

**Searches for Higgs Bosons decay-
ing to τ leptons at $\sqrt{s} = 7$ TeV**

By

EVAN KLOSE FRIIS

B.S. (University of California at San Diego) 2005

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Physics

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Professor John Conway (Chair)

Professor Robin Erbacher

Professor Mani Tripathi

Committee in Charge

2011

for Alvin A Klose

Abstract

blah blah blah

Acknowledgments

Hooray for everybody.

Table of Contents

1	The Standard Model and Beyond	1
1.1	The Standard Model	1
1.1.1	Quantum Electrodynamics and Gauge Invariance	2
1.1.2	The Weak Interactions	4
1.1.3	Spontaneous Symmetry Breaking	6
1.1.4	The Higgs Mechanism	8
1.1.5	Electroweak Unification	10
1.1.6	Quantum Chromodynamics	14
1.2	Beyond the Standard Model	16
1.2.1	The Hierarchy Problem	17
1.2.2	Supersymmetry	18
1.2.3	The Minimal Supersymmetric Model and τ leptons	18
1.3	Searches for the Higgs boson	18
1.3.1	Higgs boson phenomenology	18
1.3.2	Results from LEP and Tevatron	21
1.4	The physics of the τ lepton	21
2	The Compact Muon Solenoid Experiment	24
3	Tau Identification: The Tau Neural Classifier	25
3.1	Introduction	25
3.2	Geometric Tau Identification Algorithms	26
3.3	Decay Mode Tau Identification: Motivation	26
3.4	The Tau Neural Classifier	27
3.4.1	Decay mode reconstruction	28
3.4.2	Neural network classification	31
3.5	Summary	45
4	Mass Reconstruction: The Secondary Vertex Fit	49
4.1	Existing mass reconstruction algorithms	49
4.2	The Secondary Vertex fit	51
4.3	Parametrization of tau decays	51
4.4	Likelihood for tau decay	53
4.4.1	Likelihood for reconstructed missing transverse momentum	54
4.4.2	Likelihood for tau lepton transverse momentum balance	54
4.4.3	Secondary vertex information	56
4.5	Performance	57
5	Analysis Selections	60

6	Data-Driven Background Estimation	63
6.1	Introduction	63
6.2	The Fake-rate Method	63
6.2.1	Parameterization of fake-rates	63
6.2.2	Measurement of fake-rates	65
6.2.3	The Fake-rate method	65
6.2.4	k-Nearest Neighbor fake-rate calculation	71
6.2.5	Results of Background Estimation	71
7	Systematics	74
8	Results	75
9	Conclusions	76
	Bibliography	76

List of Figures

1.1	Fermi contact interaction diagram	5
1.2	Muon decaying through intermediate gauge boson	5
1.3	QCD Feynman Diagrams	15
1.4	Loop corrections to Higgs mass	17
1.5	Higgstrahlung production diagram at e^+e^- colliders	18
1.6	Gluon fusion Higgs production diagram	19
1.7	Vector boson fusion Higgs production diagram	19
1.8	SM Higgs cross sections at the LHC	20
1.9	SM Higgs branching fractions	21
1.10	Cross sections of interest at hadron colliders	22
3.1	The invariant mass of the visible decay products in hadronic tau decays. The decay mode $\tau^- \rightarrow \pi^- \nu_\tau$ is omitted. The different decay modes have different invariant masses corresponding to the intermediate resonance in the decay.	27
3.2	Invariant mass of the photon pair for reconstructed tau-candidates with two reconstructed photons in the signal region that are matched to generator level $\tau^- \rightarrow \pi^- \pi^0 \nu_\tau$ decays.	29
3.3	Fraction of total τ -candidate transverse momenta carried by the photon for reconstructed taus containing a single photons for two benchmark cases. On the left, the reconstructed tau-candidate is matched to generator level $\tau^- \rightarrow \pi^- \nu_\tau$ decays, for which no photon is expected. On the right, the reconstructed tau-candidate is matched to generator level $\tau^- \rightarrow \pi^- \pi^0 \nu_\tau$ decays and the photon is expected to correspond to a true π^0 meson. The requirement on the P_T fraction of the lowest P_T photon improves the purity of the decay mode reconstruction.	30
3.4	Correlations between reconstructed tau decay mode and true tau decay mode for hadronic tau decays in $Z \rightarrow \tau^+ \tau^-$ events. The correlation when no photon merging or filtering is applied is shown on the right, and the correlation for the algorithm described in section 3.4.1 is on the right. The horizontal and vertical axis are the decay mode indices of the true and reconstructed decay mode, respectively. The decay mode index N_{DM} is defined as $N_{DM} = (N_{\pi^\pm} - 1) \cdot 5 + N_{\pi^0}$. The area of the box in each cell is proportional to the fraction of tau-candidates that were reconstructed with the decay mode indicated on the vertical axis for the true tau decay on the horizontal axis. The performance of a decay mode reconstruction algorithm can be determined by the spread of the reconstructed number of π^0 mesons about the true number (the diagonal entries) determined from the generator level Monte Carlo information. If the reconstruction was perfect, the correlation would be exactly diagonal.	32

3.5	Kinematic dependence of reconstructed decay mode for tau-candidates from $Z \rightarrow \tau^+\tau^-$ (left) and QCD di-jets (right) versus transverse momentum (top) and pseudo-rapidity (bottom). Each curve is the probability for a tau-candidate to be reconstructed with the associated decay mode after the leading pion and decay mode preselection has been applied.	33
3.6	Neural network classification error for training (solid red) and testing (dashed blue) samples at ten epoch intervals over the 500 training epochs for each decay mode neural network. The vertical axis represents the classification error, defined by equation 3.2. N.B. that the choice of hyperbolic tangent for neuron activation functions results in the desired outputs for signal and background to be 1 and -1, respectively. This results in the computed neural network error being larger by a factor of four than the case where the desired outputs are (0, 1). Classifier over-training would be evidenced by divergence of the classification error of the training and testing samples, indicating that the neural net was optimizing about statistical fluctuations in the training sample.	36
3.7	Transverse momentum spectrum of signal and background tau-candidates used in neural net training before (left) and after (right) the application of $P_T - \eta$ dependent weight function. Application of the weights lowers the training significance of tau-candidates in regions of $P_T - \eta$ phase space where either the signal or background samples has an excess of events.	37
3.8	Neural network output distributions for the five reconstructed tau-candidate decay modes used in the TaNC for $Z \rightarrow \tau^+\tau^-$ events (red) and QCD di-jet events (blue).	40
3.9	Performance curves for the five neural networks used by the TaNC for tau-candidates with transverse momentum greater than 20 GeV/c. Each curve represents the signal efficiency (on the horizontal axis) and background misidentification rate (vertical axis) for a scan of the neural network selection requirement for a single neural network. The efficiency (or misidentification rate) for each neural network performance curve is defined with respect to the pre-selected tau-candidates that have the reconstructed decay mode associated with that neural network. Each neural network has a different ability to separate signal and background as each classifier uses different observables as inputs.	42
3.10	Tau Neural Classifier performance curves for tau-candidates with $P_T < 20 \text{ GeV}/c$ (left) and $20 \text{ GeV}/c < P_T < 50 \text{ GeV}/c$ (right). The vertical axis represents the expected fake-rate of QCD jets and the horizontal axis the expected signal efficiency for hadronic tau decays. The performance curve for the low transverse momentum range is worse due to leading pion selection. While both true taus and QCD are removed by this cut, the selection preferentially keeps the QCD tau-candidates with low multiplicities, which increases the number of QCD tau-candidates passing the decay mode selection.	43

3.11	Tau Neural Classifier performance curves for tau-candidates with $20 \text{ GeV}/c < P_T < 50 \text{ GeV}/c$. The figure on the left compares the optimal performance curve determined by the Monte Carlo sampling method (red) to the performance curve obtained by scanning the “transform cut” (blue) defined in equation 3.3 from zero to one. The figure on the right is the same set of cuts (and cut transformation values) applied on an independent sample to remove any biases introduced by the Monte Carlo sampling. The four dashed lines indicate the performance for the four benchmark points.	44
3.12	Transformed TaNC neural network output for tau-candidates with transverse momentum between 20 and 50 GeV/c that pass the pre-selection criteria. The neural network output for each tau-candidate has been transformation according to equation 3.3. The decay mode probabilities ρ_i^{bkg} , ρ_i^{signal} are computed using the entire transverse momentum range of the sample. . . .	45
3.13	Performance curve (red) of the TaNC tau identification for various requirements on the output transformed according to equation 3.3. The horizontal axis is the efficiency for true taus with transverse momentum between 20 and 50 GeV/c to satisfy the tau identification requirements. The vertical axis gives the rate at which QCD di-jets with generator-level transverse momentum between 20 and 50 GeV/c are incorrectly identified as taus. The performance point for the same tau-candidates using the isolation based tau-identification [?] used in many previous CMS analyses is indicated by the black star in the figure. An additional requirement that the signal cone contain one or three charged hadrons (typical in a final physics analysis) has been applied to the isolation based tau-identification to ensure a conservative comparison.	47
3.14	Comparison of the identification efficiency for hadronic tau decays from $Z \rightarrow \tau^+\tau^-$ decays (bottom row) and the misidentification rate for QCD di-jets (top row) versus tau-candidate transverse momentum (left) and pseudo-rapidity (right) for different tau identification algorithms. The efficiency (fake-rate) in a given bin is defined as the quotient of the number of true tau hadronic decays (generator level jets) in that bin that are matched to a reconstructed tau-candidate that passes the identification algorithm divided by the number of true tau hadronic decays (generator level jets) in that bin. In the low transverse momentum region both the number of tau-candidates in the denominator and the algorithm acceptance vary rapidly with respect to P_T for both signal and background; a minimum transverse momentum requirement of 20 GeV/c is applied to the pseudorapidity plots to facilitate interpretation of the plots.	48
4.1	Illustration of the coordinate system used by the SVfit to describe the decays of tau leptons.	53
4.2	Distribution of di-tau invariant mass reconstructed by the SVfit algorithm in simulated Higgs events with $M_A = 130 \text{ GeV}/c^2$. The SVfit algorithm is run in two configurations, with (blue) and without (red) the P_T -balance likelihood term included in the fit.	56

4.3	Normalized distributions of the fraction of total tau decay energy carried by the muon (left) and hadronic constituents (right) in simulated Higgs events with $M_A = 130 \text{ GeV}/c^2$. The distribution is shown before (blue) and after (red) the requirement on the P_T of the visible decay products described in section ??.	57
4.4	Normalized distributions of the fraction of total tau decay energy carried by the muon in simulated $Z \rightarrow \tau^+\tau^-$ (left) and Higgs events with $M_A = 200 \text{ GeV}/c^2$ (right). The distribution is shown before (blue) and after (red) the requirement that the P_T of the muon be greater than $15 \text{ GeV}/c$.	58
4.5	Comparison of SVfit with the Collinear Approximation algorithm	58
4.6	Comparison of SVfit with the visible mass observable	59
6.1	P_T and η dependency of tau ID performance	64
6.2	Distribution of visible mass (left) and “full” invariant mass reconstructed by the SVfit algorithm (right) observed in the like-sign charge control region compared to the background estimate obtained by the fake-rate method.	73

List of Tables

1.1	Resonances and branching ratios of the dominant decay modes of the τ lepton. The decay products listed correspond to a negatively charged τ lepton; the table is identical under charge conjugation.	23
3.1	Decay mode correlation table for the selected dominant decay modes for the naive approach. The percentage in a given row and column indicates the fraction of hadronic tau decays from $Z \rightarrow \tau^+\tau^-$ events that are matched to a generator level decay mode given by the row and are reconstructed with the decay mode given by the column. Entries in the "Other" column are immediately tagged as background.	31
3.2	Decay mode correlation table for the selected dominant decay modes for the merging and filtering approach. The percentage in a given row and column indicates the fraction of hadronic tau decays from $Z \rightarrow \tau^+\tau^-$ events that are matched to a generator level decay mode given by the row and are reconstructed with the decay mode given by the column. Entries in the "Other" column are immediately tagged as background.	31
3.3	Number of events used for neural network training and validation for each selected decay mode.	34
3.4	Input observables used for each of the neural networks implemented by the Tau Neural Classifier. The columns represents the neural networks associated to various decay modes and the rows represent the superset of input observables (see section 3.4.2) used in the neural networks. A dot in a given row and column indicates that the observable in that row is used in the neural network corresponding to that column.	41
5.1	Event selection criteria applied in the muon + tau-jet channel.	61
5.2	Muon and muon + tau-jet "cross-channel" trigger paths utilized to trigger events in the muon + tau-jet channel in different data-taking periods. . . .	62
6.1	Estimate for background contributions in mu + tau channel obtained by weighting events passing all selection criteria listed in table 5.1 except for the requirement for tau-jet candidates to pass the "medium" tight TaNC discriminator and have unit charge by fake-rates measured in QCD multi-jet, QCD muon enriched and W + jets data samples.	72

6.2	Number of events observed in like-sign control region compared to estimate obtained by fake-rate method. The number of observed events as well as the number of background events predicted by the fake-rate method is on good agreement with the Monte Carlo expectation of $XX.X$ events for the sum of $Z \rightarrow \mu^+\mu^-$, $W + \text{jets}$, $t\bar{t} + \text{jets}$ and QCD background contributions in the like-sign control region.	73
-----	---	----

Chapter 1

The Standard Model and Beyond

§1.1 The Standard Model

The Standard Model (SM) is a “theory of almost everything” that describes the interactions of elementary particles. The Standard Model is a *quantum field theory*, first appearing in its modern form in the middle of the 20th century. The model is the synthesis of the independent theories of electromagnetism, and the weak and strong nuclear forces. Each of these theories was used to describe different phenomena, which each have extremely different strengths and act at different scales. The interaction of light and matter is described by Quantum Electrodynamics (QED), a relativistic field extension of the theory of electromagnetism. The physics of radioactivity and nuclear decay was described by the Fermi theory of weak interactions and the forces that strong nuclear force binds the nuclei of atoms was described by Yukawa. An overview of these theories will be presented in this chapter.

The feature that united the disparate theories into the Standard Model was the application of the principle of *local gauge invariance*. The principle of gauge invariance first found success in QED, which predicted electromagnetic phenomenon with astounding accuracy. Local gauge invariance is now believed to a fundamental feature of nature that underpins all theories of elementary particles. Furthermore, the development of the complete Standard Model as it is known today was precipitated by Goldstones’s work on spontaneous symmetry breaking [1, 2], which produces an effective Lagrangian with additional massless “Goldstone” bosons. Higgs (and others) [3, 4, 5] developed these ideas into what is ultimately called the “Higgs Mechanism,” which uses a combination of new fields with broken symmetry to give mass to the Goldstone bosons.

In the 1960s, Glashow [6], Weinberg [7], and Salam [8] developed the above ideas into the electroweak model, which unified QED with the weak force using intermediate weak bosons in a gauge theory whose symmetry was spontaneously broken using the Higgs mechanism. This unified theory has been incredibly experimentally successful and is the foundation of modern particle theory.

§1.1.1 Quantum Electrodynamics and Gauge Invariance

The theory of QED is a modern extension of Maxwell’s theory of electromagnetism, describing the interaction of matter with light. The development of QED is a result of efforts to develop a quantum mechanical formulation of electromagnetism compatible with the theory of Special Relativity. QED is a *gauge* theory, which means that the physical observables are invariant under local gauge transformations. Requiring local gauge invariance gives rise to a “gauge” field, which can be interpreted as particles that are exchanged during an interaction.

In the following, we first describe the Dirac equation for a free electron, which is the relativistic extension of the Schroedinger equation for spin 1/2 particles. We then show that requiring the corresponding Lagrangian of the free charged particle to be invariant under local gauge transformations creates an effective gauge boson field. This “gauge field” creates terms in the Lagrangian that represent interactions between the particles.

The Dirac equation is the equation of motion of a free spin 1/2 particle of mass m and is derived from the energy–momentum relationship of relativity

$$p^\mu p_\mu - m^2 c^2 = 0 \tag{1.1}$$

. Dirac sought to express this relationship in the framework of quantum mechanics by applying the transformation

$$p_\mu \rightarrow i\hbar\partial_\mu \tag{1.2}$$

to equation Equation 1.1, but with the requirement that the resulting equation be first order in time.¹ To achieve this, Dirac factorized Equation 1.1 into

$$(\gamma^\kappa p_\kappa + mc)(\gamma^\mu p_\mu - mc) = 0, \tag{1.3}$$

¹A detailed discussion of this topic is available in [9].

where γ^μ is a set of four 4×4 matrices referred to as the Dirac matrices. The equation of motion is obtained by choosing either term (they are equivalent) from the left hand side of Equation 1.4 and making the substitution in Equation 1.2.

$$i\hbar\gamma^\mu\partial_\mu\psi - mc\psi = 0 \quad (1.4)$$

The solutions, ψ , to the Dirac equation are called “Dirac spinors,” and represent the quantum mechanical state of spin 1/2 particles.

The Lagrangian corresponding to the Dirac equation (1.4) is

$$\mathcal{L} = \bar{\psi}(i\hbar c\gamma^\mu\partial_\mu - mc^2)\psi, \quad (1.5)$$

where ψ is the spinor field of the particle in question, \hbar is Planck’s constant, c the speed of light, and γ^μ are the Dirac matrices. As $\bar{\psi}$ is the Hermitian conjugate of ψ , the Lagrangian is invariant under the global gauge transformation

$$\psi' \rightarrow e^{i\theta}\psi \quad (1.6)$$

. The Lagrangian is invariant under *local* gauge translations if θ can be defined differently at each point in space, i.e. if $\theta = \theta(x)$ in equation 1.6. However, as the derivative operator ∂_μ in equation 1.5 does not commute with $\theta(x)$, the Lagrangian must be modified to satisfy local gauge invariance. This modification is accomplished with the use of a “gauge covariant derivative.” By making the replacement

$$\partial_\mu \rightarrow D_\mu = \partial_\mu - \frac{ie}{\hbar}A^\mu \quad (1.7)$$

in equation 1.5, where **don’t think this is right** $A^\mu = \partial^\mu\theta(x)$ and e is the electric charge, the Lagrangian becomes locally gauge invariant:

$$\mathcal{L} = \bar{\psi}(i\hbar c\gamma^\mu D_\mu - mc^2)\psi \quad (1.8)$$

The difference between the locally (1.8) and the globally (1.5) gauge invariant Lagrangians is then

$$\mathcal{L}_{int} = \frac{e}{\hbar}\bar{\psi}\gamma^\mu\psi A_\mu \quad (1.9)$$

. This term can be interpreted as the coupling between the particle and the gauge boson (force carrier) fields. The coupling is proportional to the constant e , which is associated with the electric charge. This is consistent with the experimental observation that particles with zero electric charge do not interact electromagnetically with each other. In this interpretation, the electromagnetic force between two charged particles is caused by the exchange of gauge bosons (photons). The existence of this “minimal coupling” is *required* if

the Lagrangian is to satisfy local gauge invariance. The addition of a term with the gauge Field Strength Tensor to represent the kinetic term of the gauge (photon) field yields the QED Lagrangian:

$$\mathcal{L}_{QED} = \bar{\psi}(i\hbar c\gamma^\mu D_\mu - mc^2)\psi - \frac{1}{4\mu_0}F_{\mu\nu}F^{\mu\nu} \quad (1.10)$$

The gauge symmetry group of QED is $U(1)$, the unitary group of degree 1. This symmetry can be visualized as a rotation of a two-dimensional unit vector. (The application of the gauge transformation $e^{i\theta}$ rotates a number in the complex plane.) In gauge theory that the symmetry group defines the behavior of the gauge bosons and thus the interactions of the theory.

§1.1.2 The Weak Interactions

The theory of Weak Interactions was created to describe the physics of radioactive decay. The first formulation of the theory was done by Fermi [?] to explain the phenomenon of the β decay of neutrons. The initial theory was a four-fermion “contact” theory. In a contact theory, all four fermions come involved in the β -decay are connected at a single vertex. The Fermi theory Hamiltonian for the β -decay of a proton is then [10]

$$H = \frac{G_\beta}{\sqrt{2}} [\bar{\psi}_p \gamma_\mu (1 - g_A \gamma_5) \psi_n] [\bar{\psi}_e \gamma^\mu (1 - \gamma_5) \psi_\nu] + h.c., \quad (1.11)$$

where G_β is the Fermi constant and g_A is the relative fraction of the interaction with axially Lorentz structure. The value of g_A was determined experimentally to be 1.26. One of the most notable things discovered about the weak force is that weak interactions violate parity; that is, the physics of the interaction change (or become disallowed) under inversion of the spatial coordinates. This is evidenced by the $(1 - \gamma_5)$ term in Equation 1.11. This term is the “helicity operator”; the left and right “handed” helicity states are eigenstates of this term.

$$h = (1 - \gamma_5)/2$$

$$h\psi_R = \frac{1}{2}\psi_R$$

$$h\psi_L = -\frac{1}{2}\psi_L$$

It is observed that only left-handed neutrinos (or right-handed anti-neutrinos) participate in the weak interaction.

FixMe: *check handedness is correct*

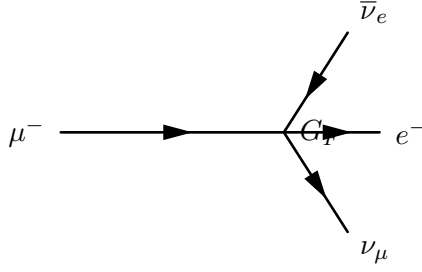


Figure 1.1: Feynman diagram of muon decay in Fermi contact interaction theory.

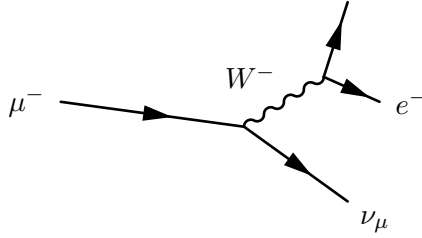


Figure 1.2: Feynman diagram of muon decay proceeding through an intermediate gauge boson W^- .

The Fermi interaction can describe both nuclear β decay ($p \rightarrow n + e^+ + \bar{\nu}_e$) as well as the decay of a muon into an electron ($\mu \rightarrow \nu_\mu + e + \bar{\nu}_e$, Figure 1.1). Furthermore, the coupling constant G is found to be a *universal* constant in weak interactions, in that it is the same for interactions regardless of the particle species participating in the interaction. That is, $G_\mu = G_e = G_F$. Using an Hamiltonian analogous to Equation 1.11 for muon decay, the decay amplitude M is found to be

$$M = \frac{G_F}{\sqrt{2}} \left[\bar{u}_{\nu_\mu} \gamma_\rho \frac{1 - \gamma_5}{2} u_\mu \right] \left[\bar{u}_{\nu_e} \gamma_\rho \frac{1 - \gamma_5}{2} u_e \right]. \quad (1.12)$$

However, the contact interaction form of Fermi's theory is not complete. When applied to scattering processes, the interaction violates unitarity: the calculated cross section grows with the center of mass energy, so that for some energy the probability for an interaction is greater than one. Furthermore, the techniques successfully used to “renormalize”² QED fail when applied to the Fermi interaction.

The first attempt to solve the problems with the Fermi theory was made by introducing an intermediate weak boson [6]. The contact interaction is replaced by a massive propagator, the W^\pm bosons. The decay of a muon to an electron and two neutrinos then proceeds as

²Renormalization of quantum field theories is a broad topic beyond the scope of this thesis. Briefly, the process involves “absorbing” infinite divergences that occur in higher-order interactions into physical observables [9].

pictured in Figure 1.2 with an amplitude given by [10]

$$M = - \left[\frac{g}{\sqrt{2}} \bar{u}_{\nu_\mu} \gamma_\rho \frac{1 - \gamma_5}{2} u_\mu \right] \frac{-g^{\rho\sigma} + \frac{q^\rho q^\sigma}{M_W^2}}{q^2 - M_W^2} \left[\frac{g}{\sqrt{2}} \bar{u}_{\nu_e} \gamma_\rho \frac{1 - \gamma_5}{2} u_e \right] \quad (1.13)$$

The presence of the large mass of gauge boson in the denominator of the central term of Equation 1.13 is the reason why the contact interaction originally formulated by Fermi is effectively describes of low-energy weak phenomenon. When the momentum transfer q in the interaction is small compared to M_W , the effect of the propagator is an effective constant. In the low energy limit, the full propagator in equation 1.13 is effectively equivalent to the Fermi contact interaction in 1.12 as

$$\lim_{q/M_W \rightarrow 0} \frac{g^2}{8(q^2 - M_W^2)} = \frac{G_F}{\sqrt{2}} \quad (1.14)$$

Unfortunately, the weak boson exchange model did not solve the problems of unitarity and renormalizability in the weak interaction. However, the form of the boson-exchange propagator in Equation 1.14 suggests the observed “weakness” of the weak interactions is an artifact of the presence of the massive propagator (M_W) and that the fundamental scale of the interaction g is the same order of magnitude as that of QED, $g \approx e$. This observation lead to the unification of the electromagnetic and weak forces, which we describe in the next sections.

§1.1.3 Spontaneous Symmetry Breaking

In the early 1960s Glashow, Weinberg, and Salam published a series of papers describing how the electromagnetic and weak forces could be unified into a common “electroweak” force. The fact that at low energy the electromagnetic and weak forces appear to be separate phenomena at low energy is due to the fact that the symmetry of the electroweak gauge group is “spontaneously broken.” Modern field theories (both the Standard Model and beyond) are predicated on the idea that the all interactions are part of a single, unified symmetry group and the differences between various scales (electromagnetic, weak, etc.) at lower energies are due to the unified symmetry being spontaneously broken.

A symmetry of a Lagrangian is spontaneously broken when the ground state, or vacuum, is at a value which about which the Lagrangian is not symmetric. In quantum field theories, a particle is interpreted as quantized fluctuations of its corresponding field about some constant (vacuum) ground state. The “effective” Lagrangian that we observe in the

(low energy) laboratory would be the expansion of the Lagrangian about this stable point. The effective Lagrangian no longer obeys the original symmetry, which has been “broken”. We give a brief example of the phenomenological effects of a spontaneously broken symmetry in a toy model, following the treatment in [10].

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi_1\partial^\mu\phi_1 + \frac{1}{2}\partial_\mu\phi_2\partial^\mu\phi_2 - V(\phi_1^2 + \phi_2^2) \quad (1.15)$$

The toy Lagrangian in Equation 1.15 has a global $U(1)$ ³ symmetry and consists of two real-valued fields, ϕ_1 and ϕ_2 . The particle mass spectra of the theory is given by expanding the field potential $V(\phi_1, \phi_2)$ about its minimum, $(\phi_1^{min}, \phi_2^{min})$. The first three terms in the series are found by

$$V(\phi_1, \phi_2) = V(\phi_1^{min}, \phi_2^{min}) + \sum_{a=1,2} \left(\frac{\partial V}{\partial \phi_a} \right)_0 (\phi_a - \phi_a^{min}) \quad (1.16)$$

$$+ \frac{1}{2} \sum_{a,b=1,2} \left(\frac{\partial^2 V}{\partial \phi_a \partial \phi_b} \right)_0 (\phi_a - \phi_a^{min})(\phi_b - \phi_b^{min}) + \dots \quad (1.17)$$

Since at the minimum the partial derivative of V is zero with respect to all fields, the second term in equation 1.17 is zero. The third term determines the masses of the particles in the theory. Since a mass term for a particle corresponding to a field ϕ_n in the Lagrangian appears as $\frac{1}{2}m^2\phi_n\phi_n$, we can identify

$$\left(\frac{\partial^2 V}{\partial \phi_a \partial \phi_b} \right)_{\phi^{min}} \quad (1.18)$$

as the a th row and b th column in the “mass matrix”. Off diagonal terms in this matrix indicate mixing terms between the fields. By diagonalizing the matrix, the combinations of fields which correspond to the physical particles (the “mass eigenstates”) are found. The m^2 of each particle is then the corresponding entry in the diagonal of the mass matrix.

The particle spectra of the model depends heavily on the form of the potential. An illustrative form (that is renormalizable and bounded from below) of a possible configuration for the potential V in Equation 1.15 is

$$V(\phi_1^2\phi_2^2) = \frac{m^2}{2}(\phi_1^2 + \phi_2^2) + \frac{\lambda}{4}(\phi_1^2 + \phi_2^2)^2. \quad (1.19)$$

If the parameters m^2 and λ are both positive, then the minimum of V is at the origin

³Technically, the symmetric transformation is

$$\begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \rightarrow \begin{pmatrix} \phi'_1 \\ \phi'_2 \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix},$$

which is $O(2)$. However, this transformation is equivalent to $U(1)$, as the two real fields ϕ_1 and ϕ_2 can be seen to correspond to the real and imaginary parts of a complex field ϕ that does transform according to $U(1)$.

($\phi_1 = \phi_2 = 0$). In this case, the mass matrix term in Equation 1.17 takes the form $\left(\frac{\partial^2 V}{\partial \phi_a \partial \phi_b}\right)_0 = \frac{m^2}{2} \delta_{ab}$, where δ_{ab} is the Kronecker delta function. Therefore the mass matrix is already diagonalized, and the ϕ_1 and ϕ_2 both correspond to particles with mass m . If the m^2 parameter in Equation 1.19 is negative, the spectrum is dramatically different. After making the replacement $m^2 = -\mu^2$ ($\mu^2 > 0$), the extrema of V are no longer unique. The requirement of $\frac{\partial V}{\partial \phi_i} = 0$ for all i is satisfied in two cases:

$$(\phi_1^{min}, \phi_2^{min}) = (0, 0) \quad (1.20)$$

$$(\phi_1^{min})^2 + (\phi_2^{min})^2 = \frac{\mu^2}{\lambda} = \nu^2. \quad (1.21)$$

If the vacuum state is defined at the point in Equation 1.20, the symmetry is unbroken and the mass spectra is unchanged. However, the system is unstable at this point, as it is a local maximum. The true global minimum is defined as the set of points which satisfy Equation 1.21, which form a continuous circle in $\phi_1 - \phi_2$ space (and is therefore infinitely degenerate). We can choose any point on the circle as the vacuum expectation value (VEV). If the point $(\phi_1^{min} = \nu, \phi_2^{min} = 0)$ ⁴ is chosen, evaluating Equation 1.18 yields the mass matrix

$$\begin{pmatrix} v^2 & 0 \\ 0 & 0 \end{pmatrix}.$$

Breaking the symmetry has changed the mass spectrum of the physical particles in the model. There is now a massive particle with $m = v$ and a massless particle. This massless particle is called the “Goldstone boson.” Goldstone found [1] that a massless particle appears for each generator in the symmetry group that is broken.

FixMe:

check matrix

FixMe:

check ref

§1.1.4 The Higgs Mechanism

As in section 1.1.1, extending the gauge symmetry requirement to be *locally* invariant creates interesting consequences for models that have spontaneously broken symmetry. This gives rise to the “Higgs Mechanism,” which we overview here. For simplicity we will again consider a model with $U(1)$ symmetry. The model is identical to the one presented in section 1.1.3, with two exceptions. First, we express the two real fields ϕ_1 and ϕ_2 as a single complex–

⁴The point chosen for the VEV here is not arbitrary. One can chose any point thats satisfies Equation 1.21 as the VEV. However, after the mass matrix is diagonalized, there will always be one physical field with a VEV= ν and one with a VEV= 0. Therefore the physical content of the theory does not depend on the choice of VEV.

valued field ϕ . Second, the model is required to be locally $U(1)$ invariant, and so uses the gauge-covariant derivatives, minimal coupling to the gauge field, and contains the kinetic term for the gauge field, as discussed in section 1.1.1. The unbroken Lagrangian is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + (D_\mu\phi^*)(D^\mu\phi) - V(\phi^*\phi) \quad (1.22)$$

$$V(\phi^*\phi) = -\mu^2\phi^*\phi + \lambda(\phi^*\phi)^2, \quad (1.23)$$

where $F_{\mu\nu}$ is related to the gauge field by $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. The Lagrangian is invariant under the local $U(1)$ gauge transformation

$$\begin{aligned} \phi \rightarrow \phi' &= e^{-i\alpha(x)}\phi \\ A_\mu \rightarrow A'_\mu &= A_\mu - \frac{1}{2}\partial_\mu\alpha(x) \end{aligned} \quad (1.24)$$

The potential is minimized when $\phi^*\phi = \frac{\mu^2}{2\lambda}$. To simplify the algebra, we can reparameterize the field into a real part $\eta(x)$ defined about ν , the minimum of V , and a complex phase $\theta(x)/\nu$

$$\phi(x) = \frac{1}{\sqrt{2}}(\nu + \eta(x))e^{i\theta(x)/\nu} \quad (1.25)$$

If the gauge transform is chosen to be $\alpha(x) = \theta(x)/\nu$, the fields are defined in the so-called “unitary gauge” and have the special forms

$$\begin{aligned} \phi(x) \rightarrow \phi'(x) &= \frac{1}{\sqrt{2}}(\nu + \eta(x)) \\ A_\mu(x) \rightarrow B_\mu(x) &= A_\mu(x) - \frac{1}{e\nu}\partial_\mu\theta(x) \end{aligned} \quad (1.26)$$

The kinetic term of the gauge field $F_{\mu\nu}$ is invariant under this transformation. If the gauge transformations of Equation 1.26 are substituted into the Lagrangian (1.22) the effective Lagrangian at the minimum of V is

$$\mathcal{L} = \frac{1}{2}\partial_\mu\eta\partial^\mu\eta - \mu^2\eta^2 - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}(e\nu)^2B_\mu B^\mu + \frac{1}{2}e^2B_\mu B^\mu\eta(\eta + 2\nu) - \lambda\nu\eta^3 - \frac{\lambda}{4}\eta^4 \quad (1.27)$$

The breaking of the original symmetry has dramatically altered the physical consequences of the model. In its unbroken form, the model described by Equation 1.22 would produce two real massive particles and one massless gauge boson mandated by local gauge invariance. After symmetry breaking, the effective Lagrangian in Equation 1.27 contains a massive scalar η with $m = \sqrt{2\mu^2}$ and a *massive* gauge boson B_μ with mass $m = \sqrt{2}e\nu$. By acquiring a mass, the gauge boson B_μ has acquired the degree of freedom (as it can now be longitudinally polarized) previously associated to the second degree of freedom in the scalar ϕ field. This phenomenon, known as the “Higgs Mechanism,” is a simplified version

of the techniques successfully used to unify the electromagnetic and weak forces that we will discuss in the next section.

§1.1.5 Electroweak Unification

In the 1960s, the ideas of local gauge invariance in field theories, spontaneous symmetry breaking, and the Higgs mechanism were combined by Glashow [?], Weinberg [?] and Salam [?] to form the unified theory of electroweak interactions, the nucleus of the Standard Model. This model successfully unified the electromagnetic and weak interactions into a unified theory with a larger symmetry group. The reason for the empirically observed difference in scales between two interactions is due to the larger, unified symmetry group being broken. This broken symmetry creates heavy gauge bosons via the Higgs mechanism, whose large mass decreases the strength of “weak” interactions at low energy, as discussed in Section 1.1.2. The model successfully predicted the existence and approximate masses of the weak force carriers, the W^\pm and Z bosons. These particles were later observed [?] with the predicted masses at the SPS experiment [?].

To provide a simple introduction to the mechanisms of the model, we will start with a model that includes only one family of leptons, the electron e and its associated neutrino ν_e . Following once again the treatment of [10], we describe the representation of the e and ν_e in, and the choice of, the symmetry group of the model. We then construct the gauge invariant Lagrangian, with spontaneously broken symmetry, and examine the particle content of the resulting model.

The form of the charged current $J_\mu(x) = \bar{u}_{\nu_e} \gamma_\mu \frac{1-\gamma_5}{2} u_e$ in the weak interaction amplitudes (1.12 indicates that the left-handed electron and neutrino (remember that the $(1 - \gamma_5)$ kills any right-handed spinors) can be combined into a doublet L of $SU(2)$.

$$L = \frac{1 - \gamma_5}{2} \begin{pmatrix} \nu_e \\ e^- \end{pmatrix} = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L \quad (1.28)$$

Defining operators that operate on “weak isospin,”

$$\tau^+ = \frac{\tau^1 + i\tau^2}{2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad (1.29)$$

$$\tau^- = \frac{\tau^1 - i\tau^2}{2} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad (1.30)$$

where the τ^i are the Pauli matrices, the weak currents J_μ^\pm can be written by combining equations 1.28–1.30

$$J_\mu^\pm = \bar{L}\gamma_\mu\tau^\pm L. \quad (1.31)$$

Since τ^1 , τ^2 , and τ^3 are the generators of the $SU(2)$ group, we can complete the group by adding a neutral current to the charged currents of Equation ???. The τ^3 generator is diagonal, so the charge of the current is zero and no mixing of the fields occurs:

$$\begin{aligned} J_\mu^3 &= \bar{L}\gamma_\mu\frac{\tau^3}{2}L \\ &= \bar{L}\gamma_\mu\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}L \\ &= \frac{1}{2}\bar{\nu}_e\gamma_\mu\nu_e - \frac{1}{2}\bar{e}_L\gamma_\mu e_L \end{aligned} \quad (1.32)$$

Naively one might hope that the neutral current of Equation 1.32 would correspond to the electromagnetic (photon) current of QED. However, this is impossible for two reasons. First, the right-handed component e_R does not appear in the current, so this interaction violates parity, a known symmetry of the electromagnetic interactions. Second, the current couples to neutrinos, which have no electric charge. Therefore, the “charge” corresponding to the $SU(2)$ gauge symmetry generators $T^i = \int J_0^i(x)d^3x$ cannot be that of the QED, and the gauge group must be enlarged to include an additional $U(1)$ symmetry. The generator of the new symmetry must commute with the generators of the $SU(2)_L$ group. The symmetry cannot be directly extended with $U(1)_{em}$ as the electromagnetic charge $Q = \int (e_L^\dagger e_L + e_R^\dagger e_R)d^3x$ does not commute with T^i . The solution is to introduce the “weak hypercharge” $\frac{Y}{2} = Q - T^3$, which commutes the generators of $SU(2)_L$. Thus the symmetry group of the electroweak model is $SU(2)_L \times U(1)_Y$.

The $SU(2)_L \times U(1)_Y$ gauge invariant Lagrangian is written

$$\begin{aligned}\mathcal{L} = & \bar{L}i\gamma^\mu(\partial_\mu - ig\frac{\vec{\tau}}{2} \cdot \vec{A}_\mu + \frac{i}{2}g'B_\mu)L \\ & + \bar{R}i\gamma^\mu(\partial_\mu + \frac{i}{2}g'B_\mu)R \\ & + -\frac{1}{4}F_{\mu\nu}^i F^{i\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}.\end{aligned}\quad (1.33)$$

As R is a singlet in $SU(2)$, it does not couple to the $SU(2)$ gauge bosons A_μ^i . For this Lagrangian to correspond to empirical observations at low energy, the $SU(2)_L \times U(1)_Y$ must be broken. As $U(1)_{em}$ symmetry is observed to be good symmetry at all scales the broken Lagrangian must be invariant under $U(1)_{em}$.

To accomplish the symmetry breaking, we introduce a new $SU(2)$ doublet of complex Higgs fields ϕ that have hypercharge $Y = 1$, and contribute \mathcal{L}_S to the Lagrangian:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \quad (1.34)$$

$$\mathcal{L}_S = (D_\mu\phi)^\dagger(D^\mu\phi) - V(\phi^\dagger\phi), \quad (1.35)$$

where D_μ is the gauge covariant derivative containing couplings to both the $SU(2)_L$ and $U(1)_Y$ gauge fields, and V has a form analogous to V in Equation 1.23. At this point we also add $SU(2)_L \times U(1)_Y$ invariant ‘‘Yukawa’’ terms

$$\mathcal{L}_Y = -G_e(\bar{L}\phi R + \bar{R}\phi^\dagger L) + h.c.) \quad (1.36)$$

to the Lagrangian which couple the fermions (L and R) to the Higgs field. After symmetry breaking these terms will allow the fermions to acquire masses. By choosing the m^2 and λ parameters of V appropriately, the new ϕ field acquires a non-zero VEV and the symmetry is spontaneously broken.

At the minimum of V , the Higgs field satisfies $\phi^\dagger\phi = \frac{\nu^2}{2}$ and the Higgs fields has a VEV of

$$\phi_{min} = \begin{pmatrix} 0 \\ \nu/\sqrt{2} \end{pmatrix}. \quad (1.37)$$

The new symmetry of the model can be confirmed by looking at the action of different symmetry generators on the VEV. If the generator acting on the vacuum state has a non-zero value, then the corresponding symmetry is broken. It can then be seen that the generators of $SU(2)_L \times U(1)_Y$, $T^{+, -}$,³ and Y are all broken. The vacuum *is* invariant under

Q , the generator of $U(1)_{em}$

$$Q\phi_{min} = (T^3 + \frac{Y}{2}) \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix} = 0.$$

The gauge boson content of the electroweak interaction is obtained by parameterizing the Higgs field in the magnitude–phase notation of Equation 1.25 and using the unitary gauge (see Section 1.1.4), where the gauge transformation is chosen so Higgs field is real. The Higgs scalar doublet is then

$$\phi' = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}(\nu + H(x)) \end{pmatrix} = \frac{1}{\sqrt{2}}(\nu + H(x))\chi. \quad (1.38)$$

In the unit The mass spectrum of the gauge bosons of the electroweak interaction (the photon, W , and Z) is determined by the interaction of the gauge field terms in the covariant derivative with the non-zero vacuum expectation value of the scalar Higgs field ϕ

$$(D_\mu\phi)' = (\partial_\mu - ig\frac{\vec{\tau}}{2} \cdot \vec{A}'_\mu) - \frac{i}{2}g'B'_\mu \frac{1}{\sqrt{2}}(\nu + H)\chi.$$

The terms in the expansion of the kinetic term of the Higgs field that are quadratic in ν^2 and a gauge boson field give the mass associated to that boson, and can be written as

$$\mathcal{L}_{mass} = \frac{\nu^2}{8}(g^2 A'^1_\mu A'^1{}^\mu + g^2 A'^2_\mu A'^2{}^\mu + (gA'^3_\mu - g'B'_\mu)^2). \quad (1.39)$$

The A'^1_μ and A'^2_μ fields can be combined such that the first two terms in Equation 1.39 into the mass term of the charged W^\pm boson

$$W^\pm_\mu = \frac{A'^1_\mu \mp iA'^2_\mu}{2}, \quad (1.40)$$

which then has mass $M_W = \frac{1}{2}g\nu$. The third term in Equation 1.39 can be written in matrix form and then diagonalized into mass eigenstates

$$\begin{aligned} & \frac{\nu^2}{8} \begin{pmatrix} A'^3_\mu & B'_\mu \end{pmatrix} \begin{pmatrix} g^2 & -gg' \\ -gg' & g'^2 \end{pmatrix} \begin{pmatrix} A'^3{}^\mu \\ B'^\mu \end{pmatrix} \\ \rightarrow & \frac{\nu^2}{8} \begin{pmatrix} Z_\mu & A_\mu \end{pmatrix} \begin{pmatrix} g^2 + g'^2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} Z^\mu \\ A^\mu \end{pmatrix}, \end{aligned} \quad (1.41)$$

giving a massive Z boson with $M_Z = \frac{\nu}{2}\sqrt{g^2 + g'^2}$ and a massless photon. The mass of the Z is related to the mass of the W^\pm by the

$$M_Z \equiv \frac{M_W}{\cos\theta_W}, \quad (1.42)$$

where θ_W is the “Weinberg angle,” which must be determined from experiment. As the Fermi contact interaction of Section 1.1.2 is an effective theory of the weak sector, the value

of G_F obtained from β and muon decay experiments give clues to the masses of the W and Z .

$$M_W = \frac{1}{2} \left(\frac{e^2}{\sqrt{2}G_F} \right)^{(1/2)} \frac{1}{\sin \theta_W} \approx \frac{38 \text{ GeV}}{\sin \theta_W} > 37 \text{ GeV} \quad (1.43)$$

$$M_Z \approx \frac{76 \text{ GeV}}{\sin 2\theta_W} > 76 \text{ GeV}. \quad (1.44)$$

The discovery of the W [11, 12] and Z [13, 14] at the CERN SPS was a huge triumph for the electroweak model.

The model that is presented in this section assumes only one species of leptons, the electron and its associated neutrino. The electroweak model is trivially extended [10] to include the other species (μ , τ) of leptons and the three families of quarks. The masses of the fermions are determined by the Yukawa terms in Equation 1.36. Each particle species has a Yukawa term relating the Higgs VEV to its mass that is not constrained by the theory, and must be determined by experiment.

§1.1.6 Quantum Chromodynamics

After electroweak unification, the Standard Model is completed by the theory of Quantum Chromodynamics (QCD), which describes the interactions between quarks and gluons. QCD is a broad field and only a brief introduction to its motivations and the phenomenology relevant to the analysis presented in this thesis is contained in this section. The existence of quarks as the composite was first proposed by Gell–Man and Zweig to explain the spectroscopy of hadrons. QCD is an $SU(3)$ non–Abelian gauge theory which is invariant under *color* transformations. Color is the charge of QCD and comes in three types: red, green and blue. The gauge boson that carries the force of QCD is called the gluon, which is massless as the $SU(3)_c$ color symmetry is unbroken.

There are three marked differences between the photon of QED and the gluon of QCD. First, the gluon itself carries a color charge, where the electron is neutral. This has the consequence that a gluon can couple to other gluons. Secondly, it is found that no colored object exists in nature. The corollary of this is that it is believed to be impossible for a single quark or gluon to be observed. The mechanism that gives rise to effect is called “color confinement.” The strength of the force between two interacting colored objects increases with distance. If two colored objects in a hadron are pulled apart, the energy required to

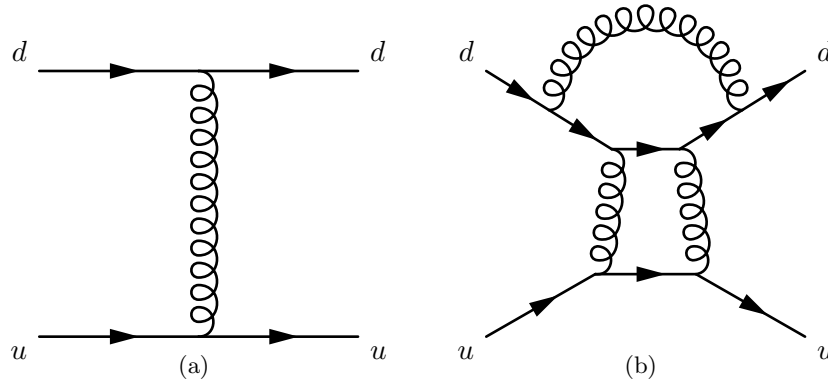


Figure 1.3: Feynman diagrams of a first-order (a) QCD interaction and a multi-loop (b) QCD interaction that have the same initial and final states. Each internal gluon propagator contributes a factor of g_s , the strong coupling constant, to the amplitude. Since $g_s > 1$, multi-loop diagrams have a larger contribution than simpler diagrams.

separate them will eventually be large enough to produce new colored objects, resulting in two (or more) color-less hadrons. Finally, at low energy, QCD is non-perturbative. What this means in practice is that when computing an amplitude from a QCD Feynman diagram, additional gluon interactions contribute a value greater than one. The dominance of multi-loop diagrams is illustrated in Figure 1.1.6. Thus higher order diagrams with many internal loops cannot be ignored in QCD. In practice what is done is to “factorize” QCD interaction amplitudes into a perturbative (high-energy) part and a non-perturbative part. The perturbative portion is calculable using the Feynman calculus; the non-perturbative must be estimated from parameterization functions that are experimentally measured.

The practical consequence of color confinement to a physicist at a high-energy particle physics experiment is the production of quark and gluon “jets,” which are high multiplicity sprays of particles observed in the detector. In a proton-proton collision, quarks and gluons can be knocked off the incident protons. These quarks and gluons immediately “hadronize,” surrounding themselves with additional hadrons, the majority of which are charged and neutral pions. Heavier quarks, such as the charm, beauty, and top quarks undergo a flavor-changing weak decays, which can give rise to structure (leptons, sub-jets) within the jet. Furthermore, due to the relative strength of the strong interaction compared that of the electroweak, collision events involving only strong interactions are produced at rates many orders of magnitudes larger than that of electroweak interactions. This makes life difficult for physicists studying the electroweak force at hadron colliders. Sections ??,

and Chapters 3 and ?? will discuss the techniques used to identify and remove QCD events from the data at different stages of the analysis.

§1.2 Beyond the Standard Model

The Standard model is one of the most successful theories of the natural world ever created. The predictions of the SM have been tested to many orders of magnitude and no experiment to date⁵ has found a result statistically incompatible with the Standard Model. However, there is a general consensus in the physics community that the Standard Model is not complete. It is believed that is only an effective theory that is valid below some energy Λ . Above this energy, there must exist some other “new physics,” which unifies the forces of the Standard Model and correctly describes the natural world at all scales, while maintaining equivalence to the Standard Model at low energy. This concept is analogous to the relationship between the effective Fermi contact theory of Section 1.1.2 and the unified electroweak theory of Section 1.1.5. The size of the cutoff scale Λ is estimated [10] to be $O(10^{15})$ GeV for a unified theory with $SU(5)$ symmetry and even larger, $O(10^{19})$ GeV = M_{planck} if the theory is unified with gravity.

There are many compelling reasons that indicate that the Standard Model is incomplete. One is the fact that the model does not include gravity, which has still not been successfully reformulated into a quantum mechanical theory. Another is that cosmological observations indicate the presence of massive amounts of “dark matter” in the universe. Dark matter is expected to be composed of a stable massive neutral particle which interacts very weakly with other matter; no Standard Model particle fits this description. Finally, there is the “hierarchy,” or fine-tuning problem. This problem strongly affects the Higgs sector, and motivated the development of Supersymmetry, which are the targets of the search presented in this thesis. An short overview of the hierarchy problem and Supersymmetry are presented in the next section.

⁵The Standard Model predicts that lepton number is a good quantum number and that the neutrinos are massless. It has recently been found that the neutrinos do have non-zero mass, and that they undergo oscillations between different neutrino species, violating lepton number.

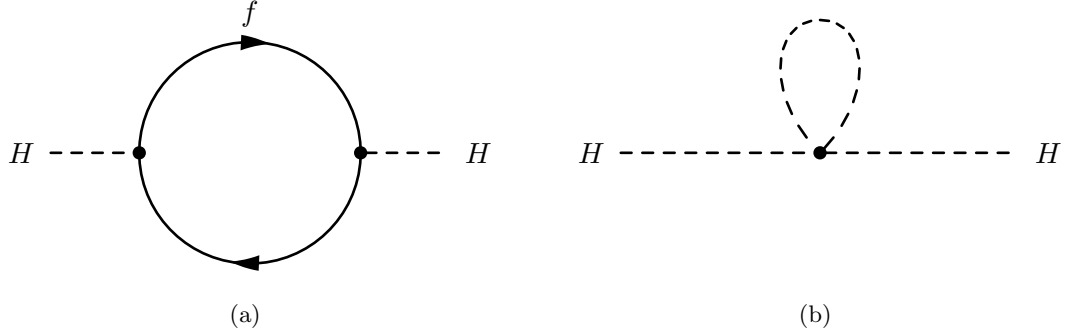


Figure 1.4: Feynman diagram of fermion (a) and scalar (b) loop corrections to Higgs mass.

§1.2.1 The Hierarchy Problem

The enormous size of the cutoff scale Λ in the Standard Model causes a major theoretical problem in the Standard Model. During renormalization of the Standard Model, amplitudes with divergent integrals are cut off at Λ . These large constant terms are “absorbed” into the physical observables. The cutoff term appears directly in quantum corrections to the Higgs mass [15]. The Yukawa term $-\lambda_f H \bar{f} f$ coupling the fermion f to the Higgs H produces loop corrections to Higgs mass. The two types of corrections due to fermion loops and scalar loops are illustrated in Figure 1.4. These contribution of the loop correction in Figure 1.4(a) is [?]

$$m_H^2 = -\frac{|\lambda_f|^2}{8\pi^2} \Lambda^2 + \dots \quad (1.45)$$

The mass of the Higgs is expected to have the same scale as M_W , $O(100)$ GeV/ c^2 . The fact that each fermion contributes a loop correction (Equation 1.45) requires that the “bare mass” of the Higgs been tuned to the precision of $(M_W/\Lambda)^2 \approx 10^{-26}$ for the renormalized mass to be correct. This is the so-called fine-tuning problem: it is believed that in a natural theory there will be only one scale. The electroweak unification analogy is in Equation 1.14, where it was noticed that the difference between the QED and weak scale was due to the massive M_W propagator term, and that the fundamental scale g of the intermediate weak boson theory was compatible with QED. The most promising solution to the hierarchy problem is the introduction of a new, “super” symmetry.

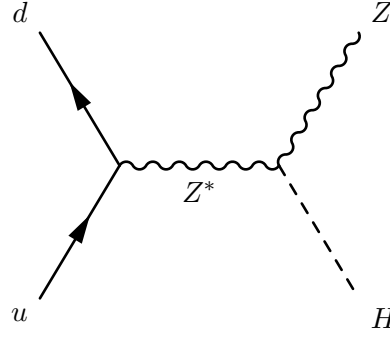


Figure 1.5: Higgstrahlung production diagram at e^+e^- colliders

§1.2.2 Supersymmetry

§1.2.3 The Minimal Supersymmetric Model and τ leptons

§1.3 Searches for the Higgs boson

§1.3.1 Higgs boson phenomenology

The phenomenology of the Higgs boson is strongly coupled to its relationship with mass. The coupling of the Higgs to the fermions is determined by the Yukawa terms (Equation 1.36) in the Lagrangian. Taking the electron as an example, after symmetry breaking, the Yukawa term is found to be

$$\mathcal{L}_e = -\frac{G_e}{\sqrt{2}}(\nu + H(x))\bar{e}e = -\frac{G_e\nu}{\sqrt{2}}\bar{e}e - \frac{G_e}{\sqrt{2}}H(x)\bar{e}e. \quad (1.46)$$

The value of G_e is a free parameter of the theory and is thus determined by the measurement of the electron mass and ν , the VEV of the Higgs field

$$\frac{G_e\nu}{\sqrt{2}} = \frac{m_e}{\nu}. \quad (1.47)$$

The left-hand side of Equation 1.47 is the same as the constant in the electron-Higgs $H(x)\bar{e}e$ coupling term in Equation 1.46. Therefore the coupling between the fermions and Higgs boson is proportional to their mass! This remarkable fact shapes the possible production modes and the branching fractions of Higgs decays.

The dominant modes of Higgs boson production depend on the type of experiment. In general, Higgs production is favored through high-mass intermediate states, due to the mass coupling. At the Tevatron and LEP experiments, which are introduced in the next section, the dominant Higgs production mode is “Higgstrahlung,” where a virtual W^\pm or

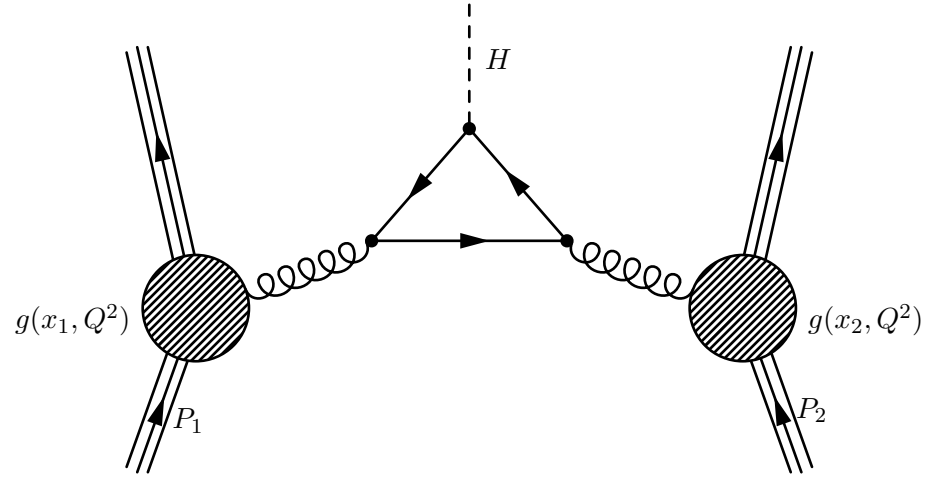


Figure 1.6: Gluon fusion Higgs production mechanism in a proton–proton collision. The Higgs mass coupling favors heavy quarks in the central loop. Image credit: [?]

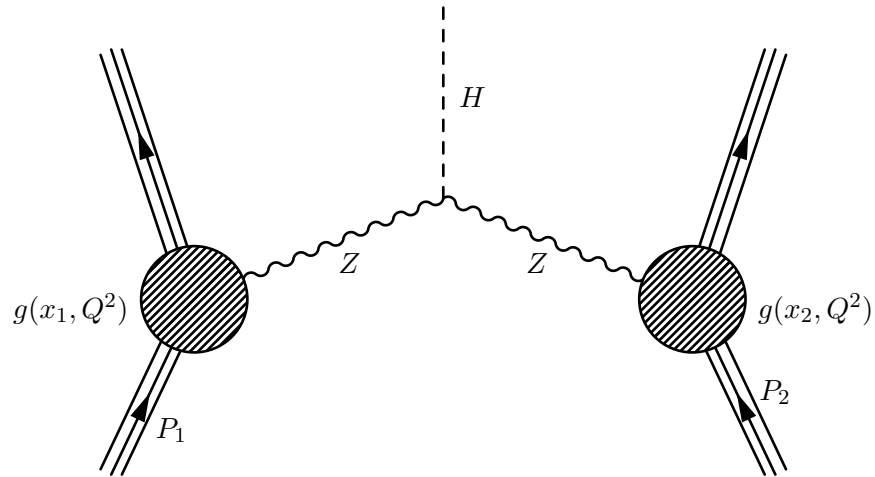


Figure 1.7: Vector boson fusion (VBF) Higgs production mechanism in proton–proton collisions. The VBF mechanism is notable for the lack of color–flow between the two incident protons.

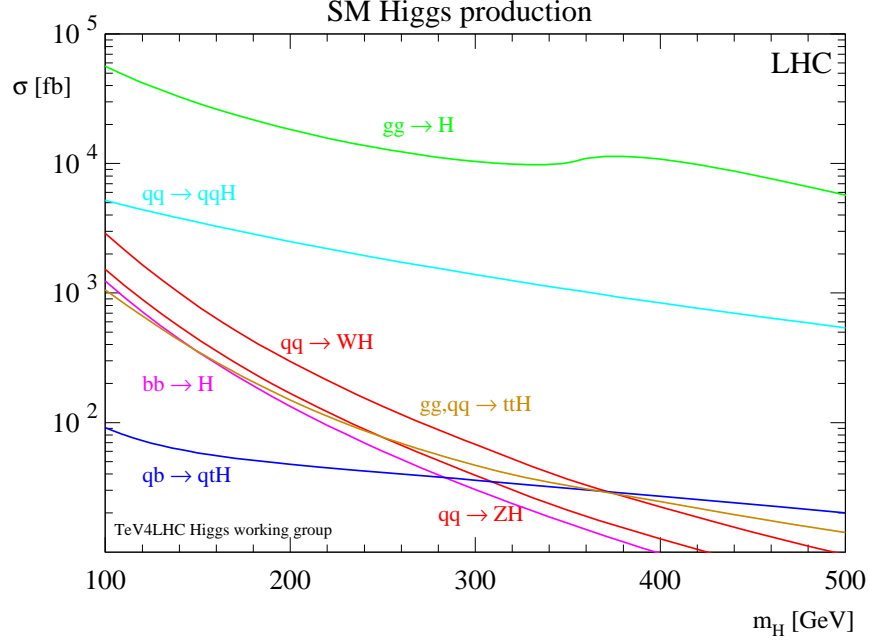


Figure 1.8: Cross section of the Standard Model Higgs boson versus the Higgs boson mass. The different curves give the contribution to the cross section from different production mechanisms. Source: [16].

Z gauge boson is produced and then radiates a Higgs boson. Higgstrahlung is illustrated in Figure ???. At the Large Hadron Collider, higher gluon luminosities result in the favored cross section being “gluon fusion,” (illustrated in Figure 1.3.1) where two gluons from the incident protons combine in a quark (dominated by the massive top quark) loop which then radiates a Higgs boson. Another important channel is “vector boson fusion,” (Figure 1.3.1) where weak gauge bosons (W^\pm or Z) are radiated from the incoming quarks and fuse to produce a Higgs. This is a notable channel due to the lack of “color-flow” (gluons) between the two protons, producing an event with low central jet activity and two “tag-jets” in the forward and backward regions. The theoretical cross sections for the SM Higgs at the LHC are shown in Figure 1.8.

The branching fractions of the different decay modes of the SM Higgs boson depend strongly on the mass of the Higgs boson. In general, the Higgs prefers (due to the Yukawa couplings) to decay pairs of the particles with the highest mass possible. Below the threshold to decay to pairs of weak bosons ($M_H < 160 \text{ GeV}/c^2$), the Higgs decays predominantly to either b -quarks ($b\bar{b}$, 90%) or a pair of τ leptons ($\tau^+\tau^-$, ≈ 10). Above the $W^\pm W^\mp$ thresh-

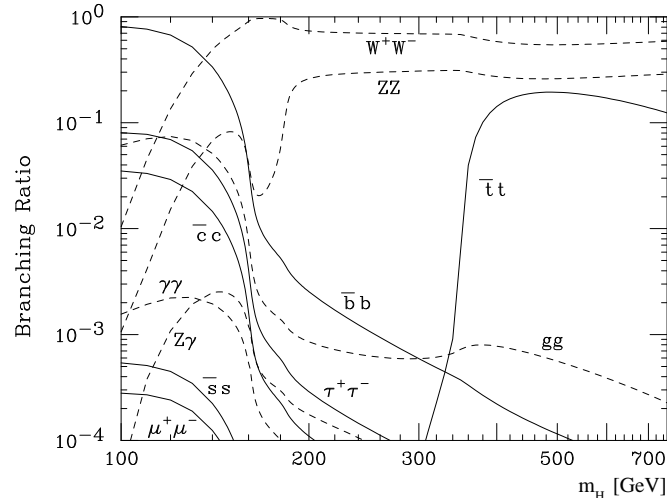


Figure 1.9: Branching fraction of the Standard Model Higgs bosons for different values of M_H . Source: [16].

old, decays to vector bosons ($H \rightarrow W^\pm W^\mp$ and $H \rightarrow ZZ$) dominate. The dependence of branching fraction on M_H and the other rare decay modes are illustrated in Figure 1.9. For low mass Higgs, the $\tau^+\tau^-$ decay mode plays a particularly important role. The dominant decay mode $Ht\bar{b}b$, suffers from enormous backgrounds from QCD jet production. It is important to understand the magnitude of difference between expected Higgs boson production and the rates of various backgrounds. Figure 1.10 illustrates the cross sections for different SM processes at hadron colliders. The rate of Higgs production is many orders of magnitude ($O(10^{-7})$) smaller than that of QCD production. It is important to therefore design searches to use handles that can reject the vast majority of the “uninteresting” events at hadron colliders.

§1.3.2 Results from LEP and Tevatron

§1.4 The physics of the τ lepton

As discussed in sections 1.3.1 and 1.2.3, the τ lepton is an important probe of Higgs physics. The τ lepton has some unusual properties which make it particularly challenging at hadron colliders. With a mass of $1.78 \text{ GeV}/c^2$, the τ lepton is heaviest of the leptons. The nominal decay distance $c\tau$ of the τ lepton is $87 \text{ } \mu\text{m}$, which in practice means that the τ will always

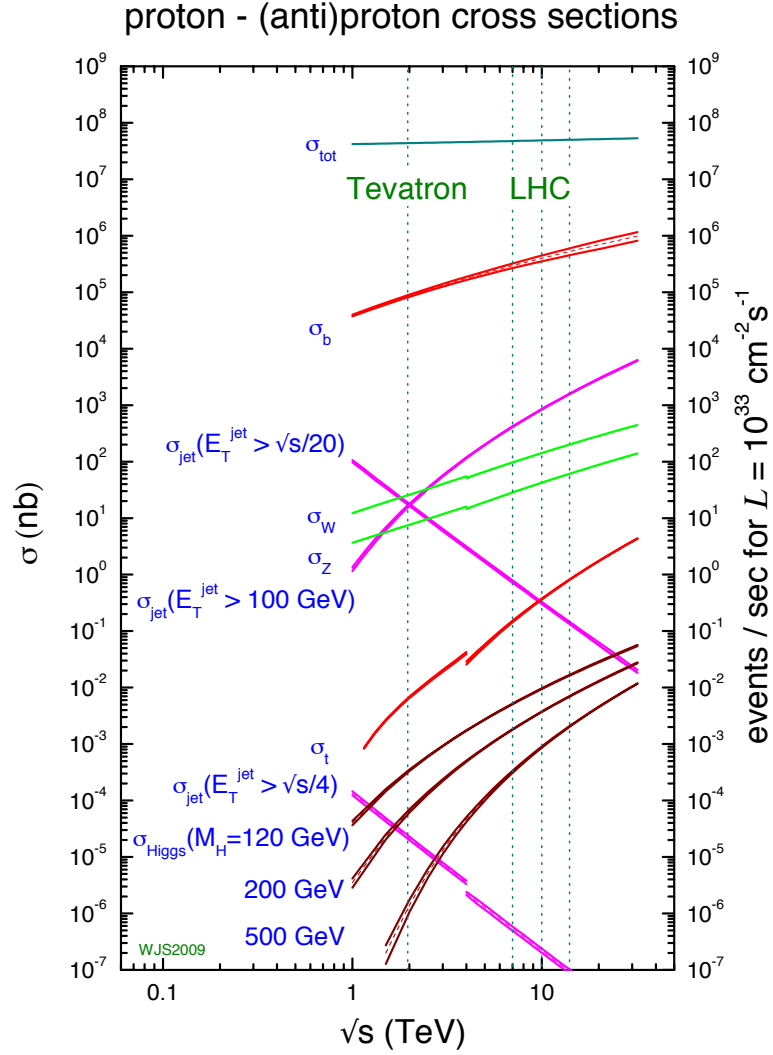


Figure 1.10: Cross sections of various processes at hadron colliders. The horizontal axis represents the center of mass energy of the collision. Of note is the vast difference in scales between Higgs production (maroon lines, $O(10^{-2})$ nb) and the QCD cross section to produce $b\bar{b}$ pairs (red line, $O(10^4)$ nb). Source: [17].

Visible Decay Products	Resonance	Mass (MeV/ c^2)	Fraction [16]
Leptonic modes			
$e^- \nu_\tau \bar{\nu}_e$	-	0.5	17.8%
$\mu^- \nu_\tau \bar{\nu}_\mu$	-	105	17.4%
Hadronic modes			
$\pi^- \nu_\tau$	-	135	10.9%
$\pi^- \pi^0 \nu_\tau$	ρ	770	25.5%
$\pi^- \pi^0 \pi^0 \nu_\tau$	a_1	1200	9.3%
$\pi^- \pi^- \pi^+ \nu_\tau$	a_1	1200	9.0%
$\pi^- \pi^- \pi^+ \pi^0 \nu_\tau$	a_1	1200	4.5%
Total			94.4%

Table 1.1: Resonances and branching ratios of the dominant decay modes of the τ lepton. The decay products listed correspond to a negatively charged τ lepton; the table is identical under charge conjugation.

decay before reaching the first layer of the detector. Tau decays can be effectively classified into two types. “Leptonic” decays consist of a τ decaying to a light lepton ($\ell = e, \mu$) and two neutrinos $\tau^+ \rightarrow \ell^+ \nu_\tau \bar{\nu}_\ell$. “Hadronic” decays consist of a low-multiplicity collimated group of hadrons, typically π^\pm and π^0 mesons. The hadronic decays of the τ lepton compose approximately 65% of the τ lepton branching fraction, with the remainder shared approximately equally by the leptonic decays. The branching fractions for the leptonic and most common hadronic decays are shown in table ??.

Chapter 2

The Compact Muon Solenoid Experiment

Chapter 3

Tau Identification: The Tau Neural Classifier

§3.1 Introduction

High tau identification performance is important for the discovery potential of many possible new physics signals at the Compact Muon Solenoid (CMS). The Standard Model background rates from true tau leptons are typically the same order of magnitude as the expected signal rate in many searches for new physics. The challenge of doing physics with taus is driven by the rate at which objects are incorrectly tagged as taus. In particular, quark and gluon jets have a significantly higher production cross-section and events where these objects are incorrectly identified as tau leptons can dominate the backgrounds of searches for new physics using taus. Efficient identification of hadronic tau decays and low misidentification rate for quarks and gluons is thus essential to maximize the significance of searches for new physics at CMS.

Tau leptons are unique in that they are the only type of leptons which are heavy enough to decay to hadrons. The hadronic decays compose approximately 65% of all tau decays, the remainder being split nearly evenly between $\tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau$ and $\tau^- \rightarrow e^- \bar{\nu}_e \nu_\tau$. The hadronic decays are typically composed of one or three charged pions and zero to two neutral pions. The neutral pions decay almost instantaneously to pairs of photons.

In this chapter, we describe a technique to identify hadronic tau decays. Tau decays to electrons and muons are difficult to distinguish from prompt production of electrons and muons in pp collisions. Analyses that use exclusively non-hadronically decaying taus typically require that the leptonic (e, μ) decays be of opposite flavor. The discrimination of hadronic tau decays from electrons and muons is described in [18]. With the Tau Neural

Classifier, we aim to improve the discrimination of true hadronic tau decays from quark and gluon jets using a neural network approach.

§3.2 Geometric Tau Identification Algorithms

The tau identification strategies used in previously published CMS analyses are fully described in [18]. A summary of the basic methods and strategies is given here. There are two primary methods for selecting objects used to reconstruct tau leptons. The CaloTau algorithm uses tracks reconstructed by the tracker and clusters of hits in the electromagnetic and hadronic calorimeter. The other method (PFTau) uses objects reconstructed by the CMS particle flow algorithm, which is described in [19]. The particle flow algorithm provides a global and unique description of every particle (charged hadron, photon, electron, etc.) in the event; measurements from sub-detectors are combined according to their measured resolutions to improve energy and angular resolution and reduce double counting. The strategies described in this paper use the particle flow objects.

Both methods typically use an “leading object” and an isolation requirement to reject quark and gluon jet background. Quark and gluon jets are less collimated and have a higher constituent multiplicity and softer constituent p_T spectrum than a hadronic tau decay of the same transverse momentum. The “leading track” requirement is applied by requiring a relatively high momentum object near the center of the jet; typically a charged track with transverse momentum greater than 5 GeV/c within $\Delta R < 0.1$ about the center of the jet axis. The isolation requirement exploits the collimation of true taus by defining an isolation annulus about the kinematic center of the jet and requiring no detector activity about a threshold in that annulus. This approach yields a misidentification rate of approximately 1% for QCD backgrounds and a hadronic tau identification efficiency of approximately 50% [18].

§3.3 Decay Mode Tau Identification: Motivation

The tau identification strategy described previously can be extended by looking at the different hadronic decay modes of the tau individually. The dominant hadronic decays of taus consist of a one or three charged π^\pm mesons and up to two π^0 mesons and are enumerated

in table 1.4. The majority of these decays proceed through intermediate resonances and each of these decay modes maps directly to a tau final state multiplicity. Each intermediate resonance has a different invariant mass (see figure 3.1). This implies that the problem of hadronic tau identification can be re-framed from a global search for collimated hadrons satisfying the tau mass constraint into an ensemble of searches for single production of the different hadronic tau decay resonances. The Tau Neural Classifier algorithm implements this approach using two complementary techniques: a method to reconstruct the decay mode and an ensemble of neural network classifiers used to identify each decay mode resonance and reject quark and gluon jets with the same final state topology.

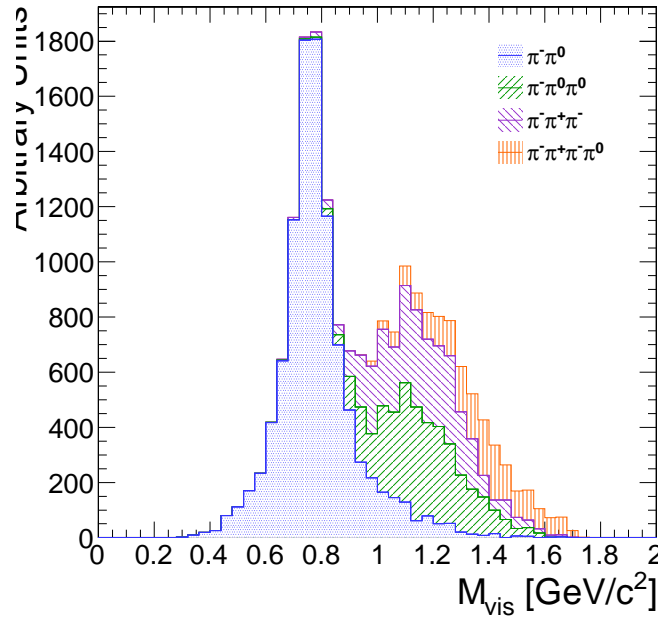


Figure 3.1: The invariant mass of the visible decay products in hadronic tau decays. The decay mode $\tau^- \rightarrow \pi^- \nu_\tau$ is omitted. The different decay modes have different invariant masses corresponding to the intermediate resonance in the decay.

§3.4 The Tau Neural Classifier

The Tau Neural Classifier algorithm reconstructs the decay mode of the tau-candidate and then feeds the tau-candidate to a discriminator associated to that decay mode to make the classification decision. Each discriminator therefore maps to a reconstructed decay mode in a one-to-one fashion. To optimize the discrimination for each of the different decay

modes, the TaNC uses an ensemble of neural nets. Each neural net corresponds to one of the dominant hadronic decay modes of the tau lepton. These selected hadronic decays constitute 95% of all hadronic tau decays. Tau-candidates with reconstructed decay modes not in the set of dominant hadronic modes are immediately tagged as background.

§3.4.1 Decay mode reconstruction

The major task in reconstructing the decay mode of the tau is determining the number of π^0 mesons produced in the decay. A π^0 meson decays almost instantaneously to a pair of photons. The photon objects are reconstructed using the particle flow algorithm [19]. The initial collection of photon objects considered to be π^0 candidates are the photons in the signal cone described by using the “shrinking-cone” tau algorithm, described in [18].

The reconstruction of photons from π^0 decays present in the signal cone is complicated by a number of factors. To suppress calorimeter noise and underlying event photons, all photons with minimum transverse energy less than 0.5 GeV are removed from the signal cone, which removes some signal photons. Photons produced in secondary interactions, pile-up events, and electromagnetic showers produced by signal photons that convert to electron-positron pairs can contaminate the signal cone with extra low transverse energy photons. Highly boosted π^0 mesons may decay into a pair of photons with a small opening angle, resulting in two overlapping showers in the ECAL being reconstructed as one photon. The π^0 meson content of the tau-candidate is reconstructed in two stages. First, photon pairs are merged together into candidate π^0 mesons. The remaining un-merged photons are then subject to a quality requirement.

Photon merging

Photons are merged into composite π^0 candidates by examining the invariant mass of all possible pairs of photons in the signal region. Only π^0 candidates (photon pairs) with a composite invariant mass less than 0.2 GeV/c are considered. The combination of the high granularity of the CMS ECAL and the particle flow algorithm provide excellent energy and angular resolution for photons; the π^0 mass peak is readily visible in the invariant mass spectrum of signal photon pairs (see figure 3.4.1). The π^0 candidates that satisfy the invariant mass requirement are ranked by the difference between the composite invariant

mass of the photon pair and the invariant mass of the π^0 meson given by the PDG [16]. The best pairs are then tagged as π^0 mesons, removing lower-ranking candidate π^0 s as necessary to ensure that no photon is included in more than one π^0 meson.

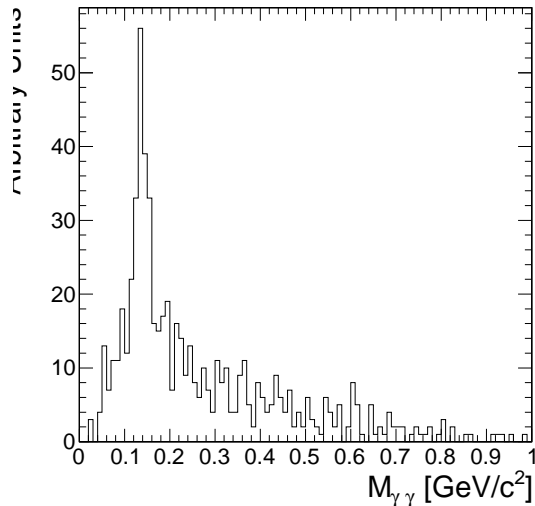


Figure 3.2: Invariant mass of the photon pair for reconstructed tau-candidates with two reconstructed photons in the signal region that are matched to generator level $\tau^- \rightarrow \pi^- \pi^0 \nu_\tau$ decays.

Quality requirements

Photons from the underlying event and other reconstruction effects cause the number of reconstructed photons to be greater than the true number of photons expected from a given hadronic tau decay. Photons that have not been merged into a π^0 meson candidate are recursively filtered by requiring that the fraction of the transverse momentum carried by the lowest P_T photon be greater than 10% with respect to the entire (tracks, π^0 candidates, and photons) tau-candidate. In the case that a photon is not merged but meets the minimum momentum fraction requirement, it is considered a π^0 candidate. This requirement removes extraneous photons, while minimizing the removal of single photons that correspond to a true π^0 meson (see 3.3). A mass hypothesis with the nominal [16] value of the π^0 is applied to all π^0 candidates. All objects that fail the filtering requirements are moved to the isolation collection.

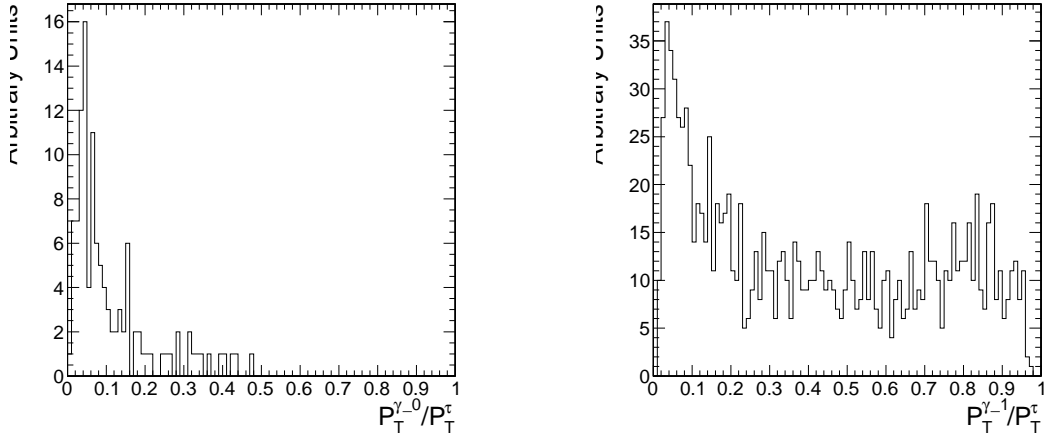


Figure 3.3: Fraction of total τ -candidate transverse momenta carried by the photon for reconstructed taus containing a single photons for two benchmark cases. On the left, the reconstructed tau-candidate is matched to generator level $\tau^- \rightarrow \pi^- \nu_\tau$ decays, for which no photon is expected. On the right, the reconstructed tau-candidate is matched to generator level $\tau^- \rightarrow \pi^- \pi^0 \nu_\tau$ decays and the photon is expected to correspond to a true π^0 meson. The requirement on the P_T fraction of the lowest P_T photon improves the purity of the decay mode reconstruction.

Performance

The performance of the decay mode reconstruction can be measured for tau-candidates that are matched to generator level hadronically decaying tau leptons by examining the correlation of the reconstructed decay mode to the true decay mode determined from the Monte Carlo generator level information. Figure 3.4 compares the decay mode reconstruction performance of a naive approach where the decay mode is determined by simply counting the number of photons to the performance of the photon merging and filtering approach described in section 3.4.1. The correlation for the merging and filtering algorithm is much more diagonal, indicating higher performance. The performance is additionally presented for comparison in tabular form in table 3.4.1 (merging and filtering approach) and table 3.4.1 (naive approach).

The performance of the decay mode reconstruction is dependent on the transverse momentum and η of the tau-candidate and is shown in figure 3.5. The P_T dependence is largely due to threshold effects; high multiplicity decay modes are suppressed at low transverse momentum as the constituents are below the minimum P_T quality requirements.

In the forward region, nuclear interactions and conversions from the increased material budget enhances modes containing π^0 mesons.

True decay mode	Reconstructed Decay Mode					
	$\pi^-\nu_\tau$	$\pi^-\pi^0\nu_\tau$	$\pi^-\pi^0\pi^0\nu_\tau$	$\pi^-\pi^+\pi^-\nu_\tau$	$\pi^-\pi^+\pi^-\pi^0\nu_\tau$	Other
$\pi^-\nu_\tau$	14.8%	1.6%	0.4%	0.1%	0.0%	0.7%
$\pi^-\pi^0\nu_\tau$	6.0%	17.1%	9.0%	0.1%	0.1%	5.5%
$\pi^-\pi^0\pi^0\nu_\tau$	0.9%	3.8%	4.2%	0.0%	0.1%	5.9%
$\pi^-\pi^+\pi^-\nu_\tau$	0.8%	0.3%	0.1%	9.7%	1.6%	6.2%
$\pi^-\pi^+\pi^-\pi^0\nu_\tau$	0.1%	0.2%	0.1%	1.7%	2.7%	4.5%

Table 3.1: Decay mode correlation table for the selected dominant decay modes for the naive approach. The percentage in a given row and column indicates the fraction of hadronic tau decays from $Z \rightarrow \tau^+\tau^-$ events that are matched to a generator level decay mode given by the row and are reconstructed with the decay mode given by the column. Entries in the "Other" column are immediately tagged as background.

True decay mode	Reconstructed Decay Mode					
	$\pi^-\nu_\tau$	$\pi^-\pi^0\nu_\tau$	$\pi^-\pi^0\pi^0\nu_\tau$	$\pi^-\pi^+\pi^-\nu_\tau$	$\pi^-\pi^+\pi^-\pi^0\nu_\tau$	Other
$\pi^-\nu_\tau$	16.2%	1.0%	0.1%	0.1%	0.0%	0.3%
$\pi^-\pi^0\nu_\tau$	10.7%	21.4%	3.6%	0.2%	0.1%	1.9%
$\pi^-\pi^0\pi^0\nu_\tau$	1.8%	7.1%	4.4%	0.1%	0.0%	1.5%
$\pi^-\pi^+\pi^-\nu_\tau$	0.9%	0.2%	0.0%	11.5%	0.6%	5.4%
$\pi^-\pi^+\pi^-\pi^0\nu_\tau$	0.1%	0.3%	0.0%	3.2%	2.9%	2.7%

Table 3.2: Decay mode correlation table for the selected dominant decay modes for the merging and filtering approach. The percentage in a given row and column indicates the fraction of hadronic tau decays from $Z \rightarrow \tau^+\tau^-$ events that are matched to a generator level decay mode given by the row and are reconstructed with the decay mode given by the column. Entries in the "Other" column are immediately tagged as background.

§3.4.2 Neural network classification

Neural Network Training

The samples used to train the TaNC neural networks are typical of the signals and backgrounds found in common physics analyses using taus. The signal-type training sample is composed of reconstructed tau-candidates that are matched to generator level hadronic tau decays coming from simulated $Z \rightarrow \tau^+\tau^-$ events. The background training sample consists of reconstructed tau-candidates in simulated QCD $2 \rightarrow 2$ hard scattering events. The QCD P_T spectrum is steeply falling, and to obtain sufficient statistics across a broad range of P_T the sample is split into different \hat{P}_T bins. Each binned QCD sample imposes a generator

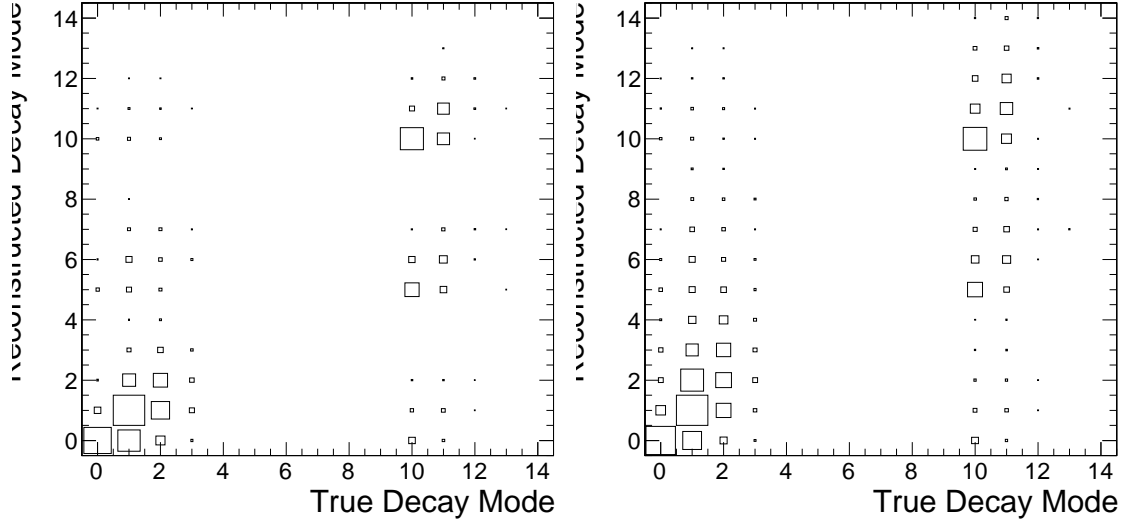


Figure 3.4: Correlations between reconstructed tau decay mode and true tau decay mode for hadronic tau decays in $Z \rightarrow \tau^+\tau^-$ events. The correlation when no photon merging or filtering is applied is shown on the right, and the correlation for the algorithm described in section 3.4.1 is on the left. The horizontal and vertical axis are the decay mode indices of the true and reconstructed decay mode, respectively. The decay mode index N_{DM} is defined as $N_{DM} = (N_{\pi^\pm} - 1) \cdot 5 + N_{\pi^0}$. The area of the box in each cell is proportional to the fraction of tau-candidates that were reconstructed with the decay mode indicated on the vertical axis for the true tau decay on the horizontal axis. The performance of a decay mode reconstruction algorithm can be determined by the spread of the reconstructed number of π^0 mesons about the true number (the diagonal entries) determined from the generator level Monte Carlo information. If the reconstruction was perfect, the correlation would be exactly diagonal.

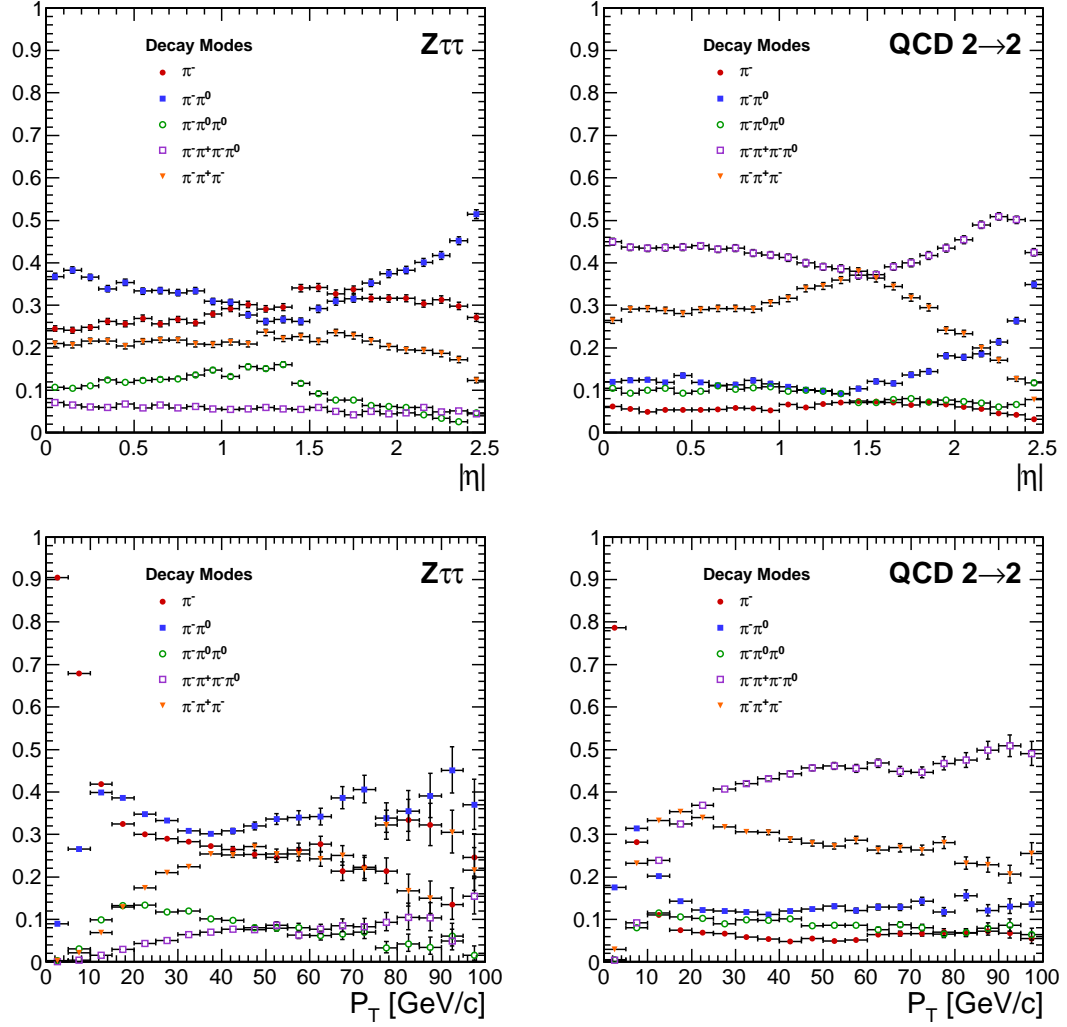


Figure 3.5: Kinematic dependence of reconstructed decay mode for tau-candidates from $Z \rightarrow \tau^+\tau^-$ (left) and QCD di-jets (right) versus transverse momentum (top) and pseudo-rapidity (bottom). Each curve is the probability for a tau-candidate to be reconstructed with the associated decay mode after the leading pion and decay mode preselection has been applied.

	Signal	Background
Total number of tau-candidates	874266	9526176
Tau-candidates passing preselection	584895	644315
Tau-candidates with $W(P_T, \eta) > 0$	538792	488917
Decay Mode	Training Events	
π^-	300951	144204
$\pi^- \pi^0$	135464	137739
$\pi^- \pi^0 \pi^0$	34780	51181
$\pi^- \pi^- \pi^+$	53247	155793
$\pi^- \pi^- \pi^+ \pi^0$	13340	135871

Table 3.3: Number of events used for neural network training and validation for each selected decay mode.

level cut on the transverse momentum of the hard interaction. During the evaluation of discrimination performance the QCD samples are weighted according to their respective integrated luminosities to remove any effect of the binning.

The signal and background samples are split into five subsamples corresponding to each reconstructed decay mode. An additional selection is applied to each subsample by requiring a “leading pion”: either a charged hadron or gamma candidate with transverse momentum greater than 5 GeV/ c . A large number of QCD training events is required as both the leading pion selection and the requirement that the decay mode match one of the dominant modes given in table 1.4 are effective discriminants. For each subsample, 80% of the signal and background tau-candidates are used for training the neural networks by the TMVA software, with half (40%) used as a validation sample used to ensure the neural network is not over-trained. The number of signal and background entries used for training and validation in each decay mode subsample is given in table 3.4.2.

The remaining 20% of the signal and background samples are reserved as a statistically independent sample to evaluate the performance of the neural nets after the training is completed. The TaNC uses the “MLP” neural network implementation provided by the TMVA software package, described in [20]. The “MLP” classifier is a feed-forward artificial neural network. There are two layers of hidden nodes and a single node in the output layer. The hyperbolic tangent function is used for the neuron activation function.

The neural networks used in the TaNC have two hidden layers and single node in the output layers. The number of nodes in the first and second hidden layers are chosen to

be $N + 1$ and $2N + 1$, respectively, where N is the number of input observables for that neural network. According to the Kolmogorov's theorem [21], any continuous function $g(x)$ defined on a vector space of dimension d spanned by x can be represented by

$$g(x) = \sum_{j=1}^{j=2d+1} \Phi_j \left(\sum_{i=1}^d \phi_i(x) \right) \quad (3.1)$$

for suitably chosen functions for Φ_j and ϕ_j . As the form of equation 3.1 is similar to the topology of a two hidden-layer neural network, Kolmogorov's theorem suggests that *any* classification problem can be solved with a neural network with two hidden layers containing the appropriate number of nodes.

The neural network is trained for 500 epochs. At ten epoch intervals, the neural network error is computed using the validation sample to check for over-training (see figure 3.6).

The neural network error E is defined [20] as

$$E = \frac{1}{2} \sum_{i=1}^N (y_{ANN,i} - \hat{y}_i)^2 \quad (3.2)$$

where N is the number of training events, $y_{ANN,i}$ is the neural network output for the i th training event, and y_i is the desired (-1 for background, 1 for signal) output the i th event. No evidence of over-training is observed.

The neural networks use as input observables the transverse momentum and η of the tau-candidates. These observables are included as their correlations with other observables can increase the separation power of the ensemble of observables. For example, the opening angle in ΔR for signal tau-candidates is inversely related to the transverse momentum, while for background events the correlation is very small [?]. In the training signal and background samples, there is significant discrimination power in the P_T spectrum. However, it is desirable to eliminate any systematic dependence of the neural network output on P_T and η , as in practice the TaNC will be presented with tau-candidates whose $P_T - \eta$ spectrum will be analysis dependent. The dependence on P_T and η is removed by applying a P_T and η dependent weight to the tau-candidates when training the neural nets.

The weights are defined such that in any region in the vector space spanned by P_T and η where the signal sample and background sample probability density functions are different, the sample with higher probability density is weighted such that the samples have identical $P_T - \eta$ probability distributions. This removes regions of $P_T - \eta$ space where the

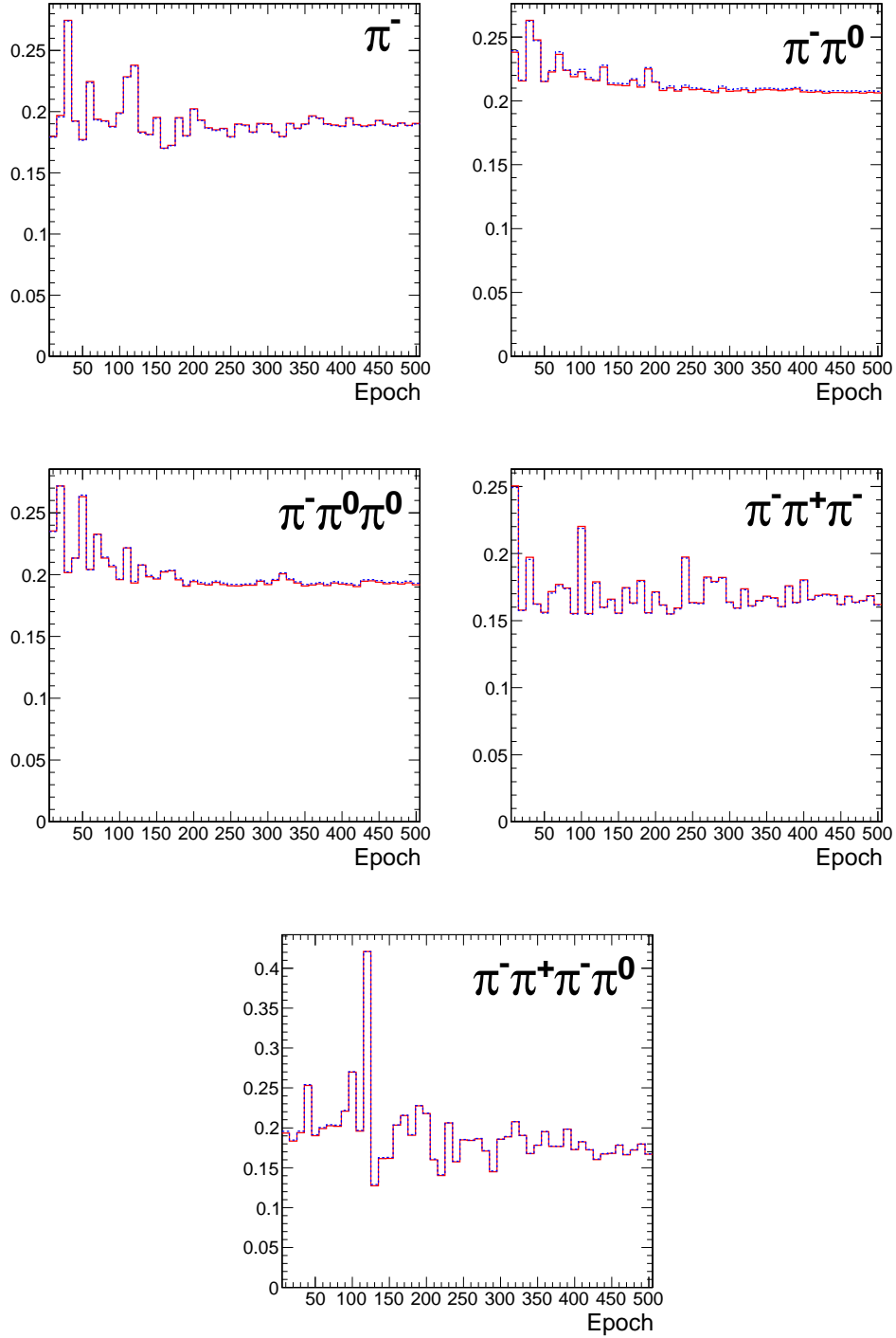


Figure 3.6: Neural network classification error for training (solid red) and testing (dashed blue) samples at ten epoch intervals over the 500 training epochs for each decay mode neural network. The vertical axis represents the classification error, defined by equation 3.2. N.B. that the choice of hyperbolic tangent for neuron activation functions results in the desired outputs for signal and background to be 1 and -1, respectively. This results in the computed neural network error being larger by a factor of four than the case where the desired outputs are (0, 1). Classifier over-training would be evidenced by divergence of the classification error of the training and testing samples, indicating that the neural net was optimizing about statistical fluctuations in the training sample.

training sample is exclusively signal or background. The weights are computed according to

$$W(P_T, \eta) = \text{less}(p_{sig}(P_T, \eta), p_{bkg}(P_T, \eta))$$

$$w_{sig}(P_T, \eta) = W(P_T, \eta)/p_{sig}(P_T, \eta)$$

$$w_{bkg}(P_T, \eta) = W(P_T, \eta)/p_{bkg}(P_T, \eta)$$

where $p_{sig}(P_T, \eta)$ and $p_{bkg}(P_T, \eta)$ are the probability densities of the signal and background samples after the “leading pion” and dominant decay mode selections. Figure 3.7 shows the signal and background training P_T distributions before and after the weighting is applied.

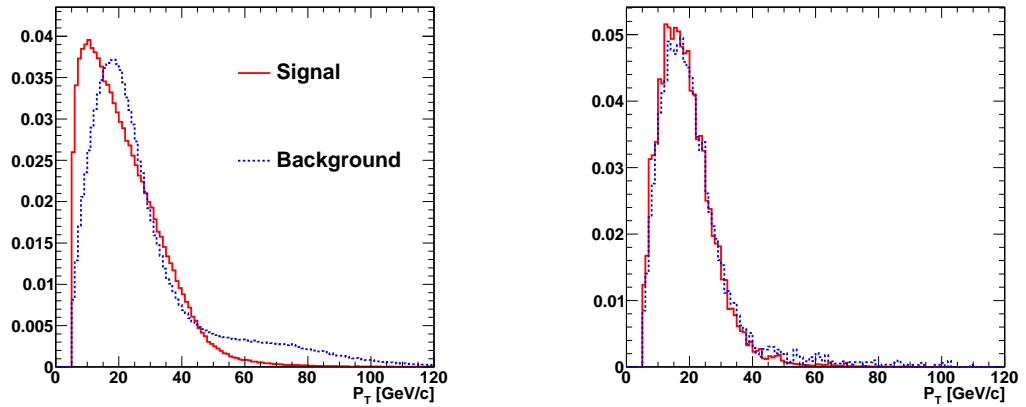


Figure 3.7: Transverse momentum spectrum of signal and background tau-candidates used in neural net training before (left) and after (right) the application of $P_T - \eta$ dependent weight function. Application of the weights lowers the training significance of tau-candidates in regions of $P_T - \eta$ phase space where either the signal or background samples has an excess of events.

Discriminants

Each neural network corresponds to a different decay mode topology and as such each network uses different observables as inputs. However, many of the input observables are used in multiple neural nets. The superset of all observables is listed and defined below. Table 3.4 maps the input observables to their associated neural networks. The signal and background distributions of the input observables for tau-candidates in the training sample are shown in appendix ???. In three prong decays, the definition of the “main track” is important. The main track corresponds to the track with charge opposite to that of the

total charge of the three tracks. This distinction is made to facilitate the use of the “Dalitz” observables, allowing identification of intermediate resonances in three-body decays. This is motivated by the fact that the three prong decays of the tau generally proceed through $\tau^- \rightarrow a1^- \nu_\tau \rightarrow \pi^- \rho^0 \nu_\tau \rightarrow \pi^- \pi^+ \pi^- \nu_\tau$; the oppositely charged track can always be identified with the ρ^0 decay.

ChargedOutlierAngleN

ΔR between the Nth charged object (ordered by P_T) in the isolation region and the tau-candidate momentum axis. If the number of isolation region objects is less than N, the input is set at one.

ChargedOutlierPtN

Transverse momentum of the Nth charged object in the isolation region. If the number of isolation region objects is less than N, the input is set at zero.

DalitzN

Invariant mass of four vector sum of the “main track” and the Nth signal region object.

Eta

Pseudo-rapidity of the signal region objects.

InvariantMassOfSignal

Invariant mass of the composite object formed by the signal region constituents.

MainTrackAngle

ΔR between the “main track” and the composite four-vector formed by the signal region constituents.

MainTrackPt

Transverse momentum of the “main track.”

OutlierNCharged

Number of charged objects in the isolation region.

OutlierSumPt

Sum of the transverse momentum of objects in the isolation region.

PiZeroAngleN

ΔR between the Nth π^0 object in the signal region (ordered by P_T) and the tau-candidate momentum axis.

PiZeroPtN

Transverse momentum of the Nth π^0 object in the signal region.

TrackAngleN

ΔR between the Nth charged object in the signal region (ordered by P_T) and the tau-candidate momentum axis, exclusive of the main track.

TrackPtN

Transverse momentum of the Nth charged object in the signal region, exclusive of the main track.

Neural network performance

The classification power of the neural networks is unique for each of the decay modes. The performance is determined by the relative separation of the signal and background distributions in the parameter space of the observables used as neural network inputs. A pathological example is the case of tau-candidates with the reconstructed decay mode of $\tau^- \rightarrow \pi^- \nu_\tau$. If there is no isolation activity, the neural net has no handle with which it can separate the signal from the background. The neural net output for tau-candidates in the testing sample (independent of the training and validation samples) for each of the five decay mode classifications is shown in figure 3.8.

When a single neural network is used for classification, choosing an operating point is relatively straightforward: the requirement on neural network output is tuned such that the desired purity is attained. However, in the case of the TaNC, multiple neural networks are used. Each network has a unique separation power (see figure 3.9) and each neural network is associated to a reconstructed decay mode that composes different relative fractions of the

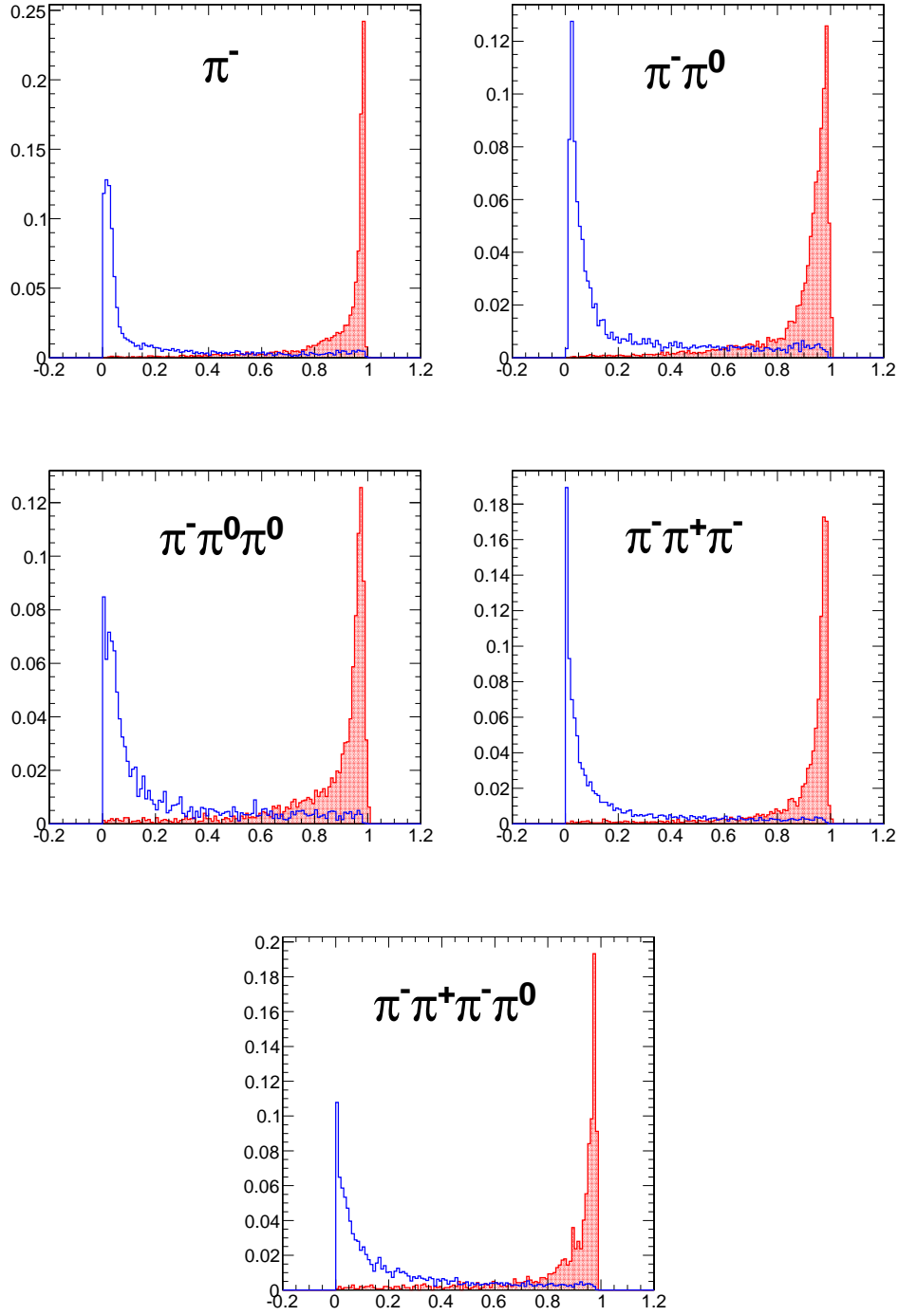


Figure 3.8: Neural network output distributions for the five reconstructed tau-candidate decay modes used in the TaNC for $Z \rightarrow \tau^+\tau^-$ events (red) and QCD di-jet events (blue).

Input observable	Neural network				
	$\pi^- \nu_\tau$	$\pi^- \pi^0 \nu_\tau$	$\pi^- \pi^0 \pi^0 \nu_\tau$	$\pi^- \pi^+ \pi^- \nu_\tau$	$\pi^- \pi^+ \pi^- \pi^0 \nu_\tau$
ChargedOutlierAngle1	•	•	•	•	•
ChargedOutlierAngle2	•	•	•	•	•
ChargedOutlierPt1	•	•	•	•	•
ChargedOutlierPt2	•	•	•	•	•
ChargedOutlierPt3	•	•	•	•	•
ChargedOutlierPt4	•	•	•	•	•
Dalitz1			•	•	•
Dalitz2			•	•	•
Eta	•	•	•	•	•
InvariantMassOfSignal		•	•	•	•
MainTrackAngle		•	•	•	•
MainTrackPt	•	•	•	•	•
OutlierNCharged	•	•	•	•	•
OutlierSumPt	•	•	•	•	•
PiZeroAngle1		•	•		•
PiZeroAngle2			•		
PiZeroPt1		•	•		•
PiZeroPt2			•		
TrackAngle1				•	•
TrackAngle2				•	•
TrackPt1				•	•
TrackPt2				•	•

Table 3.4: Input observables used for each of the neural networks implemented by the Tau Neural Classifier. The columns represents the neural networks associated to various decay modes and the rows represent the superset of input observables (see section 3.4.2) used in the neural networks. A dot in a given row and column indicates that the observable in that row is used in the neural network corresponding to that column.

signal and background tau-candidates. Therefore, a set of five numbers is required to define an “operating point” (the signal efficiency and background misidentification rate) in the TaNC output. All points in this five dimensional cut-space map to an absolute background fake-rate and signal efficiency rate. Therefore there must exist a 5D “performance curve” which for any attainable signal efficiency gives the lowest fake-rate. A direct method to approximate the performance curve is possible using a Monte Carlo technique.

The maximal performance curve can be approximated by iteratively sampling points in the five-dimensional cut space and selecting the highest performance points. The collection of points in the performance curve are ordered by expected fake rate. During each iteration, the sample point is compared to the point before the potential insertion position of the

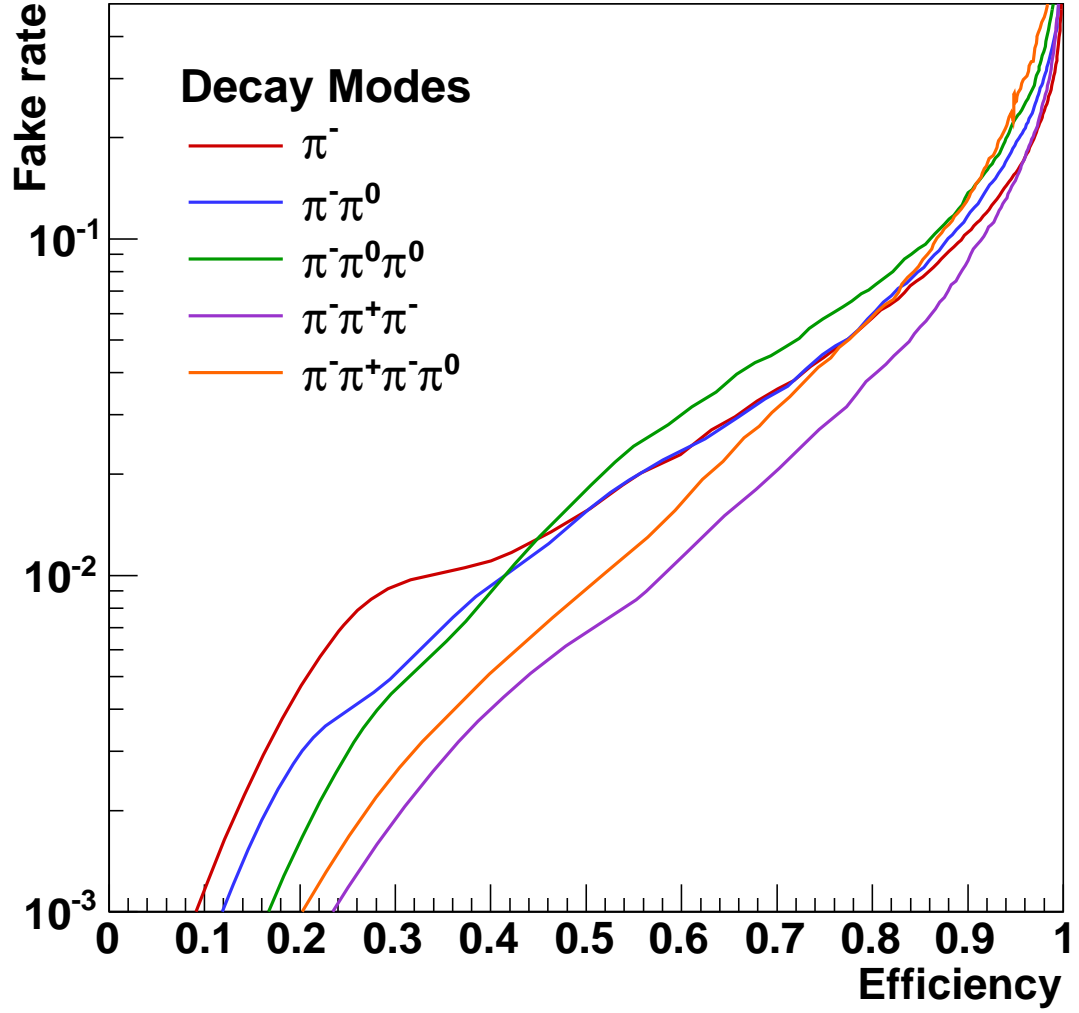


Figure 3.9: Performance curves for the five neural networks used by the TaNC for tau-candidates with transverse momentum greater than 20 GeV/ c . Each curve represents the signal efficiency (on the horizontal axis) and background misidentification rate (vertical axis) for a scan of the neural network selection requirement for a single neural network. The efficiency (or misidentification rate) for each neural network performance curve is defined with respect to the preselected tau-candidates that have the reconstructed decay mode associated with that neural network. Each neural network has a different ability to separate signal and background as each classifier uses different observables as inputs.

sample in the ordered collection. The sample point is inserted into the collection if it has a higher signal identification efficiency than the point before it. The sample point is then compared to all points in the collection after it (i.e. those with a larger fake rate); any point with a lower signal efficiency than the sample point is removed. After the performance curve has been determined, the set of cuts are evaluated on an independent validation sample to ensure that the measured performance curve is not influenced by favorable statistical fluctuations being selected by the Monte Carlo sampling. The performance curves for two different transverse momentum ranges are shown in figure 3.10.

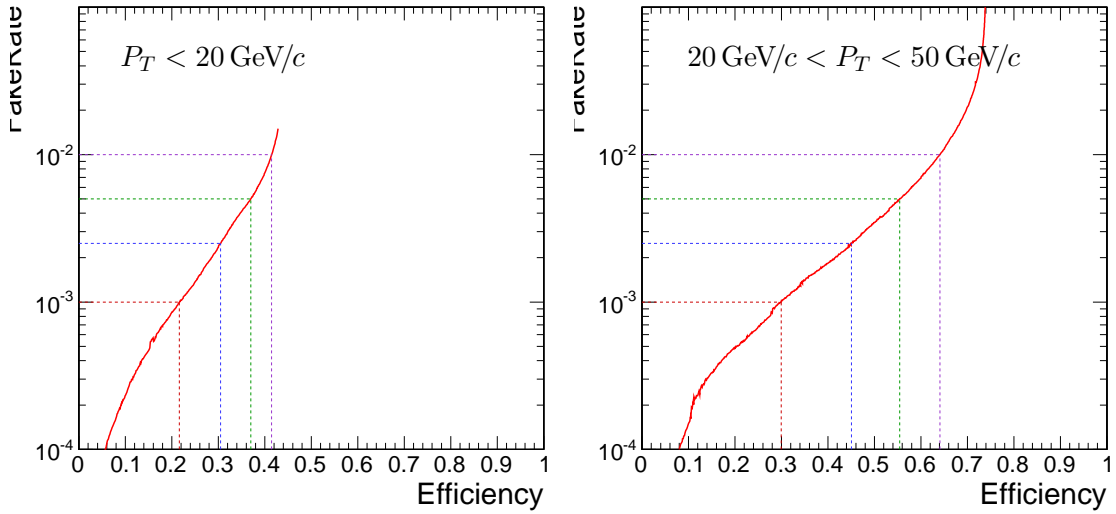


Figure 3.10: Tau Neural Classifier performance curves for tau-candidates with $P_T < 20 \text{ GeV}/c$ (left) and $20 \text{ GeV}/c < P_T < 50 \text{ GeV}/c$ (right). The vertical axis represents the expected fake-rate of QCD jets and the horizontal axis the expected signal efficiency for hadronic tau decays. The performance curve for the low transverse momentum range is worse due to leading pion selection. While both true taus and QCD are removed by this cut, the selection preferentially keeps the QCD tau-candidates with low multiplicities, which increases the number of QCD tau-candidates passing the decay mode selection.

The 5D performance curve can also be parameterized by using the probability for a tau-candidate to be identified for a given decay mode. In this manner a single number (the “transform cut”) can be used to specify any point on the performance curve. For a given decay mode i , the appropriate neural network cut x_i can be mapped from transform cut x

using the following relation:

$$x_i = \frac{\rho_i^{bkg} x}{(\rho_i^{bkg} - \rho_i^{signal})x + \rho_i^{signal}} \quad (3.3)$$

where ρ_i^{bkg} represents the probability that a preselected background tau-candidate will be reconstructed with decay mode i . The training sample neural network output after the transformation has been applied is shown in figure 3.12. The theoretical motivation for the transformation given in equation 3.3 is described in appendix ???. The performance curve for the cut on the transformed output is nearly identical to the optimal performance curve determined by the Monte Carlo sampling technique.

The discriminator output of the TaNC algorithm is a continuous quantity, enabling analysis specific optimization of the selection to maximize sensitivity. For the convenience of the user, four operating point benchmark selections are provided in addition to the continuous output. The four operating points are chosen such that for tau-candidates with transverse momentum between 20 and 50 GeV/c, the expected QCD di-jet fake rate will be 0.1%, 0.25%, 0.50% and 1.0%, respectively.

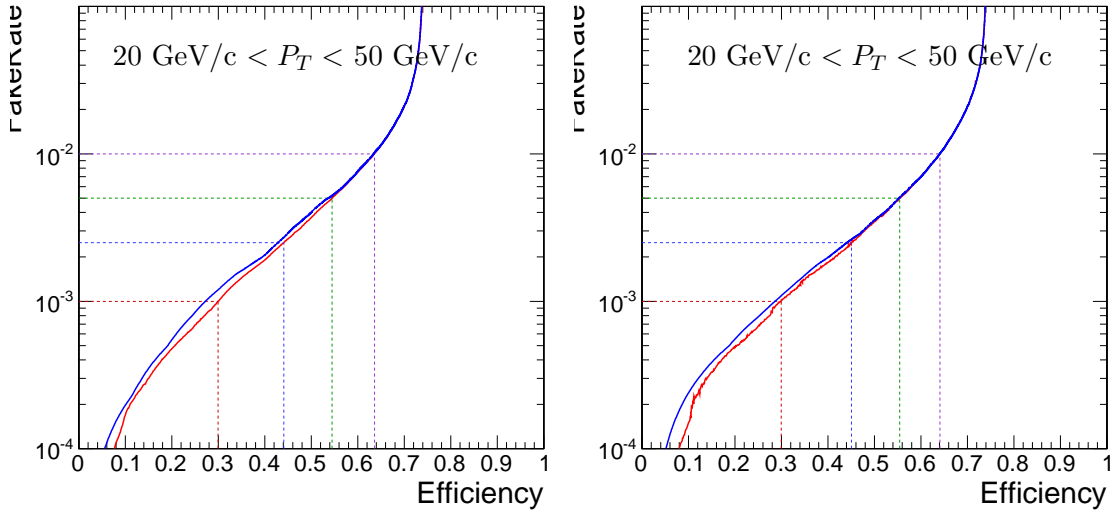


Figure 3.11: Tau Neural Classifier performance curves for tau-candidates with $20 \text{ GeV}/c < P_T < 50 \text{ GeV}/c$. The figure on the left compares the optimal performance curve determined by the Monte Carlo sampling method (red) to the performance curve obtained by scanning the “transform cut” (blue) defined in equation 3.3 from zero to one. The figure on the right is the same set of cuts (and cut transformation values) applied on an independent sample to remove any biases introduced by the Monte Carlo sampling. The four dashed lines indicate the performance for the four benchmark points.

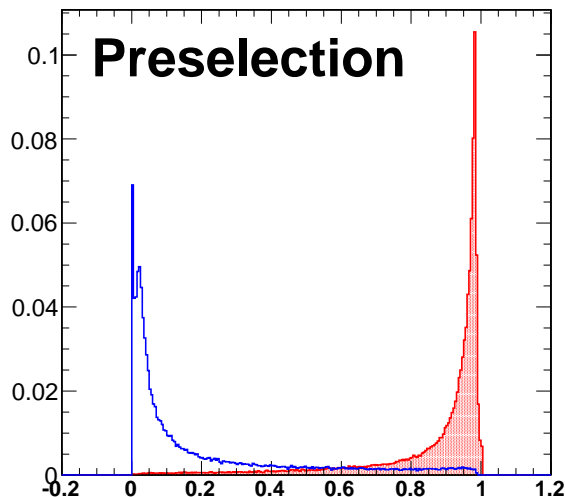


Figure 3.12: Transformed TaNC neural network output for tau-candidates with transverse momentum between 20 and 50 GeV/c that pass the pre-selection criteria. The neural network output for each tau-candidate has been transformation according to equation 3.3. The decay mode probabilities ρ_i^{bkg} , ρ_i^{signal} are computed using the entire transverse momentum range of the sample.

§3.5 Summary

The Tau Neural classifier introduces two complimentary new techniques for tau lepton physics at CMS: reconstruction of the hadronic tau decay mode and discrimination from quark and gluon jets using neural networks. The decay mode reconstruction strategy presented in section 3.4.1 significantly improves the determination of the decay mode. This information has the potential to be useful in studies of tau polarization and background estimation.

The Tau Neural classifier tau identification algorithm significantly improves tau discrimination performance compared to isolation-based approaches [?] used in previous CMS analyses. Figure 3.13 compares the performance of the “shrinking cone” isolation tau-identification algorithm [?] to the performance of the TaNC for a scan of requirements on the transformed neural network output. The signal efficiency and QCD di-jet fake rate versus tau-candidate transverse momentum and pseudo-rapidity for the four benchmark points and the isolation based tau identification are show in figure 3.14. For tau-candidates

with transverse momentum between 20 and 50 GeV/c, the TaNC operating cut can be chosen such that the two methods have identical signal efficiency; at this point the TaNC algorithm reduces the background fake rate by an additional factor of 3.9. This reduction in background will directly improve the significance of searches for new physics using tau leptons at CMS.

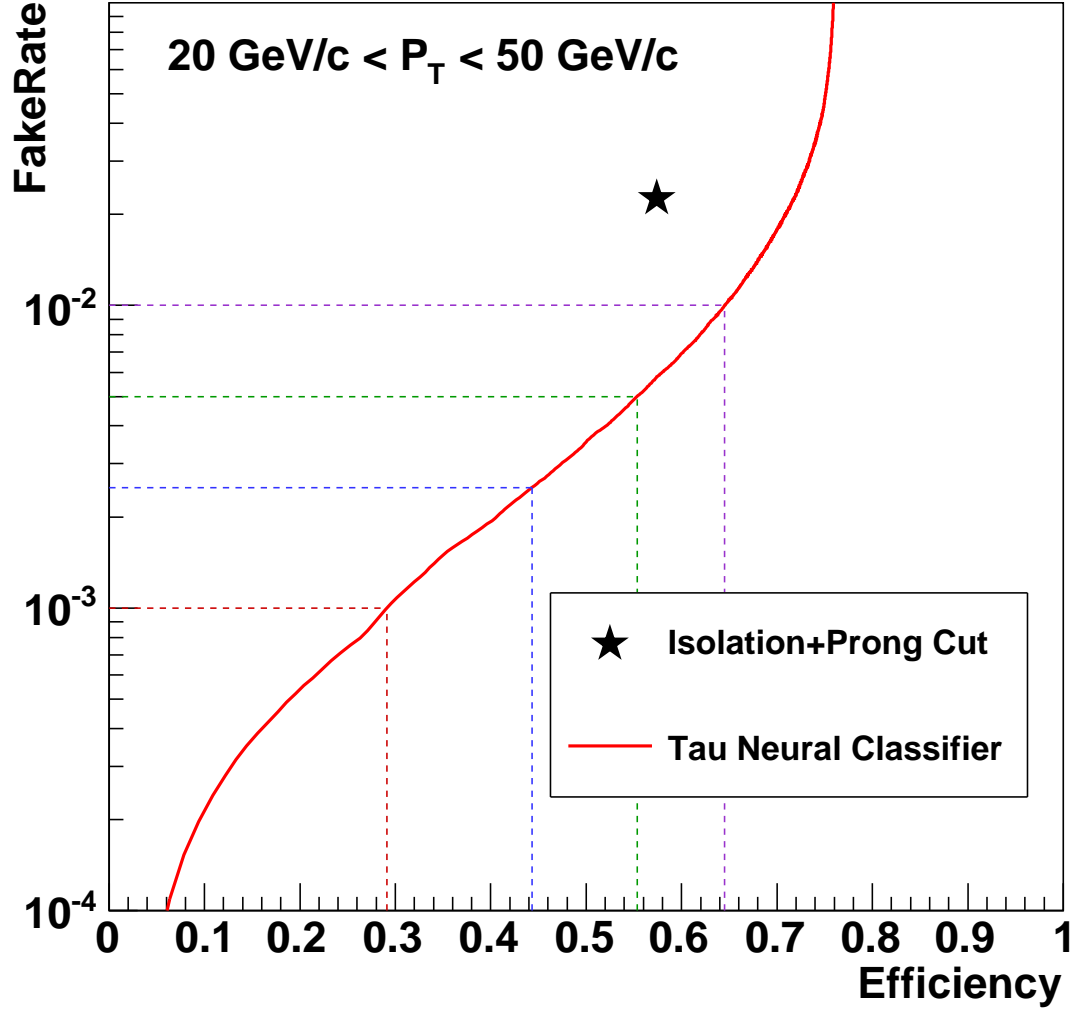


Figure 3.13: Performance curve (red) of the TaNC tau identification for various requirements on the output transformed according to equation 3.3. The horizontal axis is the efficiency for true taus with transverse momentum between 20 and 50 GeV/c to satisfy the tau identification requirements. The vertical axis gives the rate at which QCD di-jets with generator-level transverse momentum between 20 and 50 GeV/c are incorrectly identified as taus. The performance point for the same tau-candidates using the isolation based tau-identification [?] used in many previous CMS analyses is indicated by the black star in the figure. An additional requirement that the signal cone contain one or three charged hadrons (typical in a final physics analysis) has been applied to the isolation based tau-identification to ensure a conservative comparison.

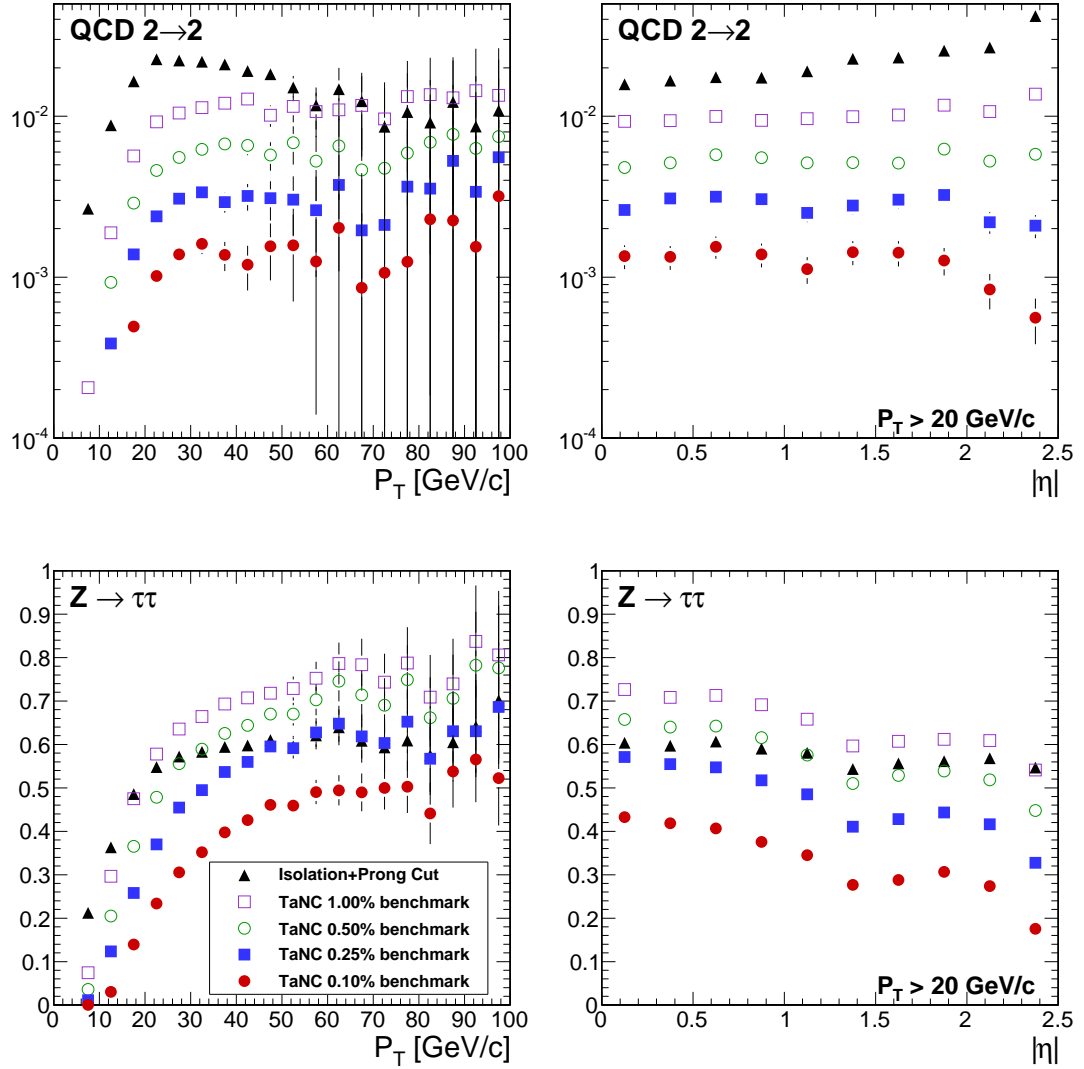


Figure 3.14: Comparison of the identification efficiency for hadronic tau decays from $Z \rightarrow \tau^+\tau^-$ decays (bottom row) and the misidentification rate for QCD di-jets (top row) versus tau-candidate transverse momentum (left) and pseudo-rapidity (right) for different tau identification algorithms. The efficiency (fake-rate) in a given bin is defined as the quotient of the number of true tau hadronic decays (generator level jets) in that bin that are matched to a reconstructed tau-candidate that passes the identification algorithm divided by the number of true tau hadronic decays (generator level jets) in that bin. In the low transverse momentum region both the number of tau-candidates in the denominator and the algorithm acceptance vary rapidly with respect to P_T for both signal and background; a minimum transverse momentum requirement of 20 GeV/c is applied to the pseudorapidity plots to facilitate interpretation of the plots.

Chapter 4

Mass Reconstruction: The Secondary Vertex Fit

The dominant background in the search for the Higgs $\rightarrow \tau^+\tau^-$ signal is due to Standard Model $Z \rightarrow \tau^+\tau^-$ events. The most “natural” observable to discriminate between Higgs signal and Z background would be the invariant mass of the di-tau system, utilizing the fact that the Z resonance is well known ($m_Z = 91.1876 \pm 0.0021 \text{ GeV}/c^2$) and has a narrow width ($\Gamma_Z = 2.4952 \pm 0.0023 \text{ GeV}$) [16]. The experimental complication in this approach is due to the neutrinos produced in the tau lepton decays, which escape detection and carry away an unmeasured amount of energy, and making it difficult to reconstruct the tau lepton four-vectors. In this chapter we give an overview of techniques used in the past

FixMe: *fuck*

§4.1 Existing mass reconstruction algorithms

The simplest observable related to the $\tau^+\tau^-$ mass is one can construct that is sensitive to new particle content is the invariant mass of the visible (reconstructible) decay products associated with each tau decays. This quantity, referred in this document as the “Visible Mass,” has the advantages of simplicity and lack of exposure to systematic errors associated with the reconstruction of the E_T^{miss} . However, no attempt is made to reconstruct the neutrinos in the event. The reconstructed mass is thus systematically smaller than mass of the resonance which produced the tau leptons. The visible mass is typically on the order of 1/2 of the resonance mass, depending on the kinematic requirements applied to the visible products of the tau decays.

The Collinear Approximation is the conventional technique to reconstruct the *full* $\tau^+\tau^-$ mass. In an event with two tau decays, there are a total of six unknowns associated with the missing energy: the three components of the momentum of each neutrino. The

Collinear Approximation makes the assumption that the neutrinos have the same direction as their associated visible decay products. This assumption reduces the number of unknown quantities to two, corresponding to the total energy of each neutrino. These two unknowns can be solved for by using the two components of the reconstructed missing transverse energy, which in the ideal case corresponds to the transverse component of the vector sum of the two neutrino's four momentum. The characteristic equation of the Collinear Approximation is

$$\begin{pmatrix} E_x^{\text{miss}} \\ E_y^{\text{miss}} \end{pmatrix} = \begin{pmatrix} \cos \phi_1 & \cos \phi_2 \\ \sin \phi_1 & \sin \phi_2 \end{pmatrix} \begin{pmatrix} E_1 \\ E_2 \end{pmatrix} \quad (4.1)$$

where $(E_x^{\text{miss}}, E_y^{\text{miss}})$ are the two components of the reconstructed missing transverse energy, $\phi_{1(2)}$ is the azimuthal angle of the visible component of the first (second) tau decay, and $E_{1(2)}$ is the reconstructed energy of neutrino of the first (second) tau decay. E_1 and E_2 can be extracted by inverting the matrix on the right hand side of Equation 4.1.

$$\begin{pmatrix} E_1 \\ E_2 \end{pmatrix} = \frac{1}{\sin(\phi_2 - \phi_1)} \begin{pmatrix} \sin \phi_2 & -\cos \phi_2 \\ -\sin \phi_1 & \cos \phi_1 \end{pmatrix} \begin{pmatrix} E_x^{\text{miss}} \\ E_y^{\text{miss}} \end{pmatrix} \quad (4.2)$$

The Collinear Approximation suffers from two problems. The approximation can fail (yielding unphysical negative energies for the reconstructed neutrinos) when the missing transverse energy is mis-measured. The events with unphysical solutions must be removed from the analysis, leading to a dramatic reduction in acceptance (on the order of 50% in this analysis). Improvements to the collinear approximation algorithm have recently been made which aim to recover part of the events with unphysical solutions [?]. But even with these improvements, no physical solution is still found for a large fraction of signal events. Additionally, the method is numerically sensitive when the two τ lepton are nearly back-to-back in azimuth. In these cases the $\sin(\phi_2 - \phi_1)^{-1}$ term in Equation 4.2 is very large and small mis-measurements of the missing transverse energy can produce a large tail on the reconstructed mass. This tail is particularly large for low-mass resonances. The large tail for low mass is predominantly due to the fact (discussed in subsection 4.4.2) that the kinematic requirements¹ applied on the visible decay products preferentially selects events where the visible decay products carry the majority of the energy of the original τ lepton,

¹The kinematic requirements on the visible decay products are necessary to reduce backgrounds and maintain compatibility with un-prescaled event triggers. This topic is discussed in detail in chapter 5.

reducing the amount of true missing energy in the event.

§4.2 The Secondary Vertex fit

A novel algorithm is presented in the following, which succeeds in finding a physical solution for every event. As an additional benefit, the new algorithm is found to improve the di-tau invariant mass resolution, making it easier to separate the Higgs signal from the $Z \rightarrow \tau^+\tau^-$ background.

The novel Secondary Vertex fit (SVfit) algorithm for di-tau invariant mass reconstruction that we present in the following utilizes a likelihood maximization to fit a $\tau^+\tau^-$ invariant mass hypothesis for each event. The likelihood is composed of separate terms which represent probability densities of:

- tau decay kinematics
- matching between the momenta of neutrinos produced in the tau decays and the reconstructed missing transverse momentum
- a regularization “ p_T -balance” term which accounts for the effects on the di-tau invariant mass of acceptance cuts on the visible tau decay products
- the compatibility of tau decay parameters with the position of reconstructed tracks and the known tau lifetime of $c\tau = 87 \mu\text{m}$ [16].

The likelihood is maximized as function of a set of parameters which fully describe the tau decay.

§4.3 Parametrization of tau decays

The decay of a tau of visible four-momentum p_{vis} measured in the CMS detector (“laboratory”) frame can be parametrized by three variables. The invisible (neutrino) momentum is fully determined by these parameters.

The “opening-angle” θ is defined as the angle between the boost direction of the tau lepton and the momentum vector of the visible decay products in the rest frame of the tau. The azimuthal angle of the tau in the lab frame is denoted as $\bar{\phi}$ (we denote quantities

defined in the laboratory frame by a overline). A local coordinate system is defined such that the \bar{z} -direction lies along the visible momentum and $\bar{\phi} = 0$ lies in the plane spanned by the momentum vector of the visible decay products and the proton beam direction. The third parameter, $m_{\nu\nu}$, denotes the invariant mass of the invisible momentum system.

Given θ , $\bar{\phi}$ and $m_{\nu\nu}$, the energy and direction of the tau lepton can be computed by means of the following equations: The energy of the visible decay products in the rest frame of the tau lepton is related to the invariant mass of the neutrino system by:

$$E^{vis} = \frac{m_\tau^2 + m_{vis}^2 - m_{\nu\nu}^2}{2m_\tau} \quad (4.3)$$

Note that for hadronic decays, $m_{\nu\nu}$ is a constant of value zero, as only a single neutrino is produced. Consequently, the magnitude of P^{vis} depends on the reconstructed mass of the visible decay products only and is a constant during the SVfit.

The opening angle $\bar{\theta}$ between the tau lepton direction and the visible momentum vector in the laboratory frame is determined by the rest frame quantities via the (Lorentz invariant) component of the visible momentum perpendicular to the tau lepton direction:

$$\begin{aligned} p_\perp^{vis} &= \bar{p}_\perp^{vis} \\ \Rightarrow \sin \bar{\theta} &= \frac{p_\perp^{vis} \sin \theta}{\bar{p}^{vis}} \end{aligned} \quad (4.4)$$

Substituting the parameters $m_{\nu\nu}$ and θ into equations 4.3 and 4.4, the energy of the tau is obtained by solving for the boost factor γ in the Lorentz transformation between tau rest frame and laboratory frame of the visible momentum component parallel to the tau direction:

$$\begin{aligned} \bar{p}^{vis} \cos \bar{\theta} &= \gamma \beta E^{vis} + \gamma p^{vis} \cos \theta \\ \Rightarrow \gamma &= \frac{E^{vis}[(E^{vis})^2 + (\bar{p}^{vis} \cos \bar{\theta})^2 - (p^{vis} \cos \theta)^2]^{1/2} - p^{vis} \cos \theta \bar{p}^{vis} \cos \bar{\theta}}{(E^{vis})^2 - (p^{vis} \cos \theta)^2}, \\ E^\tau &= \gamma m_\tau \end{aligned}$$

The energy of the tau lepton in the laboratory frame as function of the measured visible momentum depends on two of the three parameters only - the rest frame opening angle θ and the invariant mass $m_{\nu\nu}$ of the neutrino system. The direction of the tau lepton momentum vector is not fully determined by θ and $m_{\nu\nu}$, but is constrained to lie on the surface of a cone of opening angle $\bar{\theta}$ (given by equation 4.4), the axis of which is given by the visible momentum vector. The tau lepton four-vector is fully determined by the addition of the

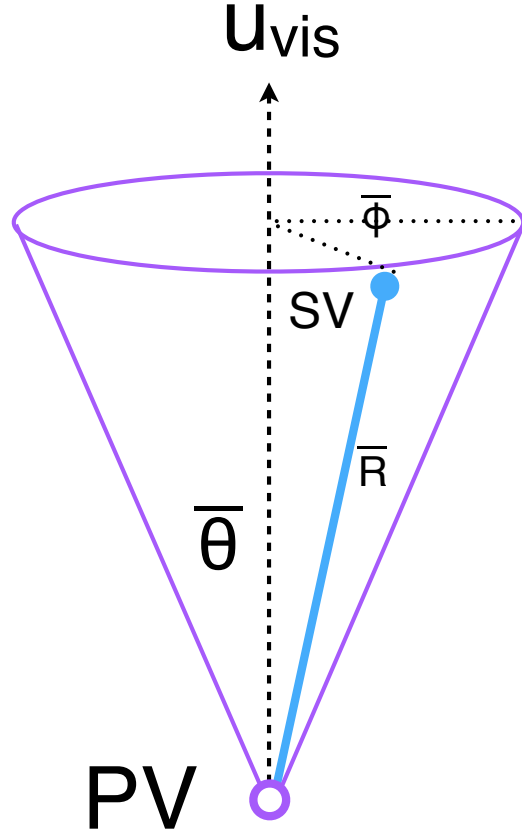


Figure 4.1: Illustration of the coordinate system used by the SVfit to describe the decays of tau leptons.

third parameter $\bar{\phi}$, which describes the azimuthal angle of the tau lepton with respect to the visible momentum vector. The spatial coordinate system used is illustrated in Figure 4.1.

§4.4 Likelihood for tau decay

The probability density functions for the tau decay kinematics are taken from the kinematics review of the PDG [16]. The likelihood is proportional to the phase-space volume for two-body ($\tau \rightarrow \tau_{had}\nu$) and three-body ($\tau \rightarrow e\nu\nu$ and $\tau \rightarrow \mu\nu\nu$) decays. For two-body decays the likelihood depends on the decay angle θ only:

$$d\Gamma \propto |\mathcal{M}|^2 \sin\theta d\theta$$

For three-body decays, the likelihood depends on the invariant mass of the neutrino system also:

$$d\Gamma \propto |\mathcal{M}|^2 \frac{((m_\tau^2 - (m_{\nu\nu} + m_{vis})^2)(m_\tau^2 - (m_{\nu\nu} - m_{vis})^2))^{1/2}}{2m_\tau} m_{\nu\nu} dm_{\nu\nu} \sin\theta d\theta \quad (4.5)$$

In the present implementation of the SVfit algorithm, the matrix element is assumed to be constant, so that the likelihood depends on the phase-space volume of the decay only ².

§4.4.1 Likelihood for reconstructed missing transverse momentum

Momentum conservation in the plane perpendicular to the beam axis implies that the vectorial sum of the momenta of all neutrinos produced in the decay of the tau lepton pair matches the reconstructed missing transverse momentum. Differences are possible due to the experimental resolution and finite P_T of particles escaping detection in beam direction at high $|\eta|$.

The E_T^{miss} resolution is measured in $Z \rightarrow \mu^+\mu^-$ events selected in the 7 TeV data collected by CMS in 2010. Corrections are applied to Monte Carlo simulated events to match the resolution measured in data. The momentum vectors of reconstructed E_T^{miss} and neutrino momenta given by the fit parameters are projected in direction parallel and perpendicular to the direction of the $\tau^+\tau^-$ momentum vector. For both components, a Gaussian probability function is assumed. The width and mean values of the Gaussian in parallel (“||”) and perpendicular (“⊥”) direction are:

$$\sigma_{||} = \max(7.54(1 - 0.00542 \cdot q_T), 5.)$$

$$\mu_{||} = -0.96$$

$$\sigma_{\perp} = \max(6.85(1 - 0.00547 \cdot q_T), 5.)$$

$$\mu_{\perp} = 0.0,$$

where q_T denotes the transverse momentum of the tau lepton pair.

§4.4.2 Likelihood for tau lepton transverse momentum balance

The tau lepton transverse momentum balance likelihood term represents the probability $p(P_T^\tau | M_{\tau\tau})$ for a tau to have a certain P_T , given that the tau is produced in the decay of

²The full matrix elements for tau decays may be added in the future, including terms for the polarization of the tau lepton pair, which is different in Higgs and Z decays [?].

a resonance of mass $M_{\tau\tau}$. The likelihood is constructed by parametrizing the shape of the tau lepton P_T distribution in simulated $\text{Higgs} \rightarrow \tau^+\tau^-$ events as a function of the Higgs mass. The functional form of the parametrization is taken to be the sum of two terms. The first term, denoted by $p^*(P_T|M)$, is derived by assuming an isotropic two-body decay, that is

$$dp^* \propto \sin\theta d\theta.$$

Performing a variable transformation from θ to $P_T \sim \frac{M}{2} \sin\theta$, we obtain

$$\begin{aligned} p^*(P_T|M) &= \frac{dp}{dP_T} = \frac{dp}{d\cos\theta} \left| \frac{d\cos\theta}{dP_T} \right| \\ &\propto \left| \frac{d}{dP_T} \sqrt{1 - \left(2\frac{P_T}{M}\right)^2} \right| \\ &= \frac{1}{\sqrt{\left(\frac{M}{2P_T}\right)^2 - 1}}. \end{aligned} \quad (4.6)$$

The first term of the P_T -balance likelihood is taken as the convolution of equation 4.6 with a Gaussian of width s . The second term is taken to be a Gamma distribution of scale parameter θ and shape parameter k , in order to account for tails in the P_T distribution of the tau lepton pair. The complete functional form is thus given by

$$p(P_T|M) \propto \int_0^{\frac{M}{2}} p^*(P'_T|M) e^{-\frac{(P_T-P'_T)^2}{2s^2}} dP'_T + a\Gamma(P_T, k, \theta). \quad (4.7)$$

Numerical values of the parameters s , θ and k are determined by fitting function 4.7 to the tau lepton P_T distribution in simulated $\text{Higgs} \rightarrow \tau^+\tau^-$ events. The relative weight a of the two terms is also determined in the fit. Replacing the integrand in equation 4.7 by its Taylor expansion, so that the integration can be carried out analytically, keeping polynomial terms up to fifth order, and assuming the fit parameters to depend at most linearly on the Higgs mass, we obtain the following numerical values for the parameters:

$$s = 1.8 + 0.018 \cdot M_{\tau\tau}$$

$$k = 2.2 + 0.0364 \cdot M_{\tau\tau}$$

$$\theta = 6.74 + 0.02 \cdot M_{\tau\tau}$$

$$a = 0.48 - 0.0007 \cdot M_{\tau\tau}.$$

The motivation to add the P_T -balance likelihood to the SVfit is to add a “regularization” term which compensates for the effect of P_T cuts applied on the visible decay products

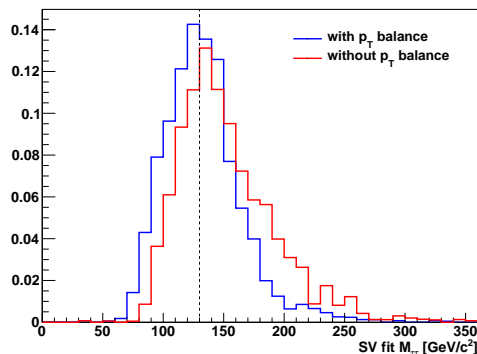


Figure 4.2: Distribution of di-tau invariant mass reconstructed by the SVfit algorithm in simulated Higgs events with $M_A = 130 \text{ GeV}/c^2$. The SVfit algorithm is run in two configurations, with (blue) and without (red) the P_T -balance likelihood term included in the fit.

of the two tau leptons. In particular for tau lepton pairs produced in decays of resonances of low mass, the visible P_T cuts significantly affect the distribution of the visible momentum fraction $x = \frac{E_{vis}}{E_\tau}$. The effect is illustrated in figures 4.3 and 4.4. If no attempt would be made to compensate for this effect, equations 4.4, 4.5 would yield likelihood values that are too high at low x , resulting in the SVfit to underestimate the energy of visible decay products (overestimate the energy of neutrinos) produced in the tau decay, resulting in a significant tail of the reconstructed mass distribution in the high mass region. The $\tau^+\tau^-$ invariant mass distribution reconstructed with and without the P_T -balance likelihood term is shown in figure 4.2. A significant improvement in resolution and in particular a significant reduction of the non-Gaussian tail in the region of high masses is seen.

§4.4.3 Secondary vertex information

The parametrization of the tau decay kinematics described in section 4.3 can be extended to describe the production and decay of the tau. As the flight direction of the tau is already fully determined by the parameters θ , $\bar{\phi}$ and $m_{\nu\nu}$, the position of the secondary (decay) vertex is hence fully determined by addition of a single parameter for the flight distance, r . The tau lifetime $c\tau = 87 \text{ }\mu\text{m}$ is large enough to allow the displacement of the tau decay vertex from the primary event vertex to be resolved by the CMS tracking detector. The resolution provided by the CMS tracking detector is utilized to improve the resolution on the

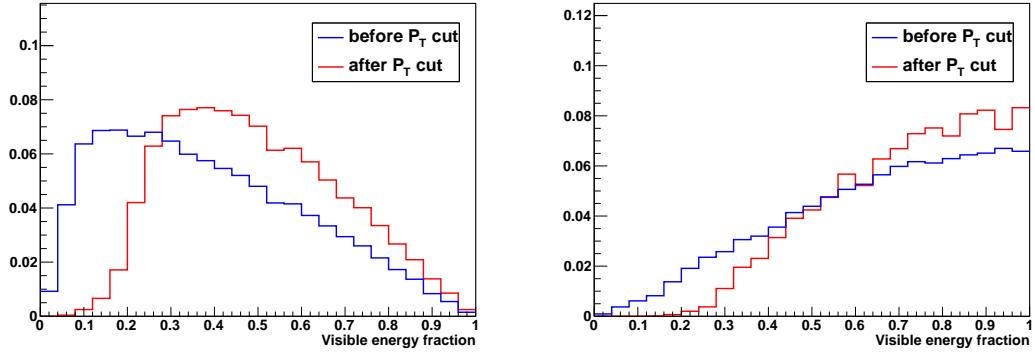


Figure 4.3: Normalized distributions of the fraction of total tau decay energy carried by the muon (left) and hadronic constituents (right) in simulated Higgs events with $M_A = 130 \text{ GeV}/c^2$. The distribution is shown before (blue) and after (red) the requirement on the P_T of the visible decay products described in section ??.

$\tau^+\tau^-$ invariant mass reconstructed by the SVfit algorithm. The likelihood term based on the secondary vertex information is based on the compatibility of the decay vertex position with the reconstructed tracks of charged tau decay products. Perhaps surprisingly, it turns out that the flight distance parameter R is sufficiently constrained even for tau decays into a single charged hadron, electron or muon.

The parameter R can be constrained further by a term which represents the probability for a tau lepton of momentum P to travel a distance d before decaying:

$$p(d|P) = \frac{m_\tau}{P c \tau} e^{-\frac{m_\tau d}{P c \tau}}$$

The likelihood terms for the secondary vertex fit have been implemented in the SVfit algorithm. In the analysis presented in this note, the decay vertex information is not used, however, because of systematic effects arising from tracker (mis-)alignment which are not yet fully understood.

§4.5 Performance

This section describes the performance of the SVfit algorithm for reconstructing the invariant mass of resonances decaying to τ lepton pairs. The performance is presented as a contrasted to the performance of the

Both of these effects are illustrated in Figure 4.5. Blah blah blah see figure 4.6.

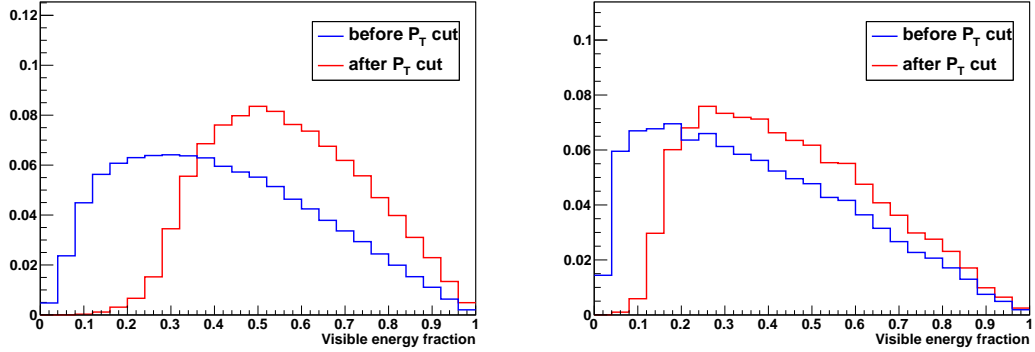


Figure 4.4: Normalized distributions of the fraction of total tau decay energy carried by the muon in simulated $Z \rightarrow \tau^+\tau^-$ (left) and Higgs events with $M_A = 200 \text{ GeV}/c^2$ (right). The distribution is shown before (blue) and after (red) the requirement that the P_T of the muon be greater than $15 \text{ GeV}/c$.

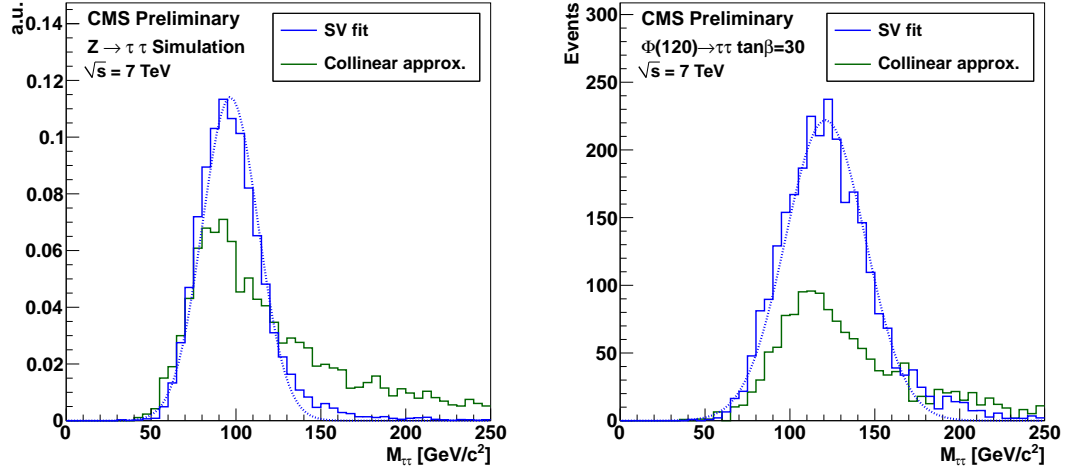


Figure 4.5: Comparison of the reconstructed tau pair mass spectrum in $Z \rightarrow \tau^+\tau^-$ (left) and MSSM $H(120) \rightarrow \tau^+\tau^-$ (right) events after the selections described in chapter 5. The mass spectrum reconstructed by the Secondary Vertex fit is shown in blue, the result of the collinear approximation algorithm is given in green. In the left plot, both distributions are normalized to unity, illustrating the improvement in resolution (shape) provided by the SVfit. In the right plot, the distributions are normalized to an (arbitrary) luminosity, illustrating the loss of events that occurs due to unphysical solutions in the application of the collinear approximation.

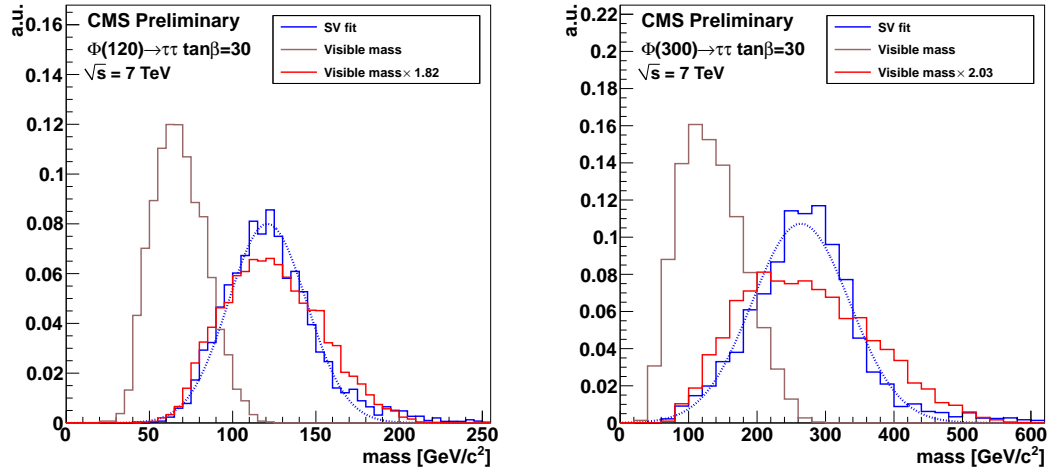


Figure 4.6: Comparison of the invariant mass of the muon and τ_{jet} (the “visible mass”) with the full $\tau^+\tau^-$ mass reconstructed by the SVfit. The spectrum is shown for two simulated MSSM Higgs samples, with $M_{A^0} = 120 \text{ GeV}/c^2$ (left), and $M_{A^0} = 200 \text{ GeV}/c^2$ (right). To illustrate that relative resolution of the SVfit is superior to that of the visible mass, the visible mass is also shown scaled up such that the mean of the two distributions are identical.

Chapter 5

Analysis Selections

Events in the muon plus tau-jet channel are selected by requiring a muon of $P_T^\mu > 15$ GeV within $|\eta_\mu| < 2.1$ and a tau-jet candidate of $P_T^{\tau\text{-jet}} > 20$ GeV within $|\eta_{\tau\text{-jet}}| < 2.3$. The muon and tau-jet candidate are required to be of opposite charge, as expected for the $Z \rightarrow \tau^+\tau^-$ signal. The muon is required to be reconstructed as global muon and pass the VBTF muon identification criteria [?]. Furthermore, the muon is required to be isolated with respect to charged hadrons of $P_T > 1.0$ GeV and neutral electromagnetic objects of $P_T > 1.5$ GeV (reconstructed by the particle-flow algorithm) within a cone of size $\Delta R = 0.4$ around the muon direction. The tau-jet candidate is required to pass the “medium” TaNC tau identification discriminator.

Additional event selection criteria are applied to reduce contributions of background processes. $Z \rightarrow \mu^+\mu^-$ background contributions are largely due to muons which failed to get reconstructed as global muons and are misidentified as tau-jet candidates. These muons are typically isolated and have a large chance to pass the TaNC tau ID discriminators. In order to reject this background, a dedicated discriminator against muons is applied [?]. Residual muon background is suppressed by rejecting tau-jet candidates which have a track of $P_T > 15$ GeV and for which the sum of energy deposits in ECAL plus HCAL is below $0.25 \cdot P$ within a cylinder of radius of radius 15 cm (ECAL) and 25 cm (HCAL), respectively. The $t\bar{t}$, $W + \text{jet}$ and QCD multi-jet backgrounds are suppressed by cuts on the transverse mass and the P_{zeta} variable. Contamination from $Z \rightarrow \tau^+\tau^-$ events in which the reconstructed tau-jet candidate is due to a $\tau \rightarrow e\nu\nu$ decay is reduced by applying a dedicated tau ID discriminator against electrons.

The complete set of event selection criteria applied in the muon + tau-jet channel are summarized in table 5.1.

Requirement	
Trigger	HLT_Mu9 for MC <i>cf.</i> table 5.2 for Data
Vertex	reconstructed with beam-spot constraint: $-24 < z_{vtx} < +24$ cm, $ \rho < 2$ cm, nDoF > 4
Muon	reconstructed as global Muon with: $P_T > 15$ GeV, $ \eta < 2.1$, VBTF Muon ID passed, isolated within $\Delta R = 0.4$ cone with respect to charged hadrons of $P_T > 1.0$ GeV and neutral electromagnetic objects of $P_T > 1.5$ GeV
Tau-jet Candidate	reconstructed by HPS + TaNC combined Tau ID algorithm TaNC “medium” Tau ID discriminator and discriminators against electrons and muons passed, calorimeter muon rejection passed
Muon + Tau-jet	charge(Muon) + charge(Tau-jet) = 0, $\Delta R(\text{Muon}, \text{Tau-jet}) > 0.5$
Kinematics	$M_T(\text{Muon-MET}) < 40$ GeV $P_\zeta - 1.5 \cdot P_\zeta^{vis} > -20$ GeV

Table 5.1: Event selection criteria applied in the muon + tau-jet channel.

The events are triggered by a combination of muon and muon + tau-jet “cross-channel” triggers. For the muon triggers, paths with lowest P_T thresholds are used as long as the path remained unrescaled (see table 5.2). The muon + tau-jet “cross-channel” trigger paths increase the trigger efficiency for events containing muons of transverse momenta close to the $P_T^\mu > 15$ GeV cut threshold. The trigger efficiency is measured in data via the tag-and-probe technique. Details of the muon trigger efficiency measurement are given in section ?? of the appendix. Monte Carlo simulated events are required to pass the HLT_Mu9 trigger path. Weights are applied to simulated events to account for the difference between the simulated HLT_Mu9 efficiency and the combined efficiency of the set HLT_Mu9, HLT_IsoMu9, HLT_Mu11, HLT_IsoMu13, HLT_Mu15, HLT_IsoMu9_PFTau15 and HLT_Mu11_PFTau15 used to trigger the data.

Trigger path	run-range
HLT_Mu9	132440 - 147116
HLT_IsoMu9	147196 - 148058
HLT_Mu11	147196 - 148058
HLT_Mu15	147196 - 149442
HLT_IsoMu13	148822 - 149182
HLT_IsoMu9_PFTau15	148822 - 149182
HLT_Mu11_PFTau15	148822 - 149182

Table 5.2: Muon and muon + tau-jet “cross-channel” trigger paths utilized to trigger events in the muon + tau-jet channel in different data-taking periods.

Chapter 6

Data–Driven Background Estimation

§6.1 Introduction

§6.2 The Fake–rate Method

In this note, we describe how knowledge of the probabilities with which quark and gluon jets get misidentified as tau–jets may be utilized to obtain an estimate of background contributions in physics analyses. As an illustrative example and in order to demonstrate the precision achievable with the method, we present results of applying the fake–rate technique to estimate the contributions of QCD, $W + \text{jets}$, $t\bar{t} + \text{jets}$ and $Z \rightarrow \mu^+\mu^-$ backgrounds in the measurement of the $Z \rightarrow \tau^+\tau^-$ cross–section, in the channel $Z \rightarrow \tau^+\tau^- \rightarrow \mu + \tau\text{-jet}$. Details of the analysis can be found in reference [?].

The results described in this note were obtained from Monte Carlo simulations of the $Z \rightarrow \tau^+\tau^-$ signal and different background processes for a centre–of–mass energy of $\sqrt{s} = 7$ TeV. Analysis of the $\sqrt{s} = 7$ TeV data recorded in 2009 [?] indicate that the probabilities for quark and gluon jets to fake the signatures of hadronic tau decays are well modeled by the Monte Carlo simulation. Once data–samples of sufficient event statistics are available at collision energies of $\sqrt{s} = 7$ TeV, fake–rates at the higher centre–of–mass energy will be measured in data and the values obtained from data will henceforth be used for the purpose of estimating background contributions via the fake–rate technique.

§6.2.1 Parameterization of fake–rates

Efficiencies and fake–rates of the tau identification algorithm based on requiring no tracks of $P_T > 1$ GeV and ECAL energy deposits of $P_T > 1.5$ GeV reconstructed within an “isolation

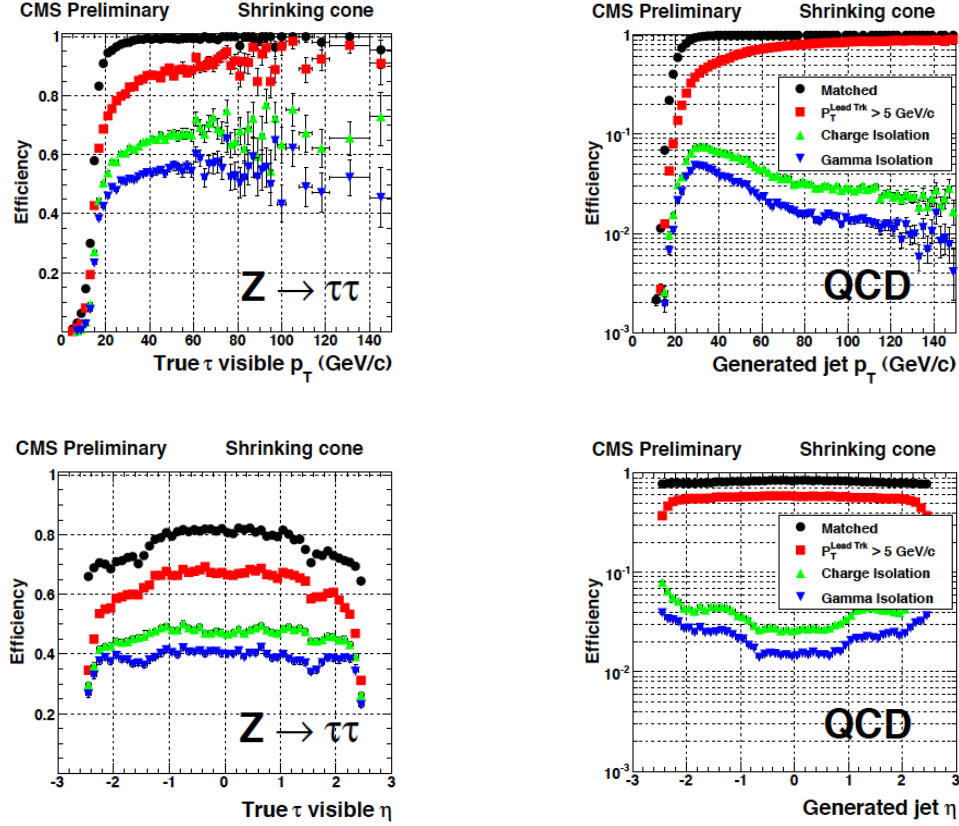


Figure 6.1: Cumulative efficiencies (left) and fake-rates (right) of successively applied tau identification cuts of the “shrinking signal cone” particle-flow based tau identification algorithm described in [?] as function of P_T^{jet} (top) and η_{jet} (bottom) of tau-jet candidates. The efficiencies/fake-rates for the complete set of tau identification criteria are represented by the blue (downwards facing) triangles.

cone” of size $dR_{iso} = 0.5$ and outside of a “shrinking signal cone” of size $dR_{sig} = 5.0/E_T$ as it is used in the $Z \rightarrow \tau^+\tau^- \rightarrow \mu + \tau$ -jet analysis are displayed in figure 6.1. In order to account for the visible P_T and η dependence, we parametrize the fake-rates in bins of transverse momentum and pseudo-rapidity. As we will show in section ??, the parametrization of the fake-rates by P_T and η makes it possible to not only estimate the total number of background events contributing to physics analyses, but to model the distributions of kinematic observables with a precision that is sufficient to extract information on the background shape.

We add a third quantity, the E_T -weighted jet-width R_{jet} , to the parametrization in order to account for differences between the fake-rates of quark and gluon jets. The jet

width is defined as

$$R_{jet} = \sqrt{E(\eta^2) + E(\phi^2)} \quad (6.1)$$

where $E(\eta^2)$ ($E(\phi^2)$) is the second η (ϕ) moment of the jet constituents, weighted by constituent transverse energy. Analyses performed by the CDF collaboration [?] found that systematic uncertainties on background estimates obtained from the fake-rate method are reduced in case differences between quark and gluon jets get accounted for in this way.

§6.2.2 Measurement of fake-rates

Efficiencies and fake-rates are then obtained by counting the fraction of tau-jet candidates passing all tau identification cuts and discriminators in a given bin of P_T^{jet} , η_{jet} and R_{jet} :

$$P_{fr} \left(P_T^{jet}, \eta_{jet}, R_{jet} \right) := \frac{N_{jets} \left(P_T^{jet}, \eta_{jet}, R_{jet} | \text{all tau ID cuts and discriminators passed} \right)}{N_{jets} \left(P_T^{jet}, \eta_{jet}, R_{jet} | \text{preselection passed} \right)} \quad (6.2)$$

The pre-selection in the denominator of equation 6.2 in general refers to P_T and η cuts, which are applied with thresholds matching those applied on the final analysis level, but may as well include loose tau identification criteria (which may be applied e.g. already during event skimming).

Different sets of fake-rates are determined for the highest P_T and for the second highest P_T jet in QCD di-jet events, for jets in a QCD event sample enriched by the contribution of heavy quarks and gluons by requiring the presence of a muon reconstructed in the final state, and for jets in “electroweak” events selected by requiring a W boson in the final state.

§6.2.3 The Fake-rate method

Knowledge of the tau identification efficiencies and fake-rates as function of the parameters P_T^{jet} , η_{jet} and R_{jet} as defined by equation 6.2 is utilized to obtain an estimate for the contributions of background processes to physics analyses involving tau lepton hadronic decays in the final state. The basic idea is to replace tau identification cuts and discriminators by appropriately chosen weights.

Application of the fake-rate technique consists of two stages. The first stage consists of loosening the tau identification cuts and discriminators and applying only the preselection requirements defined by the denominator of equation 6.2, in order to obtain an event sample

dominated by contributions of background processes, which are expected to increase by the inverse of the (average) fake-rate, typically by a factor $\mathcal{O}(100)$. In the second stage, weights are applied to all events in the background dominated control sample, according to the probabilities $P_{fr}\left(P_T^{jet}, \eta_{jet}, R_{jet}\right)$ for jets to fake the signature of a hadronic tau decay. After application of the weights, an estimate for the total number of background events passing the tau identification cuts and discriminators and thus contributing to the final analysis sample is obtained.

The fake-rate technique works best if all background contributions to the analysis arise from misidentification of quark and gluon jets as hadronic tau decays. Corrections to the estimate obtained from the fake-rate technique are needed in case of background processes contributing to the final analysis sample which either produce genuine tau leptons in the final state (e.g. $t\bar{t}$ + jets) or in which tau-jet candidates are due to misidentified electrons or muons (e.g. $Z \rightarrow \mu^+\mu^-$, $Z \rightarrow e^+e^-$), as the latter may fake signatures of hadronic tau decays with very different probabilities than quark and gluon jets.

In the “simple” fake-rate method described in more detail in the next section, the corrections are taken from Monte Carlo simulations. Corrections based on Monte Carlo are needed also to compensate for signal contributions to the background dominated control sample.

An alternative to Monte Carlo based corrections is to utilize additional information contained in the background dominated control sample. The modified version is described in section 6.2.3. It has been used to estimate background contributions in searches for Higgs boson production with subsequent decays into tau lepton pairs performed by the CDF collaboration in TeVatron run *II* data [?]. We will refer to the modified version as “CDF-type” method in the following.

“Simple” weight method

In the “simple” method all tau-jet candidates within the background dominated event sample are weighted by the probabilities of quark and gluon jets to fake the signature of a hadronic tau decay:

$$w_{jet}^{simple}\left(P_T^{jet}, \eta_{jet}, R_{jet}\right) := P_{fr}\left(P_T^{jet}, \eta_{jet}, R_{jet}\right) \quad (6.3)$$

These weights are applied to all jets in the background dominated control sample which pass the preselection defined by the denominator of equation 6.2. Note that the weights defined by equation 6.3 can be used to estimate the contributions of background processes to distributions of tau-jet related observables. They cannot be used as event weights.

In order to compare distributions of event level quantities or per-particle quantities for particles of types different from tau leptons decaying hadronically, event weights need to be defined. Neglecting the small fraction of background events in which multiple tau-jet candidates pass the complete set of all tau identification cuts and discriminators, event weights can be computed by summing up the per-jet weights defined by equation 6.3 over all tau-jet candidates in the event which pass the preselection:

$$W_{event}^{simple} := \sum w_{jet}^{simple} \quad (6.4)$$

A bit of care is needed in case one wants to compare distributions of observables related to “composite particles” the multiplicity of which depends on the multiplicity of tau-jet candidates in the event (e.g. combinations of muon + tau-jet pairs in case of the $Z \rightarrow \tau^+ \tau^- \rightarrow \mu + \tau$ -jet analysis). Per-particle weights need to be computed for such “composite particles”, depending on P_T^{jet} , η_{jet} , R_{jet} of its tau-jet candidate constituent, according to:

$$w_{comp-part}^{simple} \left(P_T^{jet}, \eta_{jet}, R_{jet} \right) := w_{jet}^{simple} \left(P_T^{jet}, \eta_{jet}, R_{jet} \right) \quad (6.5)$$

Different estimates are obtained for the fake-rate probabilities determined for the highest and second highest P_T jet in QCD di-jet events, jets in a muon enriched QCD sample and jets in W +jets events. The arithmetic average of the four estimates together with the difference between the computed average and the minimum/maximum value is given in table ??.

We take the average value as “best” estimate of the background contribution and the difference between the average and the minimum/maximum estimate as its systematic uncertainty. We obtain a value of $\mathcal{O}(15\%)$ for the systematic uncertainty and find that the true sum of QCD, W +jets, $t\bar{t}$ +jets and $Z \rightarrow \mu^+ \mu^-$ background contributions agrees well with the “best” estimate obtained by the fake-rate method within the systematic uncertainty.

Note that the estimate for the sum of background contributions which one obtains in

case one applies the “simple” fake-rate weights defined by equation 6.4 to a background dominated control sample selected in data is likely to overestimate the true value of background contributions by a significant amount. The reason is that contributions of the $Z \rightarrow \tau^+ \tau^-$ signal are non-negligible. In fact, signal contributions to the background dominated control sample are expected to be 14.9% and since the per-jet weights computed by equation 6.3 are larger on average in signal than in background events, the signal contribution increases by the weighting and amounts to 37.1% of the sum of event weights computed by equation 6.4 and given in table ??.

The contribution of the $Z \rightarrow \tau^+ \tau^-$ signal needs to be determined by Monte Carlo simulation and subtracted from the estimate obtained by applying the “simple” fake-rate method to data, in order to get an unbiased estimate of the true background contributions.

“CDF-type” weights

Instead of subtracting from the estimate obtained for the sum of background contributions a correction determined by Monte Carlo simulation, the signal contribution to the background dominated event sample selected in data can be corrected for by adjusting the weights, based solely on information contained in the analyzed data sample, hence avoiding the need to rely on Monte Carlo based corrections.

In the “CDF-type” method, additional information, namely whether or not tau-jet candidates pass or fail the tau identification cuts and discriminators, is drawn from the data. The desired cancellation of signal contributions is achieved by assigning negative weights to those tau-jet candidates which pass all tau identification cuts and discriminators, i.e. to a fair fraction of genuine hadronic tau decays, but to a small fraction of quark and gluon jets only. The small reduction of the background estimate by negative weights assigned to quark and gluon jets is accounted for by a small increase of the positive weights assigned to those tau-jet candidates for which at least one of the tau identification cuts or discriminators fails. In this way, an unbiased estimate of the background contribution is maintained.

To be specific, the “CDF-type” weights assigned to tau-jet candidates are computed

as (for a derivation of the formula, see section ?? of the appendix):

$$w_{jet}^{CDF} \left(P_T^{jet}, \eta_{jet}, R_{jet} \right) := \begin{cases} \frac{P_{fr}(P_T^{jet}, \eta_{jet}, R_{jet}) \cdot \varepsilon(P_T^{jet}, \eta_{jet}, R_{jet})}{\varepsilon(P_T^{jet}, \eta_{jet}, R_{jet}) - P_{fr}(P_T^{jet}, \eta_{jet}, R_{jet})} & \text{if all tau ID cuts and discriminators passed} \\ \frac{P_{fr}(P_T^{jet}, \eta_{jet}, R_{jet}) \cdot (1 - \varepsilon(P_T^{jet}, \eta_{jet}, R_{jet}))}{\varepsilon(P_T^{jet}, \eta_{jet}, R_{jet}) - P_{fr}(P_T^{jet}, \eta_{jet}, R_{jet})} & \text{otherwise} \end{cases} \quad (6.6)$$

The basic idea of the “CDF-type” weights is to assign negative (positive) weights to tau-jet candidates passing all tau identification cuts and discriminators (failing at least one cut or discriminator), such that signal contributions of genuine hadronic tau decays to the background dominated control sample on average cancel after application of the weights, while providing an unbiased estimate of the contribution of background processes arising from misidentification of quark and gluon jets.

For the derivation of equation 6.6 for the “CDF-type” weights assigned to tau-jet candidates, we will use the following notation: Let n_τ (n_{QCD}) denote the total number of tau-jets (quark and gluon jets) in a certain bin of transverse momentum P_T^{jet} , pseudo-rapidity η_{jet} and jet-width R_{jet} and n_τ^{sel} (n_{QCD}^{sel}) denote the number of tau-jets (quark and gluon jets) in that bin which pass all tau identification cuts and discriminators.

By definition of the tau identification efficiency $\varepsilon := \varepsilon(P_T^{jet}, \eta_{jet}, R_{jet})$ and fake-rate $f := f(P_T^{jet}, \eta_{jet}, R_{jet})$:

$$\begin{aligned} n_\tau^{sel} &= \varepsilon \cdot n_\tau \\ n_{QCD}^{sel} &= f \cdot n_{QCD}. \end{aligned} \quad (6.7)$$

Depending on whether or not a given tau-jet candidate passes all tau identification cuts and discriminators or not, we will assign a weight of value w_{passed} or w_{failed} to it.

The values of the weights w_{passed} and w_{failed} shall be adjusted such that they provide an unbiased estimate of the background contribution:

$$w_{passed} \cdot f \cdot n_{QCD} + w_{failed} \cdot (1 - f) \cdot n_{QCD} \equiv n_{QCD}^{sel} = f \cdot n_{QCD} \quad (6.8)$$

while averaging to zero for genuine hadronic tau decays:

$$w_{passed} \cdot \varepsilon \cdot n_\tau + w_{failed} \cdot (1 - \varepsilon) \cdot n_\tau \equiv 0.$$

The latter equation yields the relation:

$$w_{passed} = -\frac{1 - \varepsilon}{\varepsilon} \cdot w_{failed}, \quad (6.9)$$

associating the two types of weights. By inserting relation 6.9 into equation 6.8 we obtain:

$$\begin{aligned}
& -\frac{1-\varepsilon}{\varepsilon} \cdot w_{failed} \cdot f \cdot n_{QCD} + w_{failed} \cdot (1-f) \cdot n_{QCD} = f \cdot n_{QCD} \\
\Rightarrow & \left(\frac{-f + \varepsilon \cdot f + \varepsilon - f \cdot \varepsilon}{\varepsilon} \right) \cdot w_{failed} = f \\
\Rightarrow & w_{failed} = \frac{f \cdot \varepsilon}{\varepsilon - f}
\end{aligned}$$

and

$$w_{passed} = -\frac{f \cdot (1-\varepsilon)}{\varepsilon - f} \quad (6.10)$$

which matches exactly equation 6.6 for the “CDF-type” weights applied to tau-jet candidates given in section 6.2.3.

Event weights and the weights assigned to “composite particles” are computed in the same way as for the “simple” weights, based on the weights assigned to the tau-jet candidates:

$$\begin{aligned}
W_{event}^{CDF} &:= \Sigma w_{jet}^{CDF} \\
w_{comp-part}^{CDF} \left(P_T^{jet}, \eta_{jet}, R_{jet} \right) &:= w_{jet}^{CDF} \left(P_T^{jet}, \eta_{jet}, R_{jet} \right), \quad (6.11)
\end{aligned}$$

where the sums extend over all jets in the background dominated control sample which pass the preselection defined by the denominator of equation 6.2.

The effect of the negative weights to compensate the positive weights in case the “CDF-type” fake-rate method is applied to signal events containing genuine hadronic tau decays is shown in table ???. As expected, positive and negative weights do indeed cancel in the statistical average.

Figures ??, ?? and ?? demonstrate that an unbiased estimate of the background contribution by the “CDF-type” weights is maintained. Overall, the estimates obtained are in good agreement with the contributions expected for different background processes, indicating that the adjustment of negative and positive weights works as expected for the background as well.

Results obtained by the “CDF-type” fake-rate method are summarized in table ??, in which the total number of background events estimated by equation 6.11 is compared to the true background contributions. The “best” estimate of the background contribution obtained from the “CDF-type” method is again taken as the arithmetic average of the estimates obtained by applying the fake-rate probabilities for the highest and second highest

P_T jet in QCD di-jet events, jets in a muon enriched QCD sample and jets in W +jets events. Systematic uncertainties are taken from the difference between the computed average value and the minimum/maximum estimate. We obtain a value of $\mathcal{O}(15\text{--}20\%)$ for the systematic uncertainty of the “CDF-type” method, slightly higher than the systematic uncertainty obtained for the “simple” method. The small increase of systematic uncertainties is in agreement with our expectation for fluctuations of the jet-weights in case weights of negative and positive sign are used.

§6.2.4 k-Nearest Neighbor fake-rate calculation

§6.2.5 Results of Background Estimation

An independent estimate of the background contributions is obtained by applying the fake-rate method described in [22]. Fake-rates in QCD multi-jet events (light quark enriched sample), QCD events containing muons (heavy quark and gluon enriched sample) and W + jet events are measured in data [?, ?] and applied to events which pass all the event selection criteria listed in table 5.1 except for the requirement for tau-jet candidates to pass the “medium” tight TaNC discriminator and have unit charge.

No assumption is made on the composition of $Z \rightarrow \mu^+\mu^-$, W + jets, $t\bar{t}$ + jets and QCD backgrounds contributing to the event sample selected by the analysis. Differences between fake-rates obtained for QCD multi-jet, QCD muon enriched and W + jets background events are attributed as systematic uncertainties of the fake-rate method. Per jet and per event weights have been computed by the “simple” and “CDF-type” weights as described in [22] and the results are found to be compatible within statistical and systematic uncertainties. In the following, we present results for “CDF-type” weights. The “CDF-type” weights have the advantage that the background estimate obtained does not change, whether there is MSSM Higgs $\rightarrow \tau^+\tau^-$ signal present in the data or not.

Tau identification efficiencies need to be known when using “CDF-type” weights. Dedicated studies have checked the tau identification efficiencies in data [?, ?]. Statistical and systematic uncertainties of these studies are still sizeable at present, in the order to 20 – 30%. No indication has been found, however, that the Monte Carlo simulation would

Events weighted by:	Estimate
QCD leading jet	
QCD sub-leading jet	
QCD μ -enriched	
W + jets	
Fake-rate estimate	

Table 6.1: Estimate for background contributions in mu + tau channel obtained by weighting events passing all selection criteria listed in table 5.1 except for the requirement for tau-jet candidates to pass the “medium” tight TaNC discriminator and have unit charge by fake-rates measured in QCD multi-jet, QCD muon enriched and W + jets data samples.

not correctly model hadronic tau decays in data. For the purpose of computing fake-rate weights via the “CDF-type” method, tau identification efficiencies are taken from the Monte Carlo simulation of hadronic tau decays in $Z \rightarrow \tau^+\tau^-$ events. Systematic uncertainties on the background estimate obtained by the fake-rate method are determined by varying the tau identification efficiencies by $\pm 30\%$ relative to the value obtained from the Monte Carlo simulation.

The results of applying the fake-rate method to the mu + tau channel are summarized in table 6.1. The background prediction has been corrected for the expected contribution of $XX.X$ events from $Z \rightarrow \mu^+\mu^-$ background events in which the reconstructed tau-jet is due to a misidentified muon. The obtained estimate is in good agreement with the Monte Carlo expectation.

As an additional cross-check of the method, a sample of events containing a muon plus a tau-jet of like-sign charge is selected in data and compared to the background prediction obtained by applying the fake-rate method to the like-sign sample. The like-sign sample is expected to be dominated by the contributions of W + jets and QCD background processes and allows to verify the fake-rate method in a practically signal free event sample. The background estimate obtained by the fake-rate method is compared to the number of events observed in the like-sign data sample in table 6.2. The number of events expected in the like-sign control sample from Monte Carlo simulation is indicated in the caption. All numbers are in good agreement.

Events weighted by:	Estimate
QCD leading jet	
QCD sub-leading jet	
QCD μ -enriched	
W + jets	
Fake-rate estimate	
Observed	

Table 6.2: Number of events observed in like-sign control region compared to estimate obtained by fake-rate method. The number of observed events as well as the number of background events predicted by the fake-rate method is on good agreement with the Monte Carlo expectation of $XX.X$ events for the sum of $Z \rightarrow \mu^+\mu^-$, W + jets, $t\bar{t}$ + jets and QCD background contributions in the like-sign control region.

Figure 6.2: Distribution of visible mass (left) and “full” invariant mass reconstructed by the SVfit algorithm (right) observed in the like-sign charge control region compared to the background estimate obtained by the fake-rate method.

The fake-rate method does not only allow to estimate the total number of background events, but allows to model the distributions of background processes as well. The capability to model distributions is illustrated in figure 6.2, which shows good agreement between the distributions observed in the like-sign data sample and the predictions obtained by the fake-rate method for the distributions of muon plus tau-jet visible mass and of the “full” invariant mass reconstructed by the SVfit algorithm.

Chapter 7

Systematics

Chapter 8

Results

Conclusions

Bibliography

- [1] J. Goldstone, “Field Theories with Superconductor Solutions”, *Nuovo Cim.* **19** (1961) 154–164. doi:10.1007/BF02812722.
- [2] J. Goldstone, A. Salam, and S. Weinberg, “Broken Symmetries”, *Phys. Rev.* **127** (Aug, 1962) 965–970. doi:10.1103/PhysRev.127.965.
- [3] F. Englert and R. Brout, “Broken Symmetry and the Mass of Gauge Vector Mesons”, *Phys. Rev. Lett.* **13** (Aug, 1964) 321–323. doi:10.1103/PhysRevLett.13.321.
- [4] 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.
- [5] 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.
- [6] S. Glashow, “Partial Symmetries of Weak Interactions”, *Nucl.Phys.* **22** (1961) 579–588. doi:10.1016/0029-5582(61)90469-2.
- [7] S. Weinberg, “A Model of Leptons”, *Phys.Rev.Lett.* **19** (1967) 1264–1266. doi:10.1103/PhysRevLett.19.1264.
- [8] A. Salam, “Weak and Electromagnetic Interactions”,. Originally printed in *Svartholm: Elementary Particle Theory, Proceedings Of The Nobel Symposium Held 1968 At Lerum, Sweden*, Stockholm 1968, 367-377.
- [9] D. Griffiths, “Introduction to Elementary Particles”. Wiley-VCH, 2004.
- [10] S. M. T. Morii, C.S. Lim, “The Physics of the Standard Model and Beyond”. World Scientific, 2004.
- [11] UA1 Collaboration, “Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540 \text{ GeV}$ ”, *Phys. Lett.* **B122** (1983) 103–116.
- [12] UA2 Collaboration, “Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN $\bar{p}p$ collider”, *Phys. Lett.* **B122** (1983) 476–485. doi:10.1016/0370-2693(83)91605-2.
- [13] UA1 Collaboration, “Experimental observation of lepton pairs of invariant mass around $95 \text{ GeV}/c^2$ at the CERN SPS collider”, *Phys. Lett.* **B126** (1983) 398–410. doi:10.1016/0370-2693(83)90188-0.

- [14] UA2 Collaboration, “Evidence for $Z^0 \rightarrow e^+e^-$ at the CERN $\bar{p}p$ collider”, *Phys. Lett.* **B129** (1983) 130–140. doi:10.1016/0370-2693(83)90744-X.
- [15] S. P. Martin, “A Supersymmetry Primer”, *arXiv* **hep-ph** (sep, 1997). 128 pages. Version 5 (December 2008) contains a change in convention that flips the signs of sigma and sigmabar matrices. It also contains a total of about 2 pages of updates, mostly on supersymmetry breaking issues. Errata and a version with larger type (12 pt, 142 pages) can be found at <http://zippy.physics.niu.edu/primer.html>.
- [16] Particle Data Group Collaboration, “Review of particle physics”, *J. Phys.* **G37** (2010) 075021. doi:10.1088/0954-3899/37/7A/075021.
- [17]
- [18] CMS Collaboration, “CMS Strategies for tau reconstruction and identification using particle-flow techniques”, *CMS PAS* **CMS-PAS-PFT-08-001** (2008).
- [19] CMS Collaboration, “Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET”, *CMS PAS* **CMS-PAS-PFT-09-001** (2009).
- [20] A. Hoecker, P. Speckmayer, J. Stelzer et al., “TMVA - Toolkit for Multivariate Data Analysis”, *arXiv* **physics.data-an** (mar, 2007). Published in: PoSACAT:040,2007 TMVA-v4 Users Guide: 135 pages, 19 figures, numerous code examples and references.
- [21] A. Kolmogorov, “On the representation of continuous functions of several variables by superposition of continuous functions of one variable and addition”, *Doklady Akademiiia Nauk SSSR* **114** (1957).
- [22] J. Conway, E. Friis, and C. Veelken, “Estimation of Background contributions to Tau analyses via Fake-Rate technique”, *CMS Note* **2010/074** (2010).