

Search for Neutral Higgs Bosons Decaying to Pairs of 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

³ Abstract

⁴ blah blah blah

5 Acknowledgments

6 Hooray for everybody.

7 Table of Contents

8	1 The Standard Model and Beyond	1
9	1.1 The Standard Model	1
10	1.1.1 Quantum Electrodynamics and Gauge Invariance	2
11	1.1.2 The Weak Interactions	4
12	1.1.3 Spontaneous Symmetry Breaking	6
13	1.1.4 The Higgs Mechanism	8
14	1.1.5 Electroweak Unification	10
15	1.1.6 Quantum Chromodynamics	14
16	1.2 Beyond the Standard Model	16
17	1.2.1 The Hierarchy Problem	17
18	1.2.2 Supersymmetry	18
19	1.2.3 The Minimal Supersymmetric Model	18
20	1.3 Searches for the Higgs boson	21
21	1.3.1 Standard Model Higgs boson phenomenology	21
22	1.3.2 MSSM Higgs Phenomenology	25
23	1.3.3 Results from LEP and Tevatron	27
24	1.4 The physics of the τ lepton	27
25	2 The Compact Muon Solenoid Experiment	30
26	2.1 The Large Hadron Collider	31
27	2.2 Solenoid Magnet	33
28	2.3 Charged Particle Tracking Systems	34
29	2.4 Electromagnetic Calorimeter	36
30	2.5 Hadronic Calorimeter	38
31	2.6 Muon System	40
32	2.7 Trigger System	41
33	2.8 Particle Flow Reconstruction Algorithm	41
34	3 Tau Identification: The Tau Neural Classifier	42
35	3.1 Introduction	42
36	3.2 Geometric Tau Identification Algorithms	43
37	3.3 Decay Mode Tau Identification: Motivation	43
38	3.4 The Tau Neural Classifier	44
39	3.4.1 Decay mode reconstruction	45
40	3.4.2 Neural network classification	48
41	3.5 Summary	61
42	4 Mass Reconstruction: The Secondary Vertex Fit	67
43	4.1 Existing mass reconstruction algorithms	67

44	4.2	The Secondary Vertex fit	69
45	4.3	Parametrization of tau decays	69
46	4.4	Likelihood for tau decay	71
47	4.4.1	Likelihood for reconstructed missing transverse momentum	72
48	4.4.2	Likelihood for tau lepton transverse momentum balance	72
49	4.4.3	Secondary vertex information	74
50	4.5	Performance	75
51	5	Analysis Selections	78
52	6	Data-Driven Background Estimation	81
53	6.1	Introduction	81
54	6.2	The Fake-rate Method	81
55	6.2.1	Parameterization of fake-rates	82
56	6.2.2	Measurement of fake-rates	83
57	6.2.3	The Fake-rate method	83
58	6.2.4	k-Nearest Neighbor fake-rate calculation	89
59	6.2.5	Results of Background Estimation	89
60	7	Systematics	93
61	8	Results	94
62	9	Conclusions	95
63		Bibliography	95

64 List of Figures

65	1.1	Fermi contact interaction diagram	5
66	1.2	Muon decaying through intermediate gauge boson	5
67	1.3	QCD Feynman Diagrams	15
68	1.4	Loop corrections to Higgs mass	17
69	1.5	Higgstrahlung production diagram at e^+e^- colliders	22
70	1.6	Gluon fusion Higgs production diagram	22
71	1.7	Vector boson fusion Higgs production diagram	22
72	1.8	Parton luminosity comparison of the LHC and Tevatron	23
73	1.9	SM Higgs cross sections at the LHC	24
74	1.10	SM Higgs branching fractions	25
75	1.11	Cross sections of interest at hadron colliders	26
76	1.12	MSSM Higgs production with association b -quarks	27
77	1.13	MSSM Higgs cross sections at the LHC	28
78	2.1	Schematic drawings of the CMS detector	32
79	2.2	Material budget of the CMS tracker	35
80	2.3	Momentum and impact parameter resolutions of CMS tracker	36
81	2.4	Energy resolution of the CMS ECAL	38
82	2.5	The left figure, (a), illustrates the number of interaction lengths versus pseu-	
83		dorapidity η of material that must be traversed before reaching the different	
84		layers of the muon system. On the right, (b) shows the efficiency versus η to	
85		reconstruct a “global” muon for different transverse momenta.	41
86	3.1	Visible invariant mass of τ lepton decay products	44
87	3.2	Invariant mass photon pairs in reconstructed π^0 mesons	46
88	3.3	Neutral energy fraction in visible τ decays	47
89	3.4	Tau decay mode reconstruction performance	49
90	3.5	Kinematic dependence of decay mode reconstruction	50
91	3.6	Neural network over-training validation plots	53
92	3.7	Kinematic weighting of training sample	55
93	3.8	Neural network output in each decay mode	58
94	3.9	Performance curves for the neural networks used in the TaNC	59
95	3.10	Tau Neural Classifier performance curves for different p_T ranges	60
96	3.11	Tau Neural Classifier transformation performance	62
97	3.12	Transformed neural network output	62
98	3.13	Tau Neural Classifier performance comparison	64
99	3.14	Tau Neural Classifier kinematic performance	65
100	4.1	Coordinate system of the SVfit	71

101	4.2	Effect of p_T -balance term on SVfit performance	74
102	4.3	Effect of the visible p_T requirements on muon and hadronic τ decays	75
103	4.4	Effect of the visible p_T requirements for Z and Higgs events	76
104	4.5	Comparison of SVfit with the Collinear Approximation algorithm	76
105	4.6	Comparison of SVfit with the visible mass observable	77
106	6.1	p_T and η dependency of tau ID performance	82
107	6.2	Comparison of visible mass and SVfit mass	92

108 **List of Tables**

109	1.1	Chiral supermultiplets in the MSSM	19
110	1.2	Gauge supermultiplets in the MSSM	19
111	1.3	Decay modes of the τ lepton	29
112	3.1	Decay mode performance – naive reconstruction	48
113	3.2	Decay mode performance – TaNC reconstruction	48
114	3.3	Neural network training event statistics	52
115	3.4	Variables used in the different TaNC neural networks	66
116	5.1	Event selection criteria applied in the muon + tau-jet channel.	79
117	5.2	Muon and muon + tau-jet “cross-channel” trigger paths utilized to trigger	
118		events in the muon + tau-jet channel in different data-taking periods. . . .	80
119	6.1	Fake-rate method results	90
120	6.2	Yields in like-sign control region	91

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.

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

149 §1.1.1 Quantum Electrodynamics and Gauge Invariance

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

157 In the following, we first describe the Dirac equation for a free electron, which is the
 158 relativistic extension of the Schroedinger equation for spin 1/2 particles. We then show
 159 that requiring the corresponding Lagrangian of the free charged particle to be invariant
 160 under local gauge transformations creates an effective gauge boson field. This “gauge field”
 161 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

$$p^\mu p_\mu - m^2 c^2 = 0. \quad (1.1)$$

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

$$p_\mu \rightarrow i\hbar\partial_\mu \quad (1.2)$$

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

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

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 1.3 and making the substitution in Equation 1.2.

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

162 The solutions ψ of the Dirac equation are called “Dirac spinors,” and represent the quantum
163 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 $A^\mu = \partial^\mu\theta(x)$ 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:

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

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

169 §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]

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

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

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

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

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

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

FiXme: *check handedness is correct*

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

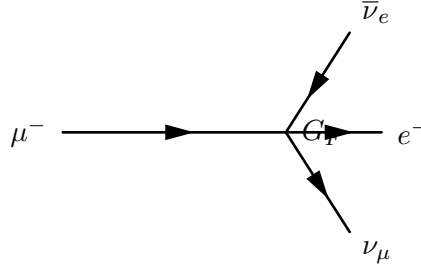


Figure 1.1: Feynman diagram of muon decay in Fermi contact interaction theory.

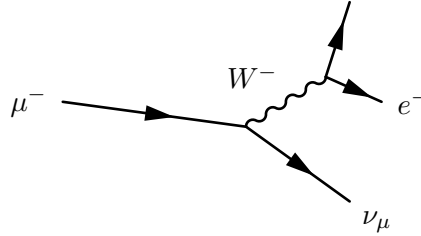


Figure 1.2: Feynman diagram of muon decay proceeding through an intermediate gauge boson W^- .

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

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

179 However, the contact interaction form of Fermi's theory is not complete. When applied
 180 to scattering processes, the interaction violates unitarity: the calculated cross section grows
 181 with the center of mass energy, so that for some energy the probability for an interaction
 182 is greater than one. Furthermore, the techniques successfully used to “renormalize”² QED
 183 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 pictured in Figure 1.2 with an amplitude given [10] by

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

The presence of the large gauge boson mass term M_W^2 in the denominator of the central

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].

term of Equation 1.13 is the reason why the contact interaction original formulated by Fermi effectively described 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 equivalent to the Fermi contact interaction in 1.12 as

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

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

191 §1.1.3 Spontaneous Symmetry Breaking

192 In the early 1960s Glashow, Weinberg, and Salam published a series of papers describing
 193 how the electromagnetic and weak forces could be unified into a common “electroweak”
 194 force. The fact that at low energy the electromagnetic and weak forces appear to be
 195 separate phenomena is due to the fact that the symmetry of the electroweak gauge group is
 196 “spontaneously broken.” Modern field theories (both the Standard Model and beyond) are
 197 predicated on the idea that the all interactions are part of a single, unified symmetry group
 198 and the differences between various scales (electromagnetic, weak, etc.) at lower energies
 199 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

$$\begin{aligned} V(\phi_1, \phi_2) &= V(\phi_1^{min}, \phi_2^{min}) + \sum_{a=1,2} \left(\frac{\partial V}{\partial \phi_a} \right)_0 (\phi_a - \phi_a^{min}) \\ &+ \frac{1}{2} \sum_{a,b=1,2} \left(\frac{\partial^2 V}{\partial \phi_a \partial \phi_b} \right)_0 (\phi_a - \phi_a^{min})(\phi_b - \phi_b^{min}) + \dots \end{aligned} \quad (1.16)$$

Since at the minimum the partial derivative of V is zero with respect to all fields, the second term in equation 1.16 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

$$\left(\frac{\partial^2 V}{\partial \phi_a \partial \phi_b} \right)_{\phi^{min}} \quad (1.17)$$

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

$$V(\phi_1^2\phi_2^2) = \frac{m^2}{2}(\phi_1^2 + \phi_2^2) + \frac{\lambda}{4}(\phi_1^2 + \phi_2^2)^2. \quad (1.18)$$

If the parameters m^2 and λ are both positive, then the minimum of V is at the origin ($\phi_1 = \phi_2 = 0$). In this case, the mass matrix term in Equation 1.16 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

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 $U(1)$.

212 m^2 parameter in Equation 1.18 is negative, the spectrum is dramatically different. After
 213 making the replacement $m^2 = -\mu^2$ ($\mu^2 > 0$), the extrema of V are no longer unique. The
 214 requirement of $\frac{\partial V}{\partial \phi_i} = 0$ for all i is satisfied in two cases:

$$(\phi_1^{min}, \phi_2^{min}) = (0, 0) \quad (1.19)$$

$$(\phi_1^{min})^2 + (\phi_2^{min})^2 = \frac{\mu^2}{\lambda} = \nu^2. \quad (1.20)$$

If the vacuum state is defined at the point in Equation 1.19, 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 1.20, 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 1.17 yields the mass matrix

$$\left(\frac{\partial^2 V}{\partial \phi_a \partial \phi_b} \right)_{\phi^{min}} = \begin{pmatrix} v^2 & 0 \\ 0 & 0 \end{pmatrix}.$$

Fixme:
check matrix

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

219 §1.1.4 The Higgs Mechanism

220 As in section 1.1.1, extending the gauge symmetry requirement to be *locally* invariant creates
 221 interesting consequences for models that have spontaneously broken symmetry. This gives
 222 rise to the “Higgs Mechanism,” which we overview here. For simplicity we will again consider
 223 a model with $U(1)$ symmetry. The model is identical to the one presented in section 1.1.3,
 224 with two exceptions. First, we express the two real fields ϕ_1 and ϕ_2 as a single complex-
 225 valued field ϕ . Second, the model is required to be locally $U(1)$ invariant, and so uses the
 226 gauge-covariant derivatives, minimal coupling to the gauge field, and contains the kinetic

⁴The point chosen for the VEV here is not arbitrary. One can choose any point that satisfies Equation 1.20 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.

term for the gauge field, as discussed in section 1.1.1. The unbroken Lagrangian is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + (D_\mu\phi^*)(D^\mu\phi) - V(\phi^*\phi) \quad (1.21)$$

$$V(\phi^*\phi) = -\mu^2\phi^*\phi + \lambda(\phi^*\phi)^2, \quad (1.22)$$

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

$$\begin{aligned} \phi \rightarrow \phi' &= e^{-i\alpha(x)}\phi \\ A_\mu \rightarrow A'_\mu &= A_\mu - \frac{1}{2}\partial_\mu\alpha(x). \end{aligned} \quad (1.23)$$

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

$$\phi(x) = \frac{1}{\sqrt{2}}(\nu + \eta(x))e^{i\theta(x)/\nu}. \quad (1.24)$$

If the gauge transform is chosen to be $\alpha(x) = \theta(x)/\nu$, the fields are defined in the so-called “unitary gauge”⁵ and have the special forms

$$\begin{aligned} \phi(x) \rightarrow \phi'(x) &= \frac{1}{\sqrt{2}}(\nu + \eta(x)) \\ A_\mu(x) \rightarrow B_\mu(x) &= A_\mu(x) - \frac{1}{e\nu}\partial_\mu\theta(x) \end{aligned} \quad (1.25)$$

The kinetic term of the gauge field $F_{\mu\nu}$ is invariant under this transformation. If the gauge transformations of Equation 1.25 are substituted into the Lagrangian (1.21) the effective Lagrangian at the minimum of V is

$$\begin{aligned} \mathcal{L} = & \frac{1}{2}\partial_\mu\eta\partial^\mu\eta - \mu^2\eta^2 \\ & - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}(e\nu)^2B_\mu B^\mu \\ & + \frac{1}{2}e^2B_\mu B^\mu\eta(\eta + 2\nu) - \lambda\nu\eta^3 - \frac{\lambda}{4}\eta^4. \end{aligned} \quad (1.26)$$

The breaking of the original symmetry has dramatically altered the physical consequences of the model. In its unbroken form, the model described by Equation 1.21 would produce two real massive particles and one massless gauge boson mandated by local gauge invariance. After symmetry breaking, the effective Lagrangian in Equation 1.26 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

⁵In the unitary gauge, the choice of gauge ensures that the mass matrix is diagonalized.

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 [6], Weinberg [7] and Salam [8] 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 [11, 12, 13, 14] with the predicted masses at the UA1 and UA2 experiments.

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 the chosen symmetry group of the model. We then construct a locally 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 $SU(2)$.

$$L = \frac{1 - \gamma_5}{2} \begin{pmatrix} \nu_e \\ e^- \end{pmatrix} = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L \quad (1.27)$$

263 The operators that operate on “weak isospin,” the quantum of $SU(2)_L$, are

$$\tau^+ = \frac{\tau^1 + i\tau^2}{2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad (1.28)$$

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

where the τ^i are the Pauli matrices. The weak currents J_μ^\pm can be written by combining equations 1.27–1.29

$$J_\mu^\pm = \bar{L}\gamma_\mu\tau^\pm L. \quad (1.30)$$

264 Since τ^1 , τ^2 , and τ^3 are the generators of the $SU(2)$ group, we can complete the group
265 by adding a neutral current to the charged currents of Equation 1.30. The τ^3 generator is
266 diagonal, so the charge of the current is zero and no mixing of the fields occur:

$$\begin{aligned} J_\mu^3 &= \bar{L}\gamma_\mu\frac{\tau^3}{2}L \\ &= \bar{L}\gamma_\mu\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}L \\ &= \frac{1}{2}\bar{\nu}_e\gamma_\mu\nu_e - \frac{1}{2}\bar{e}_L\gamma_\mu e_L. \end{aligned} \quad (1.31)$$

267 Naively one might hope that the neutral current of Equation 1.31 would correspond to
268 the electromagnetic (photon) current of QED. However, this is impossible for two reasons.
269 First, the right-handed component e_R does not appear in the current, so this interaction
270 violates parity, a known symmetry of the electromagnetic interactions. Second, the current
271 couples to neutrinos, which have no electric charge. Therefore, the “charge” corresponding
272 to the $SU(2)$ gauge symmetry generators $T^i = \int J_0^i(x)d^3x$ cannot be that of the QED, and
273 the gauge group must be enlarged to include an additional $U(1)$ symmetry. The generator of
274 the new symmetry must commute with the generators of the $SU(2)_L$ group. The symmetry
275 cannot be directly extended with $U(1)_{em}$ as the electromagnetic charge $Q = \int (e_L^\dagger e_L +$
276 $e_R^\dagger e_R)d^3x$ does not commute with T^i . The solution is to introduce the “weak hypercharge”
277 $\frac{Y}{2} = Q - T^3$, which commutes the generators of $SU(2)_L$. Thus the symmetry group of the
278 electroweak model is $SU(2)_L \times U(1)_Y$.

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

$$\begin{aligned}\mathcal{L} = & \bar{L}i\gamma^\mu(\partial_\mu - ig\frac{\vec{\tau}}{2} \cdot \vec{A}_\mu + \frac{i}{2}g'B_\mu)L \\ & + \bar{R}i\gamma^\mu(\partial_\mu + \frac{i}{2}g'B_\mu)R \\ & - \frac{1}{4}F_{\mu\nu}^i F^{i\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}.\end{aligned}\quad (1.32)$$

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

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

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \quad (1.33)$$

$$\mathcal{L}_S = (D_\mu\phi)^\dagger(D^\mu\phi) - V(\phi^\dagger\phi), \quad (1.34)$$

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 1.22. At this point we
also add $SU(2)_L \times U(1)_Y$ invariant ‘‘Yukawa’’ terms

$$\mathcal{L}_Y = -G_e(\bar{L}\phi R + \bar{R}\phi^\dagger L) + h.c. \quad (1.35)$$

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

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

$$\phi_{min} = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix}. \quad (1.36)$$

The new symmetry of the model can be confirmed by looking at the action of the 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 original
symmetry generators T^+ , T^- , T^3 , and Y are all broken. The vacuum is invariant under Q ,

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

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

so the broken Lagrangian contains the correct symmetry properties.

The gauge boson content of the electroweak interaction is obtained by parameterizing the Higgs field in the magnitude–phase notation of Equation 1.24 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

$$\phi' = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}(\nu + H(x)) \end{pmatrix} = \frac{1}{\sqrt{2}}(\nu + H(x))\chi. \quad (1.37)$$

The mass spectrum of the gauge bosons of the electroweak interaction (the photon, W^\pm , and Z) is determined by the interaction of the gauge field terms in the covariant derivative with the non–zero vacuum expectation value ν 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

$$\mathcal{L}_{mass} = \frac{\nu^2}{8}(g^2 A'^1_\mu A'^1{}^\mu + g^2 A'^2_\mu A'^2{}^\mu + (gA'^3_\mu - g'B'_\mu)^2). \quad (1.38)$$

The A'^1_μ and A'^2_μ fields can be combined such that the first two terms in Equation 1.38 are equivalent to the mass term of a charged boson

$$W^\pm_\mu = \frac{A'^1_\mu \mp iA'^2_\mu}{2}. \quad (1.39)$$

This is the familiar W^\pm boson of β and muon decay, and has mass $M_W = \frac{1}{2}g\nu$. The

third term in Equation 1.38 can be written in matrix form and then diagonalized into mass

eigenstates

$$\begin{aligned} & \frac{\nu^2}{8} \begin{pmatrix} A'^3_\mu & B'_\mu \end{pmatrix} \begin{pmatrix} g^2 & -gg' \\ -gg' & g'^2 \end{pmatrix} \begin{pmatrix} A'^3{}^\mu \\ B'^\mu \end{pmatrix} \\ & \rightarrow \frac{\nu^2}{8} \begin{pmatrix} Z_\mu & A_\mu \end{pmatrix} \begin{pmatrix} g^2 + g'^2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} Z^\mu \\ A^\mu \end{pmatrix}, \end{aligned} \quad (1.40)$$

giving a massive Z boson with

$$M_Z = \frac{\nu}{2}\sqrt{g^2 + g'^2} \quad (1.41)$$

and the massless photon A_μ of QED. The mass of the Z is related to the mass of the W^\pm

by

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

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 1.35. 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 composite particles of hadrons 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 carries a color charge, while the photon is electrically neutral. This has the consequence that a gluon can couple to other gluons. Secondly, it is found that no

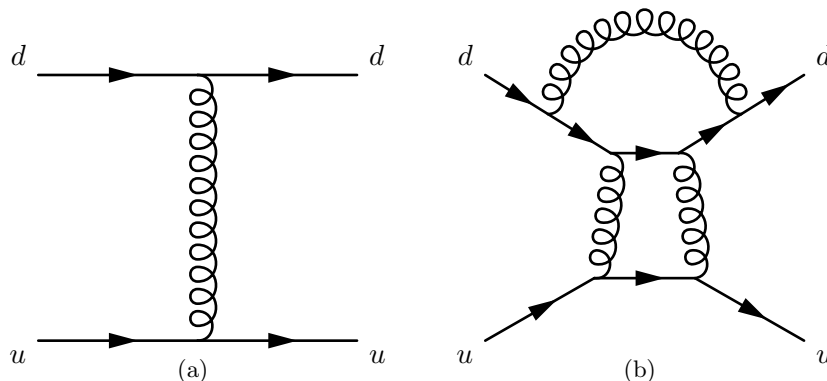


Figure 1.3: Feynman diagrams of a first-order (a) QCD interaction and a multi-loop (b) QCD interaction that have the same initial and final states. Each internal gluon propagator contributes a factor of g_s , the strong coupling constant, to the amplitude. Since $g_s > 1$, multi-loop diagrams have a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

319 colored object exists in nature. The corollary of this is that it is believed to be impossible
 320 for a single quark or gluon to be observed. The mechanism that gives rise to this effect is
 321 called “color confinement.” The strength of the strong force between two interacting colored
 322 objects increases with distance. If two colored objects in a hadron are pulled apart, the
 323 energy required to separate them will eventually be large enough to produce new colored
 324 objects, resulting in two (or more) colorless hadrons. Finally, at low energy, QCD is non-
 325 perturbative. What this means in practice is that when computing an amplitude from
 326 a QCD Feynman diagram, additional gluon interactions contribute a value greater than
 327 one. The dominance of multi-loop diagrams is illustrated in Figure 1.1.6. Thus higher
 328 order diagrams with many internal loops cannot be ignored in QCD. In practice what is
 329 done is to “factorize” QCD interaction amplitudes into a perturbative (high-energy) part
 330 and a non-perturbative part. The perturbative portion is calculable using the Feynman
 331 calculus; the non-perturbative must be estimated from parameterization functions that are
 332 experimentally measured.

333 The practical consequence of color confinement to a physicist at a high-energy particle
 334 physics experiment is the production of quark and gluon “jets,” which are high multiplic-
 335 ity sprays of particles observed in the detector. In a proton-proton collision, quarks and
 336 gluons can be knocked off the incident protons. These quarks and gluons immediately
 337 “hadronize,” surrounding themselves with additional hadrons, the majority of which are
 338 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 date⁶ 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 it is only an effective theory that is valid below some energy scale 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 presences 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 sec-

⁶The 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.

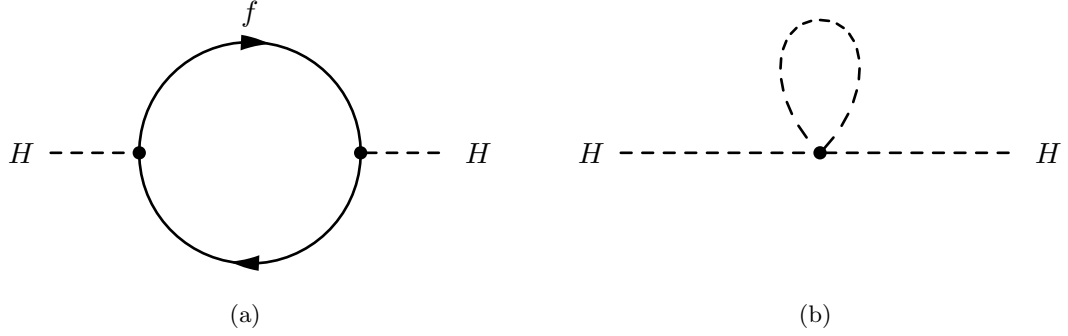


Figure 1.4: Feynman diagram of fermion (a) and scalar (b) loop corrections to Higgs mass.

366 tor, and motivated the development of Supersymmetry, which are the targets of the search
 367 presented in this thesis. An short overview of the hierarchy problem and Supersymmetry
 368 are presented in the next sections.

369 §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. The contribution of the loop correction in Figure 1.4(a) is [?]

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

370 scales with Λ^2 , which is many orders of magnitude larger than the electroweak (M_W) scale.
 371 The physical mass of the Higgs is expected to have the same scale as M_W , $O(100) \text{ GeV}/c^2$.
 372 The fact that each fermion contributes a loop correction (Equation 1.45) requires that
 373 the “bare mass” of the Higgs to be tuned to the precision of $(M_W/\Lambda)^2 \approx 10^{-26}$ for the
 374 renormalized mass to be correct! This is the so-called fine-tuning problem: it is believed
 375 that in a natural theory there will be only one scale. The electroweak unification analogy
 376 is in Equation 1.14, where it was noticed that the difference between the QED and weak
 377 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.

§1.2.2 Supersymmetry

Supersymmetry extends the Standard Model by positing that there exists a symmetry between the integer-spin bosons (γ, W^\pm, Z, H) and the half integer-spin fermions (quarks and leptons). In Supersymmetry, every particle in the Standard Model has a “superpartner” with a spin differs by $1/2$. All of the other quantum numbers (including mass) of the superpartners are the same. The introduction of this symmetry immediately solves the hierarchy problem. For every scalar loop correction (Figure 1.4(b)) to the Higgs mass there is now a corresponding fermion loop correction (Figure 1.4(a)). As the fermion and the scalar have the same quantum numbers (except for spin) it turns out that these two diagrams have the same value, but *opposite* sign. Thus the large Λ^2 superpartner loop corrections to the Higgs mass exactly cancel out the problematic Standard Model corrections. It is clear that if Supersymmetry exists, it must be broken. We have not observed a scalar charged particle with the same mass as the electron, for example. An excellent overview of possible mechanism that create spontaneous symmetry breaking in supersymmetric models is given in Chapter 6 of [15].

§1.2.3 The Minimal Supersymmetric Model

The simplest possible Supersymmetric extension to the Standard Model is the Minimal Supersymmetric Model (MSSM). The model groups superpartner pairs into chiral (a left or right-handed fermion field plus a complex scalar field) and gauge (a spin-1 vector boson and a left or right-handed *gaugino* fermion) “supermultiplets.” As the weak interactions of the Standard Model fermions are chiral, they (and their superpartners) must belong in a chiral supermultiplet. It is interesting to note that there is a different superpartner for the left and right-handed components of the fermions, even though the superpartners are spin-0 and cannot have any handedness. It is found that there must be two Higgs supermultiplets for the MSSM to be viable. As there are now fermionic particles in the Higgs sector (the Higgsinos), if only one supermultiplet is introduced the MSSM suffers from non-

Names		spin 0	spin 1/2	$SU(3)_C, SU(2)_L, U(1)_Y$
squarks, quarks ($\times 3$ families)	Q	$(\tilde{u}_L \ \tilde{d}_L)$	$(u_L \ d_L)$	$(\mathbf{3}, \mathbf{2}, \frac{1}{6})$
	\bar{u}	\tilde{u}_R^*	u_R^\dagger	$(\bar{\mathbf{3}}, \mathbf{1}, -\frac{2}{3})$
	\bar{d}	\tilde{d}_R^*	d_R^\dagger	$(\bar{\mathbf{3}}, \mathbf{1}, \frac{1}{3})$
sleptons, leptons ($\times 3$ families)	L	$(\tilde{\nu} \ \tilde{e}_L)$	$(\nu \ e_L)$	$(\mathbf{1}, \mathbf{2}, -\frac{1}{2})$
	\bar{e}	\tilde{e}_R^*	e_R^\dagger	$(\mathbf{1}, \mathbf{1}, 1)$
Higgs, higgsinos	H_u	$(H_u^+ \ H_u^0)$	$(\tilde{H}_u^+ \ \tilde{H}_u^0)$	$(\mathbf{1}, \mathbf{2}, +\frac{1}{2})$
	H_d	$(H_d^0 \ H_d^-)$	$(\tilde{H}_d^0 \ \tilde{H}_d^-)$	$(\mathbf{1}, \mathbf{2}, -\frac{1}{2})$

Table 1.1: Chiral supermultiplets in the Minimal Supersymmetric Standard Model. The spin-0 fields are complex scalars, and the spin-1/2 fields are left-handed two-component Weyl fermions. Source: [15]

Names	spin 1/2	spin 1	$SU(3)_C, SU(2)_L, U(1)_Y$
gluino, gluon	\tilde{g}	g	$(\mathbf{8}, \mathbf{1}, 0)$
winos, W bosons	$\tilde{W}^\pm \ \tilde{W}^0$	$W^\pm \ W^0$	$(\mathbf{1}, \mathbf{3}, 0)$
bino, B boson	\tilde{B}^0	B^0	$(\mathbf{1}, \mathbf{1}, 0)$

Table 1.2: Gauge supermultiplets in the Minimal Supersymmetric Standard Model. Source: [15]

renormalizable gauge anomalies.⁷ By introducing an additional Higgs supermultiplet with opposite hypercharge, the anomaly is canceled. The scalar portion of the MSSM Higgs sector then contains two complex doublet fields $H_u = (H_u^+, H_u^0)$ (up-type) and $H_d = (H_d^0, H_d^-)$ (down-type). The complete chiral and gauge supermultiplets of the MSSM are enumerated in Tables 1.1 and 1.2, respectively.

The superpotential (like the scalar potential of Section 1.1.3 but invariant under supersymmetric transformations) of the MSSM is then [15]

$$W_{\text{MSSM}} = \bar{u}_\mathbf{y} \mathbf{u} Q H_u - \bar{d}_\mathbf{y} \mathbf{d} Q H_d - \bar{e}_\mathbf{y} \mathbf{e} L H_d + \mu H_u H_d, \quad (1.46)$$

where $H_u, H_d, Q, L, \bar{u}, \bar{d}, \bar{e}$ are the superfields defined in Table 1.1. The \mathbf{y} terms are Yukawa 3×3 matrices which act on the different families. It is important to note that the up-type quarks couple to the up-type Higgs H_u , while the down-type quarks and leptons couple

⁷A gauge anomaly is a linear divergence that occurs in diagrams containing a fermion loop with three gauge bosons (total) in the initial and final states. In the Electroweak model, the sum of the fermion contributions cancel the anomaly. Interestingly, the requirement of anomaly cancellation is only achieved in the SM is achieved only by requiring there be three types of color in QCD.

414 to the down-type Higgs. This feature has large phenomenological consequences, which are
 415 discussed in 1.3.2. The scalar portion of the W_{MSSM} potential defines the spontaneous
 416 symmetry breaking. Similar to the scalar potential V symmetry breaking of Section 1.1.3,
 417 the potential of V at the minimum is found⁸ to be

$$V = (|\mu|^2 + m_{H_u}^2)|H_u^0|^2 + (|\mu|^2 + m_{H_d}^2)|H_d^0|^2 - (bH_u^0 H_d^0 + c.c.) + \frac{1}{8}(g^2 + g'^2)(|H_u^0|^2 - |H_d^0|^2)^2. \quad (1.47)$$

Under suitable choices⁹ of the parameters in Equation 1.47, the up-type and down-type neutral Higgs fields acquire a VEV, ν_u and ν_d , respectively. The VEVs are related to the VEV of electroweak symmetry breaking (Equation 1.41) in the SM,

$$\nu_u^2 + \nu_d^2 = \nu^2 = \frac{2M_Z^2}{g^2 + g'^2} \approx (174 \text{ GeV})^2.$$

The ratio of the VEVs is expressed as

$$\tan \beta \equiv \frac{\nu_u}{\nu_d},$$

418 which is an important parameter of the MSSM. As there are two complex doublets, there
 419 are a total of eight degrees of freedom in the MSSM Higgs sector. After the symmetry
 420 breaking, three of the degrees of freedom are (like the Standard Model) eaten by the W^\pm
 421 and Z weak gauge bosons. The remaining five degrees of freedom create five massive Higgs
 422 bosons: two CP-even neutral scalars h^0 and H^0 , a CP-odd neutral scalar A^0 , and two
 423 (positive and negative) charged scalars H^\pm . The masses are of the different Higgs mass
 424 eigenstates are related to each other and $\tan \beta$ at tree level by

$$m_{h^0}^2 = \frac{1}{2}(m_{A^0}^2 + m_Z^2 - \sqrt{(m_{A^0}^2 - m_Z^2)^2 + 4m_Z^2 m_{A^0}^2 \sin^2(2\beta)}) \quad (1.48)$$

$$m_{H^0}^2 = \frac{1}{2}(m_{A^0}^2 + m_Z^2 + \sqrt{(m_{A^0}^2 - m_Z^2)^2 + 4m_Z^2 m_{A^0}^2 \sin^2(2\beta)}). \quad (1.49)$$

It can be seen that the tree level mass m_{h^0} of Equation 1.48 is bounded from above by $m_{h^0} < m_Z |\cos(2\beta)| < 90 \text{ GeV}/c^2$. If this is true the model would have been excluded by LEP (see next section). However, there are important quantum corrections to m_{h^0} from the top-quark and top-squark loop diagrams which increase m_{h^0} . The Yukawa couplings in the MSSM depend on $\tan \beta$. The relationships for the most massive members of each

⁸A clever choice of the $SU(2)_L$ gauge has removed any contributions from the charged fields. The charged Higgs fields cannot have a VEV without breaking $U(1)_{em}$.

⁹See Chapter 7 of ?? for a detailed overview.

family are

$$m_t = y_t v \sin \beta, \quad m_b = y_b v \cos \beta, \quad m_\tau = y_\tau v \cos \beta. \quad (1.50)$$

The Yukawa couplings are free parameters determined by experimentally observed masses. This means that when $\tan \beta$ is large ($\beta \rightarrow \pi$), the Yukawa terms y for the b quarks and τ leptons must be enhanced to maintain the observed masses. The effect of $\tan \beta$ on the Higgs mass spectrum and couplings in the MSSM will be discussed further in Section 1.3.2.

§1.3 Searches for the Higgs boson

The potential discovery of the Higgs boson is one of the biggest prizes in science today. Dozens of experiments, thousands of scientists and billions of dollars (a human hierarchy problem. . .) have been spent in efforts to discover the Higgs. In these sections we discuss how the Higgs and MSSM could appear in modern colliders (with an emphasis on the LHC) and current limits placed on the Higgs by the LEP and Tevatron experiments.

§1.3.1 Standard Model 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 1.35) in the Lagrangian. Taking the electron as an example, after symmetry breaking, the Yukawa term is found to be

$$\mathcal{L}_e = -\frac{G_e}{\sqrt{2}}(\nu + H(x))\bar{e}e = -\frac{G_e\nu}{\sqrt{2}}\bar{e}e - \frac{G_e}{\sqrt{2}}H(x)\bar{e}e. \quad (1.51)$$

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

$$\frac{G_e\nu}{\sqrt{2}} = \frac{m_e}{\nu}. \quad (1.52)$$

The left-hand side of Equation 1.52 is the same as the constant in the electron-Higgs $H(x)\bar{e}e$ coupling term in Equation 1.51. 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

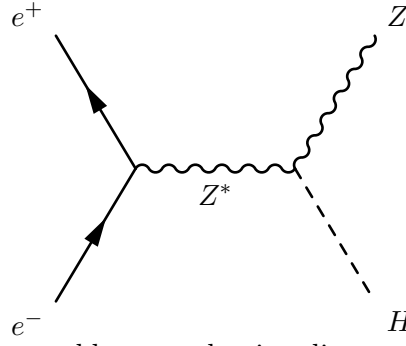


Figure 1.5: Higgstrahlung production diagram at e^+e^- colliders

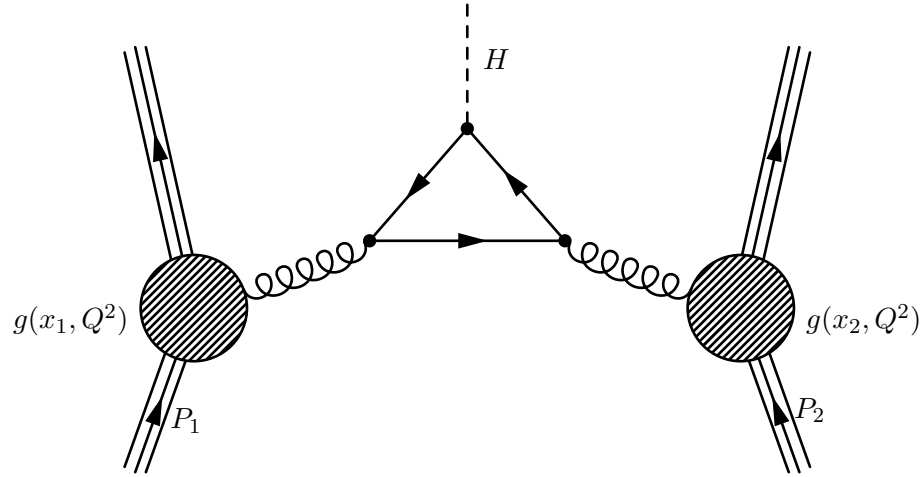


Figure 1.6: Gluon fusion Higgs production mechanism in a proton–proton collision. The Higgs mass coupling favors heavy quarks in the central loop. Image credit: [?]

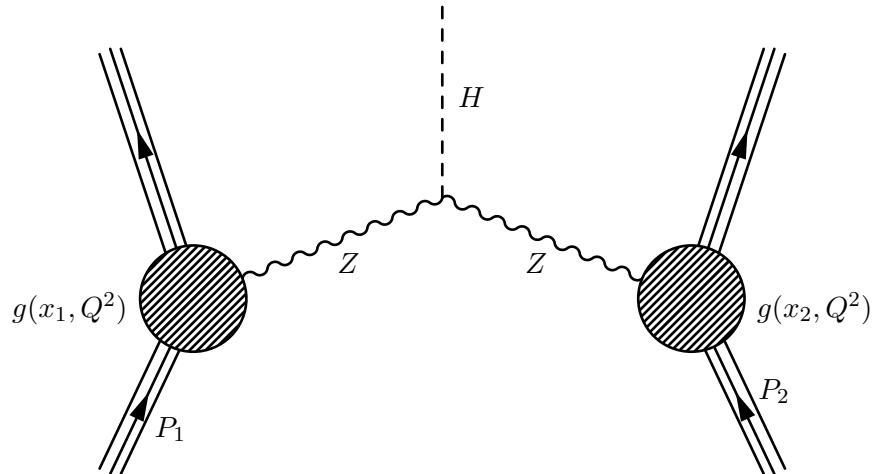


Figure 1.7: Vector boson fusion (VBF) Higgs production mechanism in proton–proton collisions. The VBF mechanism is notable for the lack of color–flow between the two incident protons.

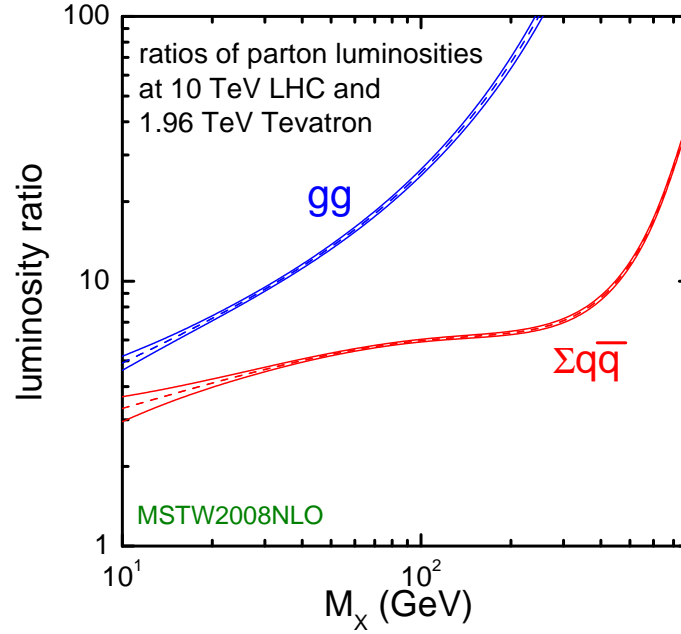


Figure 1.8: Ratio of the parton luminosity (the amount of luminosity contributed by the different species that compose the proton) of the LHC (at $\sqrt{s} = 10$ TeV) and the Tevatron. The large increase in gluon–gluon luminosity affects the favored production mechanisms of the Higgs boson.

mass-proportional coupling. At the Tevatron and LEP experiments, which will be introduced in the next section, the dominant SM Higgs production mode is “Higgstrahlung,” where a virtual W^\pm or Z gauge boson is produced and then radiates a Higgs boson. Higgstrahlung is illustrated in Figure ???. At the Large Hadron Collider, higher gluon luminosities (See Figure 1.8) 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 1.9.

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

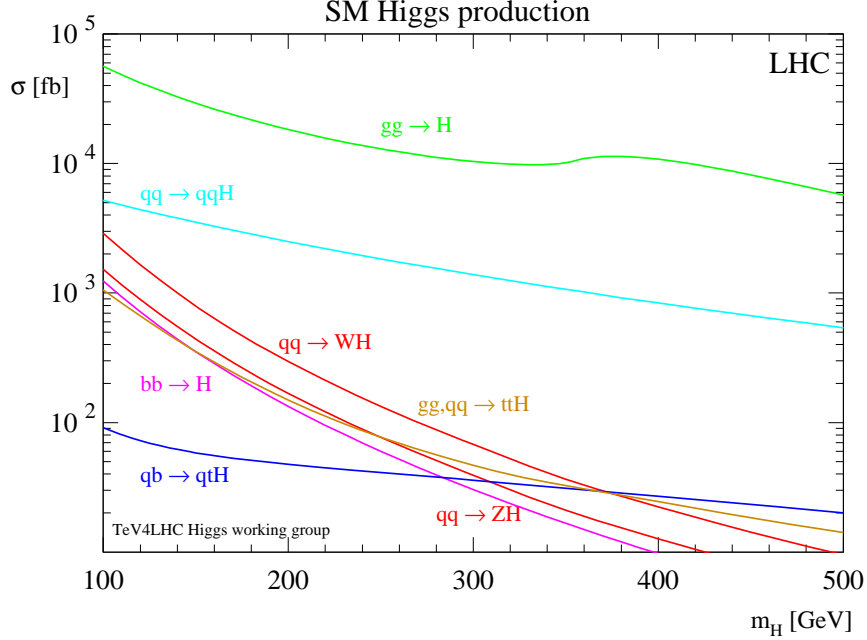


Figure 1.9: Cross section of the Standard Model Higgs boson versus the Higgs boson mass. The different curves give the contribution to the cross section from different production mechanisms. Source: [16].

457 to decay to pairs of weak bosons ($M_H < 160 \text{ GeV}/c^2$), the Higgs decays predominantly to
 458 either b -quarks ($b\bar{b}$, 90%) or a pair of τ leptons ($\tau^+\tau^-$, ≈ 10). Above the $W^\pm W^\mp$ threshold,
 459 decays to vector bosons ($H \rightarrow W^\pm W^\mp$ and $H \rightarrow ZZ$) dominate. The dependence of
 460 branching fraction on M_H and the other rare decay modes are illustrated in Figure 1.10. For
 461 low mass Higgs, the $\tau^+\tau^-$ decay mode plays a particularly important role. The dominant
 462 decay mode $H \rightarrow b\bar{b}$, suffers from enormous backgrounds from QCD jet production. It
 463 is important to understand the magnitude of difference between expected Higgs boson
 464 production and the rates of various backgrounds. Figure 1.11 illustrates the cross sections
 465 for different SM processes at hadron colliders. The rate of Higgs production is many orders
 466 of magnitude ($O(10^{-7})$) smaller than that of QCD production. It is important to therefore
 467 design searches to use handles that can reject the vast majority of the uninteresting events
 468 at hadron colliders.

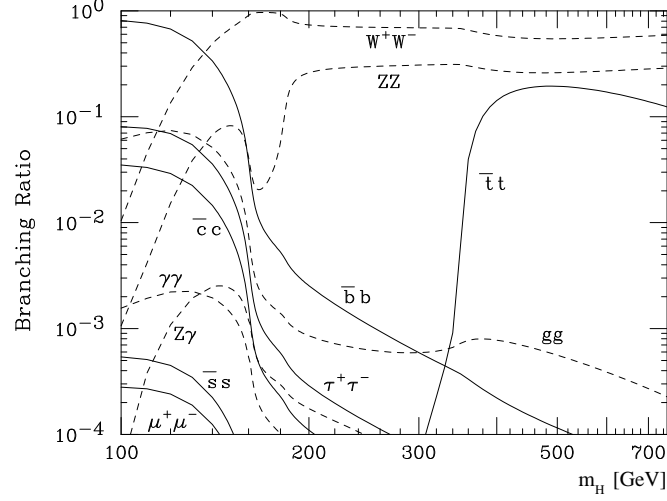


Figure 1.10: Branching fraction of the Standard Model Higgs bosons for different values of M_H . Source: [16].

§1.3.2 MSSM Higgs Phenomenology

The phenomenology of the Higgs sector of the MSSM is similar to the Standard Model in some respects, but differs in some key aspects which have important implications for final states involving τ leptons and b quarks. When the parameter $\tan \beta$ is large, the coupling factor between the Higgs and the down-type quarks and leptons (effectively the τ and b quark) is enhanced by $\tan \beta$. The gluon-gluon cross section is therefore increased by $\tan^2 \beta$, where the top quark loop in Figure 1.3.1 is replaced by a ($\tan \beta$ enhanced) b quark loop. Additionally, MSSM Higgs production with associated b -quarks, illustrated in Figure 1.3.2, becomes an important production mode. At tree-level, the MSSM can be defined by the mass of the CP-odd Higgs m_{A^0} and $\tan \beta$. For a reasonably high $\tan \beta$, there is always one CP-even Higgs (h^0 or H^0) which is mass-degenerate with the A^0 . When $\tan \beta$ and m_{A^0} are both large, associated b production dominates the total cross section [18]. The cross sections of the different MSSM neutral Higgs bosons are shown in Figure ???. The $\tan \beta$ enhancement of the MSSM Higgs coupling to the b quarks and τ leptons cause the branching fraction of all neutral MSSM Higgs to be $H \rightarrow b\bar{b}$ (90%) and $H \rightarrow \tau^+\tau^-$ (10%) across the entire range of m_{A^0} . The enhanced production rate and the high branching fraction to τ leptons make the MSSM Higgs decaying to τ leptons an exciting and promising channel to search for Higgses and Supersymmetric physics at colliders.

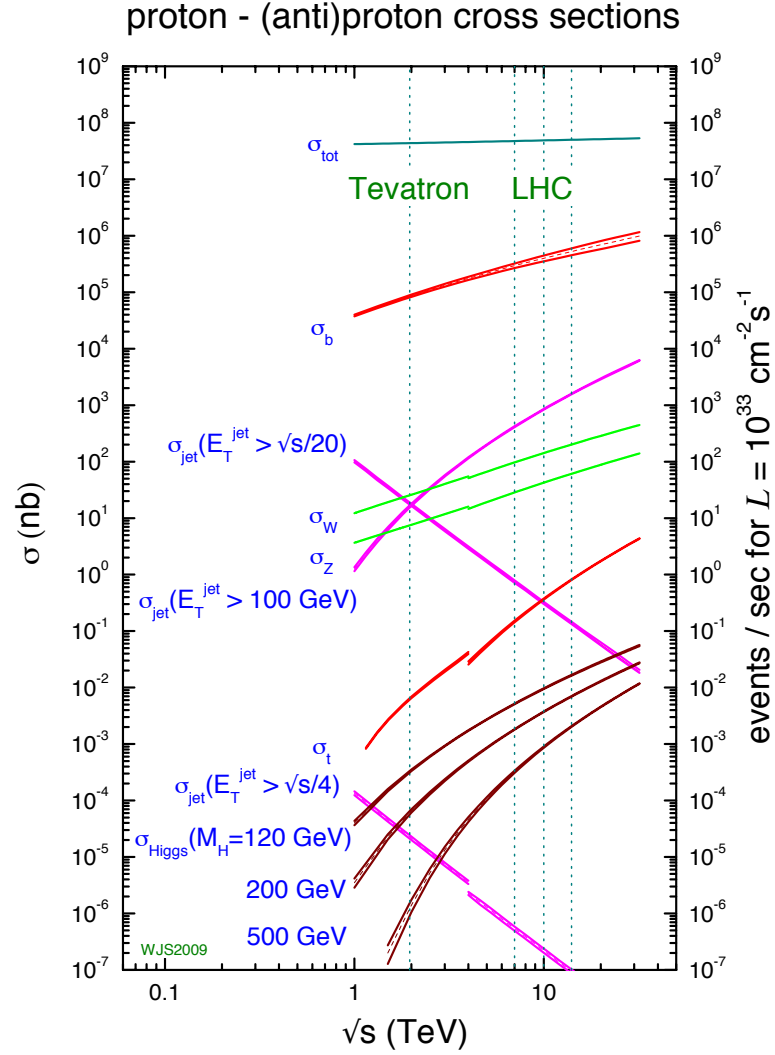


Figure 1.11: Cross sections of various processes at hadron colliders. The horizontal axis represents the center of mass energy of the collision. Of note is the vast difference in scales between Higgs production (maroon lines, $O(10^{-2})$ nb) and the QCD cross section to produce $b\bar{b}$ pairs (red line, $O(10^4)$ nb). Source: [17].

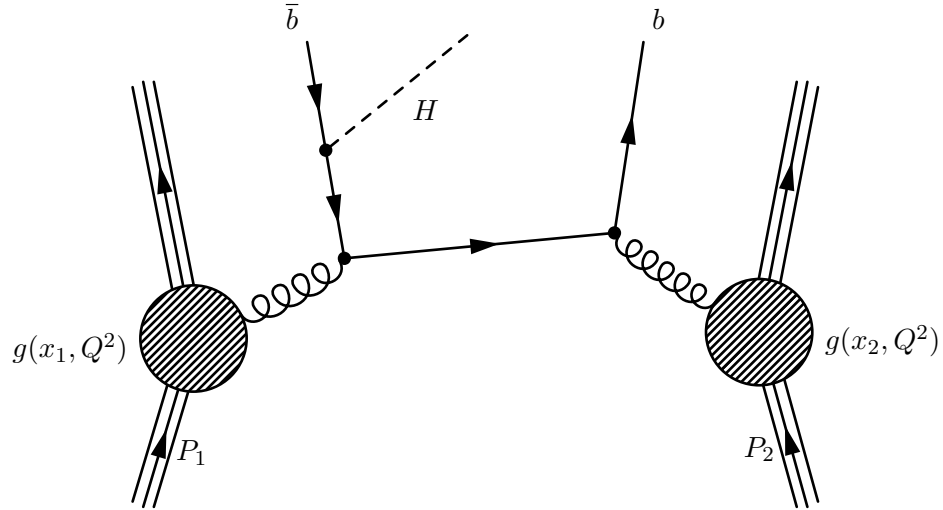
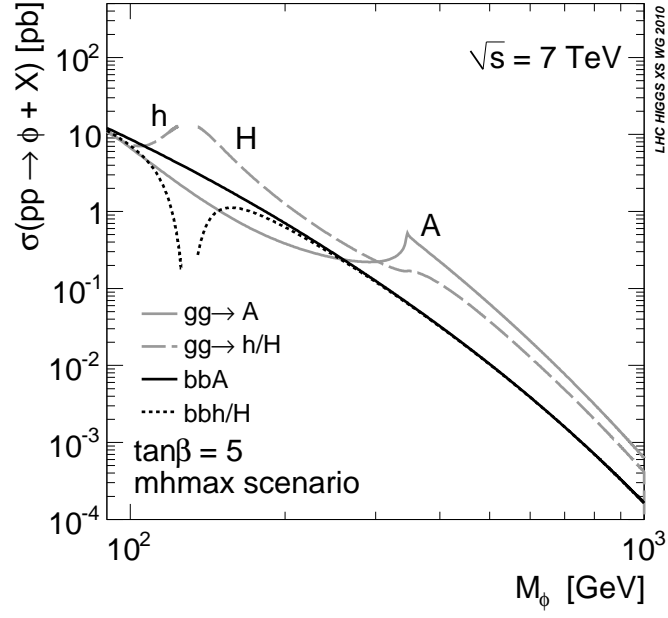


Figure 1.12: One possible diagram for an MSSM Higgs produced with associated b -quarks in a proton-proton collision.

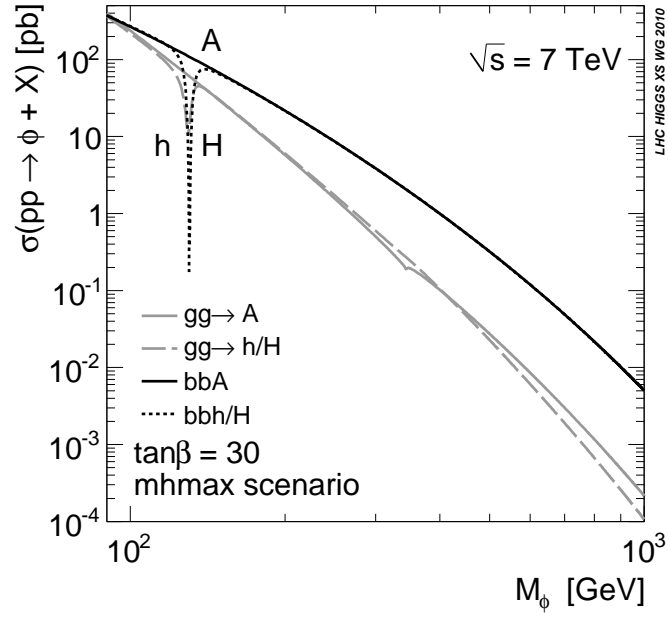
§1.3.3 Results from LEP and Tevatron

§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 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 ??.



(a)



(b)

Figure 1.13: Cross sections for the different MSSM Higgs bosons versus m_{A^0} in the $m_{h^{max}}$ benchmark scenario [19] scenario for $\tan\beta = 5$ (a) and $\tan\beta = 30$ (b). Source: [18]

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%
$\pi^- \pi^0 \pi^0 \nu_\tau$	$a1$	1200	9.3%
$\pi^- \pi^- \pi^+ \nu_\tau$	$a1$	1200	9.0%
$\pi^- \pi^- \pi^+ \pi^0 \nu_\tau$	$a1$	1200	4.5%
Total			94.4%

Table 1.3: 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.

Chapter 2

The Compact Muon Solenoid Experiment

The Compact Muon Solenoid (CMS) Experiment is a “general purpose” particle detector designed to measure collision events at the Large Hadron Collider (LHC), a proton–proton synchrotron located at the CERN laboratory in Geneva, Switzerland. The design goals of the CMS experiment are [20], in order of priority:

- Good muon identification and momentum resolution over a wide range of momenta and angles, good dimuon mass resolution ($\approx 1\%$ at 100 GeV), and the ability to determine unambiguously the charge of muons with $p < 1$ TeV;
- Good charged-particle momentum resolution and reconstruction efficiency in the inner tracker. Efficient triggering and offline tagging of τ ’s and b -jets, requiring pixel detectors close to the interaction region;
- Good electromagnetic energy resolution, good diphoton and dielectron mass resolution ($\approx 1\%$ at 100 GeV), wide geometric coverage, π^0 rejection, and efficient photon and lepton isolation at high luminosities;
- Good missing–transverse–energy and dijet–mass resolution, requiring hadron calorimeters with a large hermetic geometric coverage and with fine lateral segmentation.

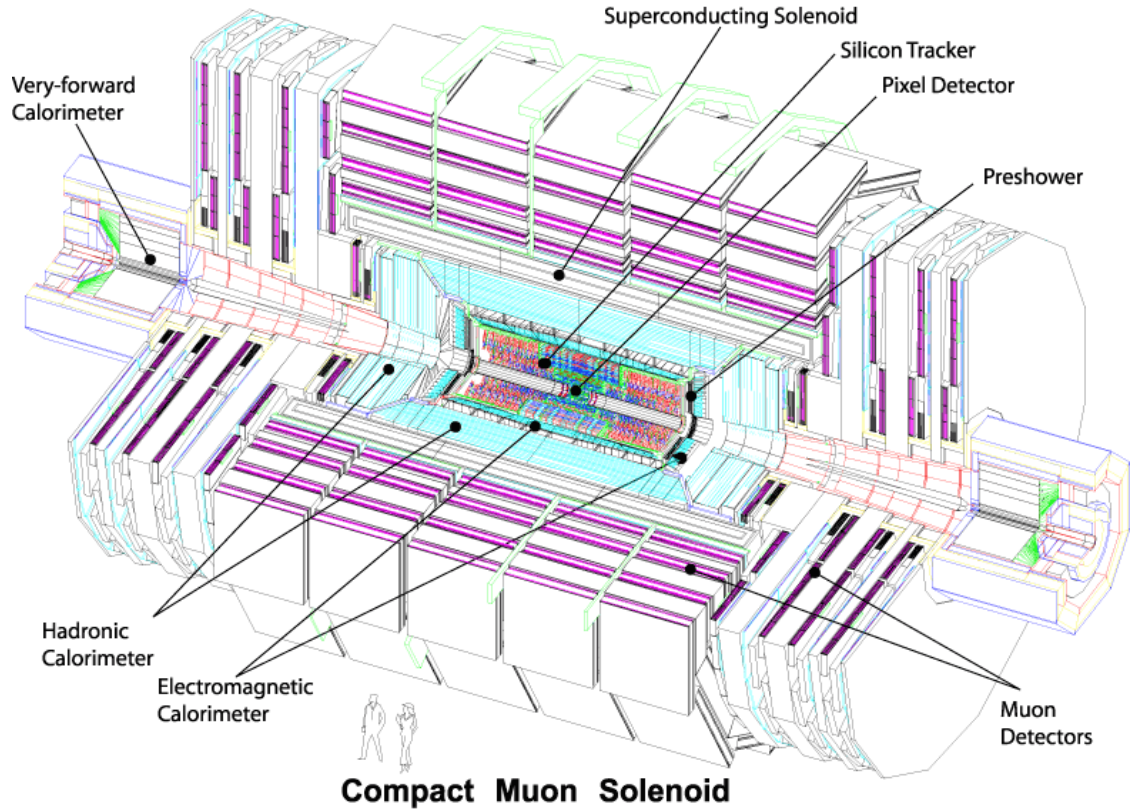
The detector uses a hermetic design that maximizes the solid–angle of the fiducial region to capture as much information about the collisions as possible. The general geometry of the detector is cylindrical. A cutaway diagram of the detector is shown in Figure 2.1. Each of the sub–detector components consists of “barrel” and “endcap” components. As its name suggests, the detector is centered around a four Tesla superconducting solenoid magnet. The individual sub–detectors of CMS are arranged in a manner that permits

523 identification of different species of particles. The central (closest to interaction point) sub-
 524 detector are the charged particle tracking systems (the “tracker”). The tracker is designed
 525 to be a *non-destructive* instrument, which means that ideally that the momentum of par-
 526 ticles are unchanged after passing through it. Outside of the tracker is the electromagnetic
 527 and hadronic calorimeters, which are abbreviated ECAL and HCAL, respectively. The
 528 calorimeter is a *destructive* detector, and is designed such that visible incident particles
 529 are completely absorbed. The outer layers of CMS are designed to measure muons, the
 530 one¹ species of particle that is immune to the effects of the calorimeter. The arrangement
 531 of destructive and non-destructive sub-detectors facilitates the identification of different
 532 types of particles. This concept is illustrated in Figure 2.1(b). In this chapter we give an
 533 brief overview of the LHC machine, and then describe the individual sub-detector systems
 534 of CMS.

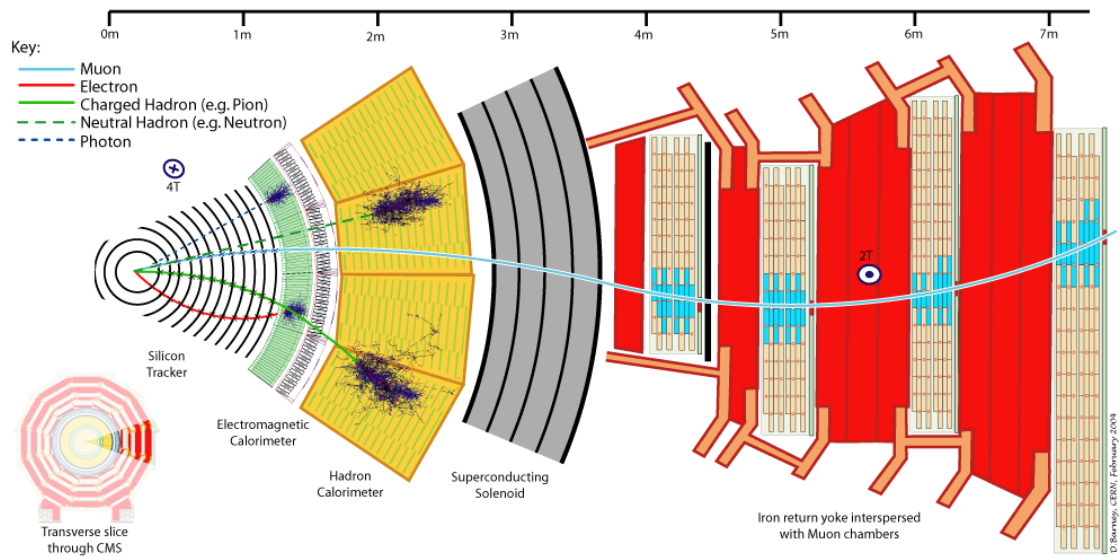
535 §2.1 The Large Hadron Collider

536 The Large Hadron Collider is a proton–proton synchrotron, with a design collision energy
 537 of 14 TeV. At the time of this writing (and for the foreseeable future), the LHC is the
 538 world’s largest and highest energy particle accelerator. A synchrotron is a machine that
 539 accelerates beams of charged particles by using magnets to steer them in a circle through
 540 radio–frequency resonating cavities which accelerate the particles. As the LHC is a collider,
 541 there are two beams that are accelerated in opposite directions. The maximum beam energy
 542 of a synchrotron is determined by its radius and the maximum strength of the magnetic
 543 fields used to bend the path of the beam. The dipole magnets used by the LHC to steer the
 544 particles are superconducting niobium–titanium. To maintain them in a superconducting
 545 state, they are cooled using superfluid liquid helium to 1.9 Kelvin. To store the beam at the
 546 injection energy of 450 GeV, the magnetic dipole fields must be maintained at 1/2 Tesla.
 547 As the energy of each beam energy is increased to its (design) maximum of 7 TeV, the
 548 dipole fields are ramped to a maximum field of over 8 Tesla.

¹Neutrinos of course fulfill this requirement as well, but are so weakly interacting that they are effectively invisible.



(a)



(b)

Figure 2.1: Figure (a), top, shows a schematic drawing of the CMS detector. The individual sub-detectors are labeled. Two humans are shown in the foreground for scale. Figure (b) shows a radial cross section of the detector and demonstrates how the (non-)destructiveness of different sub-detectors facilitates particle identification.

549 §2.2 Solenoid Magnet

The four Tesla field of the CMS solenoid magnet is a critical factor in ability of CMS to precisely measure collisions at the LHC. The momentum of charged particles is measured in the detector by examining the curvature of the particles path as it travels through the magnetic field. The radius of curvature r of a charged particle in a magnetic field is given by

$$r = \frac{p_{\perp}}{|q|B}, \quad (2.1)$$

550 where q is the charge of the particle, B is the magnetic field, and p_{\perp} is the component of the
 551 particle’s relativistic momentum perpendicular to the direction of the magnetic field. From
 552 Equation 2.1, it is evident that the ability to measure high momentum charged particles
 553 (a critical goal of CMS) requires a high magnetic field. Even at very high particle energies
 554 where the resolution becomes poor, the strength of the magnetic field is still very impor-
 555 tant for identifying the bending direction of the particle; the direction corresponds to the
 556 particle’s electric charge. Furthermore, the homogeneity of the magnetic field is important
 557 to minimize systematic errors in the measurement of tracks.

558 The CMS solenoid is extremely large. The radial bore of the magnet is 6.3 meters; the
 559 magnet is 12.5 meters in length and weighs 220 tons. The large bore of the magnet allows
 560 the tracker and calorimeter systems to be located inside the solenoid. The internal windings
 561 of solenoid is arranged in four layers to increase the total field strength and are cooled by
 562 liquid helium to a temperature of 4.5 Kelvin. The windings are magnetically coupled to
 563 the support superstructure. This coupling allows the magnetic to heat uniformly during
 564 a “quench²” event, reducing localized stresses. The nominal current at full field of the
 565 solenoid is 19.14 kA. The solenoid itself is surrounded by an iron return yoke with a total
 566 mass of 10,000 tons. The return yoke surrounding the solenoid minimizes the fringing field.
 567 The muon detector system is interspersed inside the yoke, and takes advantage of the field
 568 in the yoke to measure the momentum and charge of muons.

²A quench event occurs when some part of the magnet is suddenly no longer in a superconducting state. The coil becomes resistive and the large current in the magnet creates large amounts of heat.

§2.3 Charged Particle Tracking Systems

The charged particle tracking system measures the trajectories of charged particles emerging from the event. The tracker measures the trajectory of a charged particle by measuring “hits” along the trajectory. Each hit corresponds to the global position of the trajectory on a given surface. The trajectory can then be reconstructed by a helix to the points. The tracker is designed to have a resolution that permits the reconstruction of “secondary vertices” in b -quark and τ lepton decays. To accomplish this, there are two types of tracking detectors in CMS. The “pixel detector” composes the inner layers (three in the barrel, two in the endcaps). The pixel detector is situated as close as possible (4.4 cm) to the interaction point and has a very high resolution. Outside of the pixel detector is the silicon strip tracker, with ten layers in the barrel and 12 layers in the endcaps. A secondary vertex occurs when a particle is semi-stable, traveling some non-negligible distance in the detector, but decaying before the first layer of the tracking system. The pixel and strip tracking detectors have a fiducial region which extends to a pseudorapidity of approximately $|\eta| \approx 2.5$.

Both the pixel and strip trackers are silicon based. The principle of operation is similar to that of a charged-coupled discharge (CCD) in a modern digital camera. The sensitive portion of the detector is a silicon chip that is arranged with diode junctions formed by a p -doped layer and an n -doped layer³. Each $p - n$ junction is electrically isolated from adjacent layers. The size of each junction region determines⁴ the spatial resolution of the sensor. In the pixel detector, each sensor region “pixel” is $100 \times 150 \mu\text{m}^2$. In the strip tracker, The rear side of the chip is mounted to read-out electronics. During operation, a high-voltage reverse bias is applied to each $p - n$ junction to achieve full depletion. When a charged particle passes through the detector, the diode-junction breaks down and the readout system registers the hit.

The tracking system has been specifically designed for the high radiation environment around the interaction point. The detector is cooled to -27°C during operation to minimize

³The pixel detector actually uses a more complicated multi-layered scheme to improve radiation hardness. For details, see Section 3.2.2 of [20].

⁴Additionally, the size of the sensitive area needs to be small enough such that the hit occupancy during a typically LHC event is not too large, which would cause overlaps and spoil the ability to reconstruct tracks. The expected occupancy depends on the distance r^2 from the interaction. The expected occupancy in the pixel detector for LHC collisions is 10^{-4} .

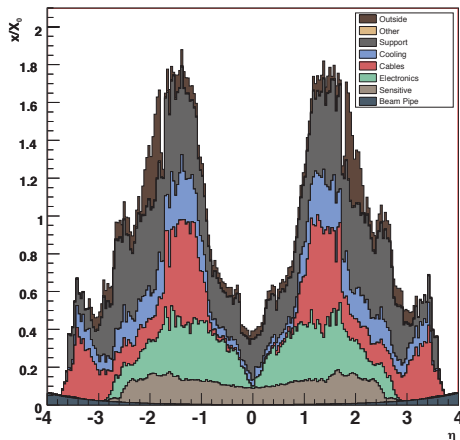


Figure 2.2: Material budget of the CMS tracker in units of radiation lengths X_0 . The material budget is broken down into the contributions from the different components of the tracker. The amount of material is largest in the “transition region” between the barrel and endcap.

595 damage. Radiation exposure produced in LHC collisions can change behavior of the tracking
 596 detector in three ways. Over time, radiation can induce positive holes in oxide layers found
 597 in the read-out electronics which increase the signal-to-noise ratio. In the sensor mass itself,
 598 radiation damage changes the doping from n to p over time. The required voltage to deplete
 599 the sensor will thus increase over time. The readout electronics, bias voltage supplies, and
 600 cooling systems are designed to scale with the radiation damage and maintain a signal-to-
 601 noise ratio of 10:1 or greater for 10 years of LHC operation. The final radiation effect
 602 is not an integrating effect. A “single event upset” is a transient effect where an ionizing
 603 charged particle passes through the readout electronics and changes the state of the digital
 604 circuitry.

605 In the ideal case, the tracker would be a non-destructive instrument. However, charged
 606 particles can interact with the mass of the tracker (and its support infrastructure). These
 607 interactions limit the resolution of the tracker. The amount of matter in the tracker is
 608 referred to as the “material budget”. The material budget of the CMS tracker depends
 609 heavily on the pseudorapidity η and is illustrated in Figure 2.2. The relatively large material
 610 budget of the CMS tracker has two effects: charged particles can “multiple scattering,”
 611 interacting with material in the tracker. This can cause “kinks” in the reconstructed track.

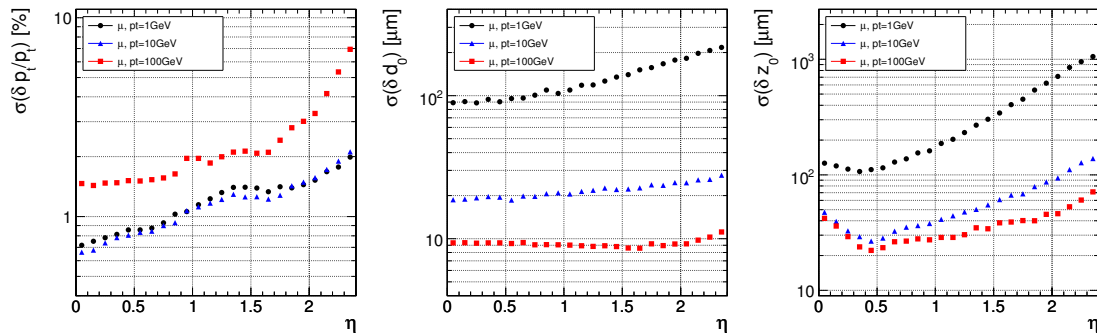


Figure 2.3: Expected resolutions of reconstructed transverse momentum (left), transverse impact parameter (center), and longitudinal impact parameter (right) versus absolute pseudorapidity $|\eta|$. The resolution is shown for three different cases of particle p_T , 1 GeV/c (black), 10 GeV/c (blue), and 100 GeV/c (red).

Hadronic particles (charged and neutral) can undergo “nuclear interactions,” which are a hard collisions between the incident particle and the tracker material. This typically produces a spray of hadrons from the point of interaction. Finally, the material budget can cause “photon conversions.” A photon conversion occurs when a photon (which typically does not interact with the tracker) converts into an electron–positron pair while passing through matter in the tracker.

The expected (from simulation) impact parameter and transverse momentum resolution of the tracker is shown in Figure 2.3. The momentum scale of the tracker has been measured [21] in 7 TeV 2010 CMS data using $J/\psi \rightarrow \mu^+\mu^-$ decays and is found to agree within 5% with the prediction from simulation. The impact parameter and vertex resolutions have also been measured [22] in data and found to be in excellent agreement with the simulation.

§2.4 Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) of CMS is designed to measure the energy of particles which interact electromagnetically with high precision.⁵ The ECAL is a *scintillation* detector, and functions by counting the number of photons produced in an electromagnetic shower inside a crystal. Upon entering the crystal, a charged particle or photon will inter-

⁵One of the design goals of the CMS experiment is to be able to conduct a search for Standard Higgs bosons decaying to pairs of photons. The branching fraction to photons is illustrated in Figure 1.10.

act electromagnetically with the crystal, producing a shower of electrons and photons. The shower will expand until it consists entirely of photons. The crystal is optically clear, so these photons travel to the rear face of the crystal where they are then counted by a photomultiplier. The number of detected photons can then be related to the energy that was deposited in the crystal. At 18°C, about 4.5 photoelectrons will be produced per MeV of deposited energy. The ECAL has excellent solid angle coverage, extending to a pseudorapidity of $|\eta| = 3.0$.

The ECAL uses lead tungstate (PbWO_4) crystals as the scintillation medium. The crystals have a very large density, which allows the calorimeter to be relatively compact. To be able to correctly measure the energy of electrons and photons, an incident photon or electron must be completely stopped by interactions with the calorimeter. The quantities that determine if an electron or photon will be completely contained is the total depth of the crystal, the crystal density, and the radiation length property X_0 of the crystal. The radiation length X_0 is defined as the mean distance (normalized to material density) after which an electron will have lost $(1 - \frac{1}{e})$ of its energy. The PbWO_4 crystals of the CMS ECAL have a density of 8.28 g/cm³ and a depth of 230 mm. A single crystal thus has a total radiation length of 25.8 X_0 , and will capture on average 99.9993% of the energy of an incident electron. The front face of the crystal is 22 mm \times 22 mm, which corresponds to an $\eta - \phi$ area of 0.00174×0.00174 . The Molière radius of a material is the average radial profile size of an electromagnetic shower, and for PbWO_4 is 2.2 cm. The fact that the Molière radius is larger than the size of the individual crystals improves the spatial resolution of the measurement. As the shower is shared between multiple crystals, the relative amounts deposited in each crystal allows the true impact point to be determined with a resolution smaller than the individual crystal size.

The transparency of the CMS ECAL crystals change as they are exposed to radiation. However, at the working temperature of the ECAL (18°C), the crystal transparency will naturally return to its nominal value. The transparency of the crystals thus decreases during the course of a run of collisions, then increases during the following period collision-less period. The changing transparency conditions need to be continuously monitored and corrected for to ensure a stable detector response. The transparency of the crystals are

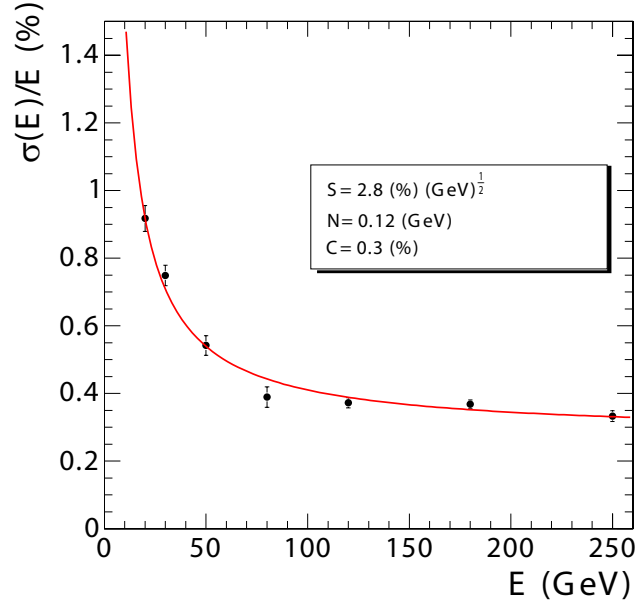


Figure 2.4: Energy resolution (in %) of the CMS ECAL measured at an electron test beam. The resolution depends on the incident energy of the electron. The points are fitted to function given in Equation ???. The fitted parameters are given in the legend.

659 measured continuously using two lasers. One laser has wavelength $\lambda = 400$ nm which
 660 corresponds to the color of light produced in the scintillations and is sensitive to changes
 661 in transparency. The other laser is in the near-infrared and is used to monitor the overall
 662 stability of the crystal. The lasers synchronized to pulse between LHC bunch trains so the
 663 transparency can be continuously monitored while collisions are occurring.

The energy resolution of the ECAL is given by

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2, \quad (2.2)$$

664 where S is a stochastic noise term (due to photon counting statistics), N is a noise term, and
 665 C is a constant term. The parameters of Equation 2.2 have been measured at an electron
 666 test-beam (see Figure 2.4). The energy resolution is better than 1% for electron energies
 667 greater than 20 GeV.

668 §2.5 Hadronic Calorimeter

669 The hadronic calorimeter (HCAL) surrounds the CMS ECAL and is located within the coil
 670 of the CMS solenoid magnet. To ensure incident particles are completely contained within

671 the calorimeter volume, in the barrel region the HCAL employs a “tail-catcher”, an extra
 672 layer of calorimetry outside of the magnet. The hadronic calorimeter measures the energy
 673 of charged and neutral hadronic particles. The HCAL is a *sampling* calorimeter. Layers of
 674 plastic scintillating tiles are interspersed between brass absorber plates. An incident hadron
 675 produces a hadronic shower as it passes through the absorber. The particles in the shower
 676 produce light as they pass through the scintillating tiles. Measuring the light produced in
 677 each layer of tile allows the reconstruction of the radial profile of the shower which can be
 678 related to the deposited energy. The response of the scintillator tiles are calibrated using a
 679 radioactive source, either Cs¹³⁷ or Co⁶⁰. Small stainless tubes permit the radioactive sources
 680 to be moved into the center of the tile during calibration. The granularity of the HCAL is
 681 0.087×0.087 and 0.17×0.17 in $\eta - \phi$ in the barrel ($|\eta| < 1.6$) and endcap ($|\eta| > 1.6$),
 682 respectively.

683 The outer HCAL (HO), or “tail catcher” is designed to capture showers which begin
 684 late in the ECAL or HCAL and ensure they do not create spurious signals in the muon
 685 system (“punch through”). The HO is installed outside of the solenoid magnet in the first
 686 layer between the first to layers of the iron return yoke. The total depth of the HCAL,
 687 including the HO is then 11.8 interaction lengths.

688 The HCAL includes a specially designed forward calorimeter (HF). The design of the
 689 forward calorimeter is constrained by the extreme amount of radiation it is exposed to,
 690 particularly at the highest rapidities. The active material of the HF are quartz fibers. The
 691 fibers are installed inside grooves inside of a steel absorber. Charged particles created in
 692 showers in the absorber create light in the fibers, provided they have energy greater than
 693 the with energy greater than the Cherenkov threshold. As Cherenkov light is created by
 694 the passage of charged particles through matter, the HF design is not sensitive to neutrons
 695 emitted by radionucleids that may be created in the absorber material durin operation.
 696 The fibers are grouped into two sets: one set of fibers are installed over the full depth of
 697 the detector, the other only cover half the depth. A crude form of particle identification
 698 is possible, as showers created by electrons and photons will deposit the majority of the
 699 energy in the front of the detector.

§2.6 Muon System

The ability to detect and measure muons is one of the most valuable tools an experimentalist has at a hadron collider experiment. Muons have particular properties that cause them to leave extremely signatures in the detectors.

- Muons are stable particles, for the typical energies and distances considered at a collider.
- Muons have non-zero charge, so their trajectories can be measured.
- Muons are heavy enough that they are “minimum ionizing particles,” in that they lose much less energy as they pass through material.

The approach to detecting muons is to build the detector to a thickness that typical particles (electrons, photons, hadrons) will not penetrate the outermost calorimeter. Any charged particle that is detected outside of this region can be identified as a muon. At CMS, the muon detection systems are built into the magnet return yoke outside of the CMS calorimeters and magnet, giving them excellent protection (illustrated in Figure 2.5(a)) against hadronic “punch-through.” The purity of particles that reach the muon system make it especially effective as a “trigger” of interesting physics. The CMS muon system has the feature that it additionally can trigger on the transverse momenta of muons. The CMS muon system is composed of three types of detectors: drift tubes (DT), resistive plate chambers (RPC), and cathode strip chambers (CSC).

A drift tube detector is of a tube filled with a mixture of argon (85%) and carbon dioxide (15%) gas with a positively charged ($V = +3.6$ kV) wire running through the middle of the tube. When a charged particle passes through the tube, it ionizes some gas. The free electrons are then drawn to the positively charged wire inside the tube, creating a signal when reach it. The speed of the detector is limited by the “drift time,” the maximum amount of time it may take for an electron to reach sensor wire. The precision of the spatial measurement can be increased by recording the time at which each wire records a signal and correlating the measurements across multiple tubes. The time resolution of the CMS DTs is on the order of a few nanoseconds, allowing the DT to provide a trigger on a given

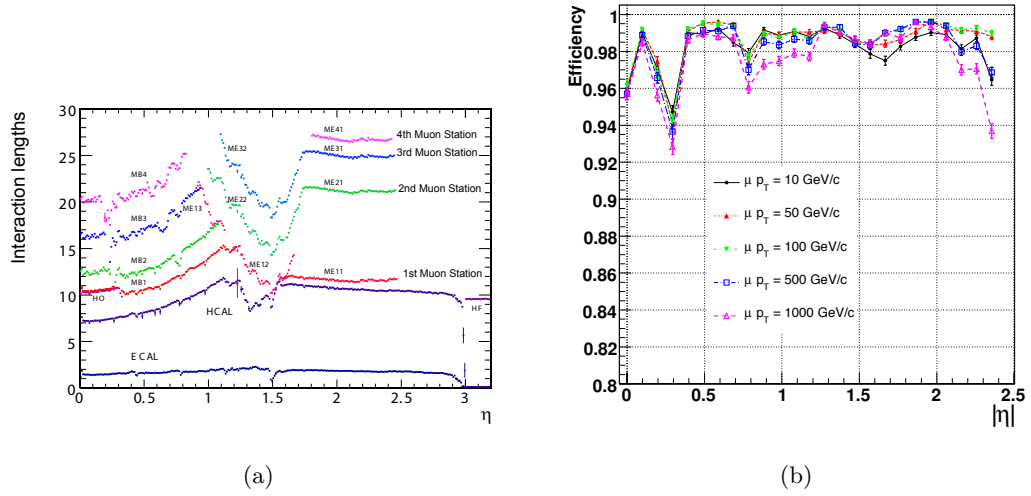


Figure 2.5: The left figure, (a), illustrates the number of interaction lengths versus pseudorapidity η of material that must be traversed before reaching the different layers of the muon system. On the right, (b) shows the efficiency versus η to reconstruct a “global” muon for different transverse momenta.

728 proton bunch crossing. The tubes in adjacent layers are offset by one half tube width to
 729 take advantage of this effect and ensure there are no gaps in the fiducial region. In CMS, the
 730 smallest unit of the DT system is the superlayer, which consists of four layers of tubes. A
 731 DT chamber consists of three or two superlayers. The tubes in the two superlayers farthest
 732 from the beam are oriented parallel to the beam and measured the bending of the muons in
 733 the magnetic field. The inner superlayer is oriented orthogonally to the beam and measures
 734 the longitudinal position of incident muons. There are four muon “stations” in the barrel
 735 which contain DT chambers. The stations correspond to available areas in the magnetic
 736 return yoke. In the barrel, the muon momentum resolution of the DTs is better than 95%.

737 §2.7 Trigger System

738 §2.8 Particle Flow Reconstruction Algorithm

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 [23]. 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 [23]. 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 [24]. 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% [23].

§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

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

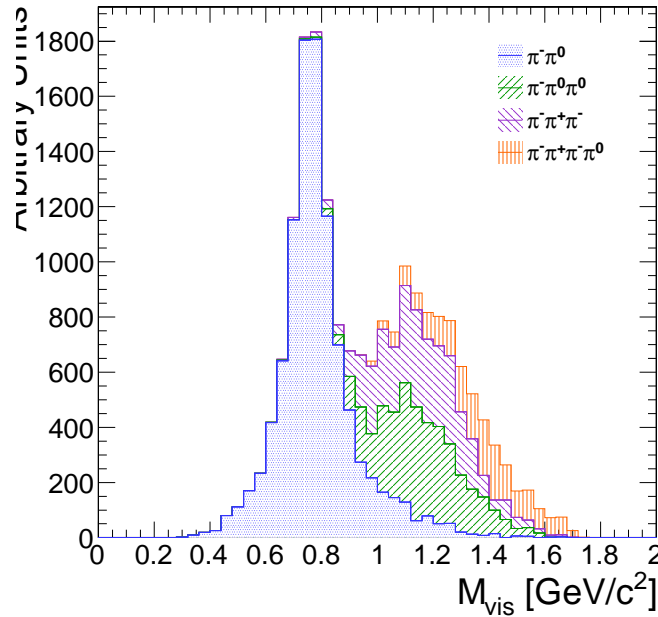


Figure 3.1: The invariant mass of the visible decay products in hadronic tau decays. The decay mode $\tau^- \rightarrow \pi^- \nu_\tau$ is omitted. The different decay modes have different invariant masses corresponding to the intermediate resonance in the decay.

800 §3.4 The Tau Neural Classifier

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

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

809 §3.4.1 Decay mode reconstruction

810 The major task in reconstructing the decay mode of the tau is determining the number of
 811 π^0 mesons produced in the decay. A π^0 meson decays almost instantaneously to a pair of
 812 photons. The photon objects are reconstructed using the particle flow algorithm [24]. The
 813 initial collection of photon objects considered to be π^0 candidates are the photons in the
 814 signal cone described by using the “shrinking-cone” tau algorithm, described in [23].

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

826 Photon merging

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

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

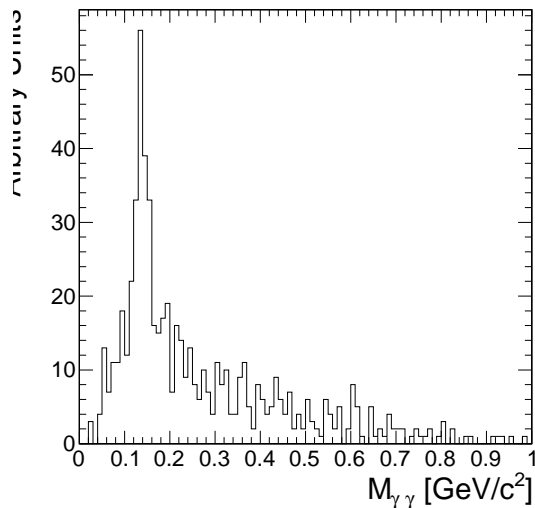


Figure 3.2: Invariant mass of the photon pair for reconstructed tau-candidates with two reconstructed photons in the signal region that are matched to generator level $\tau^- \rightarrow \pi^- \pi^0 \nu_\tau$ decays.

837 Quality requirements

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

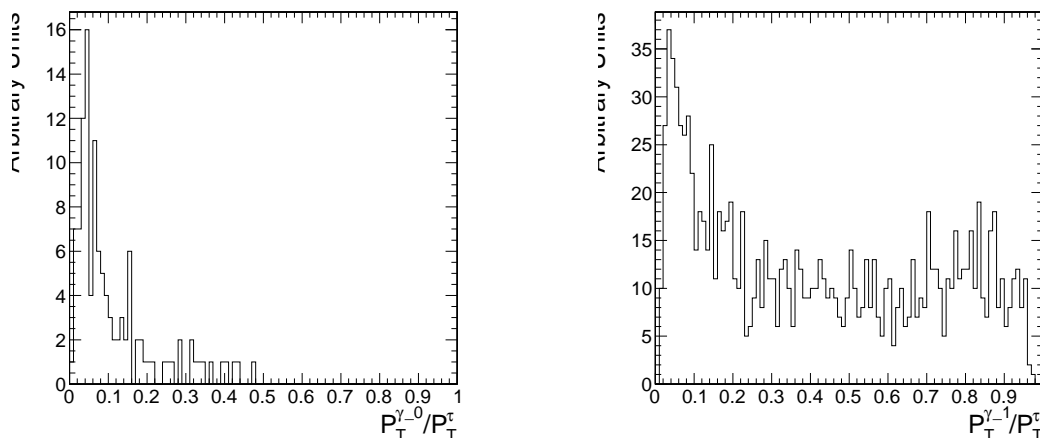


Figure 3.3: Fraction of total τ -candidate transverse momenta carried by the photon for reconstructed taus containing a single photons for two benchmark cases. On the left, the reconstructed tau-candidate is matched to generator level $\tau^- \rightarrow \pi^- \nu_\tau$ decays, for which no photon is 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.

849 Performance

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

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

864 In the forward region, nuclear interactions and conversions from the increased material
 865 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.

866 §3.4.2 Neural network classification

867 Neural Network Training

868 The samples used to train the TaNC neural networks are typical of the signals and back-
 869 grounds found in common physics analyses using taus. The signal-type training sample is

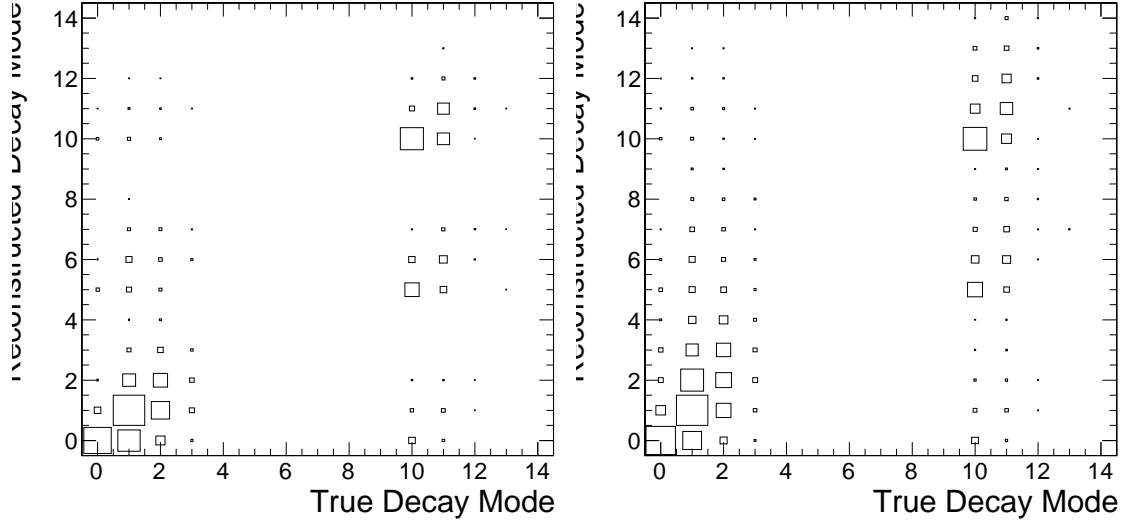


Figure 3.4: Correlations between reconstructed tau decay mode and true tau decay mode for hadronic tau decays in $Z \rightarrow \tau^+\tau^-$ events. The correlation when no photon merging or filtering is applied is shown on the left, and the correlation for the algorithm described in section 3.4.1 is 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.

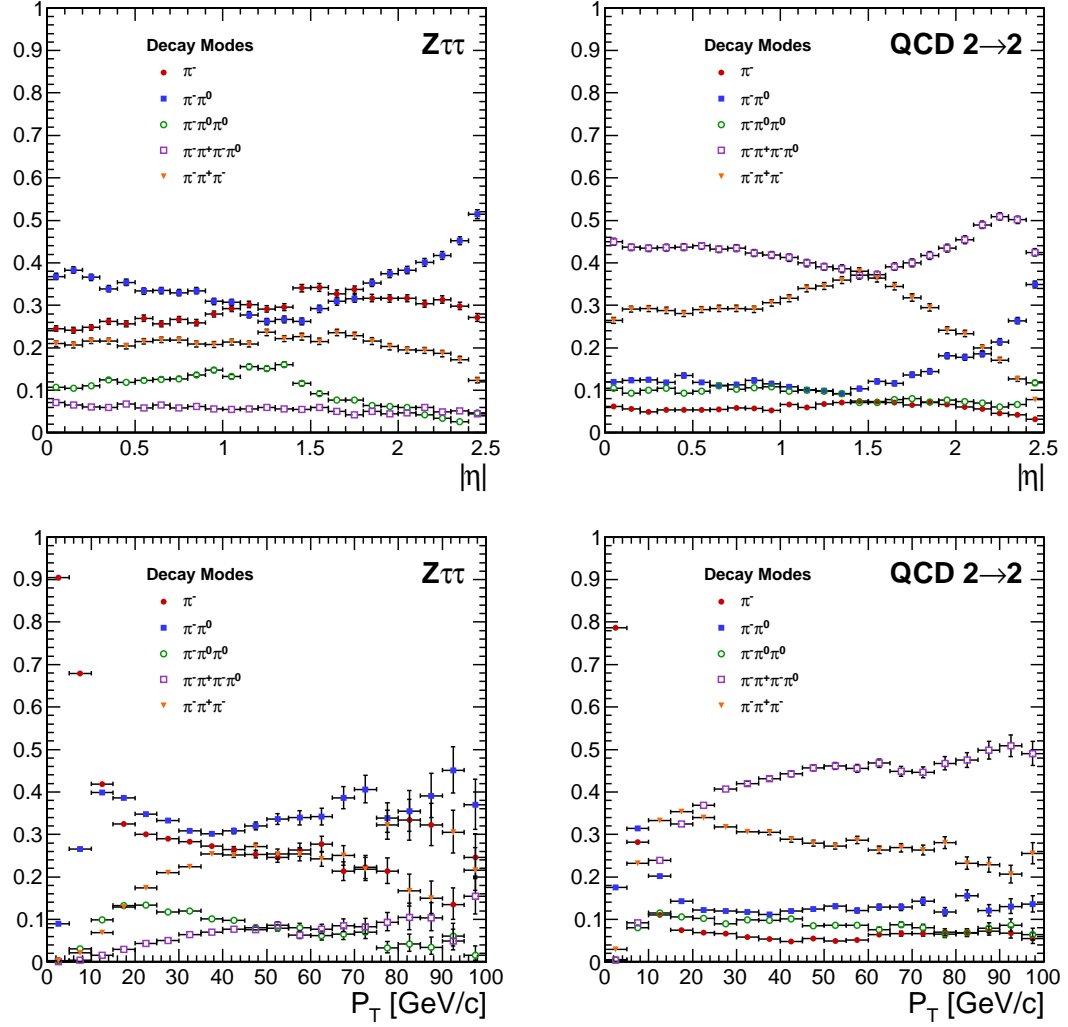


Figure 3.5: Kinematic dependence of reconstructed decay mode for tau-candidates from $Z \rightarrow \tau^+\tau^-$ (left) and QCD di-jets (right) versus transverse momentum (top) and pseudo-rapidity (bottom). Each curve is the probability for a tau-candidate to be reconstructed with the associated decay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

870 composed of reconstructed tau-candidates that are matched to generator level hadronic tau
 871 decays coming from simulated $Z \rightarrow \tau^+\tau^-$ events. The background training sample consists
 872 of reconstructed tau-candidates in simulated QCD $2 \rightarrow 2$ hard scattering events. The QCD
 873 p_T spectrum is steeply falling, and to obtain sufficient statistics across a broad range of p_T
 874 the sample is split into different \hat{P}_T bins. Each binned QCD sample imposes a generator
 875 level cut on the transverse momentum of the hard interaction. During the evaluation of
 876 discrimination performance the QCD samples are weighted according to their respective
 877 integrated luminosities to remove any effect of the binning.

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

888 The remaining 20% of the signal and background samples are reserved as a statistically
 889 independent sample to evaluate the performance of the neural nets after the training is
 890 completed. The TaNC uses the “MLP” neural network implementation provided by the
 891 TMVA software package, described in [25]. The “MLP” classifier is a feed-forward artificial
 892 neural network. There are two layers of hidden nodes and a single node in the output layer.
 893 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 [26], any continuous function $g(x)$

	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.

defined on a vector space of dimension d spanned by x can be represented by

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

894 for suitably chosen functions for Φ_j and ϕ_j . As the form of equation 3.1 is similar to the
895 topology of a two hidden-layer neural network, Kolmogorov's theorem suggests that *any*
896 classification problem can be solved with a neural network with two hidden layers containing
897 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 [25] as

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

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

901 The neural networks use as input observables the transverse momentum and η of the
902 tau-candidates. These observables are included as their correlations with other observables
903 can increase the separation power of the ensemble of observables. For example, the opening
904 angle in ΔR for signal tau-candidates is inversely related to the transverse momentum,

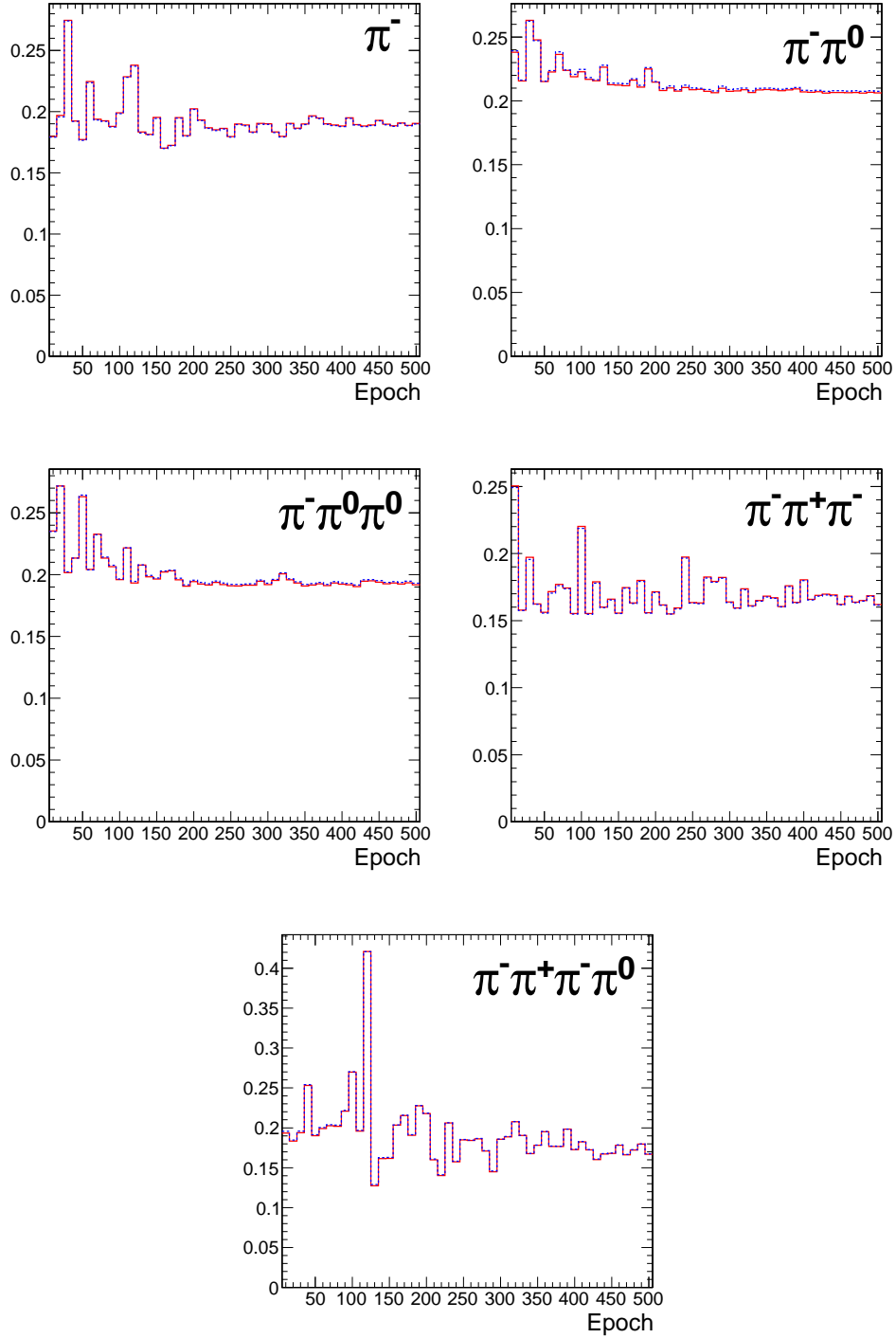


Figure 3.6: Neural network classification error for training (solid red) and testing (dashed blue) samples at ten epoch intervals over the 500 training epochs for each decay mode neural network. The vertical axis represents the classification error, defined by equation 3.2. N.B. that the choice of hyperbolic 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.

905 while for background events the correlation is very small [?]. In the training signal and
 906 background samples, there is significant discrimination power in the p_T spectrum. However,
 907 it is desirable to eliminate any systematic dependence of the neural network output on p_T
 908 and η , as in practice the TaNC will be presented with tau-candidates whose $p_T - \eta$ spectrum
 909 will be analysis dependent. The dependence on p_T and η is removed by applying a p_T and
 910 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 training sample is exclusively signal or background. The weights are computed according to

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

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

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

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

914 Discriminants

915 Each neural network corresponds to a different decay mode topology and as such each
 916 network uses different observables as inputs. However, many of the input observables are
 917 used in multiple neural nets. The superset of all observables is listed and defined below.
 918 Table 3.4 maps the input observables to their associated neural networks. The signal and
 919 background distributions of the input observables for tau-candidates in the training sample
 920 are shown in appendix ???. In three prong decays, the definition of the “main track” is
 921 important. The main track corresponds to the track with charge opposite to that of the
 922 total charge of the three tracks. This distinction is made to facilitate the use of the “Dalitz”
 923 observables, allowing identification of intermediate resonances in three-body decays. This
 924 is motivated by the fact that the three prong decays of the tau generally proceed through

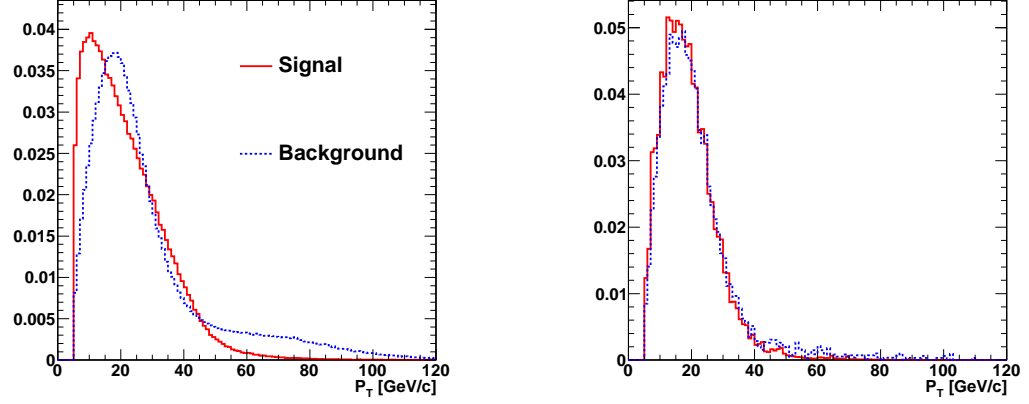


Figure 3.7: Transverse momentum spectrum of signal and background tau-candidates used in neural net training before (left) and after (right) the application of $p_T - \eta$ dependent weight function. Application of the weights lowers the training significance of tau-candidates in regions of $p_T - \eta$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

925 $\tau^- \rightarrow a1^- \nu_\tau \rightarrow \pi^- \rho^0 \nu_\tau \rightarrow \pi^- \pi^+ \pi^- \nu_\tau$; the oppositely charged track can always be identified
 926 with the ρ^0 decay.

927 **ChargedOutlierAngleN**

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

931 **ChargedOutlierPtN**

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

934 **DalitzN**

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

937 **Eta**

938 Pseudo-rapidity of the signal region objects.

939 **InvariantMassOfSignal**

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

941 **MainTrackAngle**

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

944 **MainTrackPt**

945 Transverse momentum of the “main track.”

946 **OutlierNCharged**

947 Number of charged objects in the isolation region.

948 **OutlierSumPt**

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

950 **PiZeroAngleN**

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

953 **PiZeroPtN**

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

955 **TrackAngleN**

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

958 **TrackPtN**

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

961 **Neural network performance**

962 The classification power of the neural networks is unique for each of the decay modes.

963 The performance is determined by the relative separation of the signal and background

964 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 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 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.

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

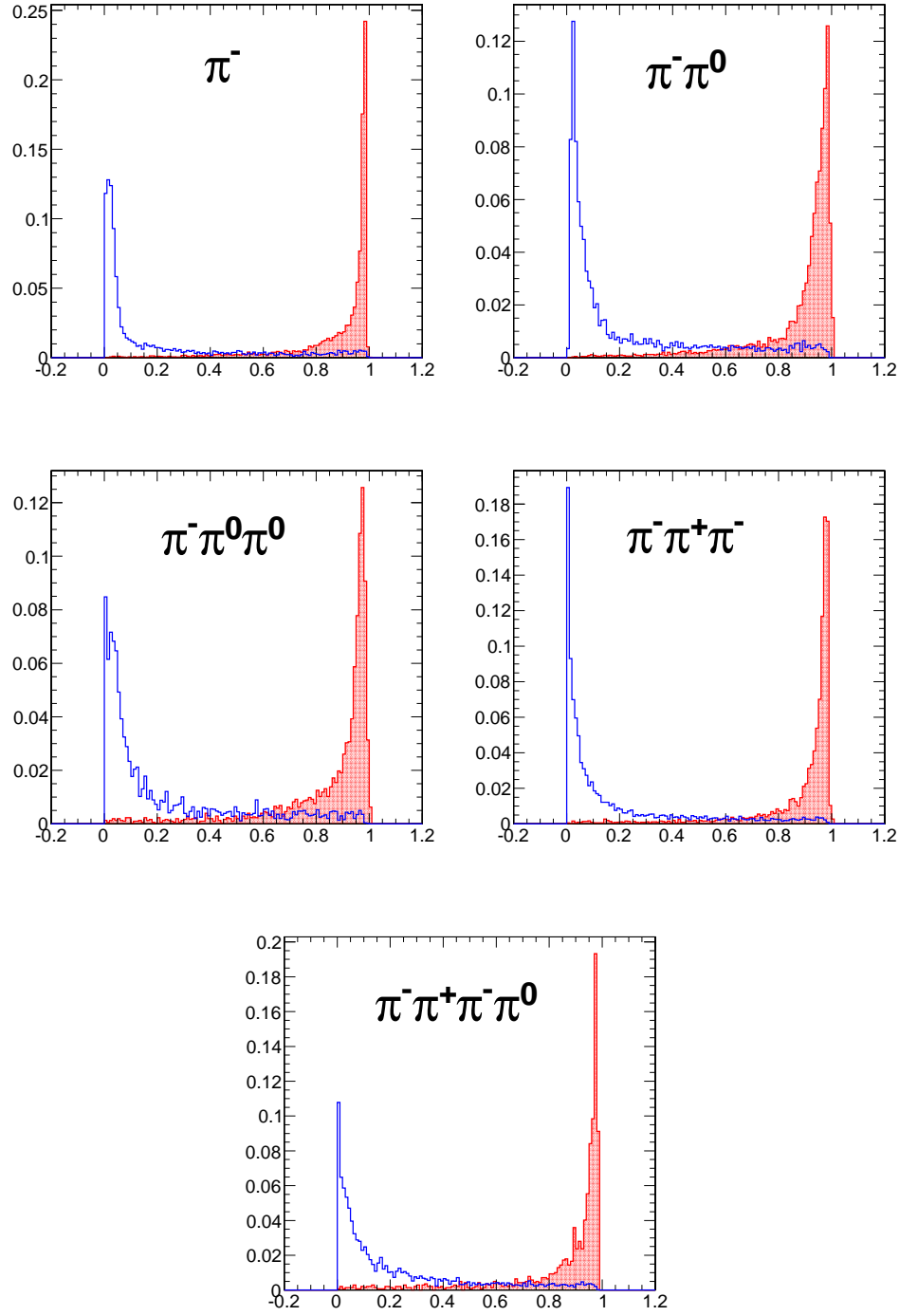


Figure 3.8: Neural network output distributions for the five reconstructed tau-candidate decay modes used in the TaNC for $Z \rightarrow \tau^+\tau^-$ events (red) and QCD di-jet events (blue).

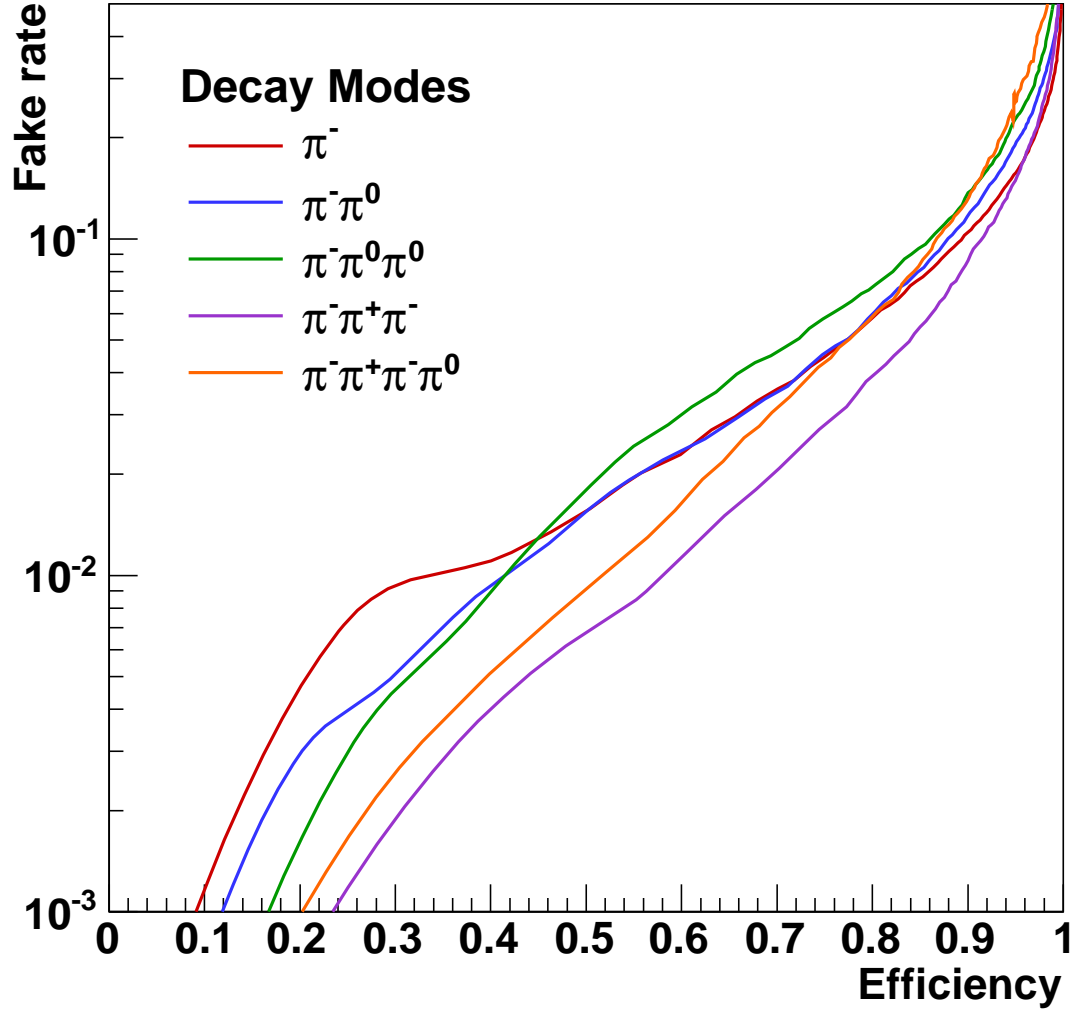


Figure 3.9: Performance curves for the five neural networks used by the TaNC for tau-candidates with transverse momentum greater than 20 GeV/ c . Each curve represents the signal efficiency (on the horizontal axis) and background misidentification rate (vertical axis) for a scan of the neural network 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.

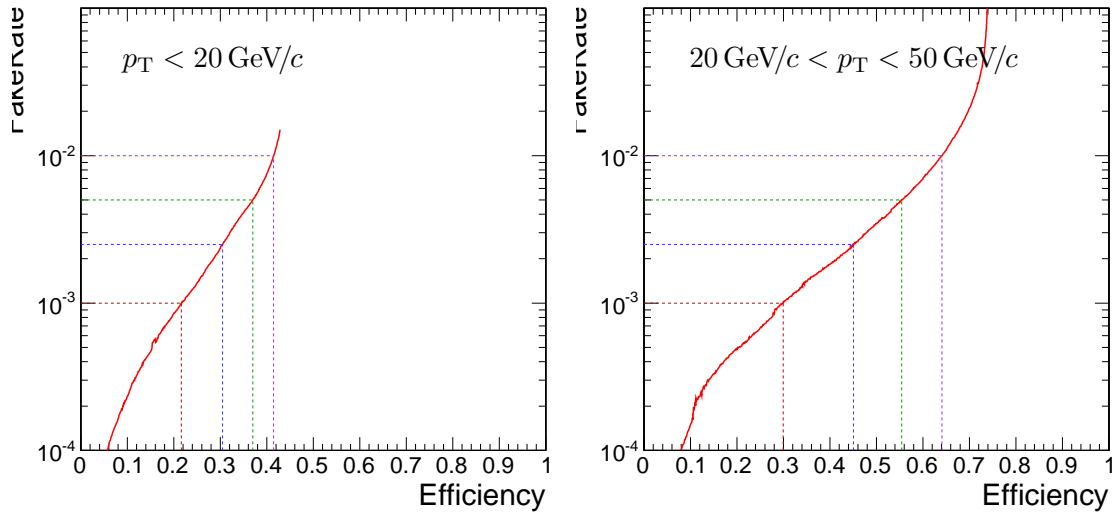


Figure 3.10: Tau Neural Classifier performance curves for tau-candidates with $p_T < 20 \text{ GeV}/c$ (left) and $20 \text{ GeV}/c < p_T < 50 \text{ GeV}/c$ (right). The vertical axis represents the expected fake-rate of QCD jets and the horizontal axis the expected signal efficiency for hadronic tau 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.

“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.

§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

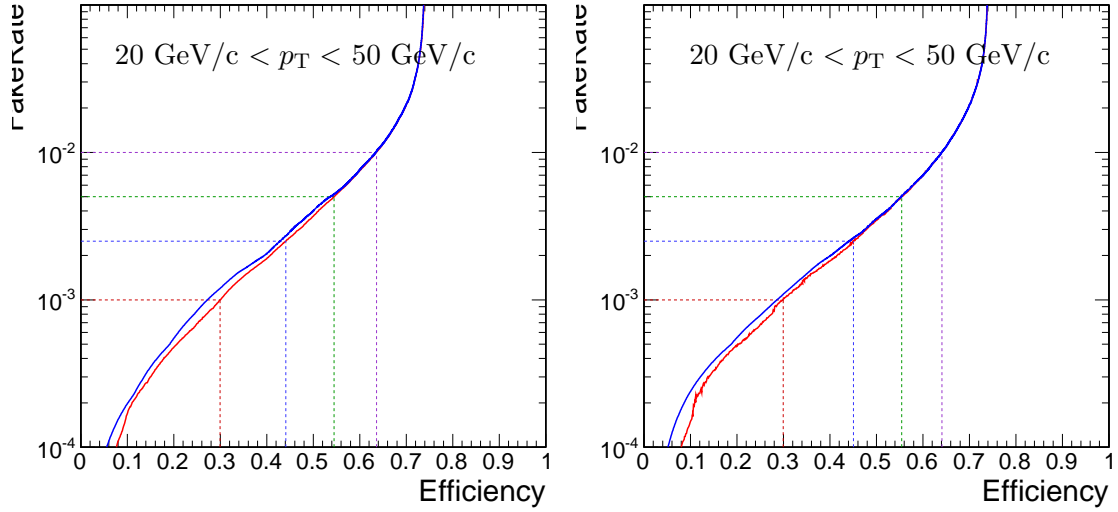


Figure 3.11: Tau Neural Classifier performance curves for tau-candidates with $20 \text{ GeV}/c < p_T < 50 \text{ GeV}/c$. The figure on the left compares the optimal performance curve determined by the Monte Carlo sampling method (red) to the performance curve obtained by scanning the “transform cut” (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.

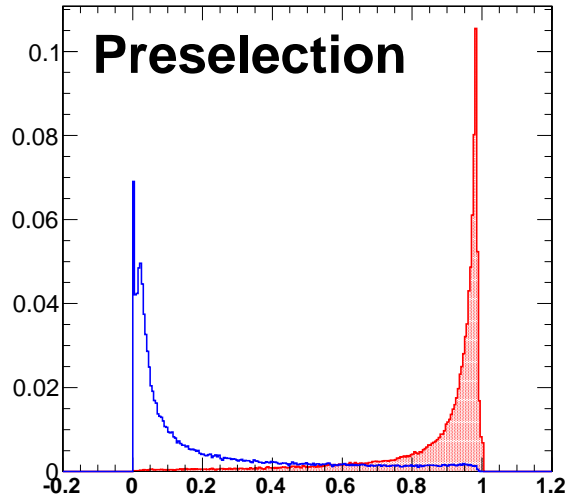


Figure 3.12: Transformed TaNC neural network output for tau-candidates with transverse momentum between 20 and 50 GeV/c that pass the pre-selection criteria. The neural network output for each tau-candidate has been transformation according to equation 3.3. The decay mode probabilities ρ_i^{bkg} , ρ_i^{signal}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

1018 points and the isolation based tau identification are show in figure 3.14. For tau-candidates
1019 with transverse momentum between 20 and 50 GeV/c, the TaNC operating cut can be
1020 chosen such that the two methods have identical signal efficiency; at this point the TaNC
1021 algorithm reduces the background fake rate by an additional factor of 3.9. This reduction
1022 in background will directly improve the significance of searches for new physics using tau
1023 leptons at CMS.

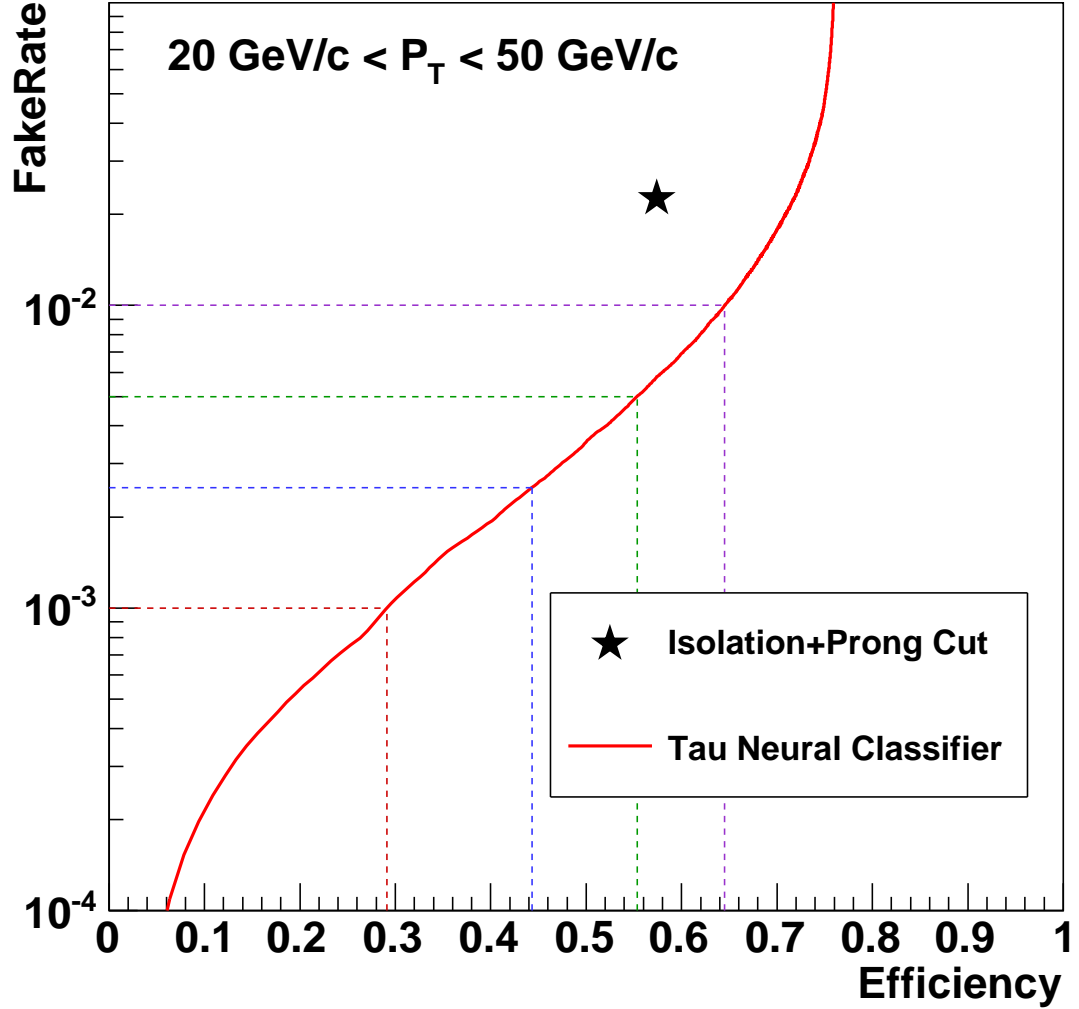


Figure 3.13: Performance curve (red) of the TaNC tau identification for various requirements on the output transformed according to equation 3.3. The horizontal axis is the efficiency for true taus with transverse momentum between 20 and 50 GeV/c to satisfy the tau identification requirements. The vertical 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.

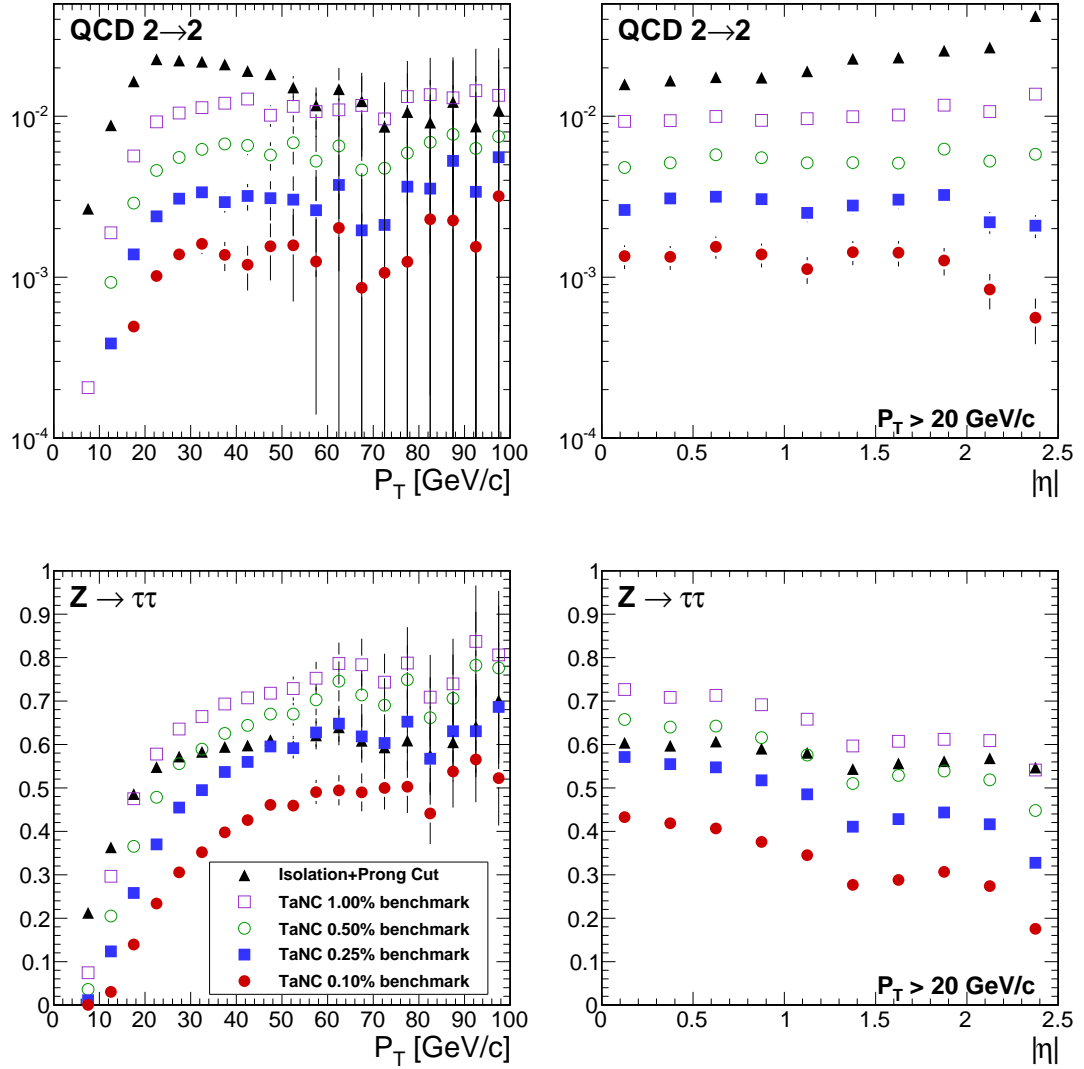


Figure 3.14: Comparison of the identification efficiency for hadronic tau decays from $Z \rightarrow \tau^+\tau^-$ decays (bottom row) and the misidentification rate for QCD di-jets (top row) versus tau-candidate transverse momentum (left) and pseudo-rapidity (right) for different tau identification 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.

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.

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 ($m_Z = 91.1876 \pm 0.0021$ GeV/ c^2) and has a narrow width ($\Gamma_Z = 2.4952 \pm 0.0023$ 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 elated 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.

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

1045 The Collinear Approximation suffers from two problems. The approximation can fail
 1046 (yielding unphysical negative energies for the reconstructed neutrinos) when the missing
 1047 transverse energy is mis-measured. The events with unphysical solutions must be removed
 1048 from the analysis, leading to a dramatic reduction in acceptance (on the order of 50% in
 1049 this analysis). Improvements to the collinear approximation algorithm have recently been
 1050 made which aim to recover part of the events with unphysical solutions [?]. But even with
 1051 these improvements, no physical solution is still found for a large fraction of signal events.
 1052 Additionally, the method is numerically sensitive when the two τ lepton are nearly back-
 1053 to-back in azimuth. In these cases the $\sin(\phi_2 - \phi_1)^{-1}$ term in Equation 4.2 is very large
 1054 and small mis-measurements of the missing transverse energy can produce a large tail on
 1055 the reconstructed mass. This tail is particularly large for low-mass resonances. The large
 1056 tail for low mass is predominantly due to the fact (discussed in subsection 4.4.2) that the

1057 kinematic requirements¹ applied on the visible decay products preferentially selects events
 1058 where the visible decay products carry the majority of the energy of the original τ lepton,
 1059 reducing the amount of true missing energy in the event.

1060 §4.2 The Secondary Vertex fit

1061 A novel algorithm is presented in the following, which succeeds in finding a physical solution
 1062 for every event. As an additional benefit, the new algorithm is found to improve the di-tau
 1063 invariant mass resolution, making it easier to separate the Higgs signal from the $Z \rightarrow \tau^+\tau^-$
 1064 background.

1065 The novel Secondary Vertex fit (SVfit) algorithm for di-tau invariant mass recon-
 1066 struction that we present in the following utilizes a likelihood maximization to fit a $\tau^+\tau^-$
 1067 invariant mass hypothesis for each event. The likelihood is composed of separate terms
 1068 which represent probability densities of:

- 1069 • tau decay kinematics
- 1070 • matching between the momenta of neutrinos produced in the tau decays and the
 1071 reconstructed missing transverse momentum
- 1072 • a regularization “ p_T -balance” term which accounts for the effects on the di-tau in-
 1073 variant mass of acceptance cuts on the visible tau decay products
- 1074 • the compatibility of tau decay parameters with the position of reconstructed tracks
 1075 and the known tau lifetime of $c\tau = 87 \mu\text{m}$ [16].

1076 The likelihood is maximized as function of a set of parameters which fully describe the tau
 1077 decay.

1078 §4.3 Parametrization of tau decays

1079 The decay of a tau of visible four-momentum p_{vis} measured in the CMS detector (“labora-
 1080 tory”) frame can be parametrized by three variables. The invisible (neutrino) momentum
 1081 is fully determined by these parameters.

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.

1082 The “opening-angle” θ is defined as the angle between the boost direction of the tau
 1083 lepton and the momentum vector of the visible decay products in the rest frame of the
 1084 tau. The azimuthal angle of the tau in the lab frame is denoted as $\bar{\phi}$ (we denote quantities
 1085 defined in the laboratory frame by a overline). A local coordinate system is defined such
 1086 that the \bar{z} -direction lies along the visible momentum and $\bar{\phi} = 0$ lies in the plane spanned
 1087 by the momentum vector of the visible decay products and the proton beam direction. The
 1088 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:

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

1089 Note that for hadronic decays, $m_{\nu\nu}$ is a constant of value zero, as only a single neutrino is
 1090 produced. Consequently, the magnitude of P^{vis} depends on the reconstructed mass of the
 1091 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:

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

1092 The energy of the tau lepton in the laboratory frame as function of the measured visible
 1093 momentum depends on two of the three parameters only - the rest frame opening angle θ and
 1094 the invariant mass $m_{\nu\nu}$ of the neutrino system. The direction of the tau lepton momentum

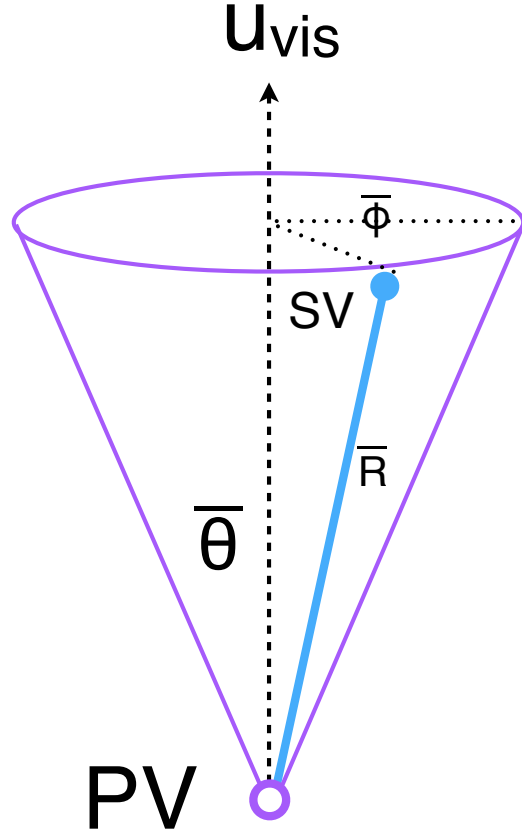


Figure 4.1: Illustration of the coordinate system used by the SVfit to describe the decays of tau leptons.

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 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:

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

1101 In the present implementation of the SVfit algorithm, the matrix element is assumed to be
1102 constant, so that the likelihood depends on the phase-space volume of the decay only ².

1103 §4.4.1 Likelihood for reconstructed missing transverse momentum

1104 Momentum conservation in the plane perpendicular to the beam axis implies that the
1105 vectorial sum of the momenta of all neutrinos produced in the decay of the tau lepton pair
1106 matches the reconstructed missing transverse momentum. Differences are possible due to
1107 the experimental resolution and finite p_T of particles escaping detection in beam direction
1108 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:

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

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

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

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

1109 where q_T denotes the transverse momentum of the tau lepton pair.

1110 §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:

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

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

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

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

1111 The motivation to add the p_T -balance likelihood to the SVfit is to add a “regulariza-
1112 tion” term which compensates for the effect of p_T cuts applied on the visible decay products

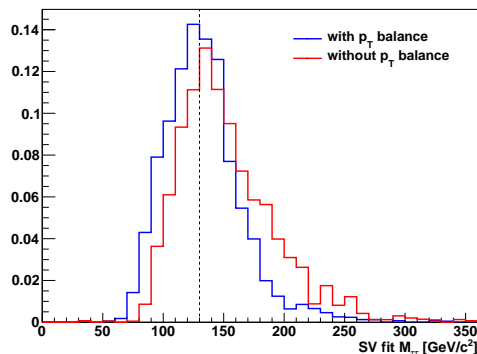


Figure 4.2: Distribution of di-tau invariant mass reconstructed by the SVfit algorithm in simulated Higgs events with $m_{A^0} = 130 \text{ GeV}/c^2$. The SVfit algorithm is run in two configurations, with (blue) and without (red) the p_T -balance likelihood term included in the fit.

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

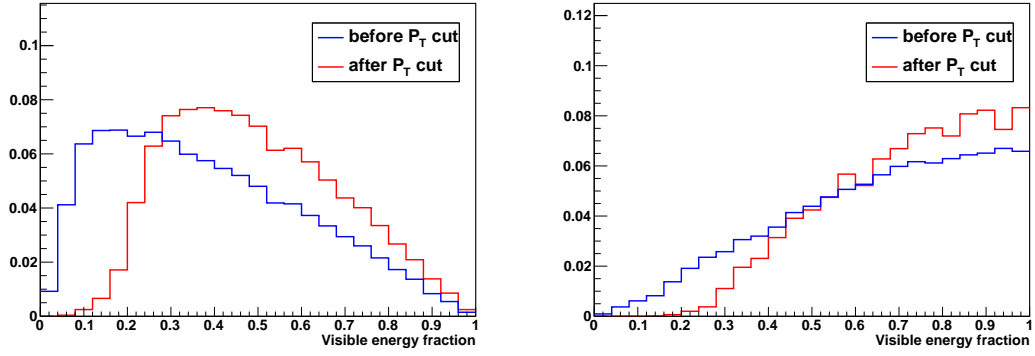


Figure 4.3: Normalized distributions of the fraction of total tau decay energy carried by the muon (left) and hadronic constituents (right) in simulated Higgs events with $m_{A^0} = 130 \text{ GeV}/c^2$. The distribution is shown before (blue) and after (red) the requirement on the p_T of the visible decay products described in section ??.

1131 $\tau^+\tau^-$ invariant mass reconstructed by the SVfit algorithm. The likelihood term based on
 1132 the secondary vertex information is based on the compatibility of the decay vertex position
 1133 with the reconstructed tracks of charged tau decay products. Perhaps surprisingly, it turns
 1134 out that the flight distance parameter R is sufficiently constrained even for tau decays into
 1135 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:

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

1136 The likelihood terms for the secondary vertex fit have been implemented in the SVfit
 1137 algorithm. In the analysis presented in this note, the decay vertex information is not used,
 1138 however, because of systematic effects arising from tracker (mis-)alignment which are not
 1139 yet fully understood.

1140 §4.5 Performance

1141 This section describes the performance of the SVfit algorithm for reconstructing the in-
 1142 variant mass of resonances decaying to τ lepton pairs. The performance is presented as a
 1143 contrasted to the performance of the

1144 Both of these effects are illustrated in Figure 4.5. Blah blah blah see figure 4.6.

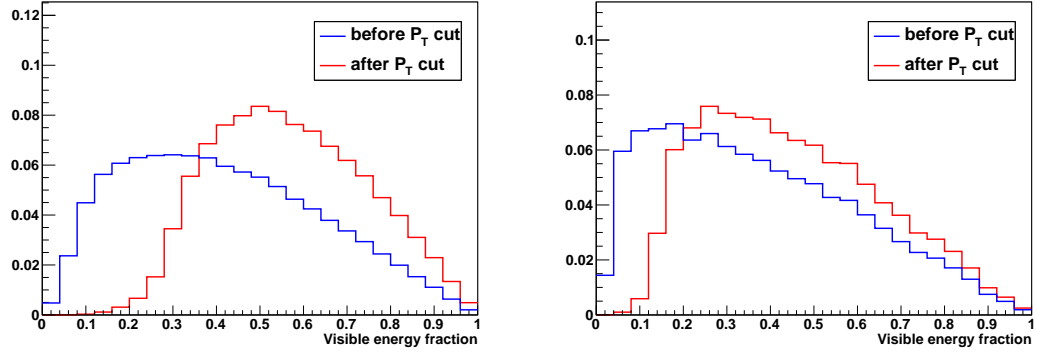


Figure 4.4: Normalized distributions of the fraction of total tau decay energy carried by the muon in simulated $Z \rightarrow \tau^+\tau^-$ (left) and Higgs events with $m_{A^0} = 200 \text{ GeV}/c^2$ (right). The distribution is shown before (blue) and after (red) the requirement that the p_T of the muon be greater than $15 \text{ GeV}/c$.

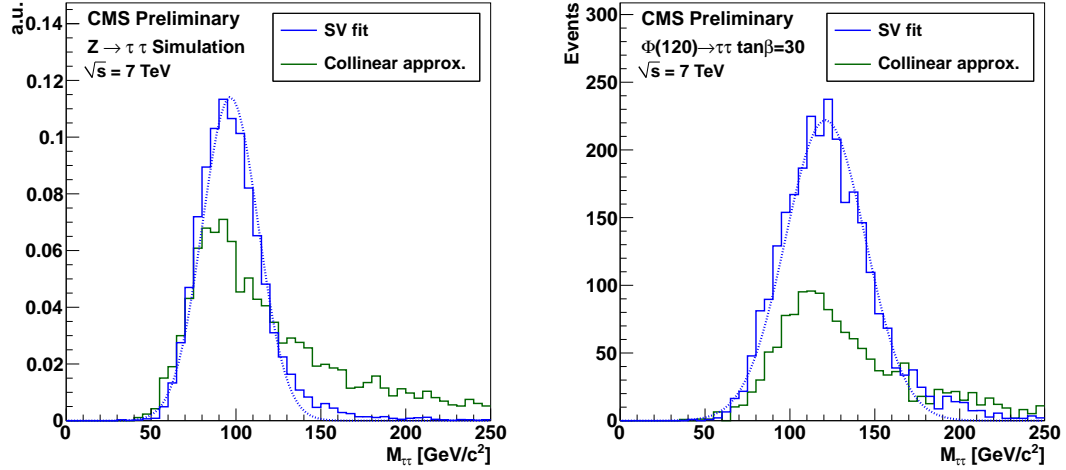


Figure 4.5: Comparison of the reconstructed tau pair mass spectrum in $Z \rightarrow \tau^+\tau^-$ (left) and MSSM $H(120) \rightarrow \tau^+\tau^-$ (right) events after the selections described in chapter 5. The mass spectrum reconstructed by the Secondary Vertex fit is shown in blue, the result of 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.

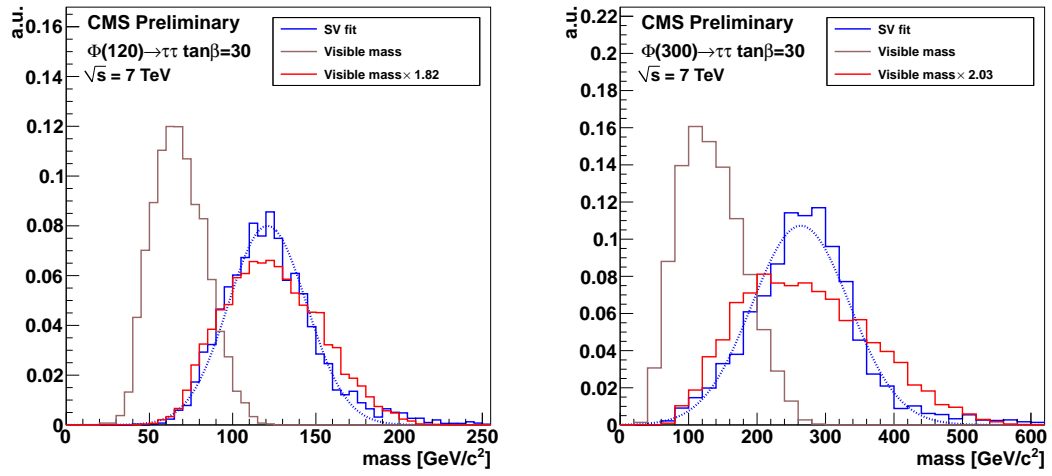


Figure 4.6: Comparison of the invariant mass of the muon and τ_{jet} (the “visible mass”) with the full $\tau^+\tau^-$ mass reconstructed by the SVfit. The spectrum is shown for two simulated MSSM Higgs samples, with $m_{A^0} = 120 \text{ GeV}/c^2$ (left), and $m_{A^0} = 200 \text{ GeV}/c^2$ (right). 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	$\text{charge}(\text{Muon}) + \text{charge}(\text{Tau-jet}) = 0$, $\Delta R(\text{Muon}, \text{Tau-jet}) > 0.5$
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.

1170 The events are triggered by a combination of muon and muon + tau-jet “cross-channel”
 1171 triggers. For the muon triggers, paths with lowest P_T thresholds are used as long as the path
 1172 remained unprescaled (see table 5.2). The muon + tau-jet “cross-channel” trigger paths
 1173 increase the trigger efficiency for events containing muons of transverse momenta close to
 1174 the $P_T^\mu > 15$ GeV cut threshold. The trigger efficiency is measured in data via the tag-and-
 1175 probe technique. Details of the muon trigger efficiency measurement are given in section ??
 1176 of the appendix. Monte Carlo simulated events are required to pass the HLT_Mu9 trigger
 1177 path. Weights are applied to simulated events to account for the difference between the sim-
 1178 ulated HLT_Mu9 efficiency and the combined efficiency of the set HLT_Mu9, HLT_IsoMu9,
 1179 HLT_Mu11, HLT_IsoMu13, HLT_Mu15, HLT_IsoMu9_PFTau15 and HLT_Mu11_PFTau15

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.

1180 used to trigger the data.

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 + jets, $t\bar{t}$ + jets and $Z \rightarrow \mu^+\mu^-$ backgrounds in the measurement of the $Z \rightarrow \tau^+\tau^-$ cross–section, in the channel $Z \rightarrow \tau^+\tau^- \rightarrow \mu + \tau$ -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.

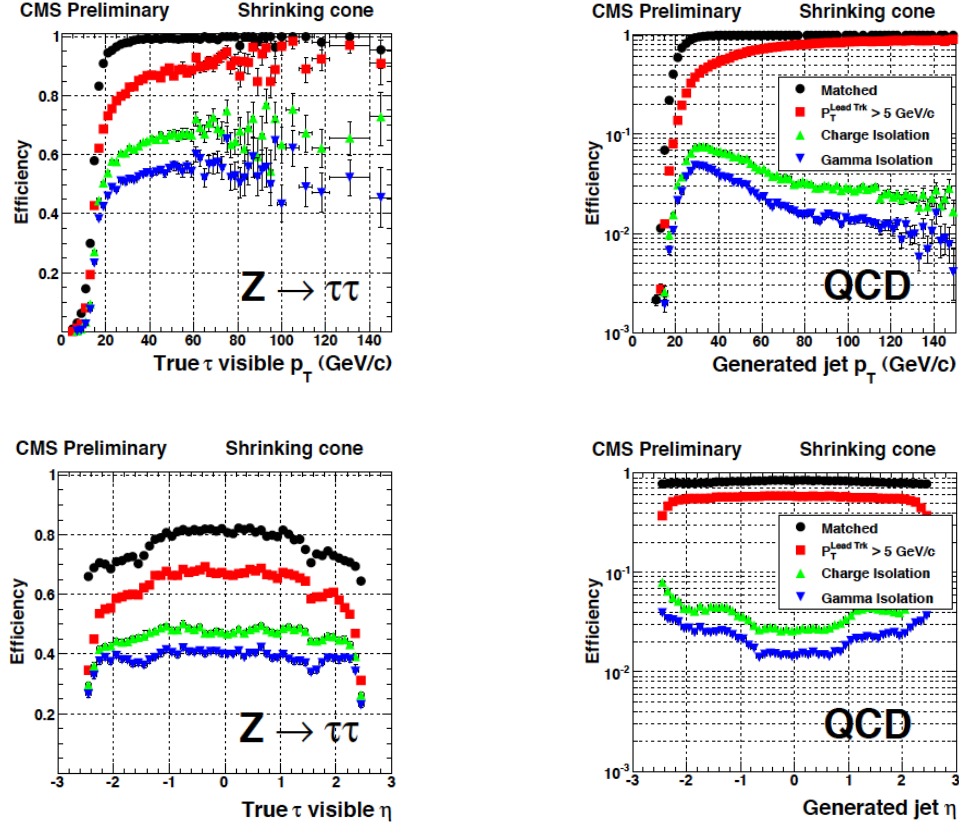


Figure 6.1: Cumulative efficiencies (left) and fake-rates (right) of successively applied tau identification cuts of the “shrinking signal cone” particle-flow based tau identification algorithm described in [?] as function of p_T^{jet} (top) and η^{jet} (bottom) of tau-jet candidates. The efficiencies/fake-rates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

§6.2.1 Parameterization of fake-rates

Efficiencies and fake-rates of the tau identification algorithm based on requiring no tracks of $p_T > 1 \text{ GeV}/c$ and ECAL energy deposits of $p_T > 1.5 \text{ GeV}$ reconstructed within an “isolation 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 $Z \rightarrow \tau^+\tau^- \rightarrow \mu + \tau\text{-jet}$ 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

1211 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

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

1212 where $E(\eta^2)$ ($E(\phi^2)$) is the second η (ϕ) moment of the jet constituents, weighted by
 1213 constituent transverse energy. Analyses performed by the CDF collaboration [?] found that
 1214 systematic uncertainties on background estimates obtained from the fake-rate method are
 1215 reduced in case differences between quark and gluon jets get accounted for in this way.

1216 §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} :

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

1217 The pre-selection in the denominator of equation 6.2 in general refers to p_T and η cuts,
 1218 which are applied with thresholds matching those applied on the final analysis level, but
 1219 may as well include loose tau identification criteria (which may be applied e.g. already
 1220 during event skimming).

1221 Different sets of fake-rates are determined for the highest p_T and for the second highest
 1222 p_T jet in QCD di-jet events, for jets in a QCD event sample enriched by the contribution
 1223 of heavy quarks and gluons by requiring the presence of a muon reconstructed in the final
 1224 state, and for jets in “electroweak” events selected by requiring a W boson in the final state.

1225 §6.2.3 The Fake-rate method

1226 Knowledge of the tau identification efficiencies and fake-rates as function of the parameters
 1227 p_T^{jet} , η_{jet} and R_{jet} as defined by equation 6.2 is utilized to obtain an estimate for the contri-
 1228 butions of background processes to physics analyses involving tau lepton hadronic decays
 1229 in the final state. The basic idea is to replace tau identification cuts and discriminators by
 1230 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 $P_{fr}(p_T^{jet}, \eta_{jet}, R_{jet})$ 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.

1258 “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:

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

1259 These weights are applied to all jets in the background dominated control sample which
 1260 pass the preselection defined by the denominator of equation 6.2. Note that the weights
 1261 defined by equation 6.3 can be used to estimate the contributions of background processes
 1262 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:

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

1263 Different estimates are obtained for the fake-rate probabilities determined for the high-
 1264 est and second highest p_T jet in QCD di-jet events, jets in a muon enriched QCD sample
 1265 and jets in W +jets events. The arithmetic average of the four estimates together with
 1266 the difference between the computed average and the minimum/maximum value is given in
 1267 table ??.

1268 We take the average value as “best” estimate of the background contribution and

1269 the difference between the average and the minimum/maximum estimate as its systematic
 1270 uncertainty. We obtain a value of $\mathcal{O}(15\%)$ for the systematic uncertainty and find that
 1271 the true sum of QCD, W +jets, $t\bar{t}$ +jets and $Z \rightarrow \mu^+\mu^-$ background contributions agrees
 1272 well with the “best” estimate obtained by the fake-rate method within the systematic
 1273 uncertainty.

1274 Note that the estimate for the sum of background contributions which one obtains in
 1275 case one applies the “simple” fake-rate weights defined by equation 6.4 to a background
 1276 dominated control sample selected in data is likely to overestimate the true value of back-
 1277 ground contributions by a significant amount. The reason is that contributions of the
 1278 $Z \rightarrow \tau^+\tau^-$ signal are non-negligible. In fact, signal contributions to the background domi-
 1279 nated control sample are expected to be 14.9% and since the per-jet weights computed by
 1280 equation 6.3 are larger on average in signal than in background events, the signal contribu-
 1281 tion increases by the weighting and amounts to 37.1% of the sum of event weights computed
 1282 by equation 6.4 and given in table ??.

1283 The contribution of the $Z \rightarrow \tau^+\tau^-$ signal needs to be determined by Monte Carlo
 1284 simulation and subtracted from the estimate obtained by applying the “simple” fake-rate
 1285 method to data, in order to get an unbiased estimate of the true background contributions.

1286 “CDF-type” weights

1287 Instead of subtracting from the estimate obtained for the sum of background contributions a
 1288 correction determined by Monte Carlo simulation, the signal contribution to the background
 1289 dominated event sample selected in data can be corrected for by adjusting the weights, based
 1290 solely on information contained in the analyzed data sample, hence avoiding the need to
 1291 rely on Monte Carlo based corrections.

1292 In the “CDF-type” method, additional information, namely whether or not tau-jet
 1293 candidates pass or fail the tau identification cuts and discriminators, is drawn from the data.
 1294 The desired cancellation of signal contributions is achieved by assigning negative weights to
 1295 those tau-jet candidates which pass all tau identification cuts and discriminators, i.e. to a
 1296 fair fraction of genuine hadronic tau decays, but to a small fraction of quark and gluon jets
 1297 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:

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

1316 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

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

1317 which matches exactly equation 6.6 for the “CDF-type” weights applied to tau-jet candi-
1318 dates given in section 6.2.3.

1319 Event weights and the weights assigned to “composite particles” are computed in the
1320 same way as for the “simple” weights, based on the weights assigned to the tau-jet candi-
1321 dates:

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

1322 where the sums extend over all jets in the background dominated control sample which pass
1323 the preselection defined by the denominator of equation 6.2.

1324 The effect of the negative weights to compensate the positive weights in case the “CDF-
1325 type” fake-rate method is applied to signal events containing genuine hadronic tau decays
1326 is shown in table ?? . As expected, positive and negative weights do indeed cancel in the
1327 statistical average.

1328 Figures ??, ?? and ?? demonstrate that an unbiased estimate of the background con-
1329 tribution by the “CDF-type” weights is maintained. Overall, the estimates obtained are
1330 in good agreement with the contributions expected for different background processes, in-
1331 dicating that the adjustment of negative and positive weights works as expected for the
1332 background as well.

1333 Results obtained by the “CDF-type” fake-rate method are summarized in table ??,

1334 in which the total number of background events estimated by equation 6.11 is compared
 1335 to the true background contributions. The “best” estimate of the background contribution
 1336 obtained from the “CDF-type” method is again taken as the arithmetic average of the
 1337 estimates obtained by applying the fake-rate probabilities for the highest and second highest
 1338 p_T jet in QCD di-jet events, jets in a muon enriched QCD sample and jets in W +jets events.
 1339 Systematic uncertainties are taken from the difference between the computed average value
 1340 and the minimum/maximum estimate. We obtain a value of $\mathcal{O}(15\text{--}20\%)$ for the systematic
 1341 uncertainty of the “CDF-type” method, slightly higher than the systematic uncertainty
 1342 obtained for the “simple” method. The small increase of systematic uncertainties is in
 1343 agreement with our expectation for fluctuations of the jet-weights in case weights of negative
 1344 and positive sign are used.

1345 §6.2.4 k-Nearest Neighbor fake-rate calculation

1346 §6.2.5 Results of Background Estimation

1347 An independent estimate of the background contributions is obtained by applying the fake-
 1348 rate method described in [27]. Fake-rates in QCD multi-jet events (light quark enriched
 1349 sample), QCD events containing muons (heavy quark and gluon enriched sample) and W
 1350 + jet events are measured in data [?, ?] and applied to events which pass all the event
 1351 selection criteria listed in table 5.1 except for the requirement for tau-jet candidates to pass
 1352 the “medium” tight TaNC discriminator and have unit charge.

1353 No assumption is made on the composition of $Z \rightarrow \mu^+\mu^-$, W + jets, $t\bar{t}$ + jets and
 1354 QCD backgrounds contributing to the event sample selected by the analysis. Differences
 1355 between fake-rates obtained for QCD multi-jet, QCD muon enriched and W + jets back-
 1356 ground events are attributed as systematic uncertainties of the fake-rate method. Per jet
 1357 and per event weights have been computed by the “simple” and “CDF-type” weights as de-
 1358 scribed in [27] and the results are found to be compatible within statistical and systematic
 1359 uncertainties. In the following, we present results for “CDF-type” weights. The “CDF-
 1360 type” weights have the advantage that the background estimate obtained does not change,
 1361 whether there is MSSM Higgs $\rightarrow \tau^+\tau^-$ signal present in the data or not.

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.

1362 Tau identification efficiencies need to be known when using “CDF-type” weights. Ded-
1363 icated studies have checked the tau identification efficiencies in data [?, ?]. Statistical
1364 and systematic uncertainties of these studies are still sizeable at present, in the order to
1365 20 – 30%. No indication has been found, however, that the Monte Carlo simulation would
1366 not correctly model hadronic tau decays in data. For the purpose of computing fake-rate
1367 weights via the “CDF-type” method, tau identification efficiencies are taken from the Monte
1368 Carlo simulation of hadronic tau decays in $Z \rightarrow \tau^+ \tau^-$ events. Systematic uncertainties on
1369 the background estimate obtained by the fake-rate method are determined by varying the
1370 tau identification efficiencies by $\pm 30\%$ relative to the value obtained from the Monte Carlo
1371 simulation.

1372 The results of applying the fake-rate method to the mu + tau channel are summarized
1373 in table 6.1. The background prediction has been corrected for the expected contribution
1374 of $XX.X$ events from $Z \rightarrow \mu^+ \mu^-$ background events in which the reconstructed tau-jet is
1375 due to a misidentified muon. The obtained estimate is in good agreement with the Monte
1376 Carlo expectation.

1377 As an additional cross-check of the method, a sample of events containing a muon plus
1378 a tau-jet of like-sign charge is selected in data and compared to the background prediction
1379 obtained by applying the fake-rate method to the like-sign sample. The like-sign sample is

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.

1380 expected to be dominated by the contributions of W + jets and QCD background processes
1381 and allows to verify the fake-rate method in a practically signal free event sample. The
1382 background estimate obtained by the fake-rate method is compared to the number of events
1383 observed in the like-sign data sample in table 6.2. The number of events expected in the
1384 like-sign control sample from Monte Carlo simulation is indicated in the caption. All
1385 numbers are in good agreement.

1386 The fake-rate method does not only allow to estimate the total number of background
1387 events, but allows to model the distributions of background processes as well. The capability
1388 to model distributions is illustrated in figure 6.2, which shows good agreement between the
1389 distributions observed in the like-sign data sample and the predictions obtained by the
1390 fake-rate method for the distributions of muon plus tau-jet visible mass and of the “full”
1391 invariant mass reconstructed by the SVfit algorithm.

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.

1392

Chapter 7

1393

Systematics

1394

Chapter 8

1395

Results

1396

Chapter 9

1397

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] LHC Higgs Cross Section Working Group Collaboration, “Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables”, *arXiv:1101.0593*.
- [19] M. Carena, S. Heinemeyer, C. Wagner et al., “MSSM Higgs boson searches at the Tevatron and the LHC: Impact of different benchmark scenarios”, *The European Physical Journal C - Particles and Fields* **45** (2006) 797–814. 10.1140/epjc/s2005-02470-y.
- [20] CMS Collaboration, “The CMS experiment at the CERN LHC”, *JINST* **3** (2008) S08004.
- [21] CMS Collaboration, “Measurement of Momentum Scale and Resolution using Low-mass Resonances and Cosmic-Ray Muons”,.
- [22] CMS Collaboration, “Tracking and Primary Vertex Results in First 7 TeV Collisions”,.
- [23] CMS Collaboration, “CMS Strategies for tau reconstruction and identification using particle-flow techniques”, *CMS PAS CMS-PAS-PFT-08-001* (2008).
- [24] CMS Collaboration, “Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET”, *CMS PAS CMS-PAS-PFT-09-001* (2009).
- [25] 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.
- [26] A. Kolmogorov, “On the representation of continuous functions of several variables by superposition of continuous functions of one variable and addition”, *Doklady Akademiia Nauk SSSR* **114** (1957).
- [27] J. Conway, E. Friis, and C. Veelken, “Estimation of Background contributions to Tau analyses via Fake-Rate technique”, *CMS Note* **2010/074** (2010).