficiencies between 0.8 % and 2.3 %. The stronger mass dependence compared

to the VBF analysis is caused by the tight momentum requirement. A large part of the light Higgs masses is rejected to suppress the background from  $Z^0$  decays. The balanced di-tau system of the GF process provides a large rejection of QCD events but an insufficient discrimination of  $Z^0$  decays. In contrast to the VBF selection, the selected sample is dominated by background events. About 17  $Z^0$  events remain at  $30 \text{ fb}^{-1}$ . The signal contributes with up to four events.

The efficiency to select a Higgs decay from GF by the VBF selection is negligible. Within the inspected datasets, no event is observed. The other way around, a few generated events from the VBF sample pass the GF selection. Due to their small cross section, this gives no significant increase of the expected number of signal events.

Table 3.15.: Matching purity  $\mathcal{P}$  of the tau pairs obtained from the VBF and the GF selection. If at least one tau decay has no matching, the whole pair is regarded as not matched. The numbers for the VBF selection include the purity of the selected jet pairs.

$m_{H^0} / \text{GeV}/c^2$	$H^0_{\text{GF}}$				$H^0_{\text{VBF}}$				Z3pr
	115	125	135	145	115	125	135	145	
$\mathcal{P}_{\text{VBF}} / \%$	-	-	-	-	95.6	94.9	94.2	94.1	-
$\mathcal{P}_{\text{GF}} / \%$	98.1	97.2	96.6	96.7	-	-	-	-	98.8

The matching purity  $\mathcal{P}$  of the final selection is given in Table 3.15. For each process, it compares the number of matched tau pairs  $n_{\text{matched}}$  to all selected pairs  $n_{\text{selected}}$ . If at least one tau decay has no matching, the whole pair is regarded as not matched. The purity is defined as

$$\mathcal{P} = n_{\text{matched}} / n_{\text{selected}} . \quad (3.6)$$

Both selections show a small mass dependence. Lighter di-tau pairs have a larger purity of up to 99 % in case of the  $Z^0$  decay. This is caused by the lower selection efficiency. The values obtained for the VBF selection include the purity of the jet selection. This slightly decreases the purity down to 94 %.

### 3.7. Invariant Mass of the Tau Pair

The event rates quoted in the last section refer to the entire mass region defined in the VBF and the GF selection. The shape of the reconstructed spectrum is not yet considered.

Figure 3.38 depicts the reconstructed di-tau masses after the selections. Each histogram contains two signal curves, the lightest and the heaviest Higgs mass, and the background distributions. To compare with the expected event rates of Table 3.13 and Table 3.14, all curves are normalized to an integrated luminosity of  $30 \text{ fb}^{-1}$ . Both selections reject all generated events of the QCD sample. No  $Z^0$  events pass the VBF selection. To obtain a suitable statistic in the selected mass window, the background contributions are modeled with shapes obtained from earlier selection steps and normalized with the proper efficiencies. For both selections, the QCD curve is obtained from tau pairs that pass the quality criteria only. In case of the VBF selection, the  $Z^0$  curve is modeled from custom tau pairs. This method induces uncertainties on the shapes, but allows for a rough approximation.

The VBF selection efficiently suppresses the background, but no signal events are expected with  $\mathcal{L} = 30 \text{ fb}^{-1}$ . The GF process can be isolated from QCD events but suffers from  $Z^0$  back-

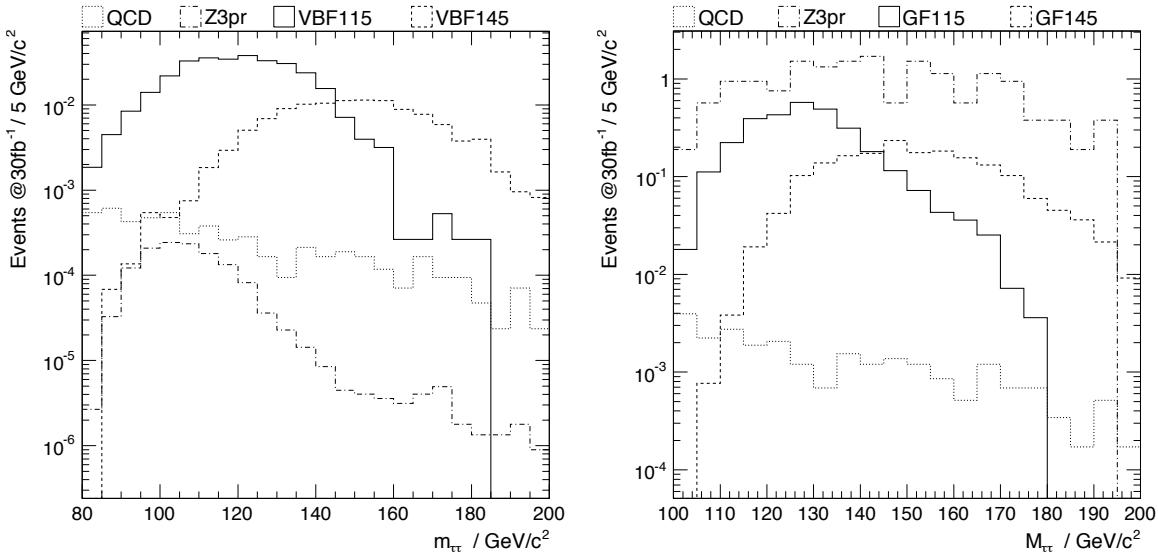


Figure 3.38.: Invariant mass of tau pairs passing the event selection for VBF (left) and GF (right). Two signal curves and the background processes are shown. All curves are scaled to an integrated luminosity of  $\mathcal{L} = 30 \text{ fb}^{-1}$ . The QCD background is modeled by the curve obtained from quality tau pairs and scaled with the efficiency of the entire selection chain. In case of the VBF selection, the Z3pr shape is obtained from custom tau pairs.

ground. This background source covers the entire mass window. The rate slightly decreases towards higher masses.

The remaining events contain off-shell  $Z^0$  bosons. The background is irreducible. Figure 3.39 compares the reconstructed and generated di-tau masses within the selected mass window. It shows a resolution of  $\sigma(m_{\text{reco}} - m_{\text{gen}}) \approx 15 \text{ GeV}$ . There is a systematic shift of 10 GeV towards larger reconstructed values. It is mainly caused by the tight requirement on the  $a_1$  momentum. The scatter plot reveals only few remaining events near the nominal  $Z^0$  mass. Most of the selected tau pairs correspond to heavy  $Z^0$  bosons.

To improve the discrimination, mass windows for every generated Higgs mass are defined by a gaussian fit to the reconstructed mass peaks. The regions cover one standard deviation around the central values. The fits for the lightest and heaviest Higgs mass are shown in Figure 3.40 for both selections.

Table 3.16.: Resolution of the reconstructed Higgs mass from VBF and GF production obtained from a gaussian fit. The number of events from signal  $s$  and background  $b$  are counted within  $\pm 1\sigma$  around the fitted mean  $\mu$  for  $\mathcal{L} = 30 \text{ fb}^{-1}$ .

$m_{H^0} / \text{GeV}/c^2$	$H^0_{\text{GF}}$				$H^0_{\text{VBF}}$			
	115	125	135	145	115	125	135	145
$\mu / \text{GeV}/c^2$	128	137	145	153	120	130	141	150
$\sigma / \text{GeV}/c^2$	11	13	15	16	15	17	18	19
$s$	2	2.3	2	1.1	0.2	0.2	0.1	0.08
$b$	5.3	6.6	7.8	7	0.002	0.001	0.001	0.001

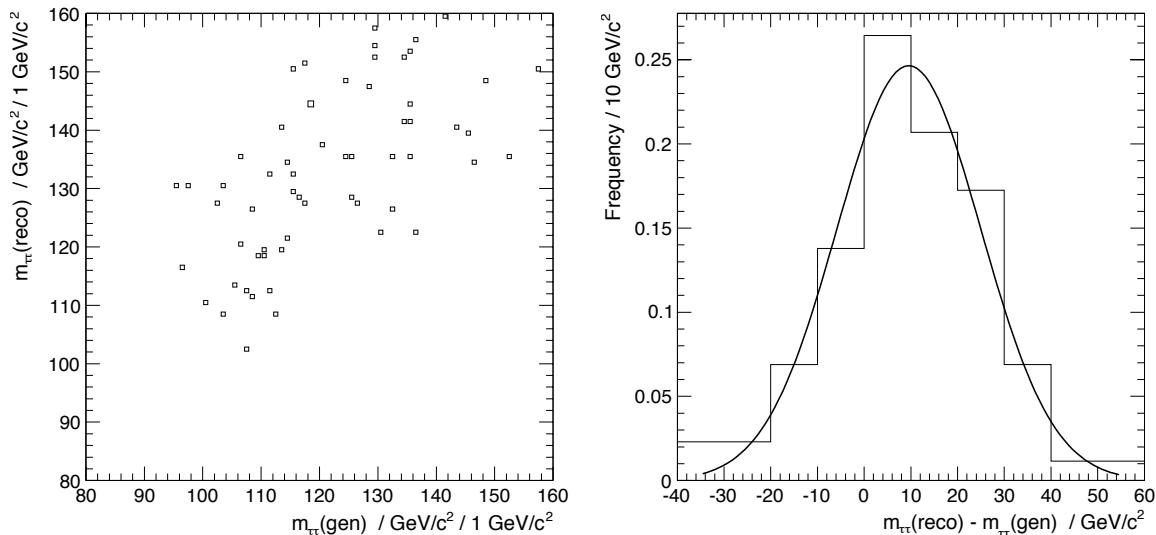


Figure 3.39.: Di-tau mass of  $Z^0$  decays from the GF selection. The generated  $Z^0$  masses scatter over the selected mass window (left). The mass resolution (right) obtained by a gaussian fit is  $\sigma(m_{\text{reco}} - m_{\text{gen}}) \approx 15 \text{ GeV}$ .

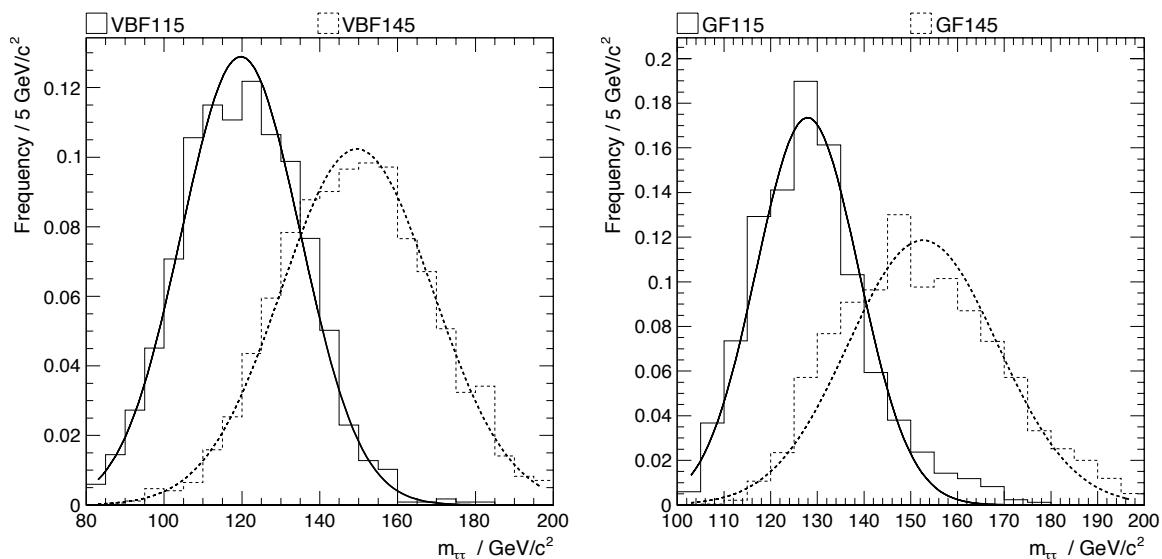


Figure 3.40.: Fit of a gaussian distribution to the reconstructed di-tau mass for the VBF (left) and GF (right) process.

The results of the fits, the central value  $\mu$  and the width  $\sigma$ , are summarized in Table 3.16. In case of the GF process, the offset to the generated mass decreases for heavier masses. Light Higgs bosons are affected by the momentum requirement. The impact on the VBF process is lower due to the significant transversal momentum of the Higgs boson. Here, the offset is about 5 GeV. For both production processes the width slightly increases with heavier masses.

The numbers of expected events that are contained in the mass windows are taken from Figure 3.38. The signal  $s$  and the total background contribution  $b$  in the last two rows of Table 3.16 are obtained. The significance of the signal will be discussed in Section 3.9.

### 3.8. Systematic Uncertainties

The results of the analysis are affected by systematic uncertainties. The main theoretical uncertainties are introduced through the cross sections of the involved processes. The precision of the Higgs cross section is in the order of 20 % for the production via gluon fusion and 3 % for vector-boson fusion [20]. The  $Z^0$ -decay rate  $\sigma \times \mathcal{B}(Z^0 \rightarrow \tau\tau)$  is known to about 4 % [100] assuming lepton universality. This value will be measured precisely at the  $Z^0$  peak. The tau decay  $\tau \rightarrow 3\pi^\pm + \nu_\tau$  is known to better than 1 %.

The largest experimental uncertainty arises from the estimation of the integrated luminosity. The instantaneous luminosity is determined with an accuracy of 11 % during the initial data taking period [101]. The real-time measurement relies on the forward hadronic calorimeters and the absolute normalization is based on Van-der-Meer scans [102]. With integrated luminosities of  $30 \text{ fb}^{-1}$ , a reduction to less than 5 % seems feasible [103].

The kinematic tau reconstruction relies on the efficiency of the initial particle-flow seeds, which is estimated with  $Z^0 \rightarrow \tau\tau$  decays to 9 % accuracy [104]. In case of the exclusive decay into three charged pions, the efficiency is less affected by the dominant calorimetry scale uncertainty. It is assumed to be 5 %. The final observable, the invariant mass of the di-tau system, depends on the kinematic tau reconstruction. It is expected to be little affected by variations of the jet-energy scale and does not rely on global event variables like the missing transversal energy. The kinematic tau reconstruction modifies the primary vertex for both hemispheres. An unambiguous assignment from both tau leptons to a common primary vertex reduces the impact of multiple vertices in pileup collisions.

The background contribution from QCD induces large uncertainties. The effective rate will be experimentally determined from well defined control regions. In both selections, the impact of the QCD background is expected to be small. Even an uncertainty of 100 % is tolerable.

### 3.9. Signal Significance

There are different methods to determine the significance of the signal. It is defined as the number of standard deviations an observed signal exceeds the background. Sophisticated approaches like the CLs method [85, 105] consider the likelihood ratio of signal and background hypotheses. They improve the statistical interpretation by exploiting the shapes of the expected curves.

Due to the small statistics of the selected sample, the significance  $S_{\text{cL}}$  is obtained by a simple

counting method. It is derived from the Poisson distribution as

$$S_{\text{cL}} = \sqrt{2[(s+b) \ln(1+s/b) - s]} \quad (3.7)$$

with the number of signal events  $s$  and expected background  $b$ . In the Gaussian limit for large  $s$  and  $b$ , this is equivalent to the simple ratio  $S = s/\sqrt{b}$ .

The systematic uncertainties from Section 3.8 are taken into account as gaussian distributions. As a conservative estimation, the expected signal is reduced and the background is increased by the systematic uncertainties.

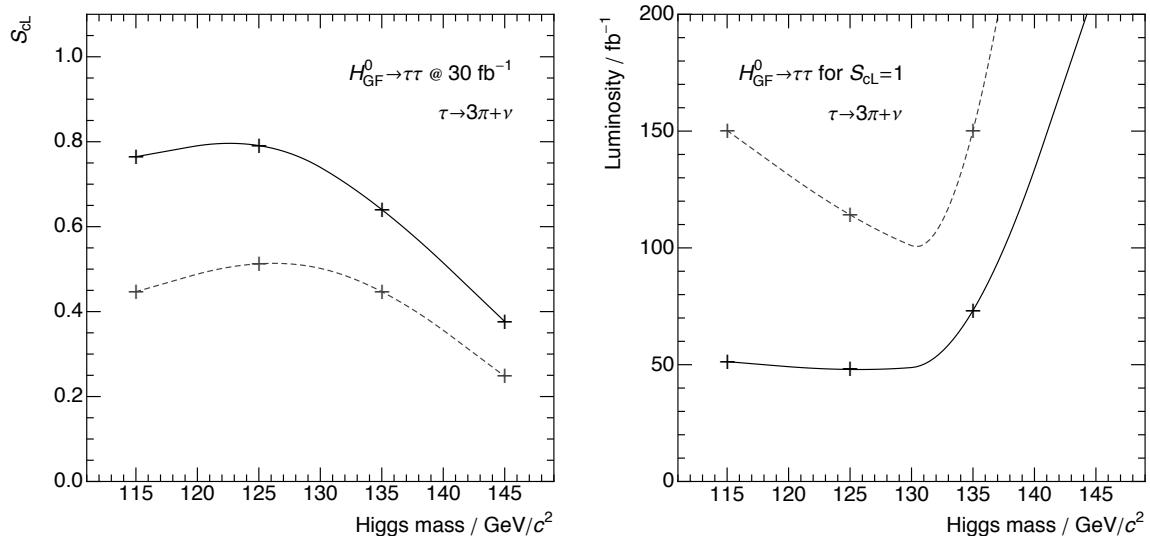


Figure 3.41.: The signal significance  $S_{\text{cL}}$  with  $\mathcal{L} = 30 \text{ fb}^{-1}$  (left) and the required luminosity for  $S_{\text{cL}} = 1$  (right) as a function of the Higgs mass in gluon-fusion production. The dashed curve considers the entire mass window of the selection. The solid curve is restricted to the peak regions from Table 3.16.

The resulting significance for the GF production with an integrated luminosity of  $\mathcal{L} = 30 \text{ fb}^{-1}$  is depicted in the left part of Figure 3.41. Two curves interpolate the four analyzed mass points. The dashed curve considers the whole mass window of the GF selection. The significance is below  $S_{\text{cL}} = 0.5$ . Restricting the mass window to the smaller peak regions defined in Table 3.16 yields the solid curve. The significance increases up to  $S_{\text{cL}} = 0.8$ . The mass dependence is mainly caused by the different cross sections and the width of the reconstructed di-tau mass. As depicted in Figure 3.38, the few remaining events of the heaviest generated Higgs mass are broadly distributed below the  $Z^0$  distribution.

The right part of Figure 3.41 gives the luminosity as a function of the Higgs mass, which is required to observe a significance of  $S_{\text{cL}} = 1.0$ . About  $50 \text{ fb}^{-1}$  are needed for Higgs masses of  $115 \text{ GeV}/c^2$  and  $125 \text{ GeV}/c^2$  assuming a constant signal-over-background ratio.

The VBF process provides a much cleaner signal selection. But the cross section of the analyzed channel is not large enough to observe events with  $\mathcal{L} = 30 \text{ fb}^{-1}$ . A calculation of the significance is meaningless.

The obtained result is limited by the statistical uncertainties. Some possibilities to enhance the event rate are discussed in the following.

### 3.9.1. Possible Enhancement

One important factor that limits the selected event rate is the branching fraction of the decay  $\tau \rightarrow 3\pi^\pm + \nu_\tau$  of less than 10 %. It enters the signal cross section quadratically

$$\sigma_{H^0}^{3\pi} \equiv \sigma(pp \rightarrow H^0) \cdot \mathcal{B}(H^0 \rightarrow \tau\tau) \cdot \mathcal{B}(\tau \rightarrow 3\pi^\pm + \nu_\tau)^2 \quad . \quad (3.8)$$

A reasonable way to increase the signal rate is to accept neutral pions in the tau decay. The efficiency for this decay mode reconstructed by the kinematic fit of three charged pions was already discussed in Section 2.4.5. Developing a dedicated fit for this additional decay mode seems feasible. The combined branching fraction  $\mathcal{B}(\tau \rightarrow 3\pi^\pm + \nu_\tau + X) \approx 15\%$  provides an increase of the signal of

$$\frac{\sigma_{H^0}^{3\pi+X}}{\sigma_{H^0}^{3\pi}} = \frac{\mathcal{B}(\tau \rightarrow 3\pi^\pm + \nu_\tau + X)^2}{\mathcal{B}(\tau \rightarrow 3\pi^\pm + \nu_\tau)^2} \approx 2.6 \quad . \quad (3.9)$$

An even larger increase will come from events where only one tau lepton decays via  $\tau \rightarrow 3\pi^\pm + \nu_\tau + X$ . The second tau lepton decays as a hadronic 1-prong, which is illustrated in Figure 3.42. The 1-prong tau decay cannot be reconstructed completely on its own, but

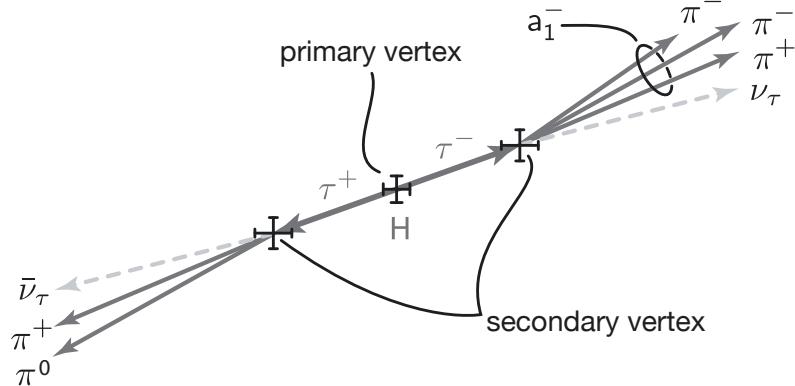


Figure 3.42.: Illustration of the Higgs decay into a tau pair. One tau decays into a 3-prong, the other one into a 1-prong.

with the assumption of a  $p_T$  balance between the tau leptons. One has to study if looser requirements on the 1-prong tau decay still provide a sufficient QCD rejection. The combined branching ratio of hadronic 1-prong and 3-prong tau decays is about  $\mathcal{B}(\tau \rightarrow \text{had}) = 65\%$ . It leads to an enhanced cross section  $\sigma_{H^0}^{\text{enhanced}}$ . The gain is

$$\begin{aligned} \frac{\sigma_{H^0}^{\text{enhanced}}}{\sigma_{H^0}^{3\pi}} &= \left[ 2 \cdot \mathcal{B}(\tau \rightarrow 3\pi^\pm + \nu_\tau + X) \cdot \mathcal{B}(\tau \rightarrow \text{had}) - \mathcal{B}(\tau \rightarrow 3\pi^\pm + \nu_\tau + X)^2 \right] \\ &\cdot \mathcal{B}(\tau \rightarrow 3\pi^\pm + \nu_\tau)^{-2} \approx 20 \quad . \end{aligned} \quad (3.10)$$

The background contribution from  $Z^0$  decays will increase with the same factor. It will be the dominant background also for the VBF selection. The di-tau system from VBF is not balanced in the transversal plane and the reconstruction may fail for this process. If it turns out to be possible, the signal gain allows for a Higgs identification with a significance of

more than  $5\sigma$  in the low mass regime. This value comes from the estimate in the left part of Figure 3.43. The expected events for signal and background from Table 3.16 are scaled with the calculated gain of (3.10) assuming a constant signal-over-background ratio.

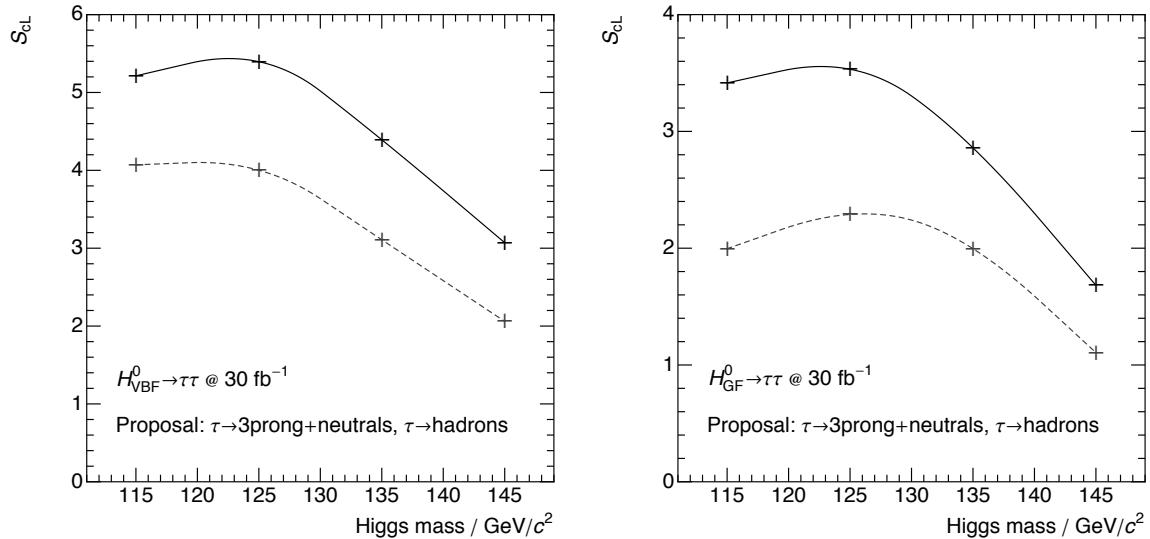


Figure 3.43.: Enhanced signal significance  $S_{cl}$  with  $\mathcal{L} = 30 \text{ fb}^{-1}$  as a function of the Higgs boson mass from VBF (left) and GF (right). The selection accepts 1-prong tau decays in one hemisphere. A constant signal-to-background ratio with regard to Table 3.16 is assumed. The dashed curve considers the entire mass window of the selection. The solid curve is restricted to the peak regions.

In case of the gluon fusion, the tau pair is well balanced and the proposed signal enhancement seems achievable. In contrast to the vector-boson fusion, the sample is dominated by background. The significance increases due to reduced statistical uncertainties. The right part of Figure 3.43 gives the resulting significance for the gluon-fusion production. Light Higgs masses exhibit a significance of more than  $3\sigma$ . This is a conservative estimate as the discrimination from  $Z^0$  decays can be improved with a larger sample. Especially the requirement on the  $a_1$  momentum can be tightened.

An independent way to increase the significance is to improve the mass resolution. In case of the GF selection, the true  $Z^0$  mass is broadly distributed over the region of interest while the width of the Higgs boson is negligible small. An improved resolution allows to collect the same amount of signal events with less contamination from backgrounds. It has to be investigated if a correction of the intrinsic resolution discussed in Section 2.4.6.1 is sufficient. Additionally, a simultaneous fit of both tau decays can exploit the momentum balance to increase the resolution.

## Chapter 4

# Summary and Outlook

The existence of Higgs bosons can be studied with the CMS experiment at the LHC. This thesis prepares a search in the di-tau channel based on Monte Carlo simulations. Light Higgs masses between  $115\text{ GeV}/c^2$  and  $145\text{ GeV}/c^2$  with the predicted cross sections of the Standard Model are considered. The reconstruction of Higgs bosons in hadronic tau decays relies on an efficient algorithm to identify the tau decay products within the hadronic environment.

A complete reconstruction of 3-prong tau decays based on a kinematic fit has been developed. The method's robustness and the achievable resolution is tested on several different tau-production processes. Its suppression of fake tau leptons from quark and gluon jets can be enhanced by quality criteria. A set of loose requirements, which cover a broad range of physics applications, has been defined. For 3-prong tau decays the efficiency is reasonably high. The fake rate is reduced by 2 orders of magnitude.

The Higgs boson can be reconstructed from pairs of tau leptons that pass the quality criteria. The event kinematics of the production processes, vector-boson fusion and gluon fusion, can be exploited to provide further suppression of backgrounds. The rejection of fake tau jets from QCD is estimated by a factorized selection. In case of the gluon fusion, the remaining background is dominated by off-shell  $Z^0$  bosons that decay into tau pairs. With an integrated luminosity of  $30\text{ fb}^{-1}$  a signal significance of  $0.8\sigma$  can be achieved. It is limited by statistics. Higgs bosons produced by vector-boson fusion are accompanied by two quark jets. Their additional signature and the significant transversal boost of the Higgs boson make a clean selection possible. Due to the small cross section, an integrated luminosity of about  $100\text{ fb}^{-1}$  has to be collected to establish a signal.

Possible enhancement may come from additional decay modes in the kinematic fit of the tau lepton. Accepting tau decays with one charged hadron in the final state increases the signal rate by a factor of 20. The impact on the signal-to-background ratio has to be studied. A constant ratio is expected, which increases the significance by a factor of 4.

To further enrich the signal from the gluon fusion, the suppression of  $Z^0$  decays has to be enhanced. An improved resolution of the di-tau mass directly affects the signal-over-background ratio. It may be achieved by a simultaneous fit of both tau decays.

The efficiencies and systematic uncertainties derived in this thesis need to be tested with real data. Especially the determination of the background rates and shapes exceed the precision of the Monte Carlo simulation.



## Appendix A

# Appendix on Kinematic Tau Reconstruction

### A.1. Helix Propagation

In an assumed solenoid magnetic field  $B$  along the  $z$ -direction, the tracks of charged particles are bend by the Lorentz force:

$$\begin{aligned}\frac{dp_T}{dt} &= qv_T B = m \frac{v_T^2}{R} \\ \Rightarrow R &= p_T/a\end{aligned}$$

$R$  is the radius of the curvature. The index T refers to the transverse component w.r.t. the orientation of the magnetic field. The factor  $a$  depends on the electric charge of the particle  $q$  and the strength of the field. Using the CMS convention, which defines distances in cm and momenta in  $\text{GeV}/c$ , one obtains  $a \approx -3 \cdot 10^{-3} qB \frac{\text{GeV}}{\text{cm}\cdot c}$ . The equation of motion can be

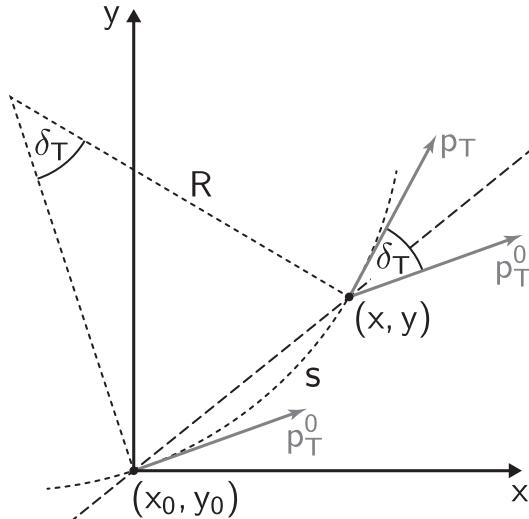


Figure A.1.: Propagation of a charged particle in a solenoid magnetic field along the  $z$ -direction.

derived from Figure A.1. A charged particle with momentum  $\vec{p}_0 = (p_{x0}, p_{y0}, p_{z0})$  defined at  $(x_0, y_0, z_0)$  propagates through the field along the pathlength  $s = R\delta$  towards  $(x, y, z)$ . The momentum is bend by the angle  $\delta$  not affecting the magnitude of the momentum vector  $p = p_0 = |\vec{p}_0|$ . A projection into the bend plane defines the transverse angle  $\delta_T = \delta \frac{p_T}{p} = \delta \frac{a}{p} R = \rho s$ , where  $\rho = a/p$  is the signed transverse curvature. Rotating the initial values of

spatial and momentum components around this angle  $\delta_T$  yields the final equations:

$$\begin{aligned} p_x &= p_{x0} \cos \rho s - p_{y0} \sin \rho s & x &= x_0 + \frac{p_{x0}}{a} \sin \rho s - \frac{p_{y0}}{a} (1 - \cos \rho s) \\ p_y &= p_{x0} \sin \rho s + p_{y0} \cos \rho s & y &= y_0 + \frac{p_{x0}}{a} (1 - \cos \rho s) + \frac{p_{y0}}{a} \sin \rho s \\ p_z &= p_{z0} & z &= z_0 + \frac{p_{z0}}{p} s \end{aligned} \quad . \quad (\text{A.1})$$

Combining these equations and eliminating terms containing the angle  $\delta_T$  defines the momentum transformation depending on the position:

$$\vec{p} = \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} = \begin{pmatrix} p_{x0} - a(y - y_0) \\ p_{y0} + a(x - x_0) \\ p_{z0} \end{pmatrix} \quad . \quad (\text{A.2})$$

## A.2. Vertex Link Significance

The significance of the three-dimensional link between two arbitrary vertices  $\vec{v}_a$  and  $\vec{v}_b$  is a measure of the vertex separation with regard to their positional uncertainties. The inaccuracies are represented by two  $3 \times 3$ -covariance matrices  $\mathbf{V}^a$  and  $\mathbf{V}^b$ . The vertex link  $\vec{v}_{ab}$  is simply defined as

$$\vec{v}_{ab} = \vec{v}_b - \vec{v}_a \quad .$$

First of all, the similarity  $\lambda^2$  of the vertex link  $\vec{v}_{ab}$  and the combination of the two covariance matrices is calculated:

$$\lambda^2 = \vec{v}_{ab}^T (\mathbf{V}^a + \mathbf{V}^b) \vec{v}_{ab} \quad .$$

As the vertex link  $\vec{v}_{ab}$  is not normalized, we have to compensate for this by constructing

$$\sigma_{\text{abs}} = \frac{\lambda}{|\vec{v}_{ab}|} \quad .$$

The resulting variable  $\sigma_{\text{abs}}$  can be identified as the combined uncertainty of both vertices in the direction of the vertex link. The unit of  $\sigma_{\text{abs}}$  corresponds to a physical length and is always positive. The vertex link significance  $\sigma$  is a dimensionless scalar and can be derived by relating the absolute uncertainty with the magnitude of the vertex link:

$$\sigma = \frac{|\vec{v}_{ab}|}{\sigma_{\text{abs}}} \quad .$$

The quantitative value of  $\sigma$  is equal to the number of standard deviations, which separate both vertices. Therefore, a vertex significance equal to zero corresponds to two vertices, which lie on top of each other. Accordingly, large values of  $\sigma$  characterize well separated vertices.

## Appendix B

# Appendix on Higgs Boson Reconstruction

### B.1. Jet Matching

Reconstructed jet candidates are matched to generated quarks in three steps. First every reconstructed object is assigned to the closest generator jet in the  $(\eta, \phi)$ -plane. Jets at generator level, referred to as gen-jets, are created by the same algorithms that are applied to the reconstruction. All generated stable particles, except for neutrinos, potentially result in a gen-jet. Their direction is expected to be closer to the reconstructed jets than directly comparing quarks to the reconstruction. As no maximal distance is defined, every reconstructed jet is assigned to a gen-jet.

The gen-jets store a collection of the contained particles. This collection gives access to the initial quark or gluon that initiated the decay chain. In a second step duplicate assignments to the same generator particle are deleted. Only the closest pair in  $\Delta R$  is accepted.

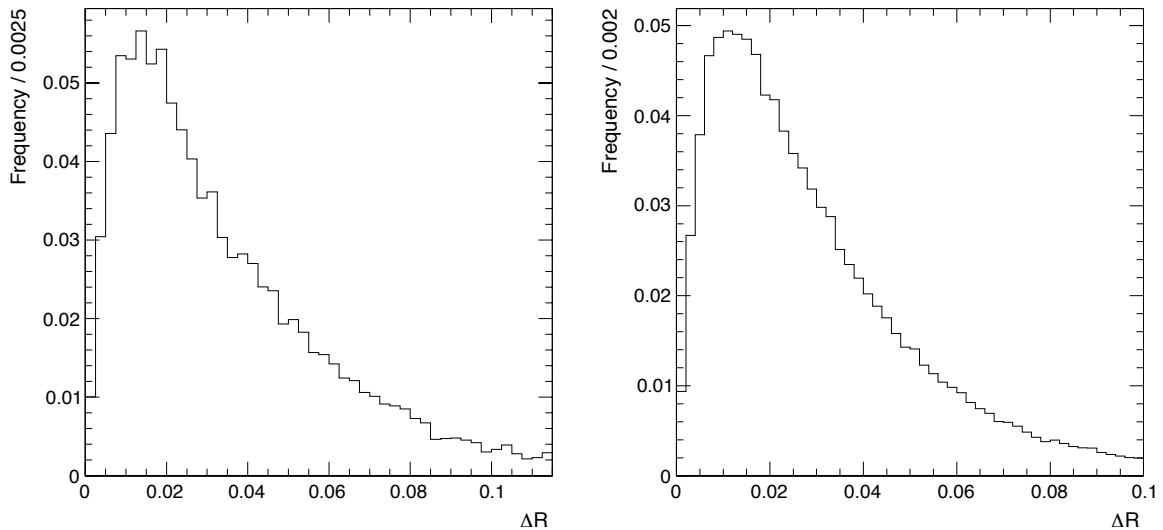


Figure B.1.: Distance in terms of  $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$  between the reconstructed jets and generated quarks from vector-boson fusion. Jets are reconstructed from calorimetry information (left) or by the particle-flow reconstruction (right). The plots are zoomed to show 95 % of the assigned jets.

Figure B.1 depicts the  $\Delta R$  distribution of the remaining assignments. The jets are obtained from a dataset containing quark jets from vector-boson fusion. The expected kinematics

of this process are described in Section 3.2.4. The jets already passed a soft filter, which requires a minimum of transversal momentum,  $p_T(\text{jet}) > 5 \text{ GeV}$ . The left curve derives from a jet reconstruction from calorimetry towers, the right curve considers candidates from the particle-flow reconstruction. In both cases an iterative-cone algorithm with an annulus of  $\Delta R = 0.5$  is applied. The calo jets show slightly larger differences to the generated quarks than the particle-flow jets. In both cases 95 % of the matched pairs satisfy  $\Delta R < 0.12$ .

In the third step a maximal separation between the assigned objects is defined. According to the  $\Delta R$  distribution in Figure B.1 a reconstructed jet is regarded as matched if its distance to the particle at generator level holds

$$\Delta R(\text{quark}, \text{jet}) < 0.1 . \quad (\text{B.1})$$

Signal-like jets from the vector-boson fusion have a preferred pseudorapidity of  $\eta = \pm 2.5$ . In this forward region the maximal distance in the  $(\eta, \phi)$ -plane corresponds to a maximal angular difference of  $\Delta\theta < 1^\circ$ . In events from vector-boson fusion each of the two quarks can be unambiguously matched with an efficiency of 87.2 % to calo jets and with 89.6 % to particle-flow jets. Excluding the reconstruction efficiency yields 93 % (calo jets) and 95 % (particle-flow jets).

## B.2. Basic Kinematic Distributions

This Section contains some distribution of basic kinematic observables. They are referred to in Chapter 3.

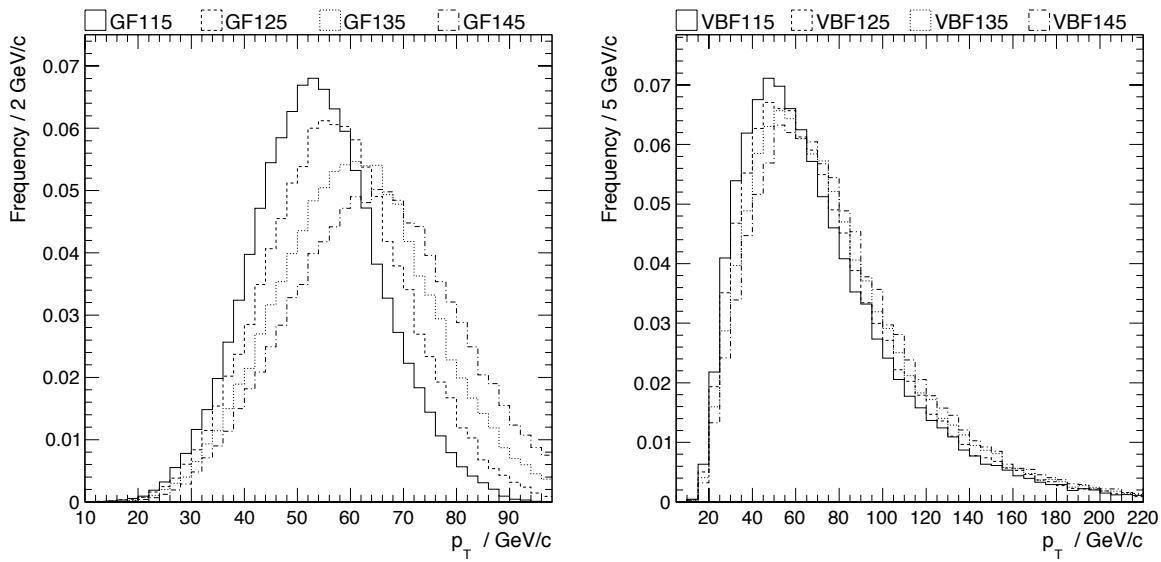


Figure B.2.: Transversal momentum of tau decays that pass the quality discrimination. The tau leptons arise from Higgs decays produced by vector-boson fusion (left) and gluon fusion (right). The distributions are referred to in Section 3.5.1.1 and Section 3.5.3.1.

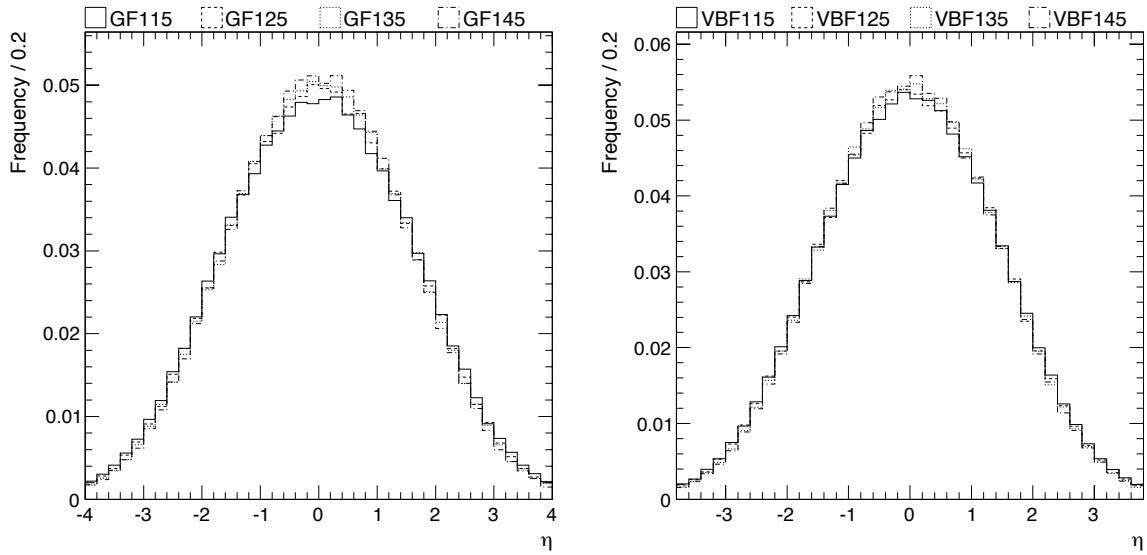


Figure B.3.: Pseudorapidity of generated tau leptons. The tau leptons arise from Higgs decays produced by vector-boson fusion (left) and gluon fusion (right). The distributions are discussed in Section 3.2.3

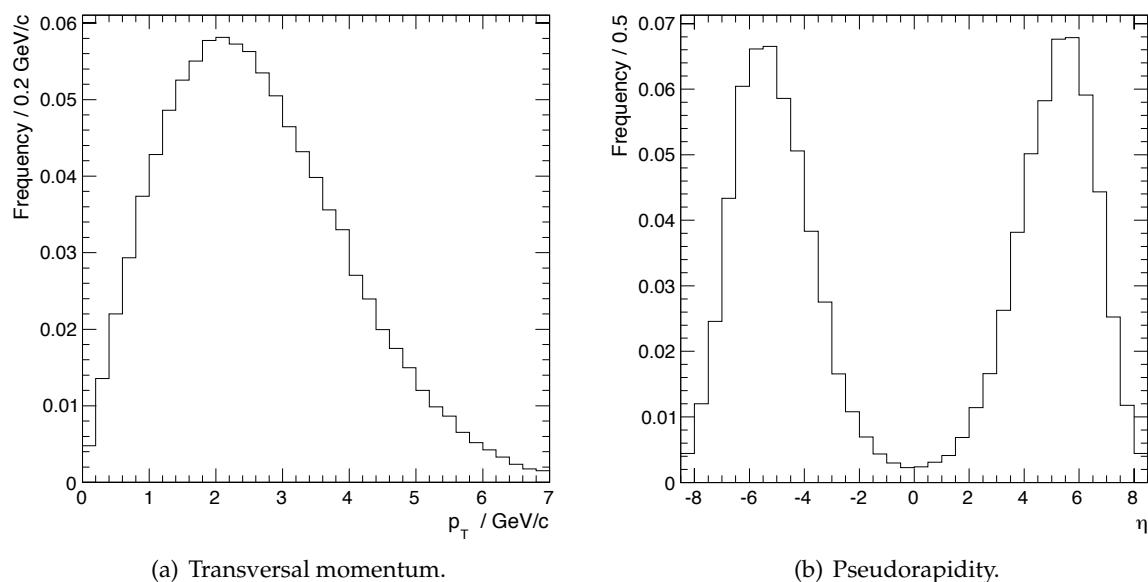


Figure B.4.: Basic kinematics of the generated  $Z^0$  boson, which are discussed in Section 3.2.5.



# References

## Notes:

- a) CMS Notes are available at <http://cms.cern.ch/iCMS/> unless otherwise noted.
  - b) References marked **doi** should be prefixed with <http://dx.doi.org/>.
- 

- [1] S. L. Glashow, "Partial-symmetries of weak interactions", *Nuclear Physics* **22** (1961), no. 4, 579 – 588. doi:[10.1016/0029-5582\(61\)90469-2](https://doi.org/10.1016/0029-5582(61)90469-2).
- [2] S. Weinberg, "A Model of Leptons", *Phys. Rev. Lett.* **19** (Nov, 1967) 1264–1266. doi:[10.1103/PhysRevLett.19.1264](https://doi.org/10.1103/PhysRevLett.19.1264).
- [3] A. Salam, "Relativistic Groups and Analyticity", in *Elementary Particle Theory*, Svartholm, Almquist, and Wiksell, eds. Stockholm, 1968.
- [4] C. N. Yang and R. L. Mills, "Conservation of Isotopic Spin and Isotopic Gauge Invariance", *Phys. Rev.* **96** (Oct, 1954) 191–195. doi:[10.1103/PhysRev.96.191](https://doi.org/10.1103/PhysRev.96.191).
- [5] M. Gell-Mann, "A schematic model of baryons and mesons", *Physics Letters* **8** (1964), no. 3, 214 – 215. doi:[10.1016/S0031-9163\(64\)92001-3](https://doi.org/10.1016/S0031-9163(64)92001-3).
- [6] G. Zweig, "An SU<sub>3</sub> model for strong interaction symmetry and its breaking; Part I".
- [7] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, "Advantages of the color octet gluon picture", *Physics Letters B* **47** (1973), no. 4, 365 – 368. doi:[10.1016/0370-2693\(73\)90625-4](https://doi.org/10.1016/0370-2693(73)90625-4).
- [8] D. J. Gross and F. Wilczek, "Ultraviolet Behavior of Non-Abelian Gauge Theories", *Phys. Rev. Lett.* **30** (Jun, 1973) 1343–1346. doi:[10.1103/PhysRevLett.30.1343](https://doi.org/10.1103/PhysRevLett.30.1343).
- [9] H. D. Politzer, "Reliable Perturbative Results for Strong Interactions?", *Phys. Rev. Lett.* **30** (Jun, 1973) 1346–1349. doi:[10.1103/PhysRevLett.30.1346](https://doi.org/10.1103/PhysRevLett.30.1346).
- [10] P. W. Higgs, "Broken symmetries, massless particles and gauge fields", *Phys. Lett.* **12** (1964) 132–133.
- [11] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons", *Phys. Rev. Lett.* **13** (Oct, 1964) 508–509. doi:[10.1103/PhysRevLett.13.508](https://doi.org/10.1103/PhysRevLett.13.508).
- [12] F. Englert and R. Brout, "Broken symmetry and the masses of gauge vector mesons", *Phys. Rev. Lett.* **13** (1964) 321–323.
- [13] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global Conservation Laws and Massless Particles", *Phys. Rev. Lett.* **13** (Nov, 1964) 585–587.

- doi:10.1103/PhysRevLett.13.585.
- [14] P. W. Higgs, "Spontaneous Symmetry Breakdown without Massless Bosons", *Phys. Rev.* **145** (May, 1966) 1156–1163. doi:10.1103/PhysRev.145.1156.
  - [15] G. 't Hooft and M. J. G. Veltman, "Regularization and renormalization of gauge fields", *Nucl. Phys. B* **44** (1972), no. 1-3, 189–213.
  - [16] J. Goldstone, "Field theories with Superconductor solutions", *Il Nuovo Cimento (1955-1965)* **19** (1961) 154–164. 10.1007/BF02812722.
  - [17] J. Goldstone, A. Salam, and S. Weinberg, "Broken Symmetries", *Phys. Rev.* **127** (Aug, 1962) 965–970. doi:10.1103/PhysRev.127.965.
  - [18] B. W. Lee, C. Quigg, and H. B. Thacker, "Weak interactions at very high energies: The role of the Higgs-boson mass", *Phys. Rev. D* **16** (Sep, 1977) 1519–1531. doi:10.1103/PhysRevD.16.1519.
  - [19] A. Djouadi, "The Anatomy of Electro-Weak Symmetry Breaking. I: The Higgs boson in the Standard Model", Technical Report hep-ph/0503172, Paris 11. Lab. Phys. Théor., Orsay, Mar, 2005.
  - [20] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti et al., "Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables", arXiv:1101.0593.
  - [21] H. M. Georgi, S. L. Glashow, M. E. Machacek et al., "Higgs Bosons from Two-Gluon Annihilation in Proton-Proton Collisions", *Phys. Rev. Lett.* **40** (Mar, 1978) 692–694. doi:10.1103/PhysRevLett.40.692.
  - [22] H. P. Nilles, "Supersymmetry, supergravity and particle physics", *Physics Reports* **110** (1984), no. 1-2, 1 – 162. doi:10.1016/0370-1573(84)90008-5.
  - [23] H. E. Haber and G. L. Kane, "The search for supersymmetry: Probing physics beyond the standard model", *Physics Reports* **117** (1985), no. 2-4, 75 – 263. doi:10.1016/0370-1573(85)90051-1.
  - [24] S. P. Martin, "A Supersymmetry Primer", Technical Report hep-ph/9709356, Sep, 1997.
  - [25] A. Djouadi, "The Anatomy of Electro-Weak Symmetry Breaking. II: The Higgs bosons in the Minimal Supersymmetric Model", Technical Report hep-ph/0503173, Paris 11. Lab. Phys. Théor., Orsay, Mar, 2005.
  - [26] K. Inoue, A. Kakuto, H. Komatsu et al., "Low Energy Parameters and Particle Masses in a Supersymmetric Grand Unified Model", *Progress of Theoretical Physics* **67** (1982), no. 6, 1889–1898. doi:10.1143/PTP.67.1889.
  - [27] R. A. Flores and M. Sher, "Higgs masses in the standard, multi-Higgs and supersymmetric models", *Annals of Physics* **148** (1983), no. 1, 95 – 134. doi:10.1016/0003-4916(83)90331-7.
  - [28] S. Heinemeyer, W. F. L. Hollik, and G. Weiglein, "The Masses of the Neutral CP-even Higgs Bosons in the MSSM: Accurate Analysis at the Two-Loop Level", *Eur. Phys. J. C* **9** (Dec, 1998) 343–366. 45 p.
  - [29] G. Degrassi, S. Heinemeyer, W. F. L. Hollik et al., "Towards High-Precision

- Predictions for the MSSM Higgs Sector”, *Eur. Phys. J. C* **28** (Dec, 2002) 133–143. 19 p.
- [30] H. E. Haber, “Non-minimal Higgs sectors: the decoupling limit and its phenomenological implications”, Technical Report hep-ph/9501320, Jan, 1995.
- [31] M. S. Carena, S. Heinemeyer, C. E. M. Wagner et al., “Suggestions for Benchmark Scenarios for MSSM Higgs Boson Searches at Hadron Colliders”, *Eur. Phys. J. C* **26** (Feb, 2002) 601–7. 8 p.
- [32] E. L. Berger, C.-W. Chiang, J. Jiang et al., “Higgs boson decay into hadronic jets”, *Phys. Rev. D* **66** (Nov, 2002) 095001. doi:10.1103/PhysRevD.66.095001.
- [33] C. Wyss, “LEP design report, v.3: LEP2”. CERN, Geneva, 1996. Vol. 1-2 publ. in 1983-84.
- [34] G. Abbiendi, “Search for the Standard Model Higgs Boson at LEP”, *Phys. Lett. B* **565** (Mar, 2003) 61–75. 22 p.
- [35] ALEPH, DELPHI, L3, and OPAL Collabs., The LEP Working Group for Higgs Boson Searches, “Search for neutral MSSM Higgs bosons at LEP”, *The European Physical Journal C - Particles and Fields* **47** (Jan, 2006) 547–587. doi:10.1140/epjc/s2006-02569-7.
- [36] ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration, the LEP Higgs Working Group, “Search for Charged Higgs bosons: Preliminary Combined Results Using LEP data Collected at Energies up to 209 GeV. oai:cds.cern.ch:508860”, Technical Report hep-ex/0107031, CERN, Geneva, Jul, 2001.
- [37] Tevatron I Group, “Design report Tevatron 1 Project”, *FERMILAB-DESIGN-1984-01* (1984).
- [38] The Tevatron New Physics, Higgs Working Group, “Combined CDF and D0 Upper Limits on Standard Model Higgs-Boson Production with up to 6.7 fb-1 of Data”, Technical Report arXiv:1007.4587, Jul, 2010.
- [39] M. Carena, S. Heinemeyer, C. Wagner et al., “MSSM Higgs boson searches at the Tevatron and the LHC: Impact of different benchmark scenarios”, *The European Physical Journal C - Particles and Fields* **45** (2006) 797–814. doi:10.1140/epjc/s2005-02470-y.
- [40] The ALEPH, CDF, D0, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the Tevatron Electroweak Working Group, and the SLD electroweak and heavy flavour groups, “Precision Electroweak Measurements and Constraints on the Standard Model”, Technical Report arXiv:1012.2367, Dec, 2010.
- [41] G. L. Bayatian, S. Chatrchyan, G. Hmayakyan et al., “CMS Physics Technical Design Report, Volume II: Physics Performance”, *J. Phys. G* **34** (2006), no. CERN-LHCC-2006-021. CMS-TDR-008-2, 995–1579. 669 p.
- [42] S. Abdullin, S. Banerjee, L. Bellucci et al., “Summary of the CMS potential for the Higgs boson discovery”, *The European Physical Journal C - Particles and Fields* **39** (2005) 41–61. doi:10.1140/epjcd/s2004-02-003-9.
- [43] The CMS Collaboration, “The CMS physics reach for searches at 7 TeV”, Technical Report CMS-NOTE-2010-008, CERN, Geneva, May, 2010.

- [44] CERN Web Communications, "CERN Web Pages". <http://public.web.cern.ch>.
- [45] P. Cortese, C. W. Fabjan, L. Riccati et al., "ALICE physics performance: Technical Design Report". Technical Design Report ALICE. CERN, Geneva, 2005.
- [46] The ATLAS Collaboration, "ATLAS detector and physics performance: Technical Design Report, 1". Technical Design Report ATLAS. CERN, Geneva, 1999.
- [47] The ATLAS Collaboration, "ATLAS detector and physics performance: Technical Design Report, 2". Technical Design Report ATLAS. CERN, Geneva, 1999.
- [48] G. L. Bayatian, S. Chatrchyan, G. Hmayakyan et al., "CMS Physics Technical Design Report Volume I: Detector Performance and Software". Technical Design Report CMS. CERN, Geneva, 2006.
- [49] R. Antunes-Nobrega, A. França-Barbosa, I. Bediaga et al., "LHCb reoptimized detector design and performance: Technical Design Report". Technical Design Report LHCb. CERN, Geneva, 2003.
- [50] V. Berardi, M. G. Catanesi, E. Radicioni et al., "Total cross-section, elastic scattering and diffraction dissociation at the Large Hadron Collider at CERN: TOTEM Technical Design Report". Technical Design Report TOTEM. CERN, Geneva, 2004.
- [51] O. Adriani, L. Bonechi, M. Bongi et al., "LHCf experiment: Technical Design Report". Technical Design Report LHCf. CERN, Geneva, 2006.
- [52] "OpenStreetMap". <http://www.openstreetmap.org>.
- [53] C. Lefevre, "LHC: the guide (English version). Guide du LHC (version anglaise)". CERN-CERN-Brochure-2009-003-Eng, Feb, 2009.
- [54] "LHC Design Report". <http://lhc.web.cern.ch/lhc/LHC-DesignReport.html>.
- [55] M. L. Mangano and G. Altarelli, eds., "1999 CERN Workshop on Standard Model Physics (and more) at the LHC", number CERN-2000-004. CERN, Geneva, 2000.
- [56] "CMS Outreach". <http://cmsinfo.cern.ch>.
- [57] LHC Experiments Committee, "The CMS magnet project Technical Design Report". Technical Design Report CMS. CERN, Geneva, 1997.
- [58] LHC Experiments Committee, "The CMS electromagnetic calorimeter project: Technical Design Report". Technical Design Report CMS. CERN, Geneva, 1997.
- [59] LHC Experiments Committee, "The CMS hadron calorimeter project: Technical Design Report". Technical Design Report CMS. CERN, Geneva, 1997.
- [60] LHC Experiments Committee, "The CMS muon project Technical Design Report". Technical Design Report CMS. CERN, Geneva, 1997.
- [61] R. Adolphi, G. Anagnostou, R. Brauer et al., "The CMS experiment at the CERN LHC", *J. Instrum.* **3** (2008) S08004.
- [62] "CMS Luminosity Public Results".  
<https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults2010>.
- [63] V. Karimäki, "The CMS tracker system project: Technical Design Report". Technical Design Report CMS. CERN, Geneva, 1997.

- [64] P. Azzurri, "Track Reconstruction Performance in CMS", Technical Report CMS-CR-2008-110. CERN-CMS-CR-2008-110, CERN, Geneva, Dec, 2008.
- [65] R. Frühwirth, "Application of Kalman filtering to track and vertex fitting", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **262** (1987), no. 2-3, 444 – 450.  
doi:10.1016/0168-9002(87)90887-4.
- [66] W. Adam, B. Mangano, T. Speer et al., "Track Reconstruction in the CMS tracker", Technical Report CMS-NOTE-2006-041, CERN, Geneva, Dec, 2006.
- [67] R. Frühwirth, W. Waltenberger, and P. Vanlaer, "Adaptive Vertex Fitting", Technical Report CMS-NOTE-2007-008, CERN, Geneva, March, 2007.
- [68] S. Cucciarelli, M. Konecki, D. Kotlinski et al., "Track reconstruction, primary vertex finding and seed generation with the Pixel Detector", Technical Report CMS-NOTE-2006-026, CERN, Geneva, Jan, 2006.
- [69] G. L. Bayatyan, N. Grigorian, V. G. Khachatryan et al., "CMS TriDAS project: Technical Design Report; 1, the trigger systems". Technical Design Report CMS. 2000.
- [70] S. Cittolin, A. Rácz, and P. Sphicas, "CMS trigger and data-acquisition project: Technical Design Report". Technical Design Report CMS. CERN, Geneva, 2002.
- [71] T. Virdee, A. Petrilli, and A. Ball, "CMS High Level Trigger", technical report, CERN, Geneva, Jun, 2007.
- [72] L. Agostino, M. Anderson, L. Apanasevich et al., "The 2009 CMS Trigger Reviews", Technical Report CMS IN-2010/022, CMS Collaboration, Aug, 2010.
- [73] R. Brun and F. Rademakers, "ROOT – An object oriented data analysis framework", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **389** (1997), no. 1-2, 81 – 86. New Computing Techniques in Physics Research V.  
doi:10.1016/S0168-9002(97)00048-X.
- [74] L. Lönnblad, "CLHEP: a project for designing a C<sup>++</sup> class library for high energy physics. oai:cds.cern.ch:259545", *Comput. Phys. Commun.* **84** (Feb, 1994) 307–316. 14 p.
- [75] N. Metropolis and S. Ulam, "The Monte Carlo Method.", *Journal of the American Statistical Association* **44** (1949) 335–341.
- [76] T. Sjöstrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual", *J. High Energy Phys.* **05** (Mar, 2006) 026. 570 p.
- [77] G. Corcella, I. G. Knowles, G. Marchesini et al., "HERWIG 6: an event generator for Hadron Emission Reactions With Interfering Gluons (including supersymmetric processes)", *J. High Energy Phys.* **01** (Nov, 2000) 010. 93 p.
- [78] T. Gleisberg, S. Hoeche, F. Krauss et al., "Event generation with SHERPA 1.1", *J. High Energy Phys.* **02** (Dec, 2008) 007.
- [79] M. L. Mangano, M. Moretti, F. Piccinini et al., "ALPGEN, a generator for hard multiparton processes in hadronic collisions", *J. High Energy Phys.* **07** (Jun, 2002) 001. 35 p.

- [80] J. Alwall, P. Demin, S. de Visscher et al., “MadGraph/MadEvent v4: The New Web Generation”, *J. High Energy Phys.* **09** (Jun, 2007) 028. 38 p.
- [81] S. Frixione, F. Stoeckli, P. Torrielli et al., “The MCaNLO 4.0 Event Generator. oai:cds.cern.ch:1297827”, Technical Report arXiv:1010.0819, Oct, 2010.
- [82] S. Jadach, J. H. Kühn, and Z. Was, “TAUOLA: a library of Monte Carlo programs to simulate decays of polarized  $\tau$  leptons. oai:cds.cern.ch:212328”, *Comput. Phys. Commun.* **64** (Sep, 1990) 275–299. 38 p.
- [83] N. Davidson, G. Nanava, T. Przedzinski et al., “Universal Interface of TAUOLA Technical and Physics Documentation”, Technical Report arXiv:1002.0543. IFJPAN-IV-2009-10, Feb, 2010.
- [84] M. Dobbs and J. B. Hansen, “The HepMC C++ Monte Carlo event record for High Energy Physics”, *Computer Physics Communications* **134** (2001), no. 1, 41 – 46. doi:10.1016/S0010-4655(00)00189-2.
- [85] K Nakamura et al and (Particle Data Group), “Review of Particle Physics”, *Journal of Physics G: Nuclear and Particle Physics* **37** (2010), no. 7A, 075021.
- [86] S. Agostinelli, J. Allison, K. Amako et al., “GEANT4-a simulation toolkit”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** (2003), no. 3, 250 – 303. doi:10.1016/S0168-9002(03)01368-8.
- [87] G. L. Bayatyan, M. Della Negra, L. Foà et al., “CMS computing: Technical Design Report”. Technical Design Report CMS. CERN, Geneva, May, 2005.
- [88] A. Ostaþtchouk, S. Hameed-Khan, S. Schael et al., “The Alignment System of the CMS Tracker”, Technical Report CMS-NOTE-2001-053, CERN, Geneva, Nov, 2001.
- [89] CMS Collaboration, “Alignment of the CMS Silicon Tracker during Commissioning with Cosmic Rays”, *J. Instrum.* **5** (Oct, 2009) T03009. 41 p.
- [90] P. Sauerland, “Kinematic Reconstruction of Tau Leptons and Test for Lepton Universality in Charged Weak Interactions with the CMS Experiment”. PhD thesis, RWTH Aachen University, 2011.
- [91] L. Perchalla, P. Sauerland, and A. Stahl, “Kinematic Reconstruction of Tau Leptons”. CMS Analysis Note CMS AN-2010/256, December, 2010.
- [92] P. Avery, “Applied Fitting Theory VI: Formulars for Kinematic Fitting”, 1998. [www.phys.ufl.edu/~avery/fitting.html](http://www.phys.ufl.edu/~avery/fitting.html).
- [93] K. Prokofiev and T. Speer, “A Kinematic fit and a decay chain reconstruction library”. CMS Internal Note, September, 2004.
- [94] P. Avery, “Applied Fitting Theory I: General Least Squares Theory”, 1991. [www.phys.ufl.edu/~avery/fitting.html](http://www.phys.ufl.edu/~avery/fitting.html).
- [95] The CMS Collaboration, “Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET”, April, 2009.
- [96] The CMS Collaboration, “CMS Strategies for tau reconstruction and identification using particle-flow techniques”, August, 2009.

- [97] G. Bagliesi, S. Dutta, S. Gennai et al., “Tau jet reconstruction and tagging at High Level Trigger and off-line”, Technical Report CMS-NOTE-2006-028, CERN, Geneva, Jan, 2006.
- [98] J. Pumplin, D. R. Stump, J. Huston et al., “New Generation of Parton Distributions with Uncertainties from Global QCD Analysis”, *J. High Energy Phys.* **07** (Jan, 2002) 012. 44 p.
- [99] P. Govoni and C. Mariotti, “Z production via Vector Boson Fusion at LHC”, Technical Report arXiv:1001.4357, Jan, 2010.
- [100] V. Khachatryan, A. M. Sirunyan, A. Tumasyan et al., “Measurements of Inclusive W and Z Cross Sections in pp Collisions at  $\sqrt{s} = 7$  TeV. oai:cds.cern.ch:1313495”, *arXiv:1012.2466* (Dec, 2010).
- [101] The CMS Collaboration, “Measurement of CMS Luminosity”, *CMS-PAS-EWK-10-004* (2010).
- [102] S. van der Meer, “Calibration of the effective beam height in the ISR”, technical report, CERN, Geneva, 1968.
- [103] M. Dittmar, F. Pauss, and D. Zürcher, “Towards a Precise Parton Luminosity Determination at the CERN LHC”, *Phys. Rev. D* **56** (1997), no. hep-ex/9705004, 7284–7290.
- [104] A. Kalinowski and A. Nikitenko, “Measurement of the tau tag efficiency”, Technical Report CMS-NOTE-2006-074, CERN, Geneva, May, 2006.
- [105] P. Bock, J. Carr, S. De Jong et al., “Lower bound for the standard model Higgs boson mass from combining the results of the four LEP experiments”, technical report, CERN, Geneva, Apr, 1998.

