

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 This thesis describes a search for the Higgs boson, a new particle predicted by a theory called
⁵ the Minimal Supersymmetric Model (MSSM). The standard model of particle physics, the
⁶ MSSM, and Higgs phenomenology are introduced briefly. The search presented in this thesis
⁷ uses a single final state configuration, in which the Higgs boson decays to two tau leptons,
⁸ where one tau decays to a muon and neutrinos, and the other decays to pions and a single
⁹ neutrino. Two new methods are introduced in this analysis, the Tau Neural Classifier tau
¹⁰ identification algorithm, and the Secondary Vertex fit tau pair mass reconstruction method.
¹¹ Both methods are discussed in detail. The analysis uses the 2010 dataset from the Compact
¹² Muon Solenoid (CMS) experiment, which contains 36 pb^{-1} of integrated luminosity at a
¹³ center of mass energy of 7 TeV. In total, 573 events are selected in the analysis. We fit the
¹⁴ observed tau pair mass spectrum and measure the composition of the events. The result
¹⁵ is compatible with the standard model expectation. No excess of signal events is observed,
¹⁶ and we set an upper limit on cross section times branching ratio of a Higgs boson. This
¹⁷ limit is interpreted in the parameter space of the MSSM.

¹⁸ **Acknowledgments**

¹⁹

Hooray for everybody.

20 Table of Contents

21	Introduction	1
22	1 The Standard Model and Beyond	3
23	1.1 The Standard Model	3
24	1.1.1 Quantum Electrodynamics and Gauge Invariance	4
25	1.1.2 The Weak Interactions	6
26	1.1.3 Spontaneous Symmetry Breaking	8
27	1.1.4 The Higgs Mechanism	10
28	1.1.5 Electroweak Unification	12
29	1.1.6 Quantum Chromodynamics	16
30	1.2 Beyond the Standard Model	18
31	1.2.1 The Hierarchy Problem	19
32	1.2.2 Supersymmetry	20
33	1.2.3 The Minimal Supersymmetric Model	20
34	1.3 Searches for the Higgs Boson	23
35	1.3.1 Standard Model Higgs Boson Phenomenology	23
36	1.3.2 MSSM Higgs Phenomenology	26
37	1.3.3 Results from LEP and Tevatron	27
38	1.4 The Physics of the Tau Lepton	33
39	2 The Compact Muon Solenoid Experiment	37
40	2.1 The Large Hadron Collider	38
41	2.2 Solenoid Magnet	40
42	2.3 Charged Particle Tracking Systems	41
43	2.4 Electromagnetic Calorimeter	43
44	2.5 Hadronic Calorimeter	45
45	2.6 Muon System	47
46	2.7 Trigger System	49
47	3 Tau Identification: The Tau Neural Classifier	52
48	3.1 Geometric Tau Identification Algorithms	53
49	3.2 Decay Mode Tau Identification: Motivation	53
50	3.3 The Tau Neural Classifier	54
51	3.3.1 Decay Mode Reconstruction	55
52	3.3.2 Neural Network Classification	58
53	3.4 Summary	72
54	3.5 HPS+TaNC: A Hybrid Algorithm	76
55	3.5.1 Decay Mode Reconstruction	76
56	3.5.2 Hadronic Tau Discrimination	78

57	3.6 Electron and Muon Rejection	79
58	4 Mass Reconstruction: The Secondary Vertex Fit	81
59	4.1 Existing Mass Reconstruction Algorithms	81
60	4.2 The Secondary Vertex Fit	83
61	4.3 Parametrization of Tau Decays	84
62	4.4 Likelihood for Tau Decays	85
63	4.4.1 Likelihood for Reconstructed E_T^{miss}	85
64	4.4.2 Likelihood for Tau Transverse Momentum Balance	87
65	4.4.3 Secondary Vertex Information	88
66	4.5 Performance	90
67	5 Analysis Selections	94
68	5.1 High Level Trigger	94
69	5.2 Particle Identification	95
70	5.2.1 Muons	95
71	5.2.2 Hadronic Taus	96
72	5.2.3 Missing Transverse Energy	96
73	5.3 Event Selections	97
74	6 Data–Driven Background Estimation	102
75	6.1 Background Enriched Control Regions	103
76	6.2 The Fake–rate Method	104
77	6.2.1 Parameterization of Fake–rates	107
78	6.2.2 Measurement of Fake–rates	108
79	6.2.3 Application of Fake–rates	108
80	6.2.4 “Simple” Weight Method	109
81	6.2.5 “CDF–type” Weights	111
82	6.2.6 k–Nearest Neighbor Fake–rate Calculation	113
83	6.2.7 Results of Fake–Rate Background Estimation	114
84	6.3 Template Method	116
85	7 Monte Carlo Corrections	129
86	7.1 Muon Identification Efficiency	129
87	7.2 Hadronic Tau Identification Efficiency	132
88	7.3 Muon and Tau Momentum Scale	134
89	7.4 Missing Transverse Energy Correction	135
90	7.5 Pile-up Event Weighting	136
91	8 Systematics and Limit Extraction	138
92	8.1 Signal Normalization Uncertainties	140
93	8.2 Background Normalization Uncertainties	140
94	8.3 Shape Uncertainties	141
95	8.4 Theory Uncertainties	142
96	8.5 Limit Extraction Method	142
97	9 Results	149
98	9.1 Selected Events	149
99	9.2 Limits on Higgs Production	150

100	9.3 Interpretation in the MSSM	151
101	Conclusions	160
102	Bibliography	160

103 List of Figures

104	1.1	Fermi contact interaction diagram	7
105	1.2	Muon decaying through intermediate gauge boson	7
106	1.3	QCD Feynman Diagrams	17
107	1.4	Loop corrections to Higgs mass	19
108	1.5	Higgstrahlung production diagram at e^+e^- colliders	24
109	1.6	Gluon fusion Higgs production diagram	24
110	1.7	Vector boson fusion Higgs production diagram	24
111	1.8	Parton luminosity comparison of the LHC and Tevatron	25
112	1.9	SM Higgs cross sections at the LHC	26
113	1.10	SM Higgs branching fractions	27
114	1.11	Cross sections of interest at hadron colliders	28
115	1.12	MSSM Higgs production with association b -quarks	29
116	1.13	MSSM Higgs cross sections at the LHC	29
117	1.14	LEP SM Higgs limit plot	31
118	1.15	Tevatron low mass standard model Higgs limit plot	32
119	1.16	Tevatron high mass standard model Higgs limit plot	33
120	1.17	LEP MSSM exclusion limits	34
121	1.18	Tevatron MSSM exclusion limits	34
122	2.1	Schematic drawings of the CMS detector	39
123	2.2	Material budget of the CMS tracker	42
124	2.3	Momentum and impact parameter resolutions of CMS tracker	43
125	2.4	Energy resolution of the CMS ECAL	45
126	2.5	Muon system material budget and identification efficiency	47
127	3.1	Visible invariant mass of τ lepton decay products	54
128	3.2	Invariant mass photon pairs in reconstructed π^0 mesons	56
129	3.3	Neutral energy fraction in visible τ decays	57
130	3.4	Tau decay mode reconstruction performance	59
131	3.5	Kinematic dependence of decay mode reconstruction	60
132	3.6	Neural network over-training validation plots	63
133	3.7	Kinematic weighting of training sample	65
134	3.8	Neural network output in each decay mode	68
135	3.9	Performance curves for the neural networks used in the TaNC	69
136	3.10	Tau Neural Classifier performance curves for different p_T ranges	70
137	3.11	Tau Neural Classifier transformation performance	72
138	3.12	Transformed neural network output	73
139	3.13	Tau Neural Classifier performance comparison	74
140	3.14	Tau Neural Classifier kinematic performance	75

141	3.15 Invariant mass distribution of PF photon pairs	77
142	4.1 Coordinate system of the SVfit	86
143	4.2 Effect of p_T -balance term on SVfit performance	89
144	4.3 Effect of the visible p_T requirements on muon and hadronic τ decays	90
145	4.4 Effect of the visible p_T requirements for Z and Higgs events	91
146	4.5 Comparison of SVfit with the Collinear Approximation algorithm	92
147	4.6 Comparison of SVfit with the visible mass observable	93
148	5.1 Distributions of M_T and muon isolation discriminants	99
149	5.2 Reconstruction and distribution of P_ζ discriminant	100
150	6.1 Visible mass distribution of the backgrounds in the signal and control regions	105
151	6.2 SVfit mass distribution of the backgrounds in the signal and control regions	106
152	6.3 p_T and η dependency of tau ID performance	119
153	6.4 Comparison of fake-rate contribution from genuine taus in the simple and CDF methods	121
154	6.5 Muon transverse momentum in the Fake-rate method	122
155	6.6 Tau-jet transverse momentum in the Fake-rate method	123
156	6.7 Visible mass in the Fake-rate method	124
157	6.8 k -Nearest Neighbor classifier example	125
158	6.9 Comparison of visible mass and SVfit mass	126
159	6.10 Comparison of background shapes in the signal and control regions	127
160	6.11 Visible mass distribution in the final fit of the Template Method	128
161	7.1 Tag-probe muon isolation method	131
162	7.2 Muon isolation correction factors	133
163	7.3 Measurement of hadronic tau identification efficiency	134
164	7.4 Distribution of number of reconstructed primary vertices per event	137
165	9.1 Transverse momentum distributions of muon and tau in the final selected events	153
166	9.2 Distributions of the η and ϕ of the muon and tau candidates in the final selected events	154
167	9.3 Distributions of final selected events	155
168	9.4 Observed and expected limits on Higgs $\sigma \times \text{BR}$	158
169	9.5 Excluded regions of MSSM $\tan \beta - m_{A^0}$ parameter space	159
170	9.6 CMS combined exclusion of MSSM $\tan \beta - m_{A^0}$ parameter space	161

174 List of Tables

175	1.1	Chiral supermultiplets in the MSSM	21
176	1.2	Gauge supermultiplets in the MSSM	21
177	1.3	Higgs search channels at LEP	30
178	1.4	Decay modes of the τ lepton	35
179	3.1	Decay mode performance – naive reconstruction	58
180	3.2	Decay mode performance – TaNC reconstruction	58
181	3.3	Neural network training event statistics	62
182	3.4	Variables used in the different TaNC neural networks	80
183	5.1	High Level Trigger paths used to select $\mu + \tau_h$ events	95
184	5.2	Analysis backgrounds that include fake taus	97
185	5.3	Event selection summary	101
186	6.1	Criteria used to select background enriched control regions	117
187	6.2	Comparison of background control region yields in data and the prediction from simulation	118
188	6.3	Fake-rate “simple” method closure test results	120
189	6.4	Fake-rate “CDF” method closure test results	121
190	6.5	Fake-rate method results	125
191	6.6	Fake-rate Method predicted yields in like-sign control region	126
192	6.7	Background yields measured using the Template Method	126
193	7.1	Muon trigger, identification, and isolation correction factors	132
194	8.1	Effect of normalization uncertainties on signal efficiency times acceptance .	143
195	8.2	Correspondence of confidence levels and $2\Delta \ln \mathcal{L}$ intervals	145
196	8.3	Summary of systematic uncertainties	148
197	9.1	Final analysis yield and background expectations	150
198	9.2	Expected signal yields at $\tan \beta = 30$	156
199	9.3	Expected and observed 95% CL $\sigma \times \text{BR}$ upper limits	157
200	9.4	Contributions of different MSSM Higgs boson types at different m_{A^0}	157

202 Introduction

203 This thesis describes a search for the Higgs boson, a new particle predicted by the standard
 204 model of particle physics. The search is optimized for a particular extension of the standard
 205 model, a theory called the Minimal Supersymmetric Model (MSSM). The analysis uses
 206 the 2010 dataset from the Compact Muon Solenoid (CMS) experiment, which contains
 207 36 pb^{-1} of integrated luminosity at a center of mass energy of 7 TeV. The Higgs boson is
 208 hypothesized to be the catalyst of electroweak symmetry breaking, the phenomenon strongly
 209 believed to impart mass to particles that form our natural world.

210 Chapter 1 begins with an introduction to the standard model (SM) of particle physics.
 211 Emphasis is given to electroweak symmetry breaking and the Higgs mechanism, the theo-
 212 retical phenomena that motivate the presence of a Higgs boson. The theoretical issues which
 213 motivated the development of the MSSM are discussed, and a brief introduction is given.
 214 Finally, the phenomenology of Higgs bosons in the SM and MSSM is discussed, with an
 215 overview of Higgs searches performed at LEP and the Tevatron.

216 This thesis then documents the development of a complete search for MSSM Higgs
 217 bosons at CMS. The CMS experiment is introduced briefly in Chapter 2. Chapters 3 and 4
 218 document in detail two fundamental components of the search, an advanced tau identifi-
 219 cation algorithm, and a novel method for reconstructing the neutrinos associated to tau
 220 decays. The development of these algorithms was motivated by the challenges of this anal-
 221 ysis, and precipitated significant improvements in the final result. Finally, in Chapters 5-8,
 222 we describe the methods and results of the event selection, background estimation, and the
 223 systematic uncertainties, and finally compute an upper limit on the presence of an MSSM
 224 Higgs boson.

225 The studies presented herein were part of a larger effort at CMS to search for an MSSM
 226 Higgs boson decaying to tau lepton pairs. In addition to the $\mu - \tau_h$ channel described in

227 this thesis, final states with an electronic and hadronic tau decay ($e - \tau$) and electronic and
228 muonic ($e - \mu$) were considered. The combination of all three was used to set a limit on
229 the MSSM [1]. This result has recently been accepted for publication in Physical Review
230 Letters B. At the time of this writing, the CMS analysis sets the world's strongest limit on
231 the MSSM Higgs boson using a direct search.

232

Chapter 1

233

The Standard Model and Beyond

234 §1.1 The Standard Model

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

245 The feature that united the disparate theories into the SM was the application of the
246 principle of *local gauge invariance*. The principle of gauge invariance first found success in
247 QED, which predicted electromagnetic phenomenon with astounding accuracy. Local gauge
248 invariance is now believed to a fundamental feature of nature that underpins all theories of
249 elementary particles. Furthermore, the development of the complete SM as it is known today
250 was precipitated by Goldstones’s work on spontaneous symmetry breaking [2, 3], which
251 produces an effective Lagrangian with additional massless “Goldstone” bosons. Higgs (and
252 others) [4, 5, 6] developed these ideas into what is ultimately called the “Higgs Mechanism,”
253 which uses a combination of new fields with broken symmetry to give mass to the Goldstone
254 bosons.

255 In the 1960s, Glashow [7], Weinberg [8], and Salam [9] developed the above ideas into
 256 the electroweak model, which unified QED with the weak force using intermediate weak
 257 bosons in a gauge theory with symmetry that is spontaneously broken using the Higgs
 258 mechanism. This unified theory has been incredibly experimentally successful and is the
 259 foundation of modern particle theory.

260 §1.1.1 Quantum Electrodynamics and Gauge Invariance

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

268 In the following, we first describe the Dirac equation for a free electron, which is the
 269 relativistic extension of the Schroedinger equation for spin 1/2 particles. We then show that
 270 requiring the corresponding Lagrangian of the free charged particle to be invariant under
 271 local gauge transformations creates an effective gauge boson field. This “gauge field” creates
 272 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 [10].

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.

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

- ²⁷³ The solutions ψ of the Dirac equation are called “Dirac spinors,” and represent the quantum
²⁷⁴ 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

$$\partial_\mu \rightarrow D_\mu = \partial_\mu - \frac{ie}{\hbar}A^\mu$$

in Equation 1.5, where $A^\mu = \partial^\mu\theta(x)$ and e is the electric charge, the Lagrangian becomes locally gauge invariant:

$$\mathcal{L} = \bar{\psi}(i\hbar c\gamma^\mu D_\mu - mc^2)\psi. \quad (1.7)$$

The difference between the locally (1.7) and the globally (1.5) gauge invariant Lagrangian is then

$$\mathcal{L}_{int} = \frac{e}{\hbar}\bar{\psi}\gamma^\mu\psi A_\mu.$$

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:

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

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

§1.1.2 The Weak Interactions

The theory of Weak Interactions was created to describe the physics of radioactive decay. The first formulation of the theory was done by Fermi [11] 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 [12]

$$H = \frac{G_\beta}{\sqrt{2}} [\bar{\psi}_p \gamma_\mu (1 - g_A \gamma_5) \psi_n] [[\bar{\psi}_e \gamma^\mu (1 - \gamma_5) \psi_\nu] + h.c., \quad (1.8)$$

where G_β is the Fermi constant and g_A is the relative fraction of the interaction with axially Lorentz structure. The value of g_A was determined experimentally to be 1.26. One of the most notable things discovered about the weak force is that weak interactions violate parity; that is, the physics of the interaction change (or become disallowed) under inversion of the spatial coordinates. This is evidenced by the $(1 - \gamma_5)$ term in Equation 1.8. This term is the “helicity operator”; the left and right “handed” helicity states are eigenstates states of this term.

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

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

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

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

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

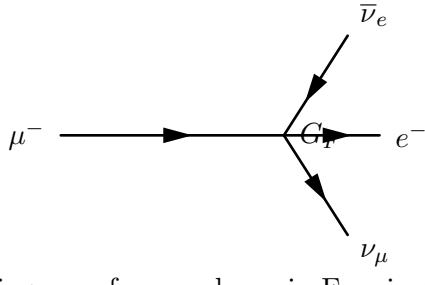


Figure 1.1: Feynmann diagram of muon decay in Fermi contact interaction theory.

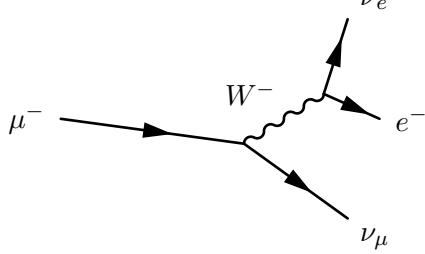


Figure 1.2: Feynmann diagram of muon decay proceeding through an intermediate gauge boson W^- .

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 1.8 for muon decay, the decay amplitude M is found to be

$$M = \frac{G_F}{\sqrt{2}} \left[\bar{u}_{\nu_\mu} \gamma_\rho \frac{1 - \gamma_5}{2} u_\mu \right] \left[\bar{u}_{\nu_e} \gamma_\rho \frac{1 - \gamma_5}{2} u_e \right]. \quad (1.9)$$

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

The first attempt to solve the problems with the Fermi theory was made by introducing an intermediate weak boson [7]. 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 [12] by

$$M = - \left[\frac{g}{\sqrt{2}} \bar{u}_{\nu_\mu} \gamma_\rho \frac{1 - \gamma_5}{2} u_\mu \right] \frac{-g^{\rho\sigma} + \frac{q^\rho q^\sigma}{M_W^2}}{q^2 - M_W^2} \left[\frac{g}{\sqrt{2}} \bar{u}_{\nu_e} \gamma_\rho \frac{1 - \gamma_5}{2} u_e \right]. \quad (1.10)$$

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 [10].

term of Equation 1.10 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 1.10 is equivalent to the Fermi contact interaction in 1.9 as

$$\lim_{q/M_W \rightarrow 0} \frac{g^2}{8(q^2 - M_W^2)} = \frac{G_F}{\sqrt{2}}. \quad (1.11)$$

295 Unfortunately, the weak boson exchange model did not solve the problems of unitarity
 296 and renormalizability in the weak interaction. However, the form of the boson-exchange
 297 propagator in Equation 1.11 suggests the observed “weakness” of the weak interactions is
 298 an artifact of the presence of the massive propagator (M_W) and that the fundamental scale
 299 of the interaction g is the same order of magnitude as that of QED, $g \approx e$. This observation
 300 lead to the unification of the electromagnetic and weak forces, which we describe in the
 301 next sections.

302 §1.1.3 Spontaneous Symmetry Breaking

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

A symmetry of a Lagrangian is spontaneously broken when the ground state, or vacuum, is at a value 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 [12].

$$\mathcal{L} = \frac{1}{2} \partial_\mu \phi_1 \partial^\mu \phi_1 + \frac{1}{2} \partial_\mu \phi_2 \partial^\mu \phi_2 - V(\phi_1^2 + \phi_2^2) \quad (1.12)$$

The toy Lagrangian in Equation 1.12 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

$$\begin{aligned} V(\phi_1, \phi_2) &= V(\phi_1^{min}, \phi_2^{min}) + \sum_{a=1,2} \left(\frac{\partial V}{\partial \phi_a} \right)_0 (\phi_a - \phi_a^{min}) \\ &\quad + \frac{1}{2} \sum_{a,b=1,2} \left(\frac{\partial^2 V}{\partial \phi_a \partial \phi_b} \right)_0 (\phi_a - \phi_a^{min})(\phi_b - \phi_b^{min}) + \dots \end{aligned} \quad (1.13)$$

Since at the minimum the partial derivative of V is zero with respect to all fields, the second term in Equation 1.13 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

$$\left(\frac{\partial^2 V}{\partial \phi_a \partial \phi_b} \right)_{\phi^{min}} \quad (1.14)$$

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 1.12 is

$$V(\phi_1^2 \phi_2^2) = \frac{m^2}{2}(\phi_1^2 + \phi_2^2) + \frac{\lambda}{4}(\phi_1^2 + \phi_2^2)^2. \quad (1.15)$$

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 1.13 takes the form $\left(\frac{\partial^2 V}{\partial \phi_a \partial \phi_b} \right)_0 = \frac{m^2}{2} \delta_{ab}$, where δ_{ab} is the Kronecker delta function. Therefore the mass matrix is already diagonalized, and the ϕ_1 and ϕ_2 both correspond to particles with mass m . If the m^2

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 $\mathcal{O}(2)$. 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 $U(1)$.

parameter in Equation 1.15 is negative, the spectrum is dramatically different. After making the replacement $m^2 = -\mu^2 (\mu^2 > 0)$, the extrema of V are no longer unique. The requirement of $\frac{\partial V}{\partial \phi_i} = 0$ for all i is satisfied in two cases:

$$(\phi_1^{min}, \phi_2^{min}) = (0, 0) \quad (1.16)$$

$$(\phi_1^{min})^2 + (\phi_2^{min})^2 = \frac{\mu^2}{\lambda} = \nu^2. \quad (1.17)$$

If the vacuum state is defined at the point in Equation 1.16, 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 1.17, 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 1.14 yields the mass matrix

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

Breaking the symmetry has changed the mass spectrum of the physical particles in the model. There is now a massive particle with $m = v^2$ and a massless particle. This massless particle is called the “Goldstone boson.” Goldstone found [2] that a massless particle appears for each generator in the symmetry group that is broken.

§1.1.4 The Higgs Mechanism

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

⁴The point chosen for the VEV here is not arbitrary. One can chose any point that satisfies Equation 1.17 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.

338 kinetic term for the gauge field, as discussed in Section 1.1.1. The unbroken Lagrangian is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + (D_\mu\phi^*)(D^\mu\phi) - V(\phi^*\phi) \quad (1.18)$$

$$V(\phi^*\phi) = -\mu^2\phi^*\phi + \lambda(\phi^*\phi)^2, \quad (1.19)$$

339 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

340 under the local $U(1)$ gauge transformation

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

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 $\theta(x)/\nu$

$$\phi(x) = \frac{1}{\sqrt{2}}(\nu + \eta(x))e^{i\theta(x)/\nu}. \quad (1.20)$$

341 If the gauge transform is chosen to be $\alpha(x) = \theta(x)/\nu$, the fields of are defined in the so-called

342 “unitary gauge”⁵ and have the special forms

$$\begin{aligned} \phi(x) \rightarrow \phi'(x) &= \frac{1}{\sqrt{2}}(\nu + \eta(x)) \\ A_\mu(x) \rightarrow B_\mu(x) &= A_\mu(x) - \frac{1}{e\nu}\partial_\mu\theta(x) \end{aligned} \quad (1.21)$$

343 The kinetic term of the gauge field $F_{\mu\nu}$ is invariant under this transformation. If the gauge

344 transformations of Equation 1.21 are substituted into the Lagrangian (1.18) the effective

345 Lagrangian at the minimum of V is

$$\begin{aligned} \mathcal{L} &= \frac{1}{2}\partial_\mu\eta\partial^\mu\eta - \mu^2\eta^2 \\ &- \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}(e\nu)^2B_\mu B^\mu \\ &+ \frac{1}{2}e^2B_\mu B^\mu\eta(\eta + 2\nu) - \lambda\nu\eta^3 - \frac{\lambda}{4}\eta^4. \end{aligned} \quad (1.22)$$

346 The breaking of the original symmetry has dramatically altered the physical consequences of

347 the model. In its unbroken form, the model described by Equation 1.18 would produce two

348 real massive particles and one massless gauge boson mandated by local gauge invariance.

349 After symmetry breaking, the effective Lagrangian in Equation 1.22 contains a massive

350 scalar η with $m = \sqrt{2\mu^2}$ and a *massive* gauge boson B_μ with mass $m = \sqrt{2}e\nu$. By ac-

351 quiring a mass, the gauge boson B_μ has acquired the degree of freedom (as it can now

352 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.

353 scalar ϕ field. This phenomenon, known as the “Higgs Mechanism,” is a simplified version
 354 of the techniques successfully used to unify the electromagnetic and weak forces that we
 355 will discuss in the next section.

356 §1.1.5 Electroweak Unification

357 In the 1960s, the ideas of local gauge invariance in field theories, spontaneous symme-
 358 try breaking, and the Higgs mechanism were combined by Glashow [7], Weinberg [8] and
 359 Salam [9] to form the unified theory of electroweak interactions, the nucleus of the SM. This
 360 model successfully unified the electromagnetic and weak interactions into a unified theory
 361 with a larger symmetry group. The reason for the empirically observed difference in scales
 362 between two interactions is due to the larger, unified symmetry group being broken. This
 363 broken symmetry creates heavy gauge bosons via the Higgs mechanism, whose large mass
 364 decreases the strength of “weak” interactions at low energy, as discussed in Section 1.1.2.
 365 The model successfully predicted the existence and approximate masses of the weak force
 366 carriers, the W^\pm and Z bosons. These particles were later observed [13, 14, 15, 16] with the
 367 predicted masses at the UA1 and UA2 experiments.

368 To provide a simple introduction to the mechanisms of the model, we will start with a
 369 model that includes only one family of leptons, the electron e and its associated neutrino
 370 ν_e . Following once again the treatment of [12], we describe the representation of the e and
 371 ν_e in the chosen symmetry group of the model. We then construct a locally gauge invariant
 372 Lagrangian with spontaneously broken symmetry, and examine the particle content of the
 373 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 (1.9) 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 $SU(2)$.

$$L = \frac{1 - \gamma_5}{2} \begin{pmatrix} \nu_e \\ e^- \end{pmatrix} = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L \quad (1.23)$$

³⁷⁴ The operators that operate on “weak isospin,” the quantum of $SU(2)_L$, are

$$\tau^+ = \frac{\tau^1 + i\tau^2}{2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad (1.24)$$

$$\tau^- = \frac{\tau^1 - i\tau^2}{2} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad (1.25)$$

where the τ^i are the Pauli matrices. The weak currents J_μ^\pm can be written by combining Equations 1.23–1.25

$$J_\mu^\pm = \bar{L}\gamma_\mu\tau^\pm L. \quad (1.26)$$

³⁷⁵ Since τ^1 , τ^2 , and τ^3 are the generators of the $SU(2)$ group, we can complete the group

³⁷⁶ by adding a neutral current to the charged currents of Equation 1.26. The τ^3 generator is
³⁷⁷ diagonal, so the charge of the current is zero and no mixing of the fields occur:

$$\begin{aligned} J_\mu^3 &= \bar{L}\gamma_\mu\frac{\tau^3}{2}L \\ &= \bar{L}\gamma_\mu\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}L \\ &= \frac{1}{2}\bar{\nu}_e\gamma_\mu\nu_e - \frac{1}{2}\bar{e}_L\gamma_\mu e_L. \end{aligned} \quad (1.27)$$

³⁷⁸ Naively one might hope that the neutral current of Equation 1.27 would correspond to the
³⁷⁹ electromagnetic (photon) current of QED. However, this is impossible for two reasons. First,
³⁸⁰ the right-handed component e_R does not appear in the current, so this interaction violates
³⁸¹ parity, a known symmetry of the electromagnetic interactions. Second, the current couples to
³⁸² neutrinos, which have no electric charge. Therefore, the “charge” corresponding to the $SU(2)$
³⁸³ gauge symmetry generators $T^i = \int J_0^i(x)d^3x$ cannot be that of the QED, and the gauge
³⁸⁴ group must be enlarged to include an additional $U(1)$ symmetry. The generator of the new
³⁸⁵ symmetry must commute with the generators of the $SU(2)_L$ group. The symmetry cannot
³⁸⁶ be directly extended with $U(1)_{em}$ as the electromagnetic charge $Q = \int(e_L^\dagger e_L + e_R^\dagger e_R)d^3x$
³⁸⁷ does not commute with T^i . The solution is to introduce the “weak hypercharge” $\frac{Y}{2} = Q - T^3$,
³⁸⁸ which commutes the generators of $SU(2)_L$. Thus the symmetry group of the electroweak
³⁸⁹ model is $SU(2)_L \times U(1)_Y$.

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

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

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

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

$$\begin{aligned}\phi &= \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \\ \mathcal{L}_S &= (D_\mu\phi)^\dagger(D^\mu\phi) - V(\phi^\dagger\phi),\end{aligned}$$

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 1.19. At this point we
also add $SU(2)_L \times U(1)_Y$ invariant “Yukawa” terms

$$\mathcal{L}_Y = -G_e(\bar{L}\phi R + \bar{R}\phi^\dagger L) + h.c. \quad (1.28)$$

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

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

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

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 $U(1)_{em}$:

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

401 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 1.20 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

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

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

$$\mathcal{L}_{mass} = \frac{\nu^2}{8}(g^2 A'_\mu{}^{1\mu} A'^{1\mu} + g^2 A'_\mu{}^{2\mu} A'^{2\mu} + (g A'_\mu{}^{3\mu} - g' B'_\mu)^2). \quad (1.29)$$

The $A'_\mu{}^{1\mu}$ and $A'_\mu{}^{2\mu}$ fields can be combined such that the first two terms in Equation 1.29 are equivalent to the mass term of a charged boson

$$W_\mu^\pm = \frac{A'_\mu{}^{1\mu} \mp i A'_\mu{}^{2\mu}}{2}.$$

402 This is the familiar W^\pm boson of β and muon decay, and has mass $M_W = \frac{1}{2}g\nu$. The third 403 term in Equation 1.29 can be written in matrix form and then diagonalized into mass 404 eigenstates

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

giving a massive Z boson with

$$M_Z = \frac{\nu}{2} \sqrt{g^2 + g'^2} \quad (1.30)$$

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

by

$$M_Z \equiv \frac{M_W}{\cos \theta_W},$$

405 where θ_W is the “Weinberg angle,” which must be determined from experiment. As the
 406 Fermi contact interaction of Section 1.1.2 is an effective theory of the weak sector, the value
 407 of G_F obtained from β and muon decay experiments give clues to the masses of the W and
 408 Z .

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

$$M_Z \approx \frac{76 \text{ GeV}}{\sin 2\theta_W} > 76 \text{ GeV}.$$

409 The discovery of the W [13, 14] and Z [15, 16] at the CERN SPS was a huge triumph for
 410 the electroweak model.

411 The model that is presented in this section assumes only one species of leptons, the
 412 electron and its associated neutrino. The electroweak model is trivially extended [12] to
 413 include the other species (μ, τ) of leptons and the three families of quarks. The masses of
 414 the fermions are determined by the Yukawa terms in Equation 1.28. Each particle species
 415 has a Yukawa term relating the Higgs VEV to its mass that is not constrained by the theory,
 416 and must be determined by experiment.

417 §1.1.6 Quantum Chromodynamics

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

427 There are three marked differences between the photon of QED and the gluon of QCD.
 428 First, the gluon carries a color charge, while the photon is electrically neutral. This has the
 429 consequence that a gluon can couple to other gluons. Secondly, it is found that no colored

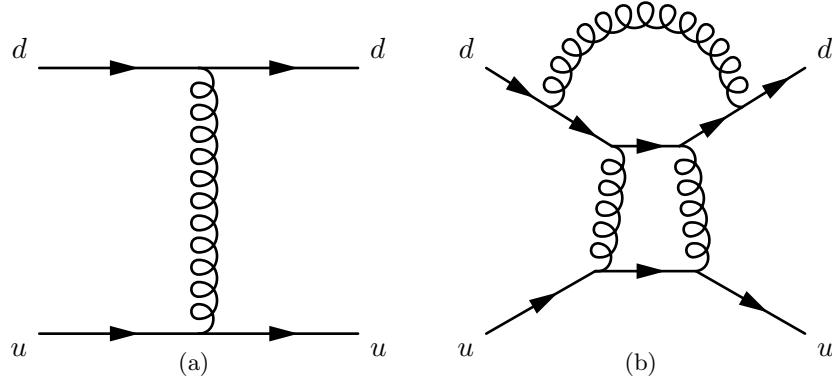


Figure 1.3: Feynman diagrams of a first–order (a) QCD interaction and a multi–loop (b) QCD interaction that have the same initial and final states. Each internal gluon propagator contributes a factor of g_s , the strong coupling constant, to the the amplitude. Since $g_s > 1$, multi–loop diagrams have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

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

The practical consequence of color confinement to a physicist studying electroweak phenomenon at a high–energy particle physics experiment is the production of quark and gluon “jets,” which are high multiplicity sprays of particles observed in the detector. In a proton–proton collision, quarks and gluons can be knocked off the incident protons. These quarks and gluons immediately “hadronize,” surrounding themselves with additional hadrons, the majority of which are charged and neutral pions. Heavier quarks, such as the charm, beauty,

450 and top quarks undergo a flavor-changing weak decays, which can give rise to structure
 451 (leptons, sub-jets) within the jet. Furthermore, due to the relative strength of the strong
 452 interaction compared that of the electroweak, collision events involving only strong inter-
 453 actions are produced at rates many orders of magnitudes larger than that of electroweak
 454 interactions. This makes life difficult for physicists studying the electroweak force at hadron
 455 colliders. Sections 2.7, and Chapters 3 and 5 will discuss the techniques used to identify and
 456 remove QCD events from the data at different stages of the analysis.

457 §1.2 Beyond the Standard Model

458 The standard model is one of the most successful theories of the natural world ever created.
 459 The predictions of the SM have been tested to many orders of magnitude and no experiment
 460 to date⁶ has found a result statistically incompatible with the SM. However, there is a
 461 general consensus in the physics community that the SM is not complete. It is believed that
 462 it is only an effective theory that is valid below some energy scale Λ . Above this energy,
 463 there must exist some other “new physics,” which unifies the forces of the SM and correctly
 464 describes the natural world at all scales, while maintaining equivalence to the SM at low
 465 energy. This concept is analogous to the relationship between the effective Fermi contact
 466 theory of Section 1.1.2 and the unified electroweak theory of Section 1.1.5. The size of
 467 the cutoff scale Λ is estimated [12] to be $\mathcal{O}(10^{15})$ GeV for a unified theory with $SU(5)$
 468 symmetry and even larger, $\mathcal{O}(10^{19})$ GeV = M_{Planck} if the theory is unified with gravity.

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

⁶The SM 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.

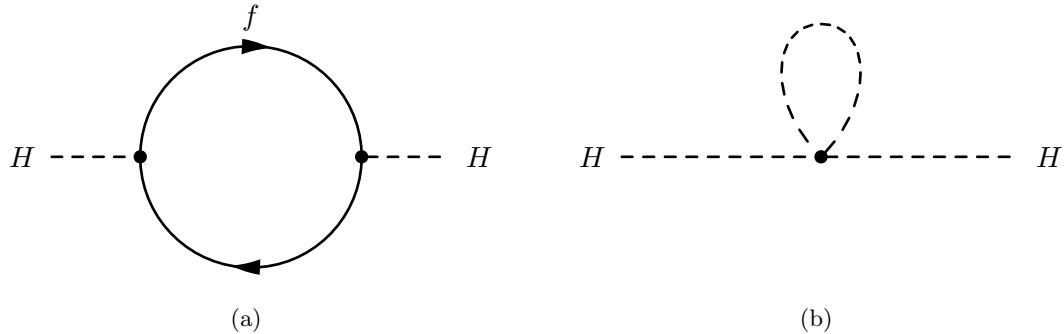


Figure 1.4: Feynman diagram of fermion (a) and scalar (b) loop corrections to Higgs mass.

⁴⁷⁷ An short overview of the hierarchy problem and Supersymmetry are presented in the next
⁴⁷⁸ sections.

479 §1.2.1 The Hierarchy Problem

The enormous size of the cutoff scale Λ in the SM causes a major theoretical problem in the SM. During renormalization of the SM, 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 [17]. 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 [17] of the loop correction in Figure 1.4(a) to the Higgs mass is

$$m_H^2 = -\frac{|\lambda_f|^2}{8\pi^2} \Lambda^2 + \dots \quad (1.31)$$

The correction scales with Λ^2 , which is many orders of magnitude larger than the electroweak (M_W) scale. The physical mass of the Higgs is expected to have the same scale as M_W , $\mathcal{O}(100 \text{ GeV}/c^2)$. The fact that each fermion contributes a loop correction (Equation 1.31) requires that the “bare mass” of the Higgs to be tuned to the precision of $(M_W/\Lambda)^2 \approx 10^{-26}$ for the renormalized mass to be correct! This is the so-called fine-tuning problem: it is believed that in a natural theory there will be only one scale. The electroweak unification analogy is in Equation 1.11, where it was noticed that the difference between the QED and weak scale was due to the massive M_W propagator term, and that the fundamental scale g of the intermediate weak boson theory was compatible with QED. The most promising solution to the hierarchy problem is the introduction of a new, “super” symmetry.

490 §1.2.2 Supersymmetry

491 Supersymmetry extends the SM by positing that there exists a symmetry between the
 492 integer-spin bosons (γ, W^\pm, Z, H) and the half integer-spin fermions (quarks and leptons).
 493 In Supersymmetry, every particle in the SM has a “superpartner” with a spin differs by 1/2.
 494 All of the other quantum numbers (including mass) of the superpartners are the same. The
 495 introduction of this symmetry immediately solves the hierarchy problem. For every scalar
 496 loop correction (Figure 1.4(b)) to the Higgs mass there is now a corresponding fermion loop
 497 correction (Figure 1.4(a)). As the fermion and the scalar have the same quantum numbers
 498 (except for spin) it turns out that these two diagrams have the same value, but *opposite*
 499 sign. Thus the large Λ^2 superpartner loop corrections to the Higgs mass exactly cancel out
 500 the problematic SM corrections. It is clear that if Supersymmetry exists, it must be broken.
 501 We have not observed a scalar charged particle with the same mass as the electron, for
 502 example. An excellent overview of possible mechanisms that create spontaneous symmetry
 503 breaking in Supersymmetric models is given in Chapter 6 of [17].

504 §1.2.3 The Minimal Supersymmetric Model

505 The simplest possible Supersymmetric extension to the SM is the Minimal Supersymmetric
 506 Model (MSSM). The model groups superpartner pairs into chiral (a left or right-handed
 507 fermion field plus a complex scalar field) and gauge (a spin-1 vector boson and a left
 508 or right-handed *gaugino* fermion) “supermultiplets.” As the weak interactions of the SM
 509 fermions are chiral, they (and their superpartners) must belong in a chiral supermultiplet.
 510 It is interesting to note that there is a different superpartner for the left and right-handed
 511 components of the fermions, even though the superpartners are spin-0 and cannot have any
 512 handedness. It is found that there must be two Higgs supermultiplets for the MSSM to be
 513 viable. As there are now fermionic particles in the Higgs sector (the Higgsinos), if only one
 514 supermultiplet is introduced the MSSM suffers from non-renormalizable gauge anomalies.⁷
 515 By introducing an additional Higgs supermultiplet with opposite hypercharge, the anomaly

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.

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	$(\overline{\mathbf{3}}, \mathbf{1}, -\frac{2}{3})$
	\bar{d}	\tilde{d}_R^*	d_R^\dagger	$(\overline{\mathbf{3}}, \mathbf{1}, \frac{1}{3})$
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: [17]

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	$\widetilde{W}^\pm \ \widetilde{W}^0$	$W^\pm \ W^0$	$(\mathbf{1}, \mathbf{3}, 0)$
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: [17]

516 is canceled. The scalar portion of the MSSM Higgs sector then contains two complex doublet
 517 fields $H_u = (H_u^+, H_u^0)$ (up-type) and $H_d = (H_d^0, H_d^-)$ (down-type). The complete chiral and
 518 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 [17]

$$W_{\text{MSSM}} = \bar{u}\mathbf{y}_{\mathbf{u}}QH_u - \bar{d}\mathbf{y}_{\mathbf{d}}QH_d - \bar{e}\mathbf{y}_{\mathbf{e}}LH_d + \mu H_u H_d ,$$

519 where H_u , H_d , Q , L , \bar{u} , \bar{d} , and \bar{e} are the superfields defined in Table 1.1. The \mathbf{y} terms are
 520 Yukawa 3×3 matrices which act on the different families. It is important to note that the
 521 up-type quarks couple to the up-type Higgs H_u , while the down-type quarks and leptons
 522 couple to the down-type Higgs. This feature has large phenomenological consequences,
 523 which are discussed in Section 1.3.2. The scalar portion of the W_{MSSM} potential defines the
 524 spontaneous symmetry breaking. Similar to the scalar potential V symmetry breaking of

525 Section 1.1.3, the potential of V at the minimum is found⁸ to be

$$\begin{aligned} V = & (|\mu|^2 + m_{H_u}^2)|H_u^0|^2 + (|\mu|^2 + m_{H_d}^2)|H_d^0|^2 \\ & - (bH_u^0 H_d^0 + c.c.) + \frac{1}{8}(g^2 + g'^2)(|H_u^0|^2 - |H_d^0|^2)^2. \end{aligned} \quad (1.32)$$

Under suitable choices⁹ of the parameters in Equation 1.32, 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 1.30) 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

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

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

$$\begin{aligned} m_{h^0}^2 &= \frac{1}{2}(m_{A^0}^2 + m_Z^2 - \sqrt{(m_{A^0}^2 - m_Z^2)^2 + 4m_Z^2 m_{A^0}^2 \sin^2(2\beta)}) \\ m_{H^0}^2 &= \frac{1}{2}(m_{A^0}^2 + m_Z^2 + \sqrt{(m_{A^0}^2 - m_Z^2)^2 + 4m_Z^2 m_{A^0}^2 \sin^2(2\beta)}). \end{aligned} \quad (1.33)$$

It can be seen that the tree level mass m_{h^0} of Equation 1.33 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 the LEP experiment (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 family are

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

533 The Yukawa couplings are free parameters determined by experimentally observed masses.
 534 This means that when $\tan \beta$ is large ($\beta \rightarrow \pi$), the Yukawa terms y for the b quarks and τ

⁸ A 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 $U(1)_{em}$.

⁹ See Chapter 7 of [17] for a detailed overview.

535 leptons must be enhanced to maintain the observed masses. The effect of $\tan \beta$ on the Higgs
 536 mass spectrum and couplings in the MSSM will be discussed further in Section 1.3.2.

537 §1.3 Searches for the Higgs Boson

538 The discovery of the Higgs boson is one of the biggest prizes in science today. Dozens of
 539 experiments, thousands of scientists and billions of dollars (a human hierarchy problem)
 540 have been spent in efforts to discovery the Higgs. In this section we discuss how the Higgs
 541 and the MSSM could appear in modern colliders (with an emphasis on the LHC) and
 542 the current limits placed on the Higgs by the Large Electron–Positron Collider (LEP) and
 543 Tevatron experiments.

544 §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 1.28) in the Lagrangian. Taking the electron as an example, after symmetry breaking, the Yukawa term is found to be

$$\mathcal{L}_e = -\frac{G_e}{\sqrt{2}}(\nu + H(x))\bar{e}e = -\frac{G_e\nu}{\sqrt{2}}\bar{e}e - \frac{G_e}{\sqrt{2}}H(x)\bar{e}e. \quad (1.34)$$

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

$$\frac{G_e\nu}{\sqrt{2}} = \frac{m_e}{\nu}. \quad (1.35)$$

545 The left-hand side of Equation 1.35 is the same as the constant in the electron Higgs
 546 coupling term ($H(x)\bar{e}e$) in Equation 1.34. Therefore the coupling between the fermions
 547 and Higgs boson is proportional to their mass! This remarkable fact shapes the possible
 548 production modes and the branching fractions of Higgs decays.

549 The dominant modes of Higgs boson production depend on the type of experiment.
 550 In general, Higgs production is favored through high-mass intermediate states, due to the
 551 mass² proportional coupling. At the Tevatron and LEP experiments, which will be intro-
 552 duced in the next section, the dominant SM Higgs production mode is “Higgstrahlung,”
 553 where a virtual W^\pm or Z gauge boson is produced and then radiates a Higgs boson. Hig-
 554 gstrahlung is illustrated in Figure 1.3.1. At the Large Hadron Collider, higher gluon lumi-

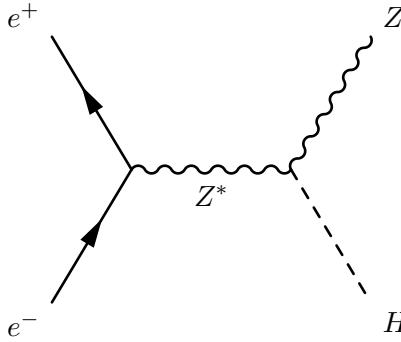


Figure 1.5: Higgstrahlung production diagram at e^+e^- colliders

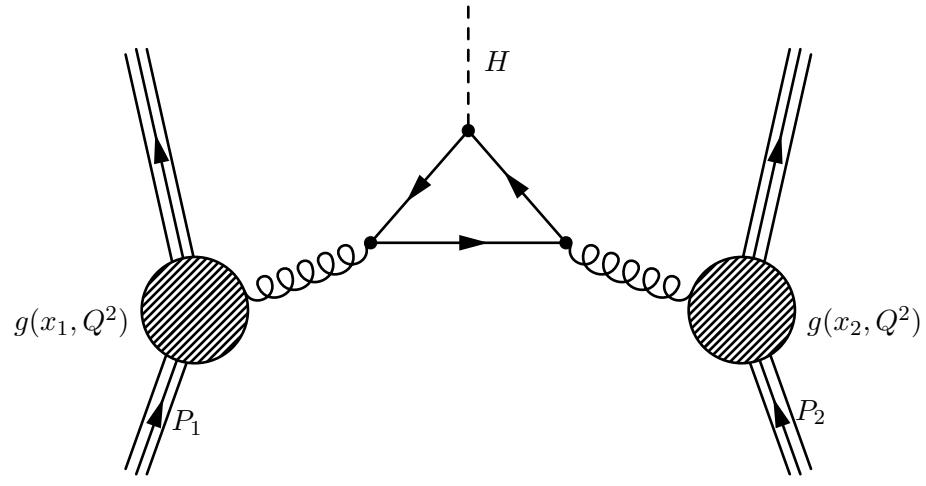


Figure 1.6: Gluon fusion Higgs production mechanism in a proton–proton collision. The Higgs mass coupling favors the heavy top quark in the central loop. Image credit: [18]

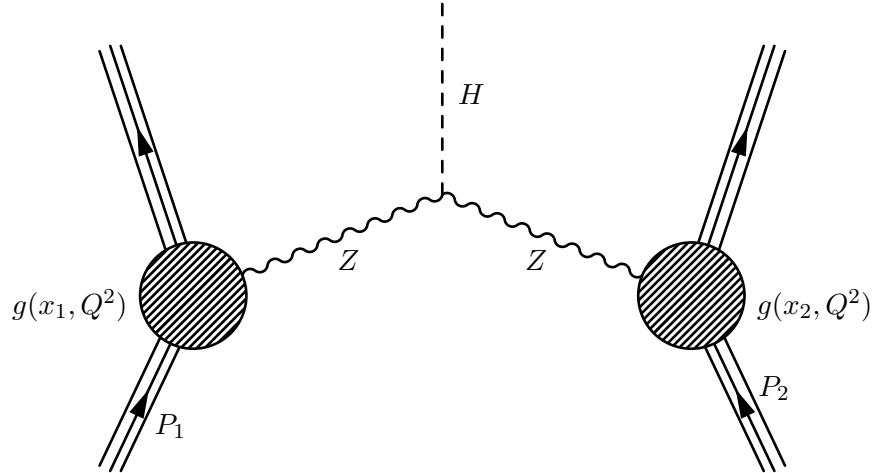


Figure 1.7: Vector boson fusion (VBF) Higgs production mechanism in proton–proton collisions. The VBF mechanism is notable for the lack of color–flow between the two incident protons, producing events with low jet activity in the central region.

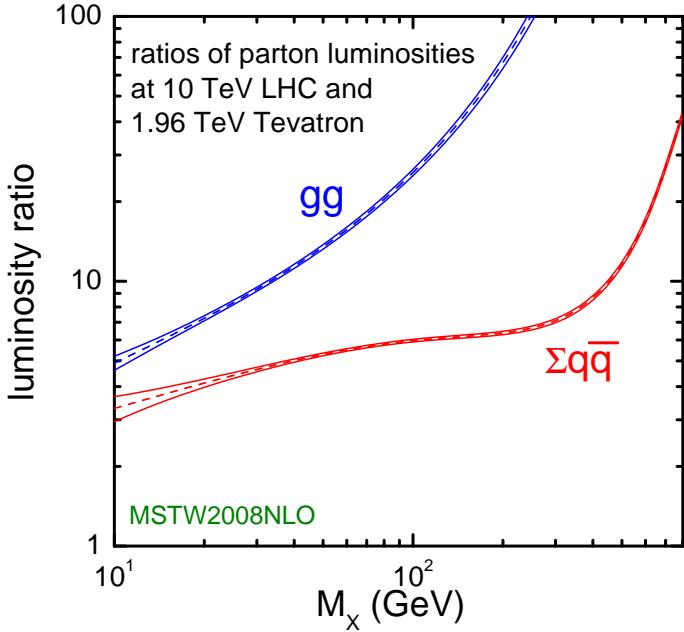


Figure 1.8: Ratio of the parton luminosity (the amount of luminosity contributed by the different species that compose the proton) of the LHC (at $\sqrt{s} = 10$ TeV) and the Tevatron. The large increase in gluon–gluon luminosity affects the favored production mechanisms of the Higgs boson.

nosities (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 [19] 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 for the various production mechanisms 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 to decay to pairs of weak bosons ($M_H < 160 \text{ GeV}/c^2$), the Higgs decays predominantly to either b -quarks ($b\bar{b}$, 90%) or a pair of τ leptons ($\tau^+\tau^-$, $\approx 10\%$). Above the $W^\pm W^\mp$ threshold, decays to vector bosons ($H \rightarrow W^\pm W^\mp$ and $H \rightarrow ZZ$) dominate. The dependence

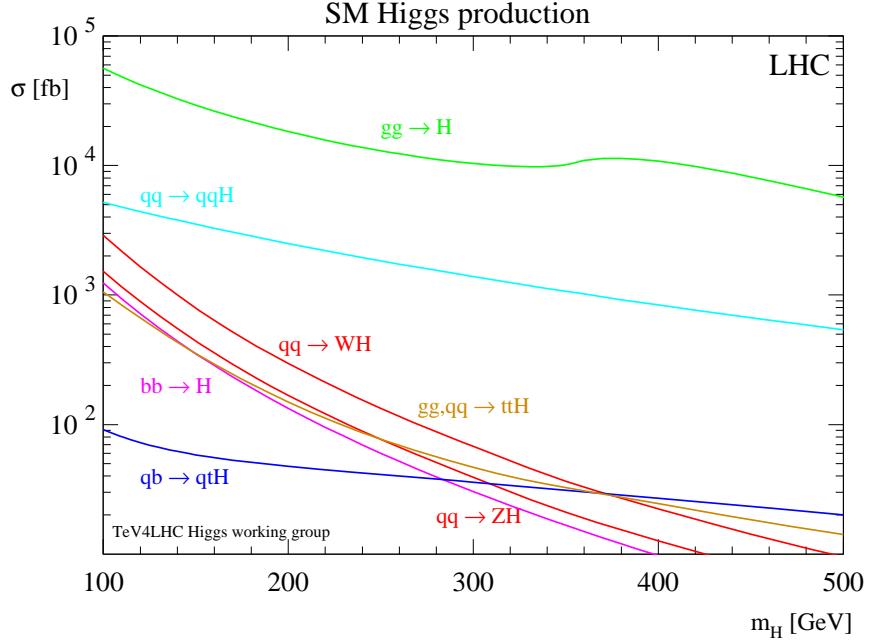


Figure 1.9: Cross section of the SM Higgs boson versus the Higgs boson mass. The different curves give the contribution to the cross section from different production mechanisms. Source: [20].

of branching fraction on M_H and the other rare decay modes are illustrated in Figure 1.10.
 For low mass Higgs, the $\tau^+ \tau^-$ decay mode plays a particularly important role. The dominant decay mode $H \rightarrow b\bar{b}$ suffers from enormous backgrounds from QCD jet production. It is important to understand the magnitude of difference between expected Higgs boson production and the rates of various backgrounds. Figure 1.11 illustrates the cross sections for different SM processes at hadron colliders. The rate of Higgs production is many orders of magnitude ($\mathcal{O}(10^{-7})$) smaller than that of QCD production. It is important to therefore design searches to use handles that can reject the vast majority of the uninteresting events at hadron colliders.

§1.3.2 MSSM Higgs Phenomenology

The phenomenology of the Higgs sector of the MSSM is similar to the SM 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

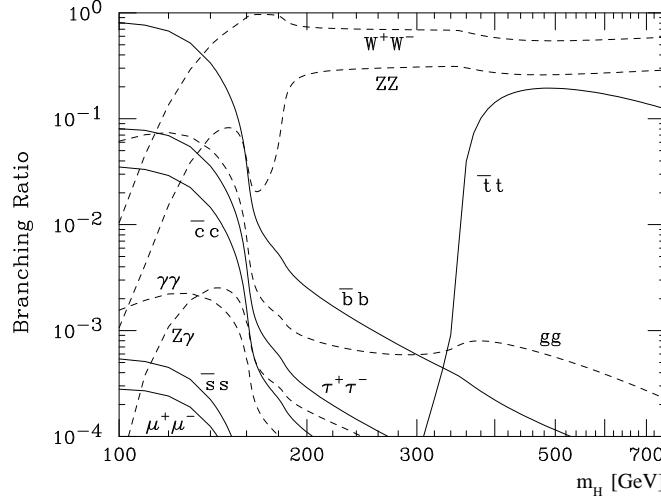


Figure 1.10: Branching fraction of the SM Higgs bosons for different values of M_H .
Source: [20].

584 by $\tan \beta$. The gluon–gluon cross section is therefore increased by $\tan^2 \beta$, where the top
 585 quark loop in Figure 1.3.1 is replaced by a ($\tan \beta$ enhanced) b quark loop. Additionally,
 586 MSSM Higgs production with associated b –quarks, illustrated in Figure 1.3.2, becomes an
 587 important production mode. At tree–level, the MSSM can be defined by the mass of the
 588 CP–odd Higgs m_{A^0} and $\tan \beta$. For a reasonably high $\tan \beta$, there is always one CP–even
 589 Higgs (h^0 or H^0) which is mass–degenerate with the A^0 . When $\tan \beta$ and m_{A^0} are both
 590 large, associated b production dominates the total cross section [22]. The cross sections of
 591 the different MSSM neutral Higgs bosons are shown in Figure 1.13. The $\tan \beta$ enhancement
 592 of the MSSM Higgs coupling to the b quarks and τ leptons causes the branching fraction of
 593 all neutral MSSM Higgs states to be $H \rightarrow b\bar{b}$ (90%) and $H \rightarrow \tau^+\tau^-$ (10%) across the entire
 594 range of m_{A^0} . The enhanced production rate and the high branching fraction to τ leptons
 595 make the MSSM Higgs decaying to τ leptons an exciting and promising channel to search
 596 for Higgses and Supersymmetric physics at colliders.

597 §1.3.3 Results from LEP and Tevatron

598 The LEP and Tevatron experiments have both set limits on the existence of the SM and
 599 MSSM Higgs boson. Additionally, precision electroweak measurements give additional hints
 600 on the prospects for both models.

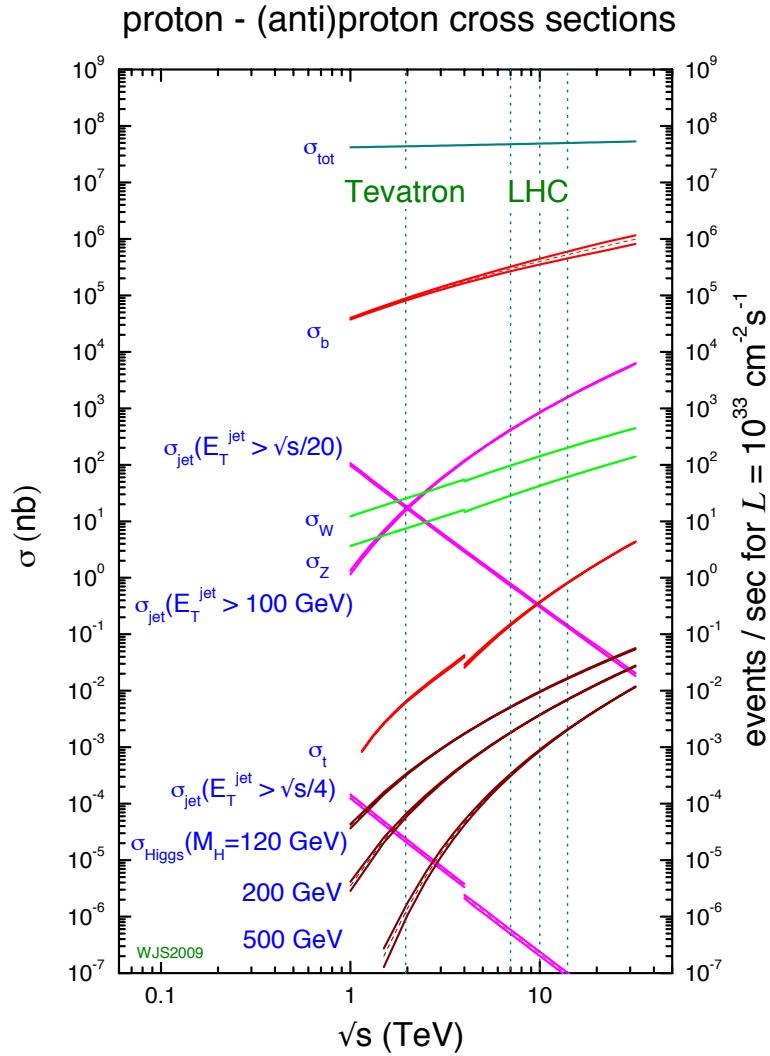


Figure 1.11: Cross sections of various processes at hadron colliders. The horizontal axis represents the center of mass energy of the collision. Of note is the vast difference in scales between Higgs production (maroon lines, $\mathcal{O}(10^{-2} \text{ nb})$) and the QCD cross section to produce $b\bar{b}$ pairs (red line, $\mathcal{O}(10^4 \text{ nb})$). Source: [21].

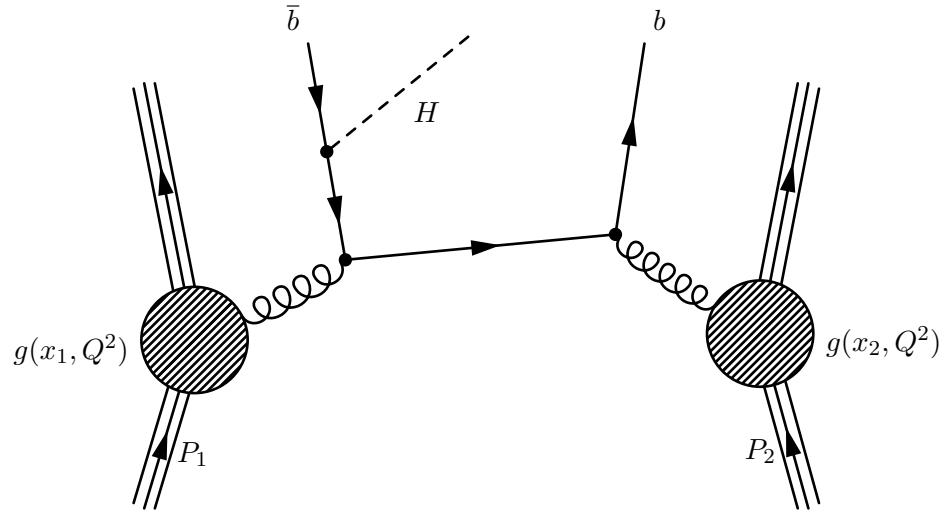


Figure 1.12: One possible diagram for an MSSM Higgs produced with associated b -quarks in a proton–proton collision.

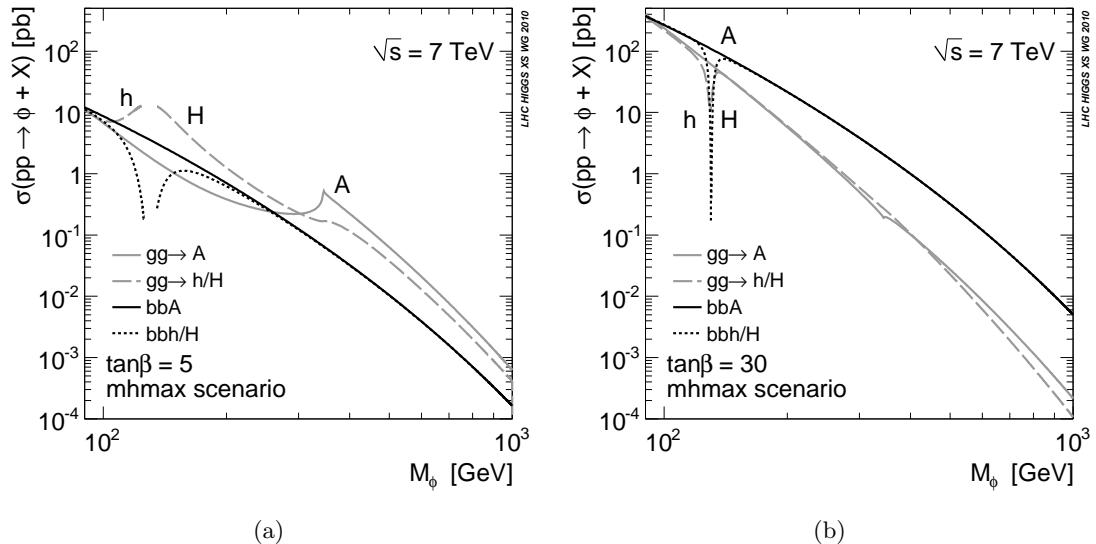


Figure 1.13: Cross sections for the different MSSM Higgs bosons versus m_{A^0} in the m_{hmax} benchmark scenario [23] scenario for $\tan\beta = 5$ (a) and $\tan\beta = 30$ (b). Source: [22]

Higgs Decay	Z Decay
$b\bar{b}$	$q\bar{q}$
$\tau^+\tau^-$	$q\bar{q}$
$b\bar{b}$	$t\bar{t}$
$b\bar{b}$	$\nu\bar{\nu}$
$b\bar{b}$	$\mu^+\mu^-$
$b\bar{b}$	e^+e^-

Table 1.3: Different channels used at LEP to search for Higgs bosons produced with the Higgstrahlung mechanism.

601 LEP was an e^+e^- collider at CERN and has effectively excluded the presence of a
 602 low (less than 114 GeV/c²) mass Higgs boson. The dominant SM Higgs production mode
 603 at LEP is Higgstrahlung, where the Higgs is produced in association with a Z boson (see
 604 Figure 1.3.1). The search at LEP utilized a number of different decay channels [20]. The
 605 decay channels used in the LEP search are summarized in Table 1.3.3.

606 The results using all channels from the four LEP experiments¹⁰ have been combined into
 607 a single limit, shown in Figure 1.14. The analysis sets a limit on the ratio $\xi^2 = (g_{HZZ}/g_{HZZ})^2$,
 608 the upper limit on the HZZ coupling divided by the predicted value of the SM. For Higgs
 609 masses below 114 GeV/c², the ratio is below unity at the 95% confidence level, ruling out a
 610 SM Higgs below that mass.

611 The Tevatron is a proton–antiproton collider with a center-of-mass energy of $\sqrt{s} = 1.96$ TeV ■
 612 There are two general purpose detectors at the Tevatron, CDF and DØ . The dominant Higgs
 613 production modes at the Tevatron are Higgstrahlung and gluon fusion (see Figure 1.3.1).
 614 For low mass ($m_H < 135$ GeV/c²) Higgs bosons the dominant channel at the Tevatron is the
 615 Higgstrahlung production mode and $H \rightarrow b\bar{b}$ decays. Large multi-jet backgrounds prevent
 616 the $H \rightarrow b\bar{b}$ decay mode from being useful for searching for Higgs bosons produced by gluon
 617 fusion. The $H \rightarrow \tau^+\tau^-$ and $H \rightarrow \gamma\gamma$ decays are additionally used in an inclusive search at
 618 low mass, but do not dominate the search sensitivity. The combined low-mass limit on the
 619 SM Higgs from both Tevatron experiments is shown in Figure 1.15. The Tevatron currently

¹⁰ALEPH, DELPHI, L3, and OPAL

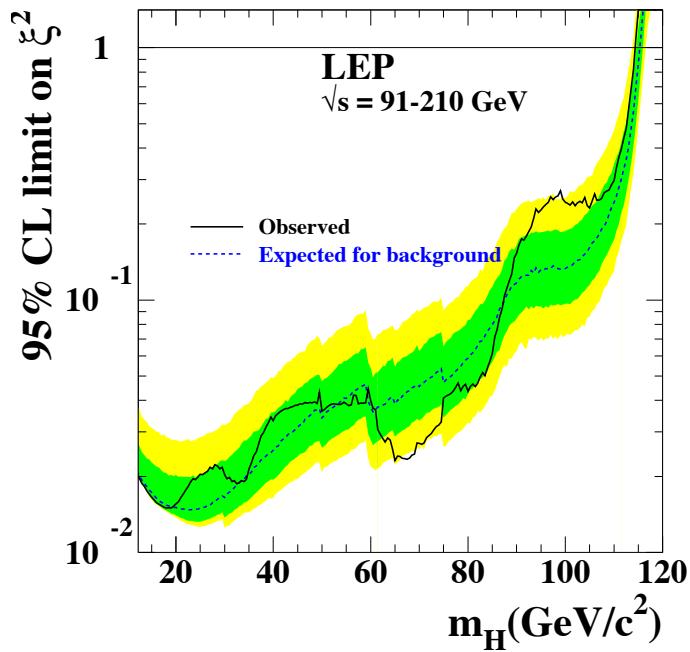


Figure 1.14: Combined LEP upper limit set on the quantity $\xi^2 = (g_{HZZ}/g_{HZZ})^2$ at 95% confidence level. Regions where the observed ratio is less than one exclude the SM. The dashed line gives the expected limit for the null (background only) hypothesis, with the green and yellow bands representing the expected variance at one and two sigma, respectively, of the limit. The solid line is the observed limit from the combined LEP data. Reference: [20]

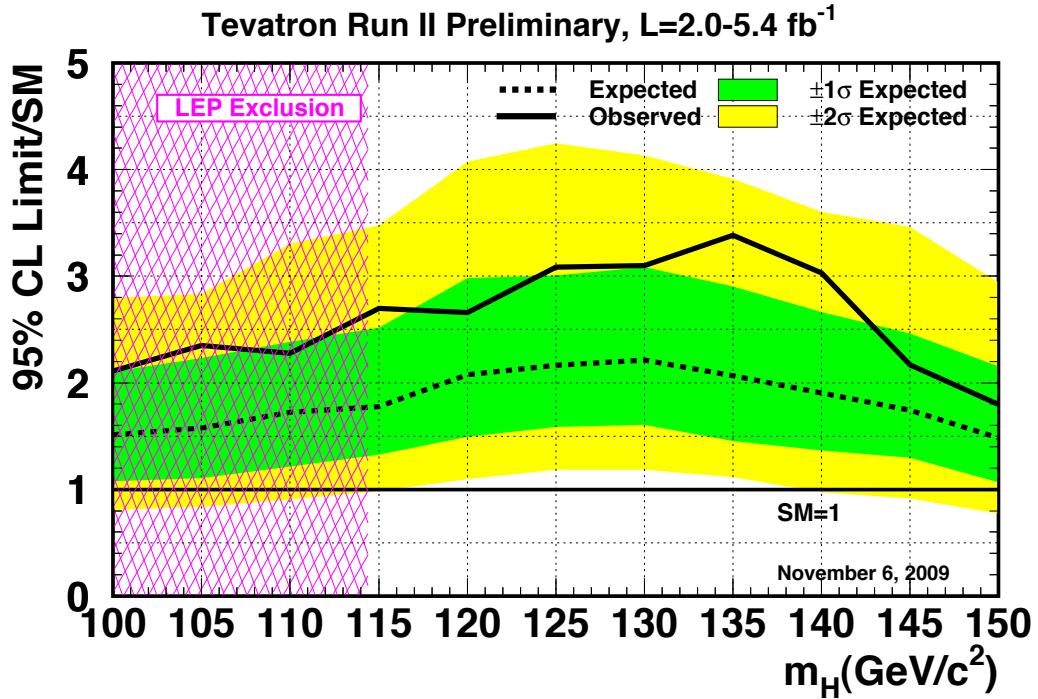


Figure 1.15: Combined CDF and DØ RunII upper limit on the cross section of a standard model-like Higgs boson. The LEP limit is shown in pink. Reference: [20]

sets an upper limit on the SM Higgs cross section of about 2.5 times the SM expectation.

When ($m_H < 135$ GeV/ c^2) the $H \rightarrow W^+W^-$ decay mode becomes significant. Low di-boson backgrounds allow this decay mode to probe both the Higgstrahlung and gluon fusion production modes. The combined results of the CDF and DØ searches using the W^+W^- have decay mode recently excluded (See Figure 1.16) a SM Higgs with a mass between 162 and 166 GeV/ c^2 . This is the first exclusion in SM Higgs mass parameter space since the LEP result.

Analyses at LEP and Tevatron have also addressed excluded regions of the MSSM. At LEP, the dominant production modes of the MSSM Higgs bosons are Higgstrahlung and pair production, where $e^+e^- \rightarrow h^0A^0$ or H^0A^0 . For the Higgstrahlung production mode, the SM search can be reinterpreted in terms of the MSSM. To address the pair production mode, searches were performed in the $e^+e^- \rightarrow h^0A^0 \rightarrow b\bar{b}b\bar{b}$ and $\tau^+\tau^-q\bar{q}$ decay modes. Finally, LEP is also sensitive to associated MSSM Higgs production at low m_{A^0} and high $\tan\beta$ to $e^+e^- \rightarrow \{\bar{\phi}$, where the associated fermions { are b -quarks or tau leptons. The

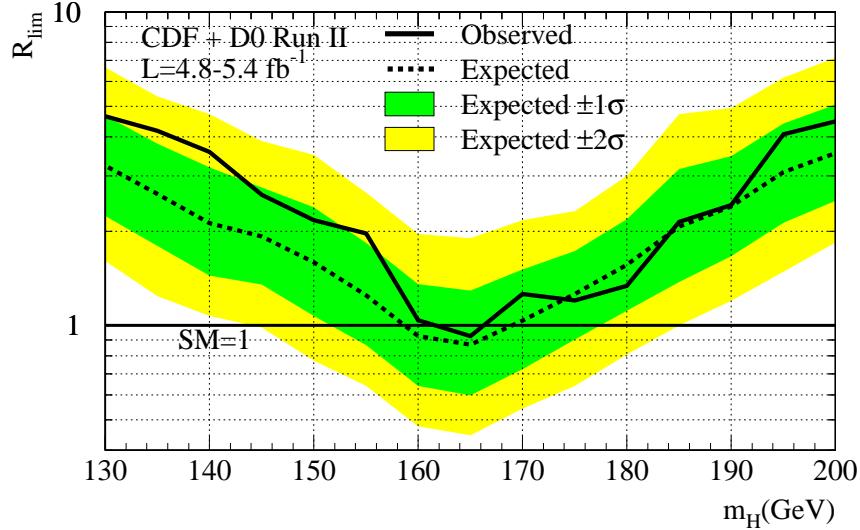


Figure 1.16: Combined CDF and DØ RunII upper limit on the cross section of a SM–like Higgs boson using the $H \rightarrow W^+W^-$ decay mode. The SM is excluded for Higgs boson masses between 162 and 166 GeV/ c^2 . Reference: [20]

634 combined limits from LEP in the $m_{A^0} - \tan \beta$ plane are shown in Figure 1.17.

635 At the Tevatron, CDF and DØ have set a combined limit on the MSSM using the
 636 inclusive $H \rightarrow \tau^+\tau^-$ channel. The analysis presented in this thesis is very similar to the
 637 approaches used at the Tevatron. Results from the Tevatron have excluded the MSSM for
 638 $\tan \beta$ greater than approximately 35 for MSSM Higgs mass $m_{A^0} < 200$ GeV/ c^2 . The full
 639 exclusion plot for the m_h –max and “no mixing” MSSM benchmark scenarios are shown in
 640 Figure 1.18.

641 §1.4 The Physics of the Tau Lepton

642 As discussed in Sections 1.3.1 and 1.2.3, the τ lepton is an important probe of Higgs physics.
 643 The τ lepton has some unusual properties which make it particularly challenging at hadron
 644 colliders. With a mass of 1.78 GeV/ c^2 , the τ lepton is heaviest of the leptons. The nominal
 645 decay distance $c\tau$ of the τ lepton is 87 μm , which in practice means that the τ will always
 646 decay before reaching the first layer of the detector. Tau decays can be effectively classified
 647 into two types. “Leptonic” decays consist of a τ decaying to a light lepton ($\ell = e, \mu$) and two
 648 neutrinos $\tau^+ \rightarrow \ell^+ \nu_\tau \bar{\nu}_\ell$. “Hadronic” decays consist of a low-multiplicity collimated group of
 649 hadrons, typically π^\pm and π^0 mesons. The hadronic decays of the τ lepton compose approx-

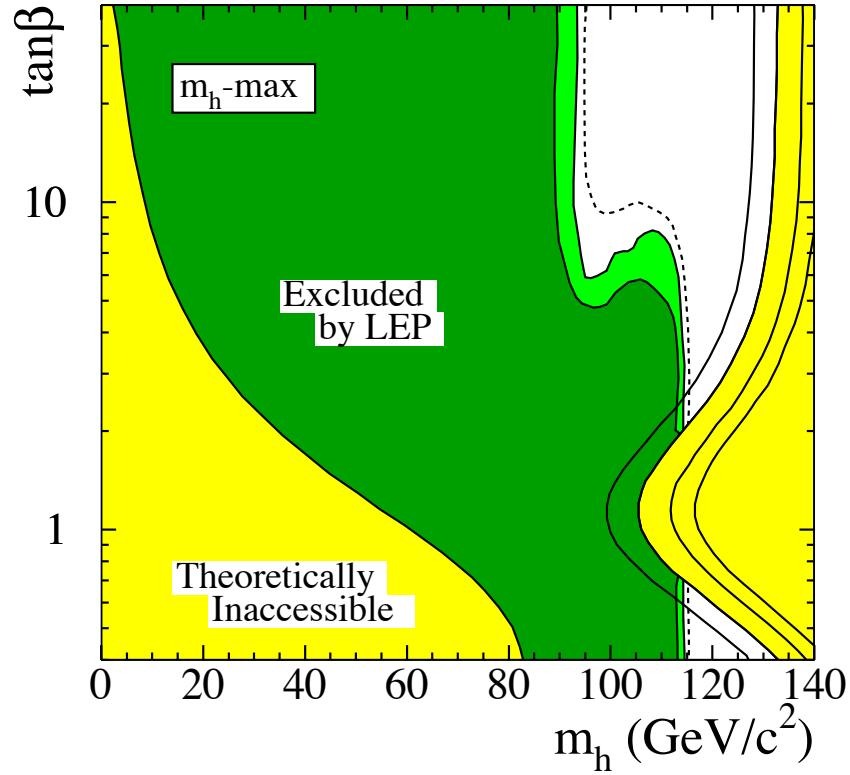


Figure 1.17: Combined LEP limits on the MSSM. The results are interpreted in the context of the m_h -max benchmark [23] scenario of the MSSM. Reference: [20]

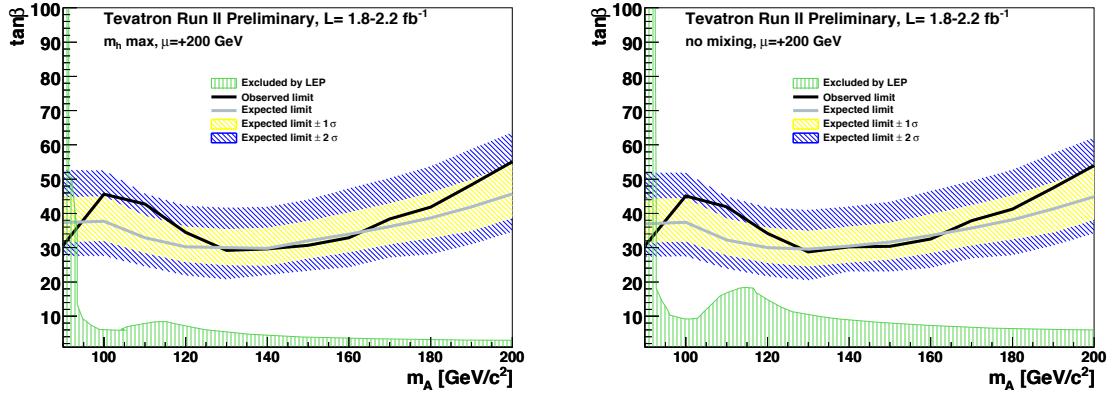


Figure 1.18: Combined Tevatron limits on the MSSM. The grey line and blue and yellow bands give the expected limit and its one and two sigma contours. The black line is the observed limit. The results are interpreted in the context of the m_h -max benchmark (left) and “no mixing” (right) MSSM scenarios. The limit from LEP is shown in green. Reference: [20]

Visible Decay Products	Resonance	Mass (MeV/ c^2)	Fraction [20]
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 1.4: 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.

650 imately 65% of the τ lepton branching fraction, with the remainder shared approximately
 651 equally by the leptonic decays. The branching fractions for the leptonic and most common
 652 hadronic decays are shown in Table 1.4.

653 The tau is also a challenging object in that the decay of the tau always includes neu-
 654 trinos. The associated neutrinos are weakly interacting and do not create a signal in any
 655 detector at CMS. The only sign that the neutrinos are there is an imbalance in the total
 656 transverse¹¹ energy in the event. This thesis will describe a novel way to reconstruct the
 657 neutrinos associated to tau decays in Chapter 4.

A tau with produced with energy E travels on average

$$\gamma c\tau = \frac{E}{1.78 \text{ GeV}} 87 \text{ } \mu\text{m}$$

658 before decaying in the detector. These lengths are comparable to the resolution of the CMS
 659 tracker, therefore it is possible to reconstruct a vertex corresponding to a tau decay that is
 660 displaced with respect to the primary vertex. This can be used as an additional discriminant
 661 against QCD, which is expected to decay promptly. Furthermore, in Chapter 4 we will see

¹¹At proton colliders, the constituent quarks/gluons of the proton share the total proton momentum. As the total fraction of momentum carried by the parton involved in a hard collision is unknown, longitudinal momentum is not conserved.

662 it may be possible to use it when reconstructing the associated neutrinos.

663

Chapter 2

664

The Compact Muon Solenoid Experiment

665 The Compact Muon Solenoid (CMS) Experiment is a “general purpose” particle detector
 666 designed to measure collision events at the Large Hadron Collider (LHC), a proton–proton
 667 synchrotron located at the CERN laboratory in Geneva, Switzerland. The design goals of
 668 the CMS experiment are [24], in order of priority:

- 669 • Good muon identification and momentum resolution over a wide range of momenta
 670 and angles, good dimuon mass resolution ($\approx 1\%$ at 100 GeV/c 2), and the ability to
 671 determine unambiguously the charge of muons with $p < 1$ TeV/c;
- 672 • Good charged-particle momentum resolution and reconstruction efficiency in the in-
 673 ner tracker. Efficient triggering and offline tagging of τ ’s and b -jets, requiring pixel
 674 detectors close to the interaction region;
- 675 • Good electromagnetic energy resolution, good diphoton and dielectron mass resolution
 676 ($\approx 1\%$ at 100 GeV/c 2), wide geometric coverage, π^0 rejection, and efficient photon and
 677 lepton isolation at high luminosities;
- 678 • Good missing-transverse-energy and dijet-mass resolution, requiring hadron calorime-
 679 ters with a large hermetic geometric coverage and with fine lateral segmentation.

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

686 of different species of particles. The central (closest to interaction point) sub-detectors are
 687 the charged particle tracking systems (the “tracker”). The tracker is designed to be a non–
 688 destructive instrument, which means that ideally that the momentum of particles are un–
 689 changed after passing through it. Outside of the tracker is the electromagnetic and hadronic
 690 calorimeters, which are abbreviated ECAL and HCAL, respectively. The calorimeters are
 691 destructive detector, and are designed such that visible incident particles are completely
 692 absorbed. The outer layers of CMS are designed to measure muons, the one¹ species of
 693 particle that is nearly immune to the stopping power of the calorimeter. The arrangement
 694 of destructive and non-destructive sub-detectors facilitates the identification of different
 695 types of particles. This concept is illustrated in Figure 2.1(b). In this chapter we give an
 696 brief overview of the LHC machine, and then describe the individual sub-detector systems
 697 of CMS.

698 §2.1 The Large Hadron Collider

699 The Large Hadron Collider is a proton–proton synchrotron, with a design collision energy
 700 of 14 TeV. At the time of this writing (and for the foreseeable future), the LHC is the
 701 world’s largest and highest energy particle accelerator. A synchrotron is a machine that
 702 accelerates beams of charged particles by using magnets to steer them in a circle through
 703 radio–frequency resonating cavities which accelerate the particles. As the LHC is a collider,
 704 there are two beams that are accelerated in opposite directions. The maximum beam energy
 705 of a synchrotron is determined by its radius and the maximum strength of the magnetic
 706 fields used to bend the path of the beam. The dipole magnets used by the LHC to steer the
 707 particles are superconducting niobium–titanium. To maintain them in a superconducting
 708 state, they are cooled using superfluid liquid helium to 1.9 Kelvin. To store the beam at the
 709 injection energy of 450 GeV, the magnetic dipole fields must be maintained at 1/2 Tesla. As
 710 the energy of each beam energy is increased to its (design) maximum of 7 TeV, the dipole
 711 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.

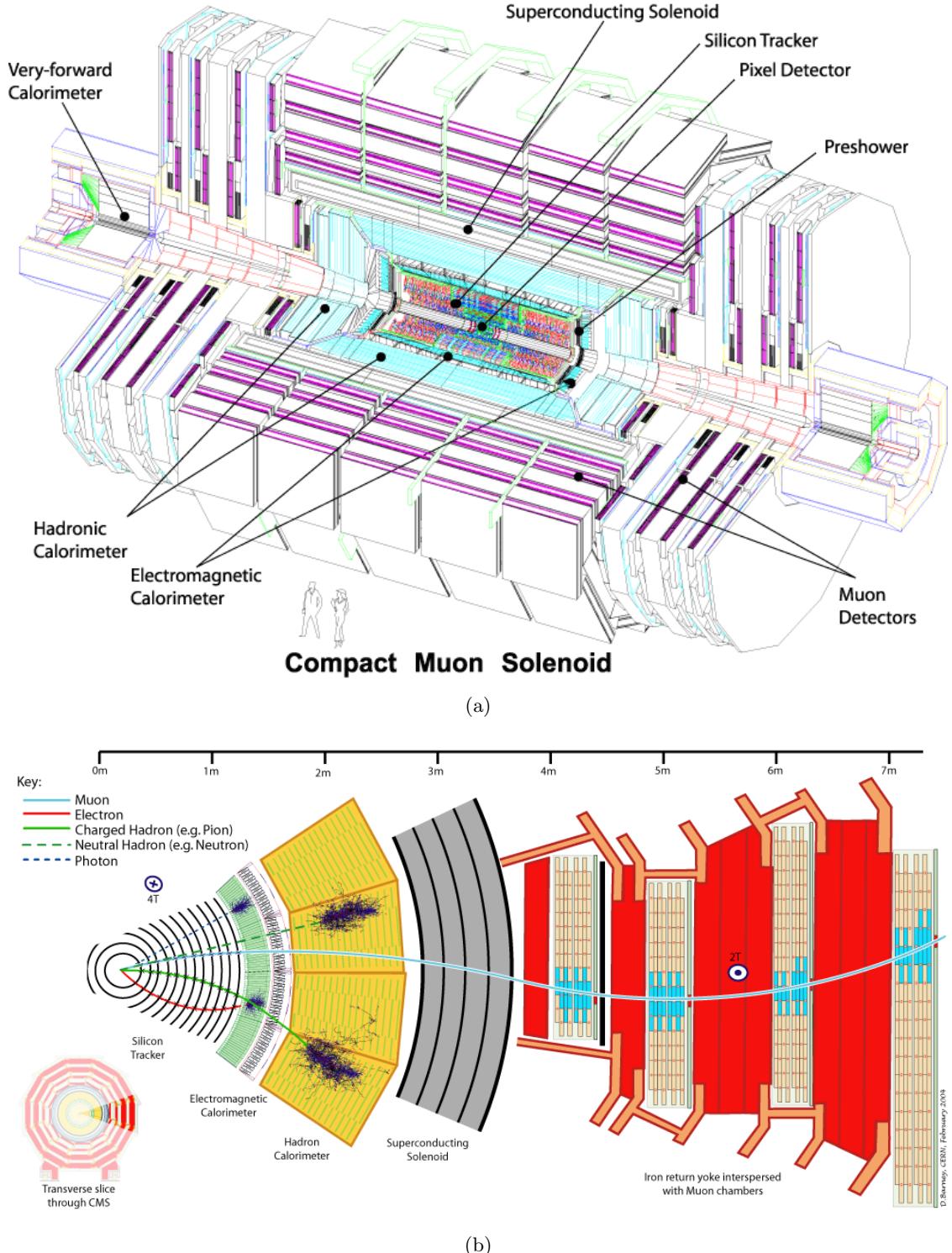


Figure 2.1: Figure (a), top, shows a schematic drawing of the CMS detector. The individual sub-detectors are labeled. Two humans are shown in the foreground for scale. Figure (b) shows a radial cross section of the detector and demonstrates how the (non-)destructiveness of different sub-detectors facilitates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

712 §2.2 Solenoid Magnet

The four Tesla field of the CMS solenoid magnet is a critical factor in ability of CMS to precisely measure the particles produced in collisions at the LHC. The momentum of charged particles is measured in the tracking 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

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

713 where q is the charge of the particle, B is strength of the magnetic field, and p_{\perp} is the
 714 component of the particle's relativistic momentum perpendicular to the direction of the
 715 magnetic field. From Equation 2.1, it is evident that the ability to measure high momentum
 716 charged particles (a critical goal of CMS) requires a high magnetic field. Even at very high
 717 particle energies where the resolution becomes poor, the strength of the magnetic field
 718 is still very important for identifying the bending direction of the particle; the direction
 719 corresponds to the particle's electric charge. Furthermore, the homogeneity of the magnetic
 720 field is important to minimize systematic errors in the measurement of tracks.

721 The CMS solenoid is extremely large. The radial bore of the magnet is 6.3 meters; the
 722 magnet is 12.5 meters in length and weighs 220 tons. The large bore of the magnet allows
 723 the tracker and calorimeter systems to be located inside the solenoid. The internal windings
 724 of solenoid is arranged in four layers to increase the total field strength and are cooled by
 725 liquid helium to a temperature of 4.5 Kelvin. The windings are magnetically coupled to
 726 the support superstructure. This coupling allows the magnetic to heat uniformly during a
 727 “quench” event², reducing localized stresses. The nominal current at full field of the solenoid
 728 is 19.14 kA. The solenoid itself is surrounded by an iron return yoke with a total mass of
 729 10,000 tons. The return yoke surrounding the solenoid minimizes the fringing field. The
 730 muon detector system is interspersed inside the yoke, and takes advantage of the return
 731 field 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.

732 §2.3 Charged Particle Tracking Systems

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

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

756 The tracking system has been specifically designed for the high radiation environment
 757 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 [24].

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} .

Fixme: right

acronym

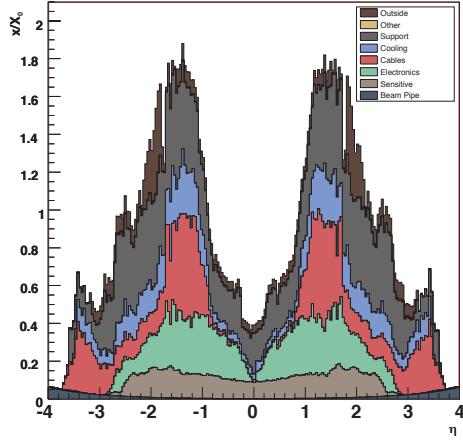


Figure 2.2: Material budget of the CMS tracker in units of radiation lengths X_0 versus pseudorapidity. The material budget is broken down into the contributions from the different components of the tracker. The amount of material is largest in the “transition region” between the barrel and endcap.

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

767 In the ideal case, the tracker would be a non-destructive instrument. However, charged
 768 particles can interact with the mass of the tracker (and its support infrastructure). These
 769 interactions limit the resolution of the tracker. The amount of matter in the tracker is
 770 referred to as the “material budget”. The material budget of the CMS tracker depends
 771 heavily on the pseudorapidity η and is illustrated in Figure 2.2. The relatively large ma-
 772 terial budget of the CMS tracker has two effects: charged particles can undergo “multiple
 773 scattering,” interacting with material in the tracker. This can cause “kinks” in the recon-
 774 structed track. Hadronic particles (charged and neutral) can undergo nuclear interactions,

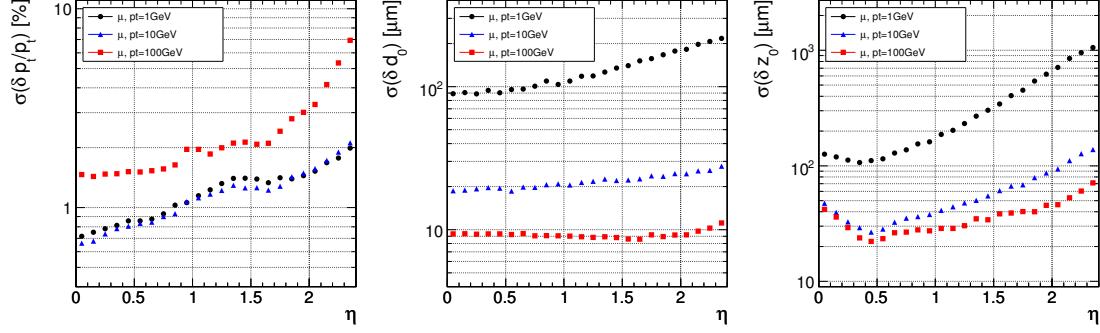


Figure 2.3: Expected resolutions of reconstructed transverse momentum (left), transverse impact parameter (center), and longitudinal impact parameter (right) versus absolute pseudorapidity $|\eta|$. The resolution is shown for three different cases of particle p_T , 1 GeV/c (black), 10 GeV/c (blue),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

which are hard collisions between the incident particle and a nucleus in 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 material 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 [25] in 7 TeV 2010 CMS data using $J/\psi \rightarrow \mu^+\mu^-$ decays and is found to agree with the prediction from simulation within 5%. The impact parameter and vertex resolutions have also been measured [26] 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 interact electromagnetically with the crystal, producing a shower of electrons and photons. The

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.

792 shower will expand until it consists entirely of photons. The crystal is optically clear, so
 793 these photons travel to the rear face of the crystal where they are then counted by a pho-
 794 tomultiplier. The number of detected photons can then be related to the energy that was
 795 deposited in the crystal. At 18°C, about 4.5 photoelectrons will be produced per MeV of de-
 796 posited energy. The ECAL has excellent solid angle coverage, extending to a pseudorapidity
 797 of $|\eta| = 3.0$.

798 The ECAL uses lead tungstate (PbWO_4) crystals as the scintillation medium. The
 799 crystals have a very large density, which allows the calorimeter to be relatively compact.
 800 To be able to correctly measure the energy of electrons and photons, an incident photon or
 801 electron must be completely stopped by interactions with the calorimeter. The quantities
 802 that determine if an electron or photon will be completely contained is the total depth of
 803 the crystal, the crystal density, and the radiation length property X_0 of the crystal. The
 804 radiation length X_0 is defined as the mean distance (normalized to material density) after
 805 which an electron will have lost $(1 - \frac{1}{e})$ of its energy. The PbWO_4 crystals of the CMS
 806 ECAL have a density of 8.28 g/cm² and a depth of 230 mm. A single crystal thus has a
 807 total radiation length of 25.8 X_0 , and will capture on average 99.9993% of the energy of an
 808 incident electron. The front face of the crystal is 22 mm × 22 mm, which corresponds to an
 809 $\eta - \phi$ area of 0.00174×0.00174 . The Molière radius of a material is the average radial profile
 810 size of an electromagnetic shower, and for PbWO_4 is 2.2 cm. The fact that the Molière
 811 radius is larger than the size of the individual crystals improves the spatial resolution of
 812 the measurement. As the shower is shared between multiple crystals, the relative amounts
 813 deposited in each crystal allows the true impact point to be determined with a resolution
 814 smaller than the individual crystal size.

815 The transparency of the CMS ECAL crystals change as they are exposed to radiation.
 816 However, at the working temperature of the ECAL (18°C), the crystal transparency will
 817 naturally return to its nominal value. The transparency of the crystals thus decreases during
 818 the course of a run of collisions, then increases during the following collision-less period.
 819 The changing transparency conditions need to be continuously monitored and corrected
 820 for to ensure a stable detector response. The transparency of the crystals are measured
 821 continuously using two lasers. One laser has wavelength $\lambda = 400$ nm which corresponds to

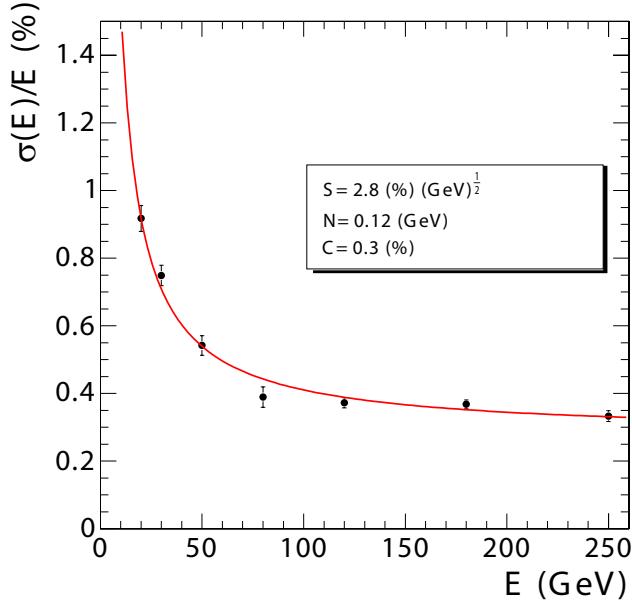


Figure 2.4: Energy resolution (in %) of the CMS ECAL measured at an electron test beam. The resolution depends on the incident energy of the electron. The points are fitted to function with the form given in Equation 2.2. The fitted parameters are given in the legend.

the color of light produced in the scintillations and is sensitive to changes in transparency.
 The other laser is in the near-infrared and is used to monitor the overall stability of the crystal. The lasers are synchronized to pulse between LHC bunch trains so the transparency can be continuously monitored while collisions are occurring.

The energy resolution of the ECAL is given by

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

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

§2.5 Hadronic Calorimeter

The hadronic calorimeter (HCAL) surrounds the CMS ECAL and is located within the coil of the CMS solenoid magnet. To ensure incident particles are completely contained within the calorimeter volume, in the barrel region the HCAL employs a “tail-catcher”, an extra

834 layer of calorimetry outside of the magnet. The hadronic calorimeter measures the energy
 835 of charged and neutral hadronic particles. The HCAL is a *sampling* calorimeter. Layers of
 836 plastic scintillating tiles are interspersed between brass absorber plates. An incident hadron
 837 produces a hadronic shower as it passes through the absorber. The particles in the shower
 838 produce light as they pass through the scintillating tiles. Measuring the light produced in
 839 each layer of tile allows the reconstruction of the radial profile of the shower which can be
 840 related to the deposited energy. The response of the scintillator tiles are calibrated using a
 841 radioactive source, either Cs¹³⁷ or Co⁶⁰. Small stainless tubes permit the radioactive sources
 842 to be moved into the center of the tile during calibration. The granularity of the HCAL is
 843 0.087 × 0.087 and 0.17 × 0.17 in $\eta - \phi$ in the barrel ($|\eta| < 1.6$) and endcap ($|\eta| > 1.6$),
 844 respectively.

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

850 The HCAL includes a specially designed forward calorimeter (HF). The design of the
 851 forward calorimeter is constrained by the extreme amount of radiation it is exposed to,
 852 particularly at the highest rapidities. The active material of the HF are quartz fibers. The
 853 fibers are installed inside grooves inside of a steel absorber. Charged particles created in
 854 showers in the absorber create light in the fibers, provided they have energy greater than
 855 the Cherenkov threshold. As Cherenkov light is created by
 856 the passage of charged particles through matter, the HF design is not sensitive to neutrons
 857 emitted by radionucleids that may be created in the absorber material during operation.
 858 The fibers are grouped into two sets: one set of fibers are installed over the full depth of
 859 the detector, the other only cover half the depth. A crude form of particle identification
 860 is possible, as showers created by electrons and photons will deposit the majority of the
 861 energy in the front of the detector.

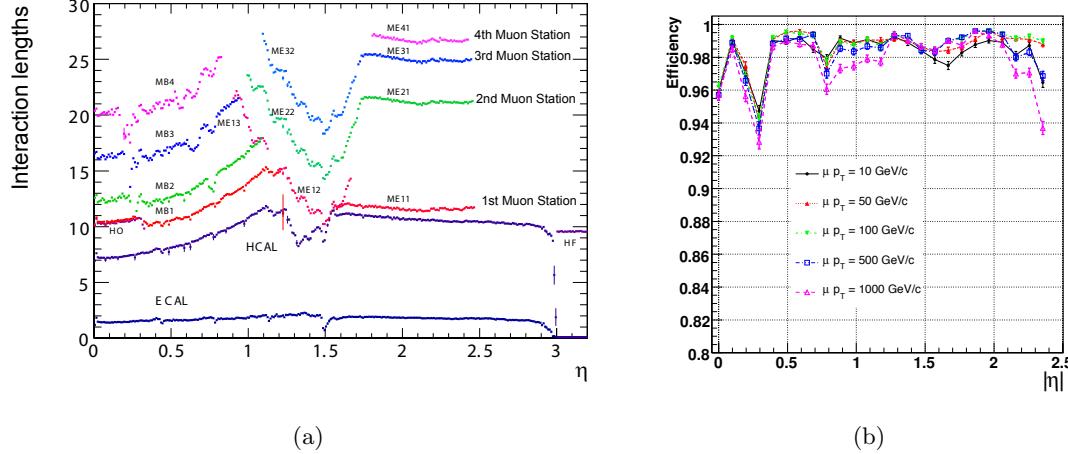


Figure 2.5: The left figure, (a), illustrates the number of interaction lengths versus pseudorapidity η of material that must be traversed before reaching the different layers of the muon system. On the right, (b) shows the efficiency versus η to reconstruct a “global” muon for different transverse momenta.

862 §2.6 Muon System

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

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

871 The approach to detecting muons is to build the detector to a thickness such that other
 872 particles (electrons, photons, hadrons) will not penetrate the outermost calorimeter. Any
 873 charged particle that is detected outside of this region can then be identified as a muon. At
 874 CMS, the muon detection systems are built into the magnet return yoke outside of the CMS
 875 calorimeters and magnet, giving them excellent protection (illustrated in Figure 2.5(a))
 876 against hadronic “punch-through.” The purity of particles that reach the muon system
 877 make it especially effective as a “trigger” of interesting physics. The CMS muon system

878 has the feature that it additionally can trigger on the transverse momenta of muons. The
 879 CMS muon system is composed of three types of detectors: drift tubes (DT), resistive plate
 880 chambers (RPC), and cathode strip chambers (CSC).

881 A drift tube detector is of a tube filled with a mixture of argon (85%) and carbon
 882 dioxide (15%) gas with a positively charged ($V = +3.6$ kV) wire running through the
 883 middle of the tube. When a charged particle passes through the tube, it ionizes some gas.
 884 The free electrons are then drawn to the positively charged wire inside the tube, creating a
 885 signal when reach it. The speed of the detector is limited by the “drift time,” the maximum
 886 amount of time it may take for an electron to reach a sensor wire. The precision of the
 887 spatial measurement can be increased by recording the time at which each wire records a
 888 signal and correlating the measurements across multiple tubes. The time resolution of the
 889 CMS DTs is on the order of a few nanoseconds, allowing the DT to provide a trigger on
 890 a given proton bunch crossing. The tubes in adjacent layers are offset by one half tube
 891 width to take advantage of this effect and ensure there are no gaps in the fiducial region. In
 892 CMS, the smallest unit of the DT system is the superlayer, which consists of four layers of
 893 tubes. A DT chamber consists of three or two superlayers. The tubes in the two superlayers
 894 farthest from the beam are oriented parallel to the beam and measured the bending of the
 895 muons in the magnetic field. The inner superlayer is oriented orthogonally to the beam and
 896 measures the longitudinal position of incident muons. There are four muon “stations” in
 897 the barrel which contain DT chambers. The stations correspond to available areas in the
 898 magnetic return yoke. In the barrel, the muon momentum resolution of the DTs is better
 899 than 95%.

900 Cathode strip chambers (CSCs) are used in the endcap muon system, providing cov-
 901 erage in the pseudorapidity range $0.9 < |\eta| < 2.4$. A cathode strip chamber consists of a
 902 chamber filled with inert gas that with a number of internal wires held at a high voltage. A
 903 number of cathode strips are installed perpendicular⁶ induced to the wires on the walls of
 904 the chamber. When a muon passes through the CSC, it ionizes some of the gas. The high
 905 voltage on a nearby wire causes this ionized gas to break down, forming a conductive pas-

⁶The wires are actually placed at an angle to the perpendicular to compensate for a shifting effect caused by the magnetic field Lorentz force.

906 sage in the gas and an “avalanche” current between the wire and a number of the cathode
 907 strips. The spatial position of the hit in two dimensions is found taking one coordinate from
 908 the wire and the other coordinate from the signal average of the cathode strips.

909 The CSCs in the CMS endcap are positioned such that a muon in the pseudorapidity
 910 range $1.2 < |\eta| < 2.4$ will cross three or four CSC detectors. The geometry of the CSC
 911 strips and wires is designed to provide a spatial $r - \phi$ resolution of 2 mm at the L1 trigger
 912 level and a final offline reconstruction resolution of 75 μm for the first layer and 150 μm for
 913 outer layers. The RMS of the response time for a CSC layer is about 11 ns, which is too
 914 long to correctly associate a signal in the CSCs to an LHC bunch crossing (25 ns) with
 915 high efficiency. By grouping the layers into chambers, and taking the shortest response, the
 916 correct bunch crossing can be identified with 98–99% efficiency.

917 The Resistive Plate Chamber (RPC) muon detectors ensure that the muon system can
 918 be used as a fast, first level trigger. The RPC detector consists of two gaps filled with gas
 919 (up and down) with a common set of strips between the two gaps. The strips are oriented
 920 parallel to the beam line to permit measurement of the transverse momentum of the muons.

921 §2.7 Trigger System

922 At the LHC, proton bunch crossings (collisions) occur every 25 ns. This corresponds to
 923 an interaction rate of 40 MHz. At this high rate, and with the huge number of channels
 924 in the CMS detector, the front-end bandwidth readout from the detector is over 1 Pb/s.
 925 Due to bandwidth and storage requirements, the rate at which events are permanently
 926 recorded must be reduced by more than a factor of a million. This reduction is achieved
 927 by CMS trigger system. As only a fraction of the total events can be stored, and the rate
 928 of diffractive and common QCD multi-jet production is many orders of magnitude larger
 929 than “interesting” new physics (see Figure 1.11). The trigger must therefore be designed to
 930 select “interesting” events. A typical requirement applied at the trigger level might be the
 931 presence of a high- p_{T} muon, an isolated ECAL deposit, or a large deposit of energy in the
 932 event.

933 The CMS trigger consists of two stages: a fast Level-1 (L1) trigger and a High-Level
 934 Trigger (HLT). The L1 trigger system is built on custom, typically reprogrammable elec-

tronics and interfaces directly to the detector subsystems. The L1 trigger has access to information from the muon and calorimeter systems. The L1 does not have access to the full granularity of the muon system and calorimeters but must make the decision based on coarse segments. The design acceptance rate of the L1 trigger is 100 kHz. The trigger typically operates at a nominal rate of 30 kHz. The maximum latency of the L1 is 3.2 μ s, requiring that the output from detector electronics be passed through memory pipelines to ensure that no bunch crossings go unanalyzed. The High-Level Trigger (HLT) runs on a farm of about 1000 commercial compute nodes and processes events that are accepted by the L1 trigger. An HLT decision (“path”) has the ability to reconstruct tracks and do a full regional unpacking of the recorded hits in a regions of the calorimeter. Each HLT path has a strict rate budget, as the total rate of the HLT is required to be less than 100 Hz. The triggers used at CMS change as the conditions change. To limit the total rate to 100 Hz as the luminosity increases, trigger paths must either increase their thresholds, or apply a “prescale.” When a prescale is applied, a fraction of events passing the trigger are thrown away randomly.

The CMS trigger is a deep subject and a complete description is beyond the scope of this thesis. A detailed description can be found in [27]. The triggers used in the analysis presented in this thesis will be briefly described. Two types of trigger selections were applied to the 2010 datasets used in this analysis. During the initial period of low luminosity running, single muon triggers were used. As the luminosity increased, the p_T threshold of the trigger was increased. In some cases, an “isolated muon” HLT trigger was required, in which a veto was applied on muons with associated energy deposits in the calorimeter. In the final period of data taking, two “cross-triggers” were used. These required the presence of both a muon and a hadronic tau decay in the event. The triggers used in this analysis in the different 2010 run periods are enumerated in Table 5.1.

The muon component of all the triggers used in this analysis is based on the “L1 seed trigger” L1_SingleMu7, which nominally selects event which contain a muon with $p_T > 7$ GeV/c. The L1 muon trigger decision is determined by the Global Muon Trigger (GMT), which combines information from the DT, CSC, and RPC sub-detectors, and is able to trigger muons up to a pseudorapidity of $|\eta| < 2.1$. Each sub-detector has a “local trigger,”

965 which can reconstruct tracks in the muon system. For the drift tubes, the Bunch Track
966 Identifiers (BTI), a custom integrated circuit, searches for aligned hits in the associated
967 DT chamber. The CSCs and RPCs employ similar strategies to detect local muon tracks.
968 The sub-detectors send the GMT the charge, p_T , η , ϕ , and a quality code of up to four
969 local muons. The measurements from the sub-detectors are combined and a final decision
970 is made by the GMT.

Chapter 3

Tau Identification: The Tau Neural Classifier

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

981 Tau leptons are unique in that they are the only type of leptons which are heavy enough
 982 to decay to hadrons. The hadronic decays compose approximately 65% of all tau decays, the
 983 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
 984 decays are typically composed of one or three charged pions and zero to two neutral pions.
 985 The neutral pions decay almost instantaneously to pairs of photons.

986 In this chapter, we describe a technique to identify hadronic tau decays. Tau decays
 987 to electrons and muons are difficult to distinguish from prompt production of electrons and
 988 muons in pp collisions. Analyses that use exclusively use the leptonic (e, μ) decays of taus
 989 typically require that the decays be of opposite flavor. With the Tau Neural Classifier, we
 990 aim to improve the discrimination of true hadronic tau decays from quark and gluon jets
 991 using a neural network approach.

995 §3.1 Geometric Tau Identification Algorithms

996 The tau identification strategies used in previously published CMS analyses are fully de-
 997 scribed in [28]. A summary of the basic methods and strategies is given here. There are
 998 two primary methods for selecting objects used to reconstruct tau leptons. The CaloTau
 999 algorithm uses tracks reconstructed by the tracker and clusters of hits in the electromag-
 1000 netic and hadronic calorimeter. The other method (PFTau) uses objects reconstructed by
 1001 the CMS particle flow algorithm, which is described in [29]. The particle flow algorithm
 1002 provides a global and unique description of every particle (charged hadron, photon, elec-
 1003 tron, etc.) in the event; measurements from sub-detectors are combined according to their
 1004 measured resolutions to improve energy and angular resolution and reduce double counting.
 1005 All of the tau identification strategies described in this thesis use the particle flow objects.

1006 Both methods typically use an “leading object” and an isolation requirement to reject
 1007 quark and gluon jet background. Quark and gluon jets are less collimated and have a higher
 1008 constituent multiplicity and softer constituent p_T spectrum than a hadronic tau decay of
 1009 the same transverse momentum. The “leading track” requirement is applied by requiring a
 1010 relatively high momentum object near the center of the jet; typically a charged track with
 1011 transverse momentum greater than 5 GeV/c within $\Delta R < 0.1$ about the center of the jet
 1012 axis. The isolation requirement exploits the collimation of true taus by defining an isolation
 1013 annulus about the kinematic center of the jet and requiring no detector activity about a
 1014 threshold in that annulus. This approach yields a misidentification rate of approximately 1%
 1015 for QCD backgrounds and a hadronic tau identification efficiency of approximately 50% [28].

1016 §3.2 Decay Mode Tau Identification: Motivation

1017 The tau identification strategy described previously can be extended by looking at the dif-
 1018 ferent hadronic decay modes of the tau individually. The dominant hadronic decays of taus
 1019 consist of a one or three charged π^\pm mesons and up to two π^0 mesons and are enumerated
 1020 in Table 1.4. The majority of these decays proceed through intermediate resonances and
 1021 each of these decay modes maps directly to a tau final state multiplicity. Each intermediate
 1022 resonance has a different invariant mass (see Figure 3.1). This implies that the problem of

hadronic tau identification can be re-framed from a global search for collimated hadrons satisfying the tau mass constraint into a ensemble of searches for single production of the different hadronic tau decay resonances. The Tau Neural Classifier algorithm implements this approach using two complimentary techniques: a method to reconstruct the decay mode and an ensemble of neural network classifiers used to identify each decay mode resonance and reject quark and gluon jets with the same final state topology.

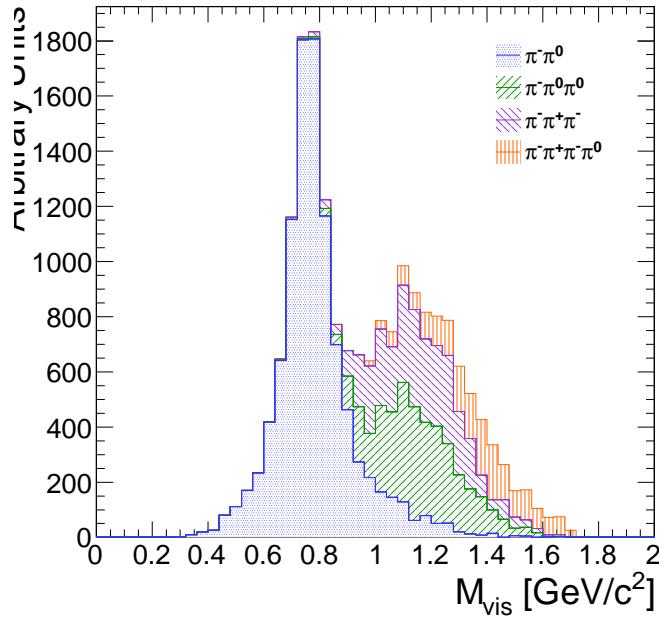


Figure 3.1: The invariant mass of the visible decay products in hadronic tau decays. The decay mode $\tau^- \rightarrow \pi^- \nu_\tau$ is omitted. The different decay modes have different invariant masses corresponding to the intermediate resonance in the decay.

§3.3 The Tau Neural Classifier

The Tau Neural Classifier (TaNC) algorithm reconstructs the decay mode of the tau-candidate and then feeds the tau-candidate to a discriminator associated to that decay mode to make the classification decision. Each discriminator therefore maps to a reconstructed decay mode in a one-to-one fashion. To optimize the discrimination for each of the different decay modes, the TaNC uses an ensemble of neural nets. Each neural net corresponds to one of the dominant hadronic decay modes of the tau lepton. These selected

hadronic decays constitute 95% of all hadronic tau decays. Tau–candidates with reconstructed decay modes not in the set of dominant hadronic modes are immediately tagged as background.

§3.3.1 Decay Mode Reconstruction

The major task in reconstructing the decay mode of the tau is determining the number of π^0 mesons produced in the decay. A π^0 meson decays almost instantaneously to a pair of photons. The photon objects are reconstructed using the particle flow algorithm [29]. The initial collection of photon objects considered to be π^0 candidates are the photons in the signal cone described by using the “shrinking–cone” tau algorithm, described in [28].

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

Photon Merging

Photons are merged into composite π^0 candidates by examining the invariant mass of all possible pairs of photons in the signal region. Only π^0 candidates (photon pairs) with a composite invariant mass less than 0.2 GeV/c are considered. The combination of the high granularity of the CMS ECAL and the particle flow algorithm provide excellent energy and angular resolution for photons; the π^0 mass peak is readily visible in the invariant mass spectrum of signal photon pairs (see figure 3.3.1). The π^0 candidates that satisfy the invariant mass requirement are ranked by the difference between the composite invariant mass of the photon pair and the invariant mass of the π^0 meson given by the PDG [20]. The

1065 best pairs are then tagged as π^0 mesons, removing lower-ranking candidate π^0 s as necessary
 1066 to ensure that no photon is included in more than one π^0 meson.

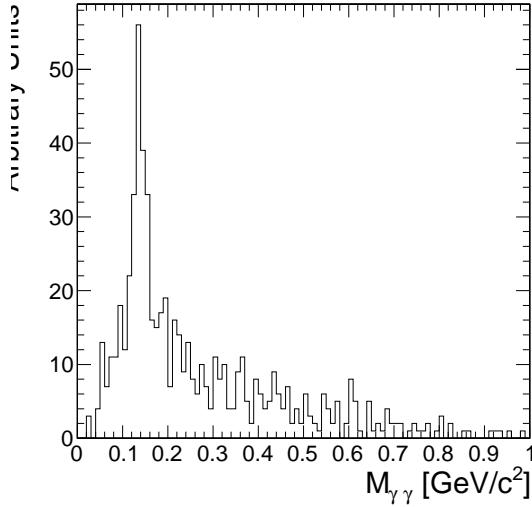


Figure 3.2: Invariant mass of the photon pair for reconstructed tau-candidates with two reconstructed photons in the signal region that are matched to generator level $\tau^- \rightarrow \pi^- \pi^0 \nu_\tau$ decays.

1067 Quality Requirements

1068 Photons from the underlying event and other reconstruction effects cause the number of
 1069 reconstructed photons to be greater than the true number of photons expected from a given
 1070 hadronic tau decay. Photons that have not been merged into a π^0 meson candidate are
 1071 recursively filtered by requiring that the fraction of the transverse momentum carried by
 1072 the lowest p_T photon be greater than 10% with respect to the entire (tracks, π^0 candidates,
 1073 and photons) tau-candidate. In the case that a photon is not merged but meets the minimum
 1074 momentum fraction requirement, it is considered a π^0 candidate. This requirement removes
 1075 extraneous photons, while minimizing the removal of single photons that correspond to a
 1076 true π^0 meson (see Figure 3.3). A mass hypothesis with the nominal [20] value of the π^0
 1077 is applied to all π^0 candidates. All objects that fail the filtering requirements are moved to
 1078 the isolation collection.

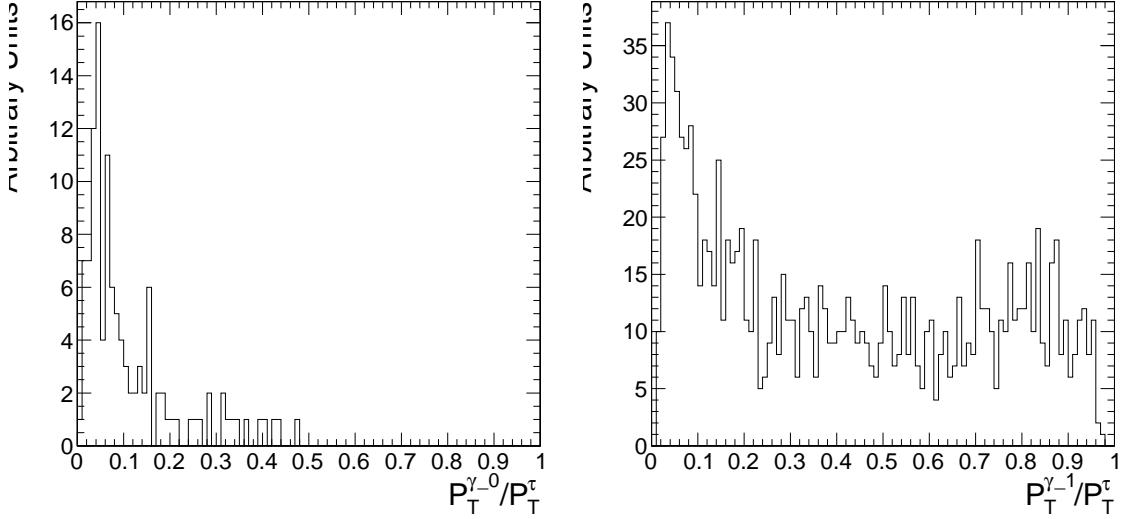


Figure 3.3: Fraction of total τ -candidate transverse momenta carried by the photon for reconstructed taus containing a single photons for two benchmark cases. On the left, the reconstructed tau-candidate is matched to generator level $\tau^- \rightarrow \pi^- \nu_\tau$ decays, for which no photon is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

1079 Performance

1080 The performance of the decay mode reconstruction can be measured for tau-candidates that
 1081 are matched to generator level hadronically decaying tau leptons by examining the correla-
 1082 tion of the reconstructed decay mode to the true decay mode determined from the Monte
 1083 Carlo generator level information. Figure 3.4 compares the decay mode reconstruction per-
 1084 formance of a naive approach where the decay mode is determined by simply counting
 1085 the number of photons to the performance of the photon merging and filtering approach
 1086 described in Section 3.3.1. The correlation for the merging and filtering algorithm is much
 1087 more diagonal, indicating higher performance. The performance is additionally presented for
 1088 comparison in tabular form in Table 3.3.1 (merging and filtering approach) and Table 3.3.1
 1089 (naive approach).

1090 The performance of the decay mode reconstruction is dependent on the transverse
 1091 momentum and η of the tau-candidate and is shown in Figure 3.5. The p_T dependence
 1092 is largely due to threshold effects; high multiplicity decay modes are suppressed at low

1093 transverse momentum as the constituents are below the minimum p_T quality requirements.
 1094 In the forward region, nuclear interactions and conversions from the increased material
 1095 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.

1096 §3.3.2 Neural Network Classification

1097 Neural Network Training

1098 The samples used to train the TaNC neural networks are typical of the signals and back-
 1099 grounds found in common physics analyses using taus. The signal-type training sample is

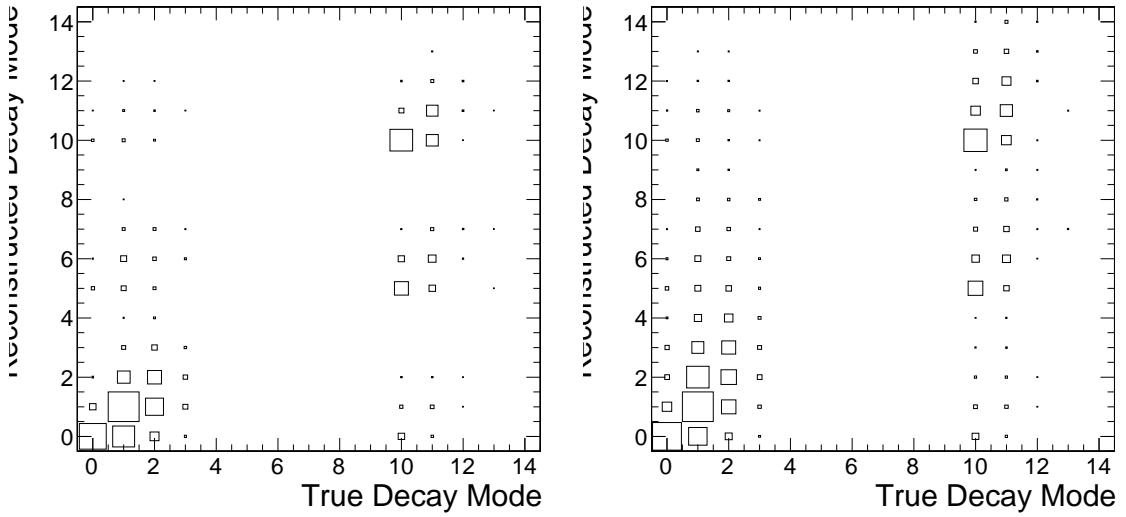


Figure 3.4: Correlations between reconstructed tau decay mode and true tau decay mode for hadronic tau decays in $Z \rightarrow \tau^+\tau^-$ events. The correlation when no photon merging or filtering is applied is shown on the right, and the correlation for the algorithm described in Section 3.3.1 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.

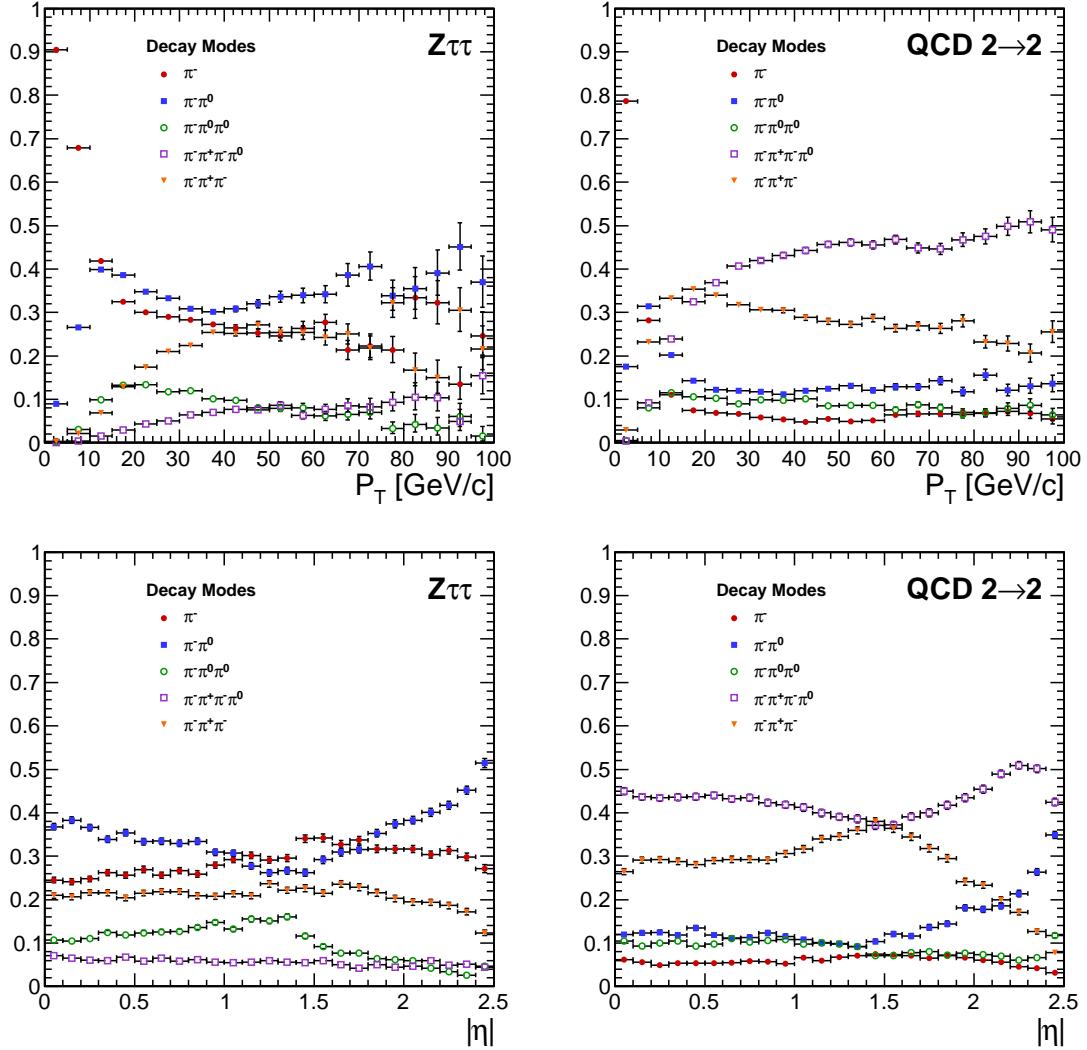


Figure 3.5: Kinematic dependence of reconstructed decay mode for tau-candidates in $Z \rightarrow \tau^+\tau^-$ (left) and QCD di-jets (right) events versus transverse momentum (top) and pseudo-rapidity (bottom). Each curve is the probability for a tau-candidate to be reconstructed with the associated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

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

The signal and background samples are split into five subsamples corresponding to each reconstructed decay mode. An additional selection is applied to each subsample by requiring a “leading pion”: either a charged hadron or gamma candidate with transverse momentum greater than 5 GeV/c. A large number of QCD training events is required as both the leading pion selection and the requirement that the decay mode match one of the dominant modes given in Table 1.4 are effective discriminants. For each subsample, 80% of the signal and background tau–candidates are used for training the neural networks, with half (40%) used as a validation sample used to ensure the neural network is not over-trained. The number of signal and background entries used for training and validation in each decay mode subsample is given in Table 3.3.2.

The remaining 20% of the signal and background samples are reserved as a statistically independent sample to evaluate the performance of the neural nets after the training is completed. The TaNC uses the Multi-layer Perceptron (MLP) neural network implementation provided by the TMVA software package, described in [30]. The MLP classifier is a feed-forward artificial neural network. There are two layers of hidden nodes and a single node in the output layer. The hyperbolic tangent function is used for the neuron activation function.

The neural networks used in the TaNC have two hidden layers and single node in the output layers. The number of nodes in the first and second hidden layers are chosen to be $N + 1$ and $2N + 1$, respectively, where N is the number of input observables for that neural network. According to the Kolmogorov’s theorem [31], any continuous function $g(x)$ defined

	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.

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

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

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

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

The neural network error E is defined [30] as

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

where N is the number of training events, $y_{ANN,i}$ is the neural network output for the i th training event, and y_i is the desired (-1 for background, 1 for signal) output the i th event.

No evidence of over-training is observed.

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

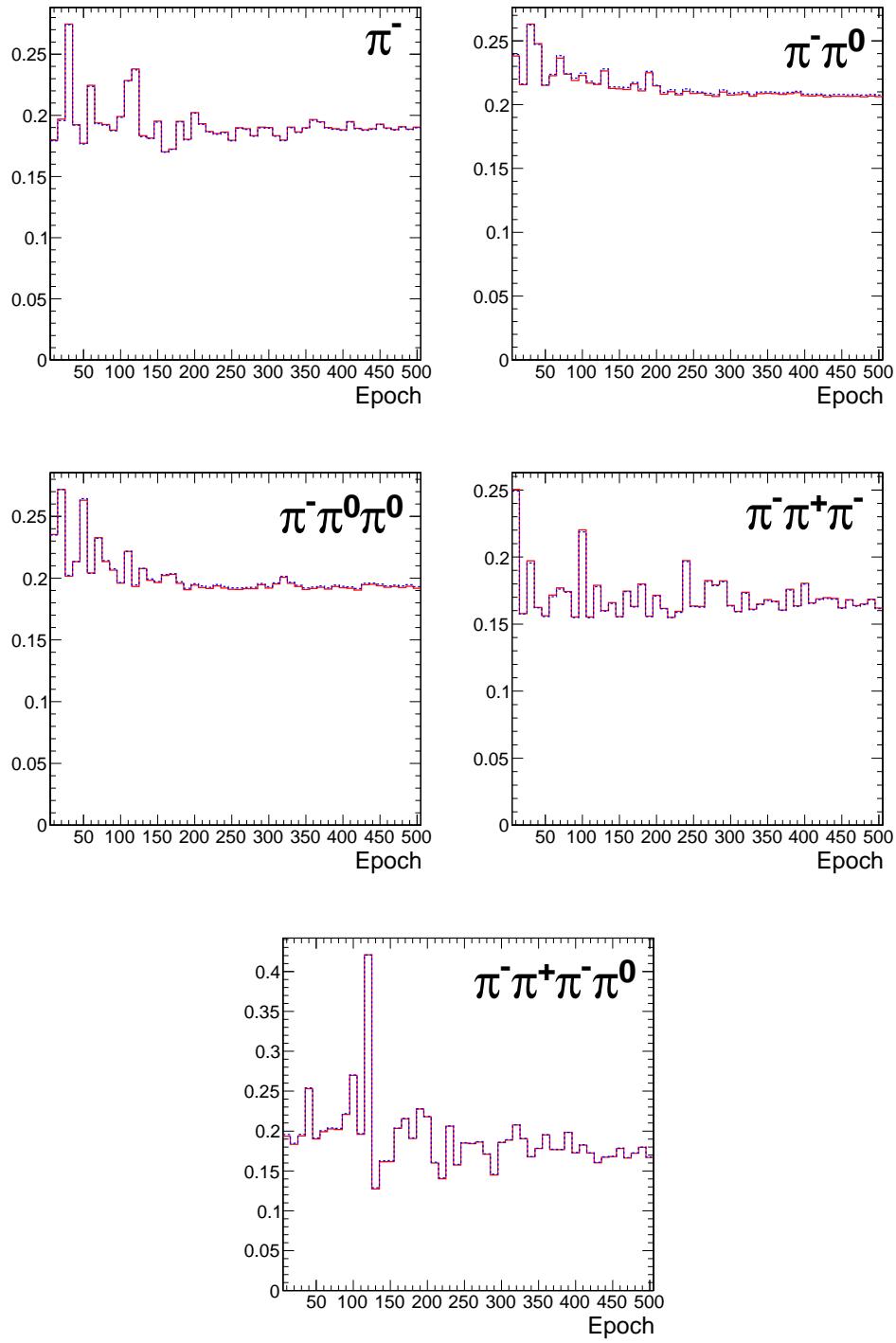


Figure 3.6: Neural network classification error for training (solid red) and testing (dashed blue) samples at ten epoch intervals over the 500 training epochs for each decay mode neural network. The vertical axis represents the classification error, defined by equation 3.2. N.B. that the choice of hyperbolic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

1136 while for background events the correlation is very small [32]. In the training signal and
 1137 background samples, there is significant discrimination power in the p_T spectrum. However,
 1138 it is desirable to eliminate any systematic dependence of the neural network output on p_T
 1139 and η , as in practice the TaNC will be presented with tau-candidates whose $p_T - \eta$ spectrum
 1140 will be analysis dependent. The dependence on p_T and η is removed by applying a p_T and
 1141 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

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

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

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

1142 where $p_{sig}(p_T, \eta)$ and $p_{bkg}(p_T, \eta)$ are the probability densities of the signal and background
 1143 samples after the “leading pion” and dominant decay mode selections. Figure 3.7 shows the
 1144 signal and background training p_T distributions before and after the weighting is applied.

1145 Discriminants

1146 Each neural network corresponds to a different decay mode topology and as such each
 1147 network uses different observables as inputs. However, many of the input observables are
 1148 used in multiple neural nets. The superset of all observables is listed and defined below.
 1149 Table 3.4 maps the input observables to their associated neural networks. In three prong
 1150 decays, the definition of the “main track” is important. The main track corresponds to the
 1151 track with charge opposite to that of the total charge of the three tracks. This distinction is
 1152 made to facilitate the use of the “Dalitz” observables, allowing identification of intermediate
 1153 resonances in three-body decays. This is motivated by the fact that the three prong decays
 1154 of the tau generally proceed through $\tau^- \rightarrow a1^- \nu_\tau \rightarrow \pi^- \rho^0 \nu_\tau \rightarrow \pi^- \pi^+ \pi^- \nu_\tau$; the oppositely
 1155 charged track can always be identified with the ρ^0 decay.

1156 **ChargedOutlierAngleN**

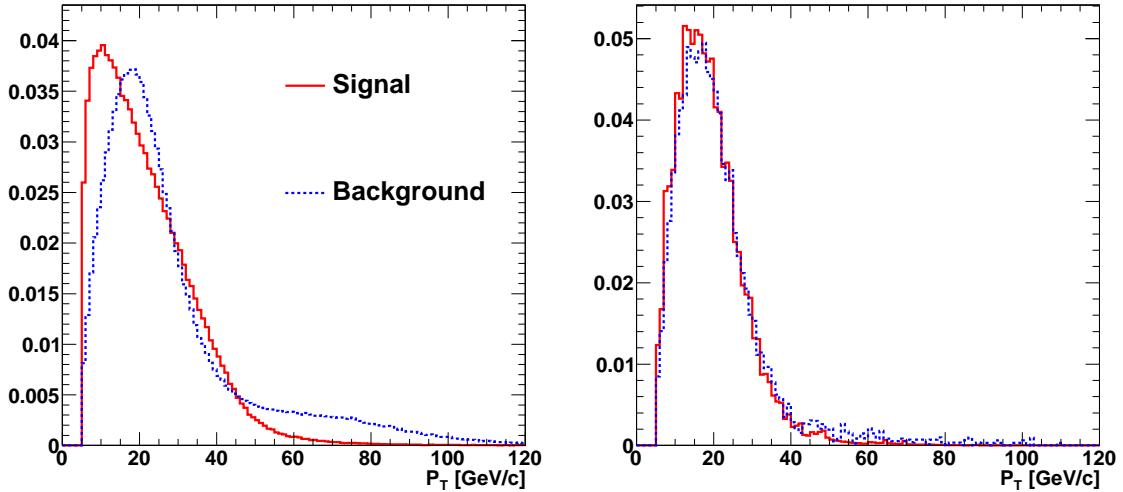


Figure 3.7: Transverse momentum spectrum of signal and background tau-candidates used in neural net training before (left) and after (right) the application of $p_T - \eta$ dependent weight function. Application of the weights lowers the training significance of tau-candidates in regions of $p_T - \eta$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

1157 ΔR between the Nth charged object (ordered by p_T) in the isolation region and the
 1158 tau-candidate momentum axis. If the number of isolation region objects is less than
 1159 N, the input is set at one.

1160 **ChargedOutlierPtN**

1161 Transverse momentum of the Nth charged object in the isolation region. If the number
 1162 of isolation region objects is less than N, the input is set at zero.

1163 **DalitzN**

1164 Invariant mass of four vector sum of the “main track” and the Nth signal region
 1165 object.

1166 **Eta**

1167 Pseudo-rapidity of the signal region objects.

1168 **InvariantMassOfSignal**

1169 Invariant mass of the composite object formed by the signal region constituents.

1170 **MainTrackAngle**

1171 ΔR between the “main track” and the composite four–vector formed by the signal
 1172 region constituents.

1173 **MainTrackPt**

1174 Transverse momentum of the “main track.”

1175 **OutlierNCharged**

1176 Number of charged objects in the isolation region.

1177 **OutlierSumPt**

1178 Sum of the transverse momentum of objects in the isolation region.

1179 **PiZeroAngleN**

1180 ΔR between the Nth π^0 object in the signal region (ordered by p_T) and the tau–
 1181 candidate momentum axis.

1182 **PiZeroPtN**

1183 Transverse momentum of the Nth π^0 object in the signal region.

1184 **TrackAngleN**

1185 ΔR between the Nth charged object in the signal region (ordered by p_T) and the
 1186 tau–candidate momentum axis, exclusive of the main track.

1187 **TrackPtN**

1188 Transverse momentum of the Nth charged object in the signal region, exclusive of the
 1189 main track.

1190 Neural Network Performance

1191 The classification power of the neural networks is unique for each of the decay modes.
 1192 The performance is determined by the relative separation of the signal and background
 1193 distributions in the parameter space of the observables used as neural network inputs. A
 1194 pathological example is the case of tau–candidates with the reconstructed decay mode of
 1195 $\tau^- \rightarrow \pi^- \nu_\tau$. If there is no isolation activity, the neural net has no handle with which it

1196 can separate the signal from the background. The neural net output for tau-candidates in
 1197 the testing sample (independent of the training and validation samples) for each of the five
 1198 decay mode classifications is shown in Figure 3.8.

1199 When a single neural network is used for classification, choosing an operating point is
 1200 relatively straightforward: the requirement on neural network output is tuned such that the
 1201 desired purity is attained. However, in the case of the TaNC, multiple neural networks are
 1202 used. Each network has a unique separation power (see Figure 3.9) and each neural network
 1203 is associated to a reconstructed decay mode that composes different relative fractions of the
 1204 signal and background tau-candidates. Therefore, a set of five numbers is required to define
 1205 an “operating point” (the signal efficiency and background misidentification rate) in the
 1206 TaNC output. All points in this five dimensional cut-space map to an absolute background
 1207 fake-rate and signal efficiency rate. Therefore there must exist a 5D “performance curve”
 1208 which for any attainable signal efficiency gives the lowest fake-rate. A direct method to
 1209 approximate the performance curve is possible using a Monte Carlo technique.

1210 The maximal performance curve can be approximated by iteratively sampling points in
 1211 the five-dimensional cut space and selecting the highest performance points. The collection
 1212 of points in the performance curve are ordered by expected fake rate. During each iteration,
 1213 the sample point is compared to the point before the potential insertion position of the
 1214 sample in the ordered collection. The sample point is inserted into the collection if it has
 1215 a higher signal identification efficiency than the point before it. The sample point is then
 1216 compared to all points in the collection after it (i.e. those with a larger fake rate); any point
 1217 with a lower signal efficiency than the sample point is removed. After the performance curve
 1218 has been determined, the set of cuts are evaluated on an independent validation sample
 1219 to ensure that the measured performance curve is not influenced by favorable statistical
 1220 fluctuations being selected by the Monte Carlo sampling. The performance curves for two
 1221 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. An artificial neural network maps a point in the space of input observables to some value of neural network output x . The neural network training error is given by Equation 3.2. A given point in the vector space

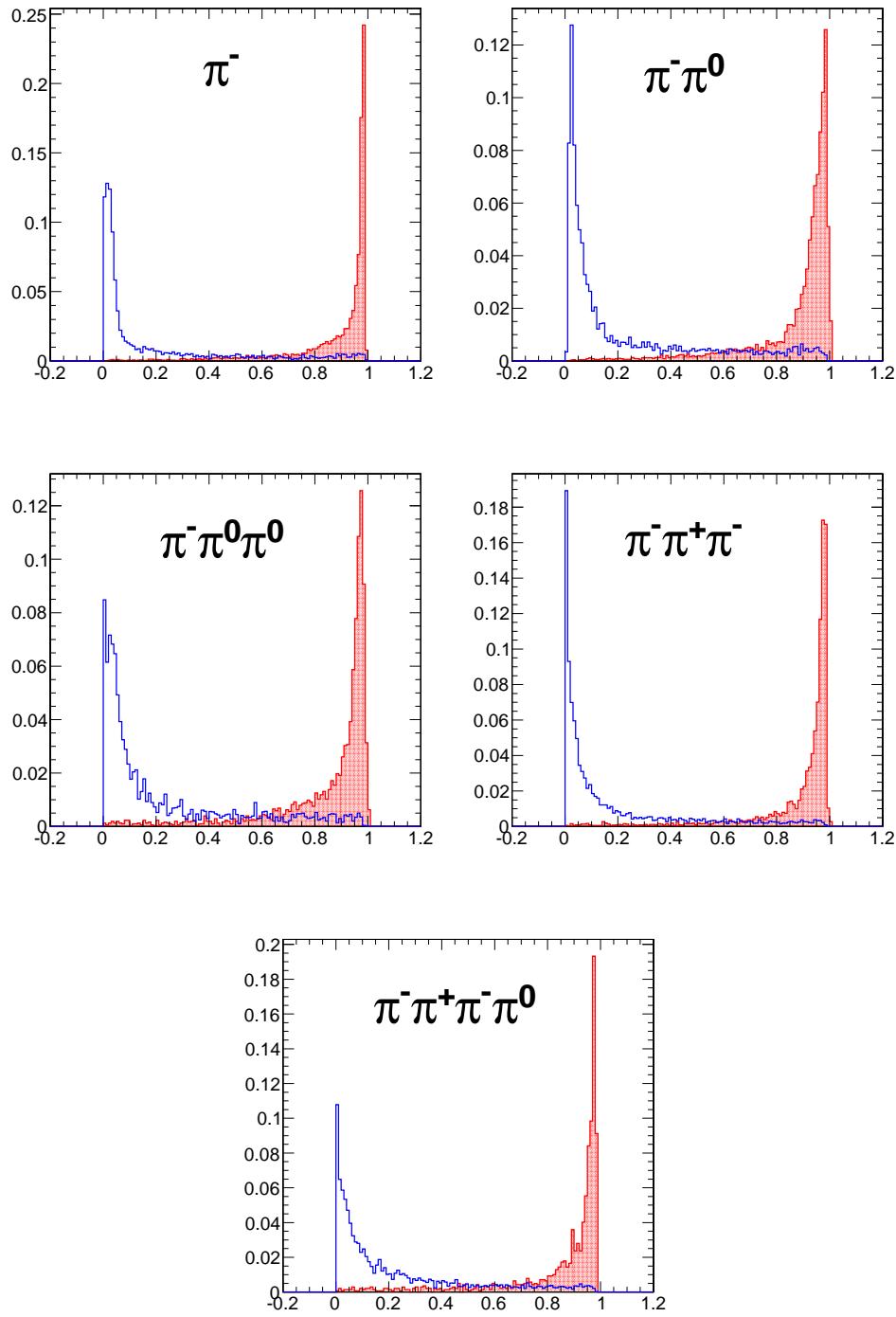


Figure 3.8: Neural network output distributions for the five reconstructed tau–candidate decay modes used in the TaNC for $Z \rightarrow \tau^+\tau^-$ events (red) and QCD di–jet events (blue).

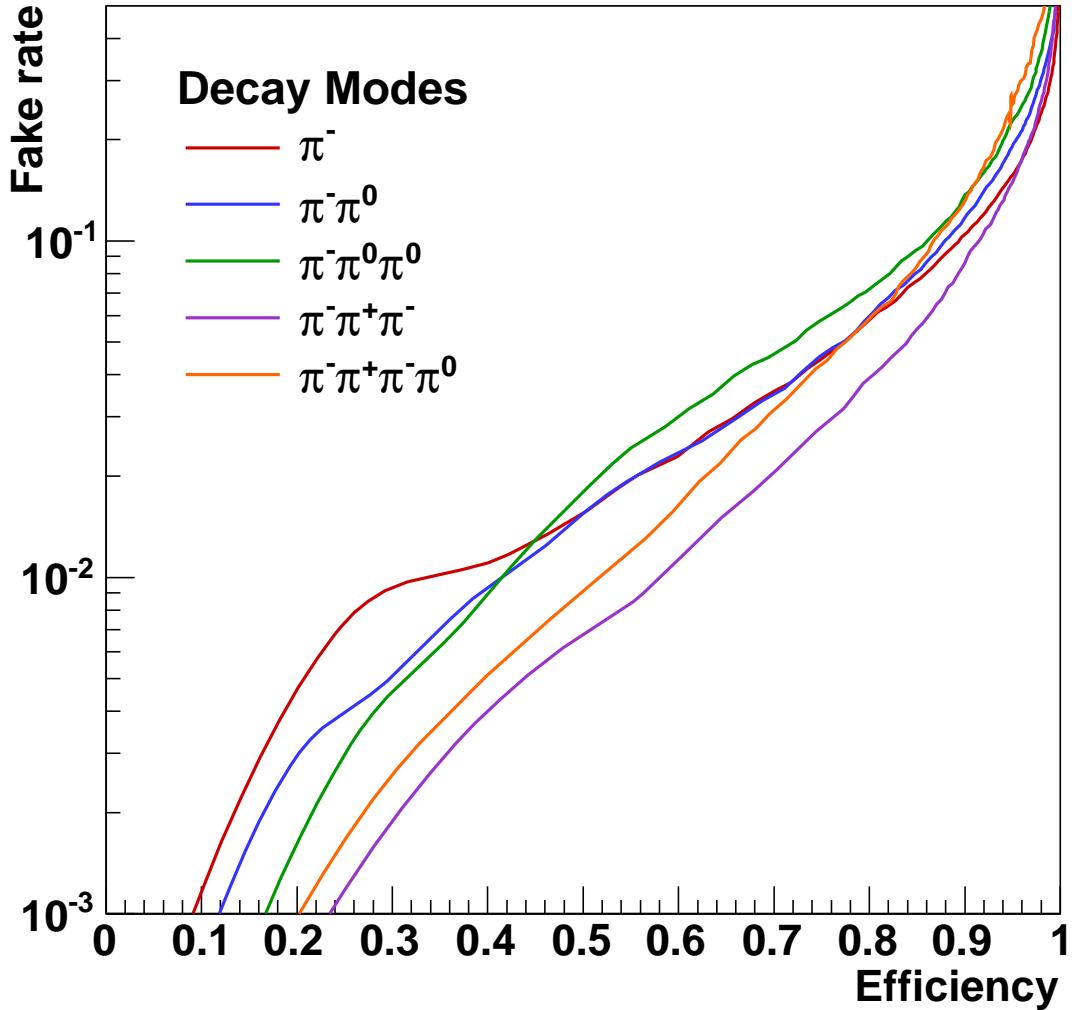


Figure 3.9: Performance curves for the five neural networks used by the TaNC for tau-candidates with transverse momentum greater than 20 GeV/c. Each curve represents the signal efficiency (on the horizontal axis) and background misidentification rate (vertical axis) for a scan of the neural network selection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

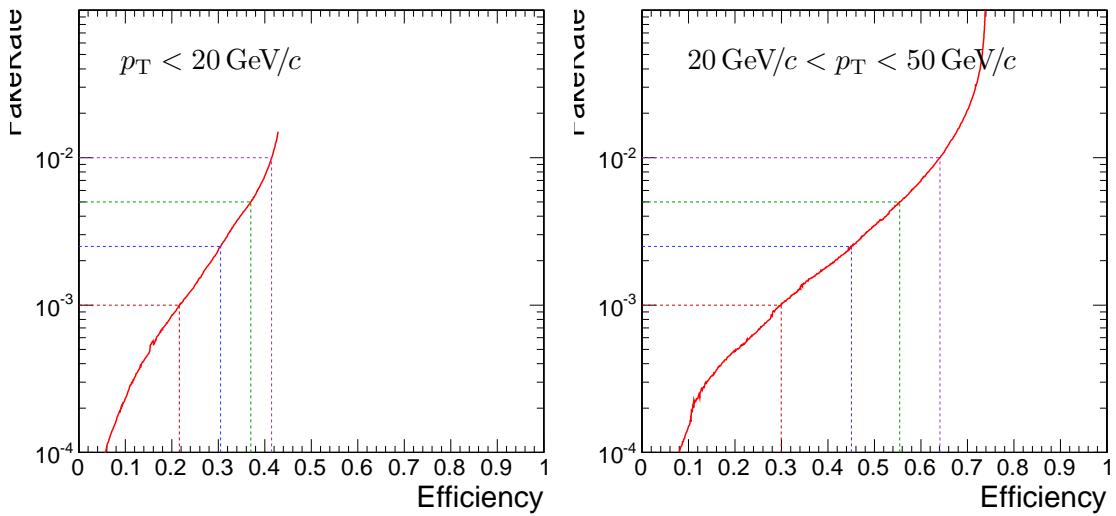


Figure 3.10: Tau Neural Classifier performance curves for tau-candidates with $p_T < 20 \text{ GeV}/c$ (left) and $20 < p_T < 50 \text{ GeV}/c$ (right). The vertical axis represents the expected fake-rate of QCD jets and the horizontal axis the expected signal efficiency for hadronic tau decays. The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

spanned by the neural network input observables (denoted as “feature space”) contributes to the neural network training error E by

$$E' = (1 - x)^2 \cdot \rho^\tau + x^2 \cdot \rho^{QCD}$$

where $\rho^\tau(\rho^{QCD})$ denotes the training sample density of the τ signal and QCD-jet background at that point in feature space.

The value x assigned by the neural network to this region in feature space should satisfy the requirement of minimal error:

$$\begin{aligned} \frac{\partial E'}{\partial x} &= 0 \\ 0 &= -2(1 - x) \cdot \rho^\tau + 2x \cdot \rho^{QCD} \\ x &= \frac{\rho^\tau}{\rho^\tau + \rho^{QCD}} \end{aligned} \tag{3.3}$$

$$\begin{aligned} \rho^{QCD} &= \frac{1}{x} - 1 \\ \frac{\rho^{QCD}}{\rho^\tau} &= \frac{1}{x} - 1 \end{aligned} \tag{3.4}$$

The ratio $\frac{\rho^{QCD}}{\rho^\tau}$ corresponds to the ratio of the normalized probability density functions of signal and background input observable distributions, i.e. $\int \rho^\tau d\vec{x} = 1$.

In the case of multiple neural networks, one can derive a formula that maps the output x_j of the neural network corresponding to decay mode j according to the “prior probabilities” $p_j^\tau(p_j^{QCD})$ for true τ lepton hadronic decays (quark and gluon jets) to pass the preselection criteria and be reconstructed with decay mode j . By substituting $\rho^s \rightarrow \rho^s p_j^s$ for $s \in \{\tau, QCD\}$ in Equation 3.3, the output x_j can be related to $p_j^\tau(p_j^{QCD})$ by

$$x'_j = \frac{\rho^\tau \cdot p_j^\tau}{\rho^\tau \cdot p_j^\tau + \rho^{QCD} \cdot p_j^{QCD}} = \frac{p_j^\tau}{p_j^\tau + \frac{\rho^{QCD}}{\rho^\tau} \cdot p_j^{QCD}} \tag{3.5}$$

Substituting Equation 3.4 into Equation 3.5 yields the transformation of the output x_j of the neural network corresponding to any selected decay mode j to a single discriminator output x'_j which for a given point on the optimal performance curve should be independent of j .

$$x'_j = \frac{p_j^\tau}{p_j^\tau + \left(\frac{1}{x_j} - 1\right) \cdot p_j^{QCD}} \tag{3.6}$$

In this manner a single number (the “transform cut”) given by Equation 3.6 can be used to specify any point on the performance curve. The training sample neural network output after the transformation has been applied is shown in Figure 3.12. The performance curve

1229 for the cut on the transformed output is nearly identical to the optimal performance curve
 1230 determined by the Monte Carlo sampling technique.

1231 The discriminator output of the TaNC algorithm is a continuous quantity, enabling
 1232 analysis specific optimization of the selection to maximize sensitivity. For the convenience
 1233 of the user, four operating point benchmark selections are provided in addition to the
 1234 continuous output. The four operating points are chosen such that for tau-candidates with
 1235 transverse momentum between 20 and 50 GeV/c, the expected QCD di-jet fake rate will be
 1236 0.1%, 0.25%, 0.50% and 1.0%, respectively.

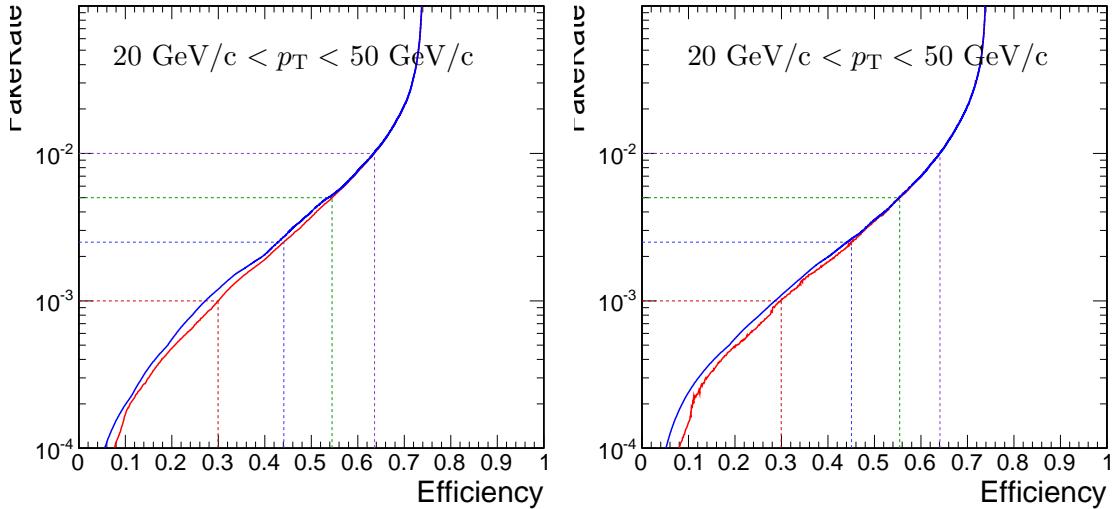


Figure 3.11: Tau Neural Classifier performance curves for tau-candidates with $20 < p_T < 50 \text{ GeV}/c$. The figure on the left compares the optimal performance curve determined by the Monte Carlo sampling method (red) to the performance curve obtained by scanning the “transform cut” (blue) defined in Equation 3.6 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.

1237 §3.4 Summary

1238 The Tau Neural classifier introduces two complimentary new techniques for tau lepton
 1239 physics at CMS: reconstruction of the hadronic tau decay mode and discrimination from
 1240 quark and gluon jets using neural networks. The decay mode reconstruction strategy pre-
 1241 sented in Section 3.3.1 significantly improves the determination of the decay mode. This

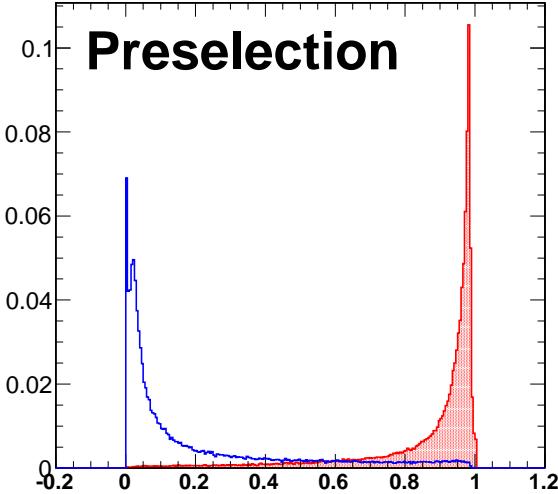


Figure 3.12: Transformed TaNC neural network output for tau-candidates with transverse momentum between 20 and 50 GeV/c that pass the pre-selection criteria. The neural network output for each tau-candidate has been transformation according to Equation 3.6. The decay mode probabilities ρ_i^{bkg} , ρ_i^{signal} are computed using the entire transverse momentum range of the sample.

1242 information has the potential to be useful in studies of tau polarization and background
1243 estimation.

1244 The Tau Neural classifier tau identification algorithm significantly improves tau dis-
1245 crimination performance compared to isolation-based approaches [28] used in previous CMS
1246 analyses. Figure 3.13 compares the performance of the “shrinking cone” isolation tau-
1247 identification algorithm [28] to the performance of the TaNC for a scan of requirements
1248 on the transformed neural network output. The signal efficiency and QCD di-jet fake rate
1249 versus tau-candidate transverse momentum and pseudo-rapidity for the four benchmark
1250 points and the isolation based tau identification are show in Figure 3.14. For tau-candidates
1251 with transverse momentum between 20 and 50 GeV/c, the TaNC operating cut can be cho-
1252 sen such that the two methods have identical signal efficiency; at this point the TaNC
1253 algorithm reduces the background fake rate by an additional factor of 3.9. This reduction
1254 in background will directly improve the significance of searches for new physics using tau
1255 leptons at CMS.

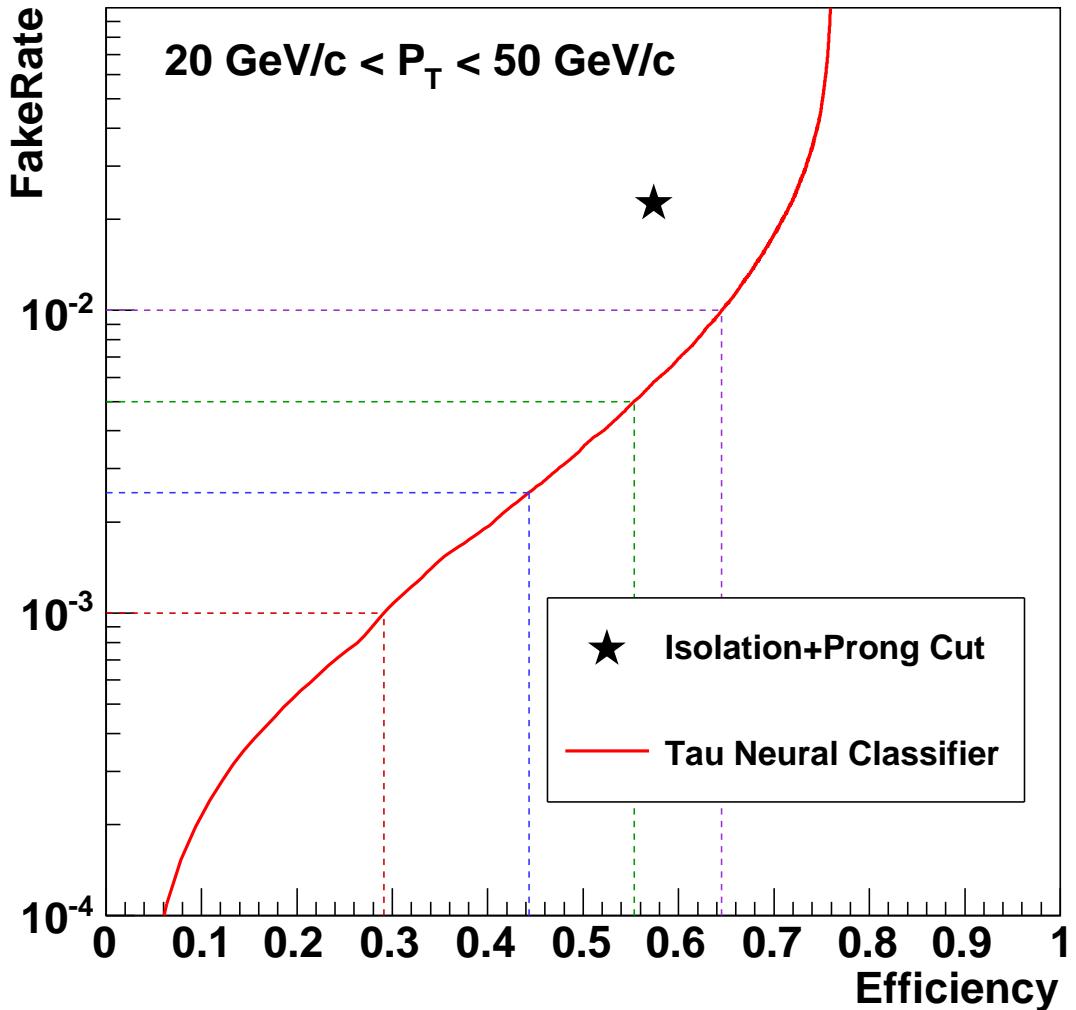


Figure 3.13: Performance curve (red) of the TaNC tau identification for various requirements on the output transformed according to Equation 3.6. The horizontal axis is the efficiency for true taus with transverse momentum between 20 and 50 GeV/c to satisfy the tau identification requirements. The vertical 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 [28] 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.

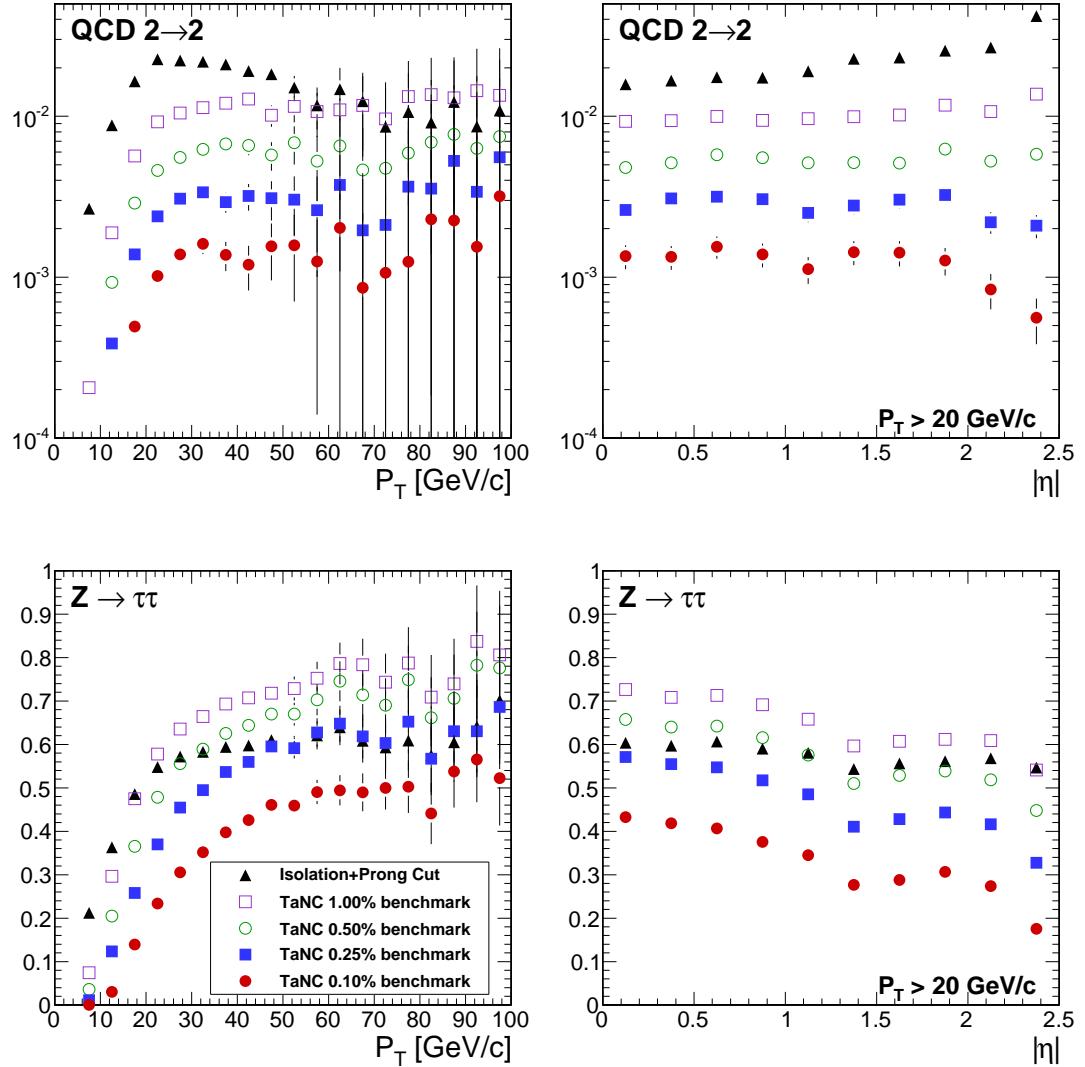


Figure 3.14: Comparison of the identification efficiency for hadronic tau decays from $Z \rightarrow \tau^+\tau^-$ decays (bottom row) and the misidentification rate for QCD di-jets (top row) versus tau-candidate transverse momentum (left) and pseudo-rapidity (right) for different tau identification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

1256 §3.5 HPS+TaNC: A Hybrid Algorithm

1257 The techniques used in the TaNC have been hybridized with techniques used by the “Hadrons
 1258 plus Strips” (HPS) algorithm [33]. The combined algorithm is referred to “Hadrons plus
 1259 Strips and Tau Neural Classifier” (HPS + TaNC) identification algorithm. The algorithm
 1260 combines the HPS methods of constructing the signal components of the tau candidate
 1261 and the discrimination methods of the TaNC algorithm. Both algorithms are based on re-
 1262 constructing individual tau lepton hadronic decay modes, which has been demonstrated to
 1263 improve the tau identification performance significantly with respect to previously used cone
 1264 isolation based algorithms [34]. The HPS + TaNC algorithm first reconstructs the hadronic
 1265 decay mode of the tau, and applies different discriminants based on the reconstructed de-
 1266 cay mode. Identification of hadronic tau decays by the HPS + TaNC algorithm proceeds in
 1267 two stages: first, the hadronic decay mode of the tau is reconstructed and then different
 1268 discriminators are applied, based on the reconstructed decay mode. In the decay mode re-
 1269 construction particular attention is paid to the reconstruction of neutral pions, which are
 1270 expected for the majority of hadronic decay modes.

1271 §3.5.1 Decay Mode Reconstruction

1272 The decay mode reconstruction algorithm is seeded by particle-flow jets reconstructed by
 1273 the anti- k_T algorithm [35]. In order to reconstruct the decay mode, the algorithm needs to
 1274 merge photon candidates into candidate π^0 mesons. The π^0 candidates are reconstructed by
 1275 two algorithms which are executed concurrently. The “combinatorial” π^0 algorithm produces
 1276 a π^0 candidate for every possible pair of photons within the jet. The “strips” algorithm clus-
 1277 ters photons strips in $\eta - \phi$. The results of both algorithms are combined and then “cleaned”,
 1278 resolving multiple hypotheses. The quality of a π^0 candidate is determined according to the
 1279 following categorical rankings:

- 1280 • The π^0 candidate is in the ECAL barrel region ($|\eta| < 1.5$) and has invariant mass
 1281 $|m_{\gamma\gamma} - m_{\pi^0}| < 0.05 \text{ GeV}/c^2$.
- 1282 • The π^0 candidate is in the ECAL endcap region ($|\eta| > 1.5$) and has invariant mass
 1283 $m_{\gamma\gamma} < 0.2 \text{ GeV}/c^2$.

- 1284 • The π^0 candidate contains two or more photons within an $\eta - \phi$ strip of size 0.05×0.20 .
- 1285 • Photons not satisfying any of the other categories are considered as unresolved π^0
- 1286 candidates in case they have $p_T > 1.0$ GeV/c.

1287 The symbol m_{π^0} denotes the nominal neutral pion mass [20]. The size of the invariant mass

1288 windows in the ECAL endcap and barrel regions is motivated by the resolution of the π^0

1289 mass (illustrated in Figure 3.15) measured during the commissioning of the particle-flow

1290 algorithm in early CMS data [36]. Multiple π^0 candidates in the same category are ranked

1291 in quality according to the difference of the reconstructed photon pair mass to the nominal

1292 π^0 mass. After the π^0 candidates are ranked, the highest ranked candidate is selected for

1293 the final collection. The photon constituents of the highest ranked candidate are removed

1294 from remaining π^0 candidates not yet selected for the final collection in order to prevent

1295 photons from entering more than one π^0 candidate. The rank of remaining π^0 candidates

1296 is reevaluated and the π^0 candidate with the next highest rank is selected for the output

collection. The process is repeated until no more π^0 candidates are remaining.

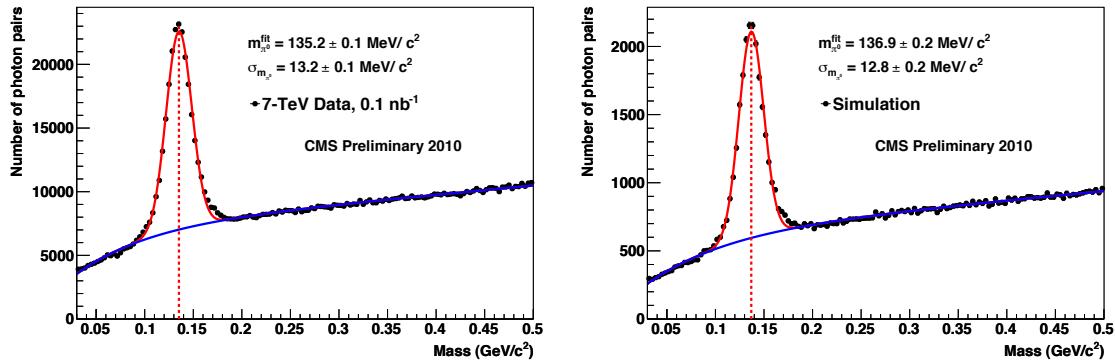


Figure 3.15: Invariant mass distribution of photon pairs reconstructed by the particle-flow in 2010 CMS minimum bias events (left), and predicted by the simulation (right). A clear resonant pick corresponding to the π_0 meson is visible above the combinatoric background. Reference: [36]

1297

1298 Once the final collection of π^0 candidates is determined, tau reconstruction in the

1299 HPS + TaNC algorithm proceeds by building tau candidates from reconstructed π^0

1300 candidates and charged hadrons reconstructed by the particle-flow algorithm. A combinatoric

1301 approach is again employed for the tau candidate building. A tau candidate hypothesis is

1302 built for every combination of jet constituents (π^0 candidates plus charged hadrons) which
 1303 has a multiplicity consistent with a hadronic tau decay. The tau candidates are ranked anal-
 1304 ogous to the ranking utilized for the π^0 reconstruction, but with the following categorical
 1305 rankings:

- 1306 • In each decay mode category, the tau candidate with the highest neural network
 1307 output is selected.
 - 1308 • The tau candidate has unit charge.
 - 1309 • The tau candidate passes the “lead pion” criteria, requiring that there is a photon or
 1310 charged pion candidate with $p_T > 5$ GeV/c.
 - 1311 • The tau candidate passes the HPS invariant mass and collimation¹ requirements.
- 1312 In case multiple tau candidates satisfy all four categorical requirements, the tau candidate
 1313 with the highest energy sum of charged and neutral pions is selected as the highest ranking
 1314 one.

1315 §3.5.2 Hadronic Tau Discrimination

1316 The final level of discrimination is performed by an ensemble of neural networks, with each
 1317 neural network corresponding to a specific decay mode, analogously to the method used
 1318 original TaNC algorithm (Section 3.3.2). The inputs of each neural network are different and
 1319 correspond to the observables (invariant mass, Dalitz masses) available for its associated
 1320 decay mode. The neural networks are trained on samples simulated $Z \rightarrow \tau^+\tau^-$ events
 1321 (“signal”) and QCD di-jet events selected in the 7 TeV data collected by CMS in 2010
 1322 (“background”). All of the tau hypothesis from a given jet reconstructed in data are used
 1323 for training. The $Z \rightarrow \tau^+\tau^-$ signal sample is generated by PYTHIA [37] which has been
 1324 interfaced to TAUOLA [38] for the purpose of generating the decays of the tau leptons using
 1325 the full matrix element information. After generation, the events passed through the “full”
 1326 GEANT [39] based simulation of the CMS detector. Only tau candidates which have been

¹The invariant mass of the signal candidates is required to be compatible with the resolution for that decay mode. The collimation selection requires the maximum ΔR between any two signal candidates to be less than $2.8/E_T$, where E_T is the total transverse energy of the signal candidates. A full description is available in [33].

1327 reconstructed in a decay mode matching the true decay mode of the tau on generator
 1328 level enter the signal training sample. The neural network implementation, network layout,
 1329 and training strategies are the same as in the original TaNC algorithm described in this
 1330 chapter. To account for differences in the input signal purity and separation power of the
 1331 neural networks between decay modes, the outputs of each neural network are transformed
 1332 according to the method described in Section 3.3.2. Multiple working-points corresponding
 1333 to different purities are provided. The “loose” working point corresponds to an approximate
 1334 fake-rate of 1%, and has slightly higher signal efficiency performance at high p_T than the
 1335 corresponding HPS isolation-only working point.

1336 §3.6 Electron and Muon Rejection

1337 Additional discriminators must be applied to prevent electrons and muons from being iden-
 1338 tified as hadronic tau decays. This is especially important for removing $Z \rightarrow e^+e^-$ and
 1339 $Z \rightarrow \mu^+\mu^-$ contributions when selecting events with two taus and requiring one of them to
 1340 decay leptonically and the other hadronically. The electron and muon discrimination algo-
 1341 rithms and performance are described in detail in [28]. A cursory overview of the techniques
 1342 used are given here. Muon removal is achieved with high purity by requiring that no track in
 1343 the signal collection of the tau candidate is matched to a segment² in the muon system. The
 1344 rejections of true electrons is more difficult. Electrons leave no signal in the muon system
 1345 and produce Bremsstrahlung photons as they travel through the magnetic field. The most
 1346 significant difference from a true hadronic tau is that an electron is not expected to deposit
 1347 any energy in the hadronic calorimeter. Electrons are thus rejected by requiring that there
 1348 is a HCAL energy deposit with a magnitude that is greater than 10% of the momentum of
 1349 the leading track in the tau.

²A track reconstructed in the DT or CSC sub-detectors.

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 3.3.2) 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.

1350 Chapter 4

1351 Mass Reconstruction: The Secondary Ver- 1352 tex Fit

1353 The dominant background in the search for a Higgs decaying to a $\tau^+\tau^-$ pair are standard
 1354 model $Z \rightarrow \tau^+\tau^-$ events. The most “natural” observable to discriminate between a Higgs
 1355 signal and the Z background is the invariant mass of the di-tau system, utilizing the fact
 1356 that the Z resonance is well known ($m_Z = 91.1876 \pm 0.0021$ GeV/ c^2) and has a narrow width
 1357 ($\Gamma_Z = 2.4952 \pm 0.0023$ GeV) [20]. The experimental complication in this approach is due to
 1358 the neutrinos produced in the tau lepton decays, which escape detection and carry away an
 1359 unmeasured amount of energy, making it difficult to reconstruct the tau lepton four-vectors.
 1360 In this chapter we give an overview of techniques used in previous literature [19, 40, 41] to
 1361 construct an observable related to the tau pair mass. We then introduce a new algorithm,
 1362 called the Secondary Vertex (SV) fit. The SVfit reconstructs the “full” tau pair mass, and
 1363 provides increased performance with respect to techniques previously used in the literature.

1364 §4.1 Existing Mass Reconstruction Algorithms

1365 The simplest observable related to the $\tau^+\tau^-$ mass is one can construct that is sensitive to
 1366 new particle content is the invariant mass of the visible (reconstructible) decay products
 1367 associated with each tau decays. This quantity, referred in this document as the “Visible
 1368 Mass,” has the advantages of simplicity and lack of exposure to systematic errors associ-
 1369 ated with the reconstruction of the E_T^{miss} . However, no attempt is made to reconstruct the
 1370 neutrinos in the event. The reconstructed mass is thus systematically smaller than mass of
 1371 the resonance which produced the tau leptons. The visible mass is typically on the order of
 1372 1/2 of the resonance mass, depending on the kinematic requirements applied to the visible
 1373 products of the tau decays.

The Collinear Approximation is a technique previously used [19] to reconstruct the full $\tau^+\tau^-$ mass. In an event with two tau decays, there are a total of six¹ unknowns associated with the missing energy: the three components of the momentum of each neutrino. The Collinear Approximation makes the assumption that the neutrinos have the same direction as their associated visible decay products. This assumption reduces the number of unknown quantities to two, corresponding to the total energy of each neutrino. These two unknowns can be solved for by using the two components of the reconstructed missing transverse energy, which in the ideal case corresponds to the transverse component of the vector sum of the two neutrino's four momentum. The characteristic equation of the Collinear Approximation is

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

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

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

The Collinear Approximation suffers from two problems. The approximation can fail (yielding unphysical negative energies for the reconstructed neutrinos) when the missing transverse energy is mis-measured. The events with unphysical solutions must be removed from the analysis, leading to a dramatic reduction in acceptance (on the order of 50% in this analysis). Improvements to the collinear approximation algorithm have recently been made which aim to recover part of the events with unphysical solutions [42]. But even with these improvements, no physical solution is still found for a large fraction of signal events. Additionally, the method is numerically sensitive when the two τ lepton are nearly back-to-back in azimuth. In these cases the $\sin(\phi_2 - \phi_1)^{-1}$ term in Equation 4.2 is very large

¹Technically, there is an extra unknown for each leptonic tau decay, which has two associated neutrinos. This is a small effect compared to the overall resolution of the collinear approximation.

1383 and small mis-measurements of the missing transverse energy can produce a large tail on
 1384 the reconstructed mass. This tail is particularly large for low-mass resonances. The large
 1385 tail for high mass is predominantly due to the fact (discussed in Section 4.4.2) that the
 1386 kinematic requirements² applied on the visible decay products preferentially selects events
 1387 where the visible decay products carry the majority of the energy of the original τ lepton,
 1388 reducing the amount of true missing energy in the event.

1389 §4.2 The Secondary Vertex Fit

1390 A novel algorithm is presented in the following, which succeeds in finding a physical solution
 1391 for every event. As an additional benefit, the new algorithm is found to improve the di-tau
 1392 invariant mass resolution, making it easier to separate the Higgs signal from the $Z \rightarrow \tau^+\tau^-$
 1393 background.

1394 The novel Secondary Vertex fit (SVfit) algorithm for di-tau invariant mass reconstruc-
 1395 tion that we present in the following utilizes a likelihood maximization to fit a $\tau^+\tau^-$ in-
 1396 variant mass hypothesis for each event. The likelihood is composed of separate terms which
 1397 represent probability densities of:

- 1398 • tau decay kinematics
- 1399 • matching between the momenta of neutrinos produced in the tau decays and the
 1400 reconstructed missing transverse momentum
- 1401 • a regularization “ p_T -balance” term which accounts for the effects on the di-tau in-
 1402 variant mass of acceptance cuts on the visible tau decay products
- 1403 • the compatibility of tau decay parameters with the position of reconstructed tracks
 1404 and the known tau lifetime of $c\tau = 87 \mu\text{m}$ [20].

1405 The likelihood is maximized as function of a set of parameters which fully describe the tau
 1406 decay.

²The 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.

1407 §4.3 Parametrization of Tau Decays

1408 The decay of a tau with visible four-momentum p_{vis} measured in the CMS detector (“laboratory”) frame can be parametrized by three variables. The invisible (neutrino) momentum
 1409
 1410 is fully determined by these parameters.

1411 The “opening-angle” θ is defined as the angle between the boost direction of the tau
 1412 lepton and the momentum vector of the visible decay products in the rest frame of the
 1413 tau. The azimuthal angle of the tau in the lab frame is denoted as $\bar{\phi}$ (we denote quantities
 1414 defined in the laboratory frame by a overline). A local coordinate system is defined such
 1415 that the \bar{z} -direction lies along the visible momentum and $\bar{\phi} = 0$ lies in the plane spanned
 1416 by the momentum vector of the visible decay products and the proton beam direction. The
 1417 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:

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

1418 Note that for hadronic decays, $m_{\nu\nu}$ is a constant of value zero, as only a single neutrino is
 1419 produced. Consequently, the magnitude of P^{vis} depends on the reconstructed mass of the
 1420 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:

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

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 the tau rest frame and laboratory frame of the visible momentum component parallel to the

tau direction:

$$\bar{p}^{vis} \cos \bar{\theta} = \gamma \beta E^{vis} + \gamma p^{vis} \cos \theta$$

$$\Rightarrow \gamma = \frac{E^{vis}[(E^{vis})^2 + (\bar{p}^{vis} \cos \bar{\theta})^2 - (p^{vis} \cos \theta)^2]^{1/2} - p^{vis} \cos \theta \bar{p}^{vis} \cos \bar{\theta}}{(E^{vis})^2 - (p^{vis} \cos \theta)^2},$$

$$E^\tau = \gamma m_\tau$$

1421 The energy of the tau lepton in the laboratory frame as function of the measured visible
1422 momentum depends on two of the three parameters only - the rest frame opening angle θ and
1423 the invariant mass $m_{\nu\nu}$ of the neutrino system. The direction of the tau lepton momentum
1424 vector is not fully determined by θ and $m_{\nu\nu}$, but is constrained to lie on the surface of a
1425 cone of opening angle $\bar{\theta}$ (given by Equation 4.4), the axis of which is given by the visible
1426 momentum vector. The tau lepton four-vector is fully determined by the addition of the
1427 third parameter $\bar{\phi}$, which describes the azimuthal angle of the tau lepton with respect to the
1428 visible momentum vector. The spatial coordinate system used is illustrated in Figure 4.1.

1429 §4.4 Likelihood for Tau Decays

The probability density functions for the tau decay kinematics are taken from the kinematics review of the PDG [20]. 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:

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

1430 In the present implementation of the SVfit algorithm, the matrix element is assumed to be
1431 constant, so that the likelihood depends on the phase-space volume of the decay only ³.

1432 §4.4.1 Likelihood for Reconstructed E_T^{miss}

1433 Momentum conservation in the plane perpendicular to the beam axis implies that the
1434 vectorial sum of the momenta of all neutrinos produced in the decay of the tau lepton pair

³The 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 [43].

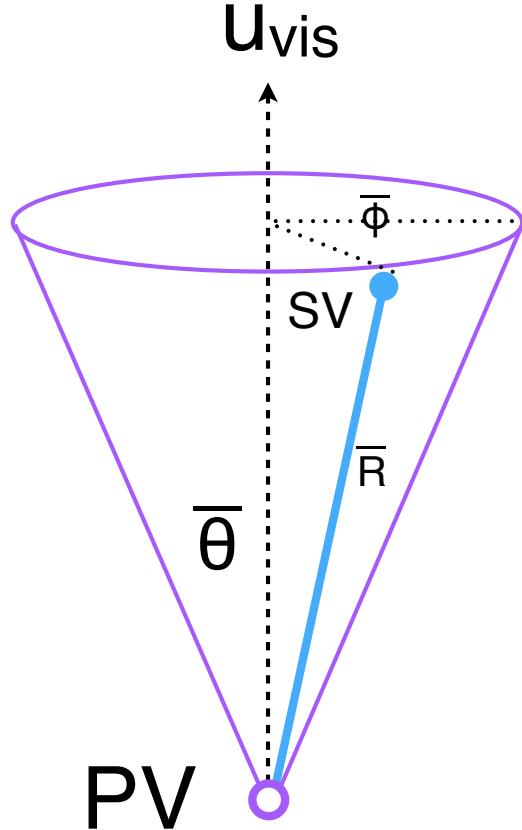


Figure 4.1: Illustration of the coordinate system used by the SVfit to describe the decays of tau leptons.

1435 matches the reconstructed missing transverse momentum. Differences are possible due to
1436 the experimental resolution and finite p_T of particles escaping detection in beam direction
1437 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 the distribution of E_T^{miss} in the Monte Carlo simulated events to match the resolution measured in data. The uncertainty on this correction factor is taken as a “shape systematic.” The treatment of this correction and its corresponding uncertainty are described in Chapters 7 and 8. 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 (“ \perp ”) direction are:

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

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

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

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

¹⁴³⁸ where q_T denotes the transverse momentum of the tau lepton pair.

¹⁴³⁹ §4.4.2 Likelihood for Tau Transverse Momentum Balance

The tau lepton transverse momentum balance likelihood term represents the probability density $p(p_T^{\tau}|M_{\tau\tau})$ of the tau transverse momentum, given the condition that the tau is produced in the decay of a resonance of mass $M_{\tau\tau}$. The likelihood is constructed by parametrizing the shape of the tau lepton p_T distribution in simulated 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

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

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 with 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

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

Numerical values of the parameters s , θ and k are determined by fitting the PDF in Equation 4.7 to the tau lepton p_T distribution in simulated 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:

$$\begin{aligned}s &= 1.8 + 0.018 \cdot M_{\tau\tau} \\k &= 2.2 + 0.0364 \cdot M_{\tau\tau} \\\theta &= 6.74 + 0.02 \cdot M_{\tau\tau} \\a &= 0.48 - 0.0007 \cdot M_{\tau\tau}.\end{aligned}$$

The motivation for the p_T -balance likelihood is to add a “regularization” term which compensates for the effect of p_T cuts applied on the visible decay products 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 $x = E_{vis}/E_\tau$. 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 and 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 $c\tau = 87 \mu\text{m}$ 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

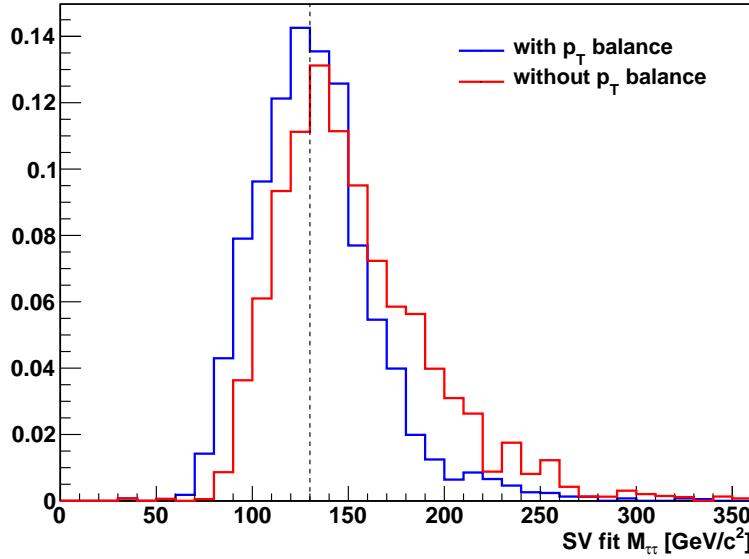


Figure 4.2: Distribution of di-tau invariant mass reconstructed by the SVfit algorithm in simulated Higgs events with $m_{A^0} = 130 \text{ GeV}/c^2$. The SVfit algorithm is run in two configurations, with (blue) and without (red) the p_T -balance likelihood term included in the fit.

1460 the $\tau^+\tau^-$ invariant mass reconstructed by the SVfit algorithm. The likelihood term based on
 1461 the secondary vertex information is based on the compatibility of the decay vertex position
 1462 with the reconstructed tracks of charged tau decay products. Perhaps surprisingly, it turns
 1463 out that the flight distance parameter R is sufficiently constrained even for tau decays into
 1464 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:

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

1465 The likelihood terms for the secondary vertex fit have been implemented in the SVfit
 1466 algorithm. In the analysis presented in this note, the decay vertex information is not used,
 1467 however, because of systematic effects arising from tracker (mis-)alignment which are not
 1468 yet fully understood.

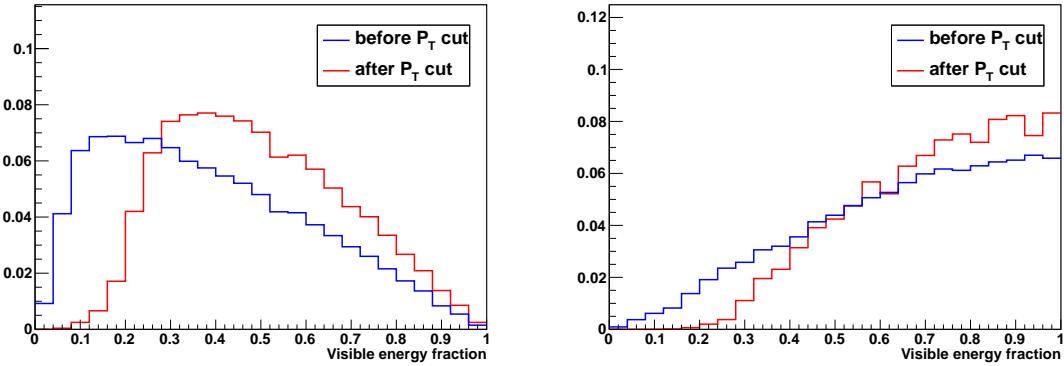


Figure 4.3: Normalized distributions of the fraction of total tau decay energy carried by the muon (left) and hadronic constituents (right) in simulated Higgs events with $m_{A^0} = 130 \text{ GeV}/c^2$. The distribution is shown before (blue) and after (red) the requirement on the p_T of the visible decay products described in Chapter 5.

1469 §4.5 Performance

1470 The tau pair mass reconstructed by the Secondary Vertex fit (“SVfit mass”) provides
 1471 the observable with the largest separation between signal Higgs events and the dominant
 1472 $Z \rightarrow \tau^+ \tau^-$ background. The mean of the SVfit mass is located at the true mass of the di-tau
 1473 pair. The SVfit algorithm has a higher acceptance and better resolution than the Collinear
 1474 Approximation algorithm. The SVfit always finds a physical solution, improving on the effi-
 1475 ciency of the collinear approximation by a factor of two. Additionally, it has a much better
 1476 resolution. The collinear approximation reconstructed mass distribution has a large tail at
 1477 high mass due to events with poorly measured E_T^{miss} . The shape of the SVfit distribution
 1478 is nearly Gaussian. The comparison is illustrated in Figure 4.5. Previous searches for Higgs
 1479 bosons decaying to tau leptons [40] have in general used the “visible mass” as the observ-
 1480 able used to search for new resonances. The SVfit method has the obvious difference that it
 1481 reconstructs the “full” tau pair mass, which is the most natural observable corresponding
 1482 to a particle decaying to tau leptons. In addition, the relative resolution⁴ of the SV fit is
 1483 superior to that of the visible mass. This feature is illustrated in Figure 4.6. In Figure 4.6,
 1484 the visible mass distribution is scaled by an arbitrary number such that the scaled mean
 1485 of the distribution matches the true mass of the tau pair (and the SVfit mass). The width

⁴We define this metric of performance as the variance of a distribution divided by its mean.

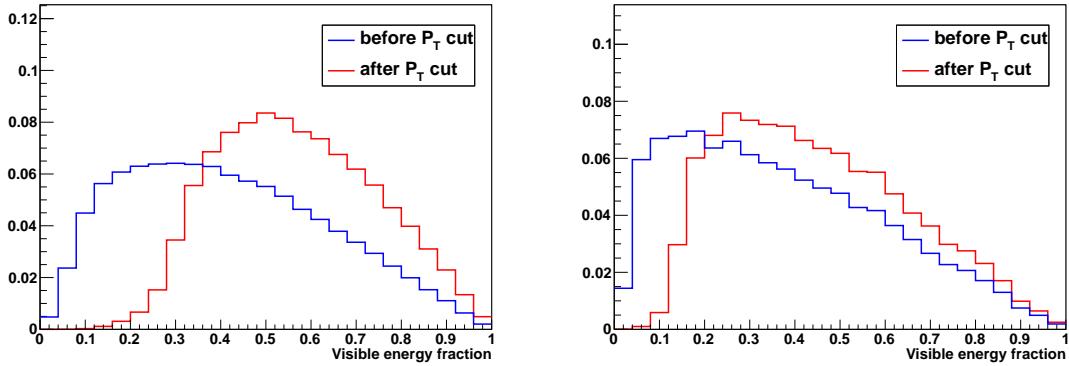


Figure 4.4: Normalized distributions of the fraction of total tau decay energy carried by the muon in simulated $Z \rightarrow \tau^+\tau^-$ (left) and Higgs events with $m_{A^0} = 200$ GeV/c^2 (right). The distribution is shown before (blue) and after (red) the requirement that the p_T of the muon be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

of the SVfit distribution is smaller than that of the scaled visible mass distribution, indicating better performance. The increase in relative resolution allows a “bump,” due to the presence of signal events, to be more easily distinguished from the $Z \rightarrow \tau^+\tau^-$ background. This increases the power of the search for the new signal.

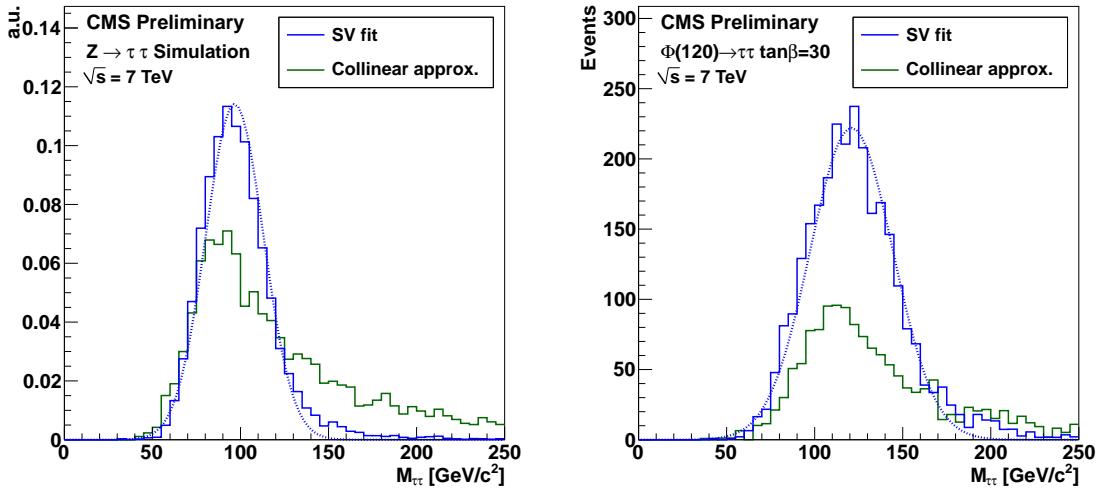


Figure 4.5: Comparison of the reconstructed tau pair mass spectrum in $Z \rightarrow \tau^+\tau^-$ (left) and MSSM $H(120) \rightarrow \tau^+\tau^-$ (right) events after the selections described in chapter 5. The mass spectrum reconstructed by the Secondary Vertex fit is shown in blue, the result of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

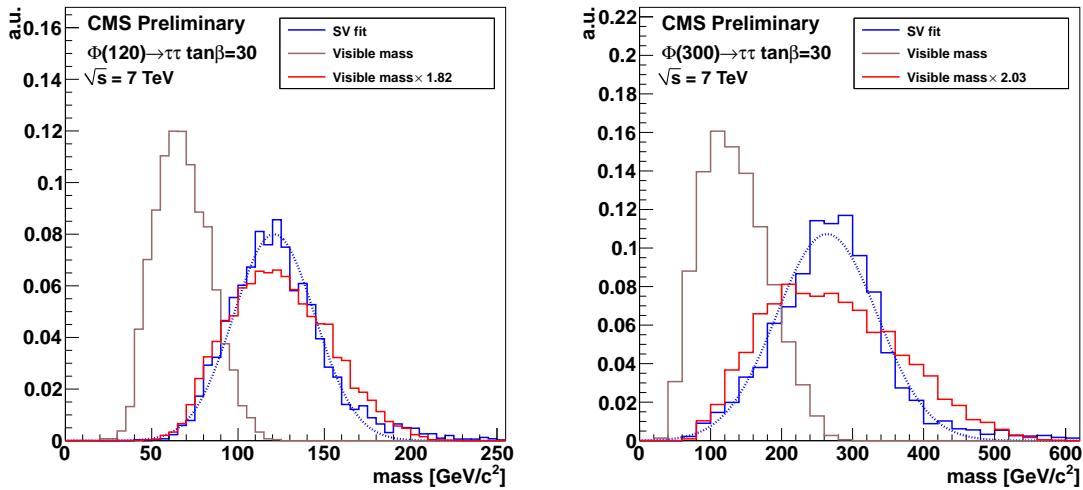


Figure 4.6: Comparison of the invariant mass of the muon and τ_{jet} (the “visible mass”) with the full $\tau^+\tau^-$ mass reconstructed by the SVfit. The spectrum is shown for two simulated MSSM Higgs samples, with $m_{A^0} = 120 \text{ GeV}/c^2$ (left), and $m_{A^0} = 200 \text{ GeV}/c^2$ (right).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

1490

Chapter 5

1491

Analysis Selections

1492 The selections applied to data in this analysis are designed to maximize the significance of
 1493 Higgs signal events in the final set of selected events. This analysis presented in this thesis
 1494 is an inclusive analysis, meaning that no preference is given to any single Higgs production
 1495 mechanism. The analysis looks specifically at the channel in which one tau decays to a
 1496 muon and the other decays to hadrons. Therefore the first step in the analysis selection is
 1497 to find High Level Trigger selection that is highly efficient for the signal and is not highly
 1498 prescaled¹. After the trigger selection, events are required to contain at least a good muon
 1499 and a good tau. Vetoos on extra leptons are applied to reduce backgrounds from di-muon
 1500 events. Finally, kinematic and charge selections on the are applied to the event to reduce
 1501 $W + \text{jets}$ and QCD backgrounds.

1502

§5.1 High Level Trigger

1503 As only data which passes the HLT is recorded, it is thus critical that an appropriate trig-
 1504 ger path is found. The events in this analysis are triggered by a combination of muon and
 1505 muon + tau-jet “cross-channel” triggers. For the muon triggers, the HLT paths with low-
 1506 est p_T thresholds are used as long as the path remained unprescaled (see Table 5.1). The
 1507 muon + tau-jet “cross-channel” trigger paths increase the trigger efficiency for events con-
 1508 taining muons of transverse momenta close to the $p_T^\mu > 15$ GeV/c cut threshold. The trigger
 1509 efficiency is measured in data via the tag-and-probe technique. Details of the muon trigger
 1510 efficiency measurement are given in Section 7.1. Monte Carlo simulated events are required

¹If a trigger has high background rates, it may exceed its rate budget with increasing luminosity. When this happens, it is generally “prescaled,” and some fraction of the events that pass this trigger are randomly thrown it away to reduce the rate. In general, it is better to use an unprescaled trigger with lower efficiency than a prescaled trigger.

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 5.1: Muon and muon + tau–jet “cross–channel” trigger paths utilized to trigger events in different data–taking periods.

1511 to pass the HLT_Mu9 trigger path. Weights are applied to simulated events to account for the
 1512 difference between the simulated HLT_Mu9 efficiency and the combined efficiency of the set
 1513 HLT_Mu9, HLT_IsoMu9, HLT_Mu11, HLT_IsoMu13, HLT_Mu15, HLT_IsoMu9_PFTau15
 1514 and HLT_Mu11_PFTau15 used to trigger the data.

1515 §5.2 Particle Identification

1516 §5.2.1 Muons

1517 Muon candidates are required to be reconstructed as global and as tracker muons, meaning
 1518 that a full track is reconstructed in the muon system and is well matched to a track in the
 1519 silicon strip and pixel trackers. Additionally, they are required to pass the “Vector Boson
 1520 Task Force” (VBF) muon identification criteria developed for the $Z \rightarrow \mu^+ \mu^-$ cross–section
 1521 measurement [44]:

- 1522 • ≥ 1 Pixel hits
- 1523 • ≥ 10 hits in silicon Pixel + Strip detectors
- 1524 • ≥ 1 hit(s) in muon system
- 1525 • ≥ 2 matched segments

- 1526 • $\chi^2/DoF < 10$ for global track fit
 1527 • transverse impact parameter of “inner” track $d_{\text{IP}} < 2$ mm with respect to beam-spot

1528 In order to reduce background contributions from muons originating from heavy quark
 1529 decays in QCD multi-jet events, muons are required to be isolated. Isolation is computed
 1530 as the p_{T} sum of charged and neutral hadrons plus photons reconstructed by the CMS
 1531 particle-flow algorithm [29] within a cone of size $\Delta R_{iso} = 0.4$ around the muon direction
 1532 divided by the muon p_{T} . The innermost region of size $\Delta R_{veto} = 0.08$ (0.05) is excluded
 1533 from the computation of the isolation p_{T} sum with respect to neutral hadrons (photons),
 1534 in order to avoid energy deposits in the electromagnetic and hadronic calorimeters which
 1535 are due to the muon to enter the sum. In order to reduce pile-up effects, particles entering
 1536 the isolation p_{T} sum are required to have transverse momenta $p_{\text{T}} > 1.0$ GeV/c. Charged
 1537 particles are additionally required to originate from the same vertex as the muon. The
 1538 muons are required to be isolated with respect to charged hadrons of $p_{\text{T}} > 1.0$ GeV/c and
 1539 photons of $p_{\text{T}} > 1.5$ GeV/c as reconstructed by the particle-flow algorithm [29] in a cone
 1540 of size $\Delta R = 0.4$ around the direction of the muon. The distribution of the muon isolation
 1541 discriminant is shown in Figure 5.1(a).

1542 §5.2.2 Hadronic Taus

1543 Hadronic decays of taus are identified by the HPS + TaNC hybrid algorithm described in
 1544 Section 3.5. The “loose” working point is used, corresponding to an expected QCD fake-rate
 1545 of about 1%. $Z \rightarrow \mu^+ \mu^-$ background contributions are largely due to muons which failed
 1546 to get reconstructed as global muons (thus failing the muon identification requirement) and
 1547 are misidentified as tau-jet candidates. These muons are typically isolated and have a large
 1548 chance to pass the hadronic tau ID discriminators. To reject these events, hadronic taus are
 1549 additionally required to pass an anti-muon veto described in Section 3.6.

1550 §5.2.3 Missing Transverse Energy

1551 The missing transverse energy $E_{\text{T}}^{\text{miss}}$, in the event is reconstructed based on the vectorial
 1552 momentum sum of particle candidates reconstructed by the particle-flow algorithm [29, 45].
 1553 In the ideal case, the $E_{\text{T}}^{\text{miss}}$ corresponds to the vector sum of the transverse components of

Background	Cross Section (pb)
QCD Heavy Flavor	84679 ³
$W \rightarrow \mu\nu + \text{jets}$	10435
$Z \rightarrow \mu\mu + \text{jets}$	1666
$t\bar{t} + \text{jets}$	158

Table 5.2: The different backgrounds to the analysis presented in this thesis that include misidentified hadronic taus.

1554 all neutrinos in the event. The E_T^{miss} resolution in simulated $Z \rightarrow \mu^+\mu^-$ events is found
 1555 to be smaller (better) than in the data. The reconstructed E_T^{miss} in the simulated events is
 1556 “smeared” by a correction factor such that the data and simulation are in agreement. The
 1557 “Z-recoil” E_T^{miss} correction procedure is described in Section 7.4.

1558 §5.3 Event Selections

1559 The selections applied to the analysis are designed to reject large fractions of the background
 1560 while maintaining a high efficiency for identifying signal Higgs events. The backgrounds
 1561 can be divided into two classifications: “fake” backgrounds, in which there is at least one
 1562 misidentified hadronic tau decay, and the irreducible $Z \rightarrow \tau^+\tau^-$ background, which cannot²
 1563 be distinguished from the potential presence of a Higgs boson of the same mass. Strategies
 1564 for dealing with the irreducible Z background will be discussed in the Chapter 9. The
 1565 different fake backgrounds, their cross section, and the basic removal strategies are outlined
 1566 in Table 5.2.

1567 Events are selected by requiring a muon of $p_T^\mu > 15$ GeV/c within $|\eta_\mu| < 2.1$ and a tau-
 1568 jet candidate of $p_T^{\tau-\text{jet}} > 20$ GeV/c within $|\eta_{\tau-\text{jet}}| < 2.3$. The η requirement on the muon
 1569 ensures that it is within the fiducial region of the muon trigger system. The η requirement
 1570 on the hadronic tau ensures it is well within the fiducial region of the tracker ($|\eta| < 2.5$)
 1571 and minimizes exposure to large QCD backgrounds in the very forward region.

²Due to the differences in spin between the Z (spin 1) and the Higgs (spin 0), in the future it may be possible to separate the two using spin correlations of the two tau decays.

1572 The muon and tau–jet candidate are required to be of opposite charge, as the Higgs is
 1573 neutral and charge is conserved. The muon is required to be pass the identification criteria
 1574 described in Section 5.2.1. The tau–jet candidate is required to pass the “loose” TaNC tau
 1575 identification discriminator.

1576 Additional event selection criteria are applied to reduce contributions of specific back-
 1577 ground processes. In order to reject $Z \rightarrow \mu^+ \mu^-$ background, a dedicated discriminator
 1578 against muons is applied [28]. Remaining di-muon background is suppressed by reject-
 1579 ing events which have a track of $p_T > 15$ GeV/c and for which the sum of energy deposits
 1580 in ECAL plus HCAL is below $0.25 \cdot P$ within a cylinder of radius 15 cm (ECAL) and
 1581 25 cm (HCAL), respectively. Contamination from $Z \rightarrow \tau^+ \tau^-$ events in which the recon-
 1582 structed tau–jet candidate is due to a $\tau \rightarrow e \nu \nu$ decay is reduced by applying a dedicated
 1583 tau ID discriminator against electrons.

The $t\bar{t}$ and $W + \text{jets}$ backgrounds are suppressed by cuts on the transverse mass of the $\mu - E_T^{\text{miss}}$ system and the P_ζ variable. The transverse mass (M_T) cut is defined as the quantity

$$M_T = p_T^\mu E_T^{\text{miss}} \sqrt{1 - \cos \Delta\phi},$$

1584 where $\Delta\phi$ is the angle between the muon and the reconstructed E_T^{miss} in the transverse
 1585 plane. The M_T quantity is much higher in events $W \rightarrow \mu\nu$ decays than in signal Higgs
 1586 events. In $W \rightarrow \mu\nu$ decays, the neutrino expected to be produced in the opposite to the
 1587 muon in azimuth. In signal events, there are three neutrinos produced, with the majority
 1588 (two) of the neutrinos being associated to the $\tau \rightarrow \mu\nu\nu$ decay. Accordingly, we expect
 1589 that the E_T^{miss} is on average collinear with the muon in signal events. The M_T distribution
 1590 immediately before the M_T cut is applied is illustrated in Figure 5.1(b) for the different
 1591 background sources and 2010 data.

1592 The P_ζ variable is another quantity with discriminant power against $W + \text{jets}$ and
 1593 $t\bar{t}$ backgrounds. The observable has been introduced in the search for $H \rightarrow \tau^+ \tau^-$ events
 1594 performed by the CDF collaboration [46]. The observable is motivated by the fact that in
 1595 $\tau^+ \tau^-$ signal events all neutrinos are produced in direction very close of the visible decay
 1596 products of one of the two tau leptons. It is therefore expected that the direction of the
 1597 missing transverse energy vector in these events point in a direction somewhere between

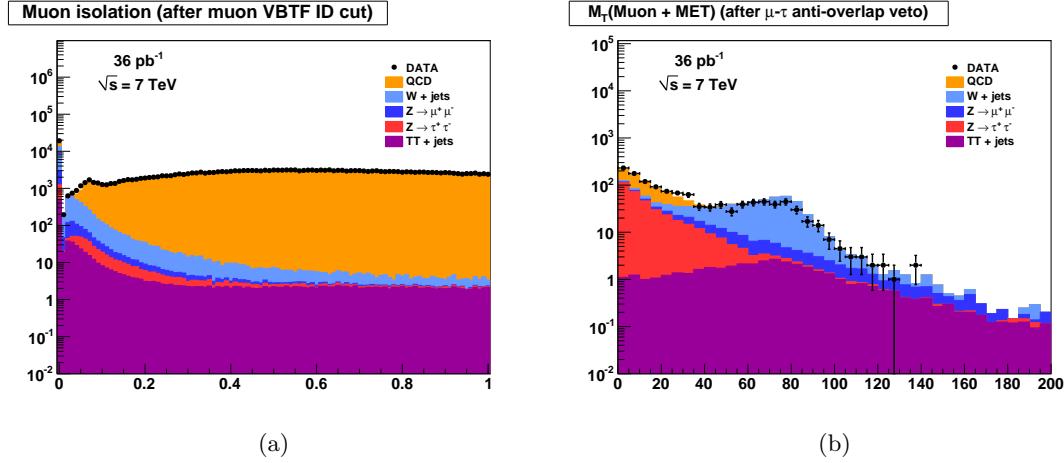


Figure 5.1: Distributions of the muon isolation (a) and M_T (b) discriminant variables. The muon isolation discriminant rejects the QCD background at a high rate. The M_T cut is designed to reject $W + \text{jets}$ and $t\bar{t} + \text{jets}$ backgrounds. The distributions shown are computed immediately before the corresponding selection is applied.

1598 the visible τ -lepton decay products. This event topology is not preferred in $W + \text{jets}$, $t\bar{t}$
 1599 and QCD background events. The observable is computed as difference of the projections:

$$\begin{aligned} P_\zeta &= \vec{P}_T^{vis_1} + \vec{P}_T^{vis_2} + E_T^{\text{miss}} \\ P_\zeta^{vis} &= \vec{P}_T^{vis_1} + \vec{P}_T^{vis_2} \end{aligned}$$

on the axis ζ bisecting the directions $\vec{P}_T^{vis_1}$ and $\vec{P}_T^{vis_2}$ of the visible τ lepton decay products in the transverse plane (see Figure 5.2(a) for an illustration). The distribution of P_ζ after the M_T selection has been applied is shown in Figure 5.2(b). The complete set of event selection criteria applied are summarized in Table 5.3.

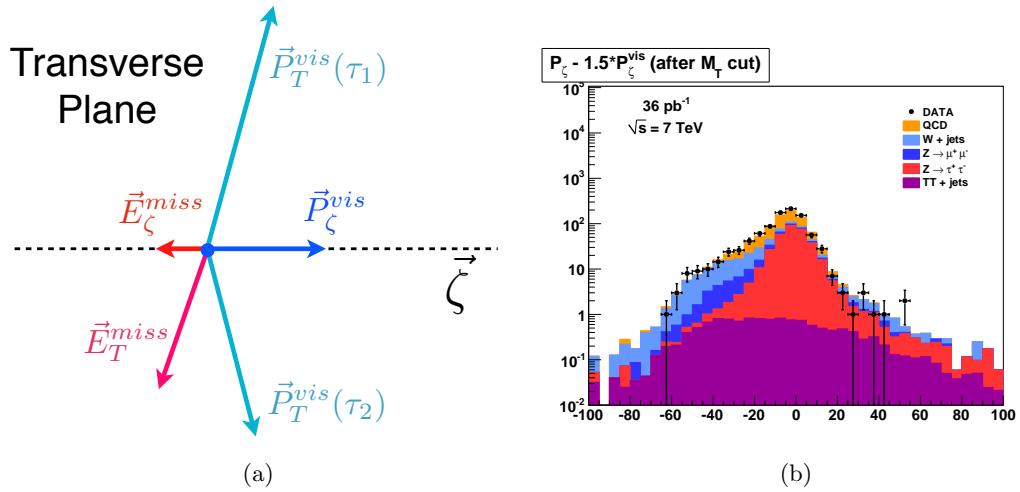


Figure 5.2: The vector quantities used in construction of the quantity $P_\zeta - 1.5 \cdot P_\zeta^{\text{vis}}$ are illustrated in Figure (a). Image credit: [46] The distribution of the P_ζ variable in the different background sources and 2010 data after the M_T cut has been applied is shown at right in (b).

Requirement	
Trigger	HLT Mu9 for MC <i>cf.</i> table 5.1 for Data
Vertex	reconstructed with beam-spot constraint:
Muon	$-24 < z_{vttx} < +24 \text{ cm}$, $ \rho < 2 \text{ cm}$, $N_{\text{DOF}} > 4$ reconstructed as global Muon with: $p_T > 15 \text{ GeV}/c$, $ \eta < 2.1$, VBF Muon ID passed, isolated within $\Delta R = 0.4$ cone with respect to charged hadrons of $p_T > 1.0 \text{ GeV}/c$ and neutral electromagnetic objects of $E_T > 1.5 \text{ GeV}$
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	$M_T(\text{Muon-MET}) < 40 \text{ GeV}$ $P_\zeta - 1.5 \cdot P_\zeta^{\text{vis}} > -20 \text{ GeV}$

Table 5.3: Event selection criteria applied to select $H \rightarrow \tau^+ \tau^- \rightarrow \mu \tau_{had}$ events.

1604

Chapter 6

1605

Data–Driven Background Estimation

1606 For the result of this analysis to be reliable, it is of paramount importance that the back-
 1607 grounds be well understood. The CMS experiment has adopted a policy that if possible, all
 1608 background processes should be measured in a “data–driven” way. By requiring that the
 1609 background comes from data, biases due to incorrectly modeling the background processes
 1610 in simulation can be minimized or eliminated. In general, the data–driven methods also have
 1611 the advantage that they are independent of the uncertainty on the integrated luminosity.
 1612 This analysis measures the backgrounds using two complementary methods, the “Template
 1613 Method” and the “Fake–rate method.” In both cases, predictions are made about back-
 1614 grounds in the signal region using measurements obtained in background enriched control
 1615 regions of the data.

1616 The Template Method fits the sum of background shape templates to the M_{vis} spec-
 1617 trum of events selected in the final analysis and is described in Section 6.3. The results of
 1618 Template Method presented in this chapter cannot be used directly in the final result of this
 1619 analysis, as the final likelihood fit (see Chapter 8) implicitly performs a template fit of the
 1620 background shapes. Nonetheless, we present the results here as an additional cross–check
 1621 of the simulation performance, and to give an overview of the method.

1622 The Fake–rate Method is based on applying probabilities for quark and gluon jets to
 1623 be misidentified as hadronic tau decays to events passing all event selection criteria except
 1624 the tau identification requirements. The probabilities with which jets fake hadronic tau
 1625 signatures are measured in data. Contrary to the Template Method, the Fake–rate Method
 1626 estimates the sum of the contributions of backgrounds that contain incorrectly identified
 1627 taus. The Fake–rate method is detailed in Section 6.2. The two methods are complementary
 1628 as the Template Method uses only information about the different visible mass distribution

1629 shapes of the backgrounds, while the Fake–rate method uses only information about the
 1630 hadronic tau fake–rate.

1631 §6.1 Background Enriched Control Regions

1632 The criteria applied to select events in the background enriched control regions for the Tem-
 1633 plate Method is based on the work described in [47]. With respect to that work, the muon
 1634 isolation criteria applied to select $Z \rightarrow \mu^+ \mu^-$, $W + \text{jets}$, $t\bar{t} + \text{jets}$ and QCD background
 1635 enriched control samples has been changed to relative isolation with respect to charged
 1636 hadrons and neutral electromagnetic objects reconstructed by the particle–flow algorithm.
 1637 The selection of the enriched backgrounds is accomplished by disabling or inverting specific
 1638 selections of Chapter 5 that were implemented to reject the given background. The selection
 1639 of control regions used to measure the fake–rates for different types of background processes
 1640 are very similar to the selections used for the Template Method. The details of the fake–rate
 1641 measurement selections may be found in [48].

1642 All control regions are selected from the 2010 CMS muon primary datasets using single
 1643 muon HLT trigger paths. The set of triggers and run–ranges used to select events in the
 1644 background enriched control samples is the same as for the analysis (see Table 5.1). The
 1645 Monte Carlo simulated events used for comparison with the control region selections are
 1646 required to pass the HLT_Mu9 trigger path and are weighted according to the description
 1647 in Chapter 7 to account for the difference in efficiency between HLT_Mu9 and the trigger
 1648 paths required to have passed in the data.

1649 QCD di–jet events containing a muon (originating from the leptonic decay of a b or
 1650 c quark) are selected by applying an anti–isolation requirement on the jet containing a
 1651 muon. $W + \text{jets}$ and $t\bar{t} + \text{jets}$ are selected by requiring an isolated muon, and inverting the
 1652 transverse mass (M_T) and P_ζ selections. Tau–jet candidates considered in the $Z \rightarrow \mu^+ \mu^-$
 1653 sample where the reconstructed tau–jet candidate is faked by a misidentified muon and in
 1654 the $t\bar{t} + \text{jets}$ control sample are required to pass the “loose” TaNC discriminator. For the
 1655 Template Method, the $Z \rightarrow \mu^+ \mu^-$ sample where the reconstructed tau–jet candidate is faked
 1656 by a misidentified quark or gluon jet, the $W + \text{jets}$ and the QCD enriched control samples
 1657 have a loose hadronic tau “preselection” applied. The tau–jet candidates are required to

1658 pass the “very loose”, but fail the “loose” TaNC discriminator. The criteria applied to select
 1659 events in the different background enriched control samples are summarized in Table 6.1.
 1660 The goal of the background enriched selection process is to select different background
 1661 processes with high purity. A highly pure background control sample improves the stability
 1662 of inferences about the signal region made using information in the enriched control region.
 1663 The purity of the control regions (estimated using simulation) are summarized in Table 6.2.

1664

1665 The number of events observed in the different control samples is compared to the
 1666 Monte Carlo expectation in Table 6.2. Except for the contribution of $Z \rightarrow \mu^+ \mu^-$ events
 1667 in which the reconstructed tau–jet candidate is due to a misidentified quark or gluon jet,
 1668 good agreement between data and Monte Carlo simulation is observed. Differences observed
 1669 between data and simulation will be accounted for as systematic uncertainties.

1670 The distributions of visible and “full” $\tau^+ \tau^-$ invariant mass reconstructed by the SVfit
 1671 algorithm (see Chapter 4) observed in the background enriched control regions is compared
 1672 to the Monte Carlo simulation in Figures 6.1 and 6.2. The template for the $W +$ jets back-
 1673 ground has been corrected for the bias on the $M_{vis}^{\mu\tau_{had}}$ shape caused by the $M_T^{\mu E_T^{\text{miss}}} <$
 1674 $50 \text{ GeV}/c^2$ and $P_\zeta - 1.5 \cdot P_\zeta^{\text{vis}} > -20 \text{ GeV}$ requirements applied in the final analysis via the
 1675 reweighting procedure described in [47]. In the $t\bar{t} +$ jets enriched control region a peak at the
 1676 Z mass is observed in data, which is not modeled by the Monte Carlo samples considered.
 1677 The peak could be due to $Z \rightarrow \mu^+ \mu^-$ events produced in association with b quarks. On
 1678 the other hand, the contribution from $t\bar{t} +$ jets events to that sample seems to be overesti-
 1679 mated. The origin of the Z mass peak merits further investigation, but overall the $t\bar{t} +$ jets
 1680 is a negligible background contribution.

1681 §6.2 The Fake–rate Method

1682 The probabilities with which quark and gluon jets get misidentified as tau–jets may be
 1683 utilized to obtain an estimate of background contributions in physics analyses. As an il-
 1684 lustrative example and in order to demonstrate the precision achievable with the method,
 1685 we introduce the method in the context of a “closure test,” using a simulated samples,

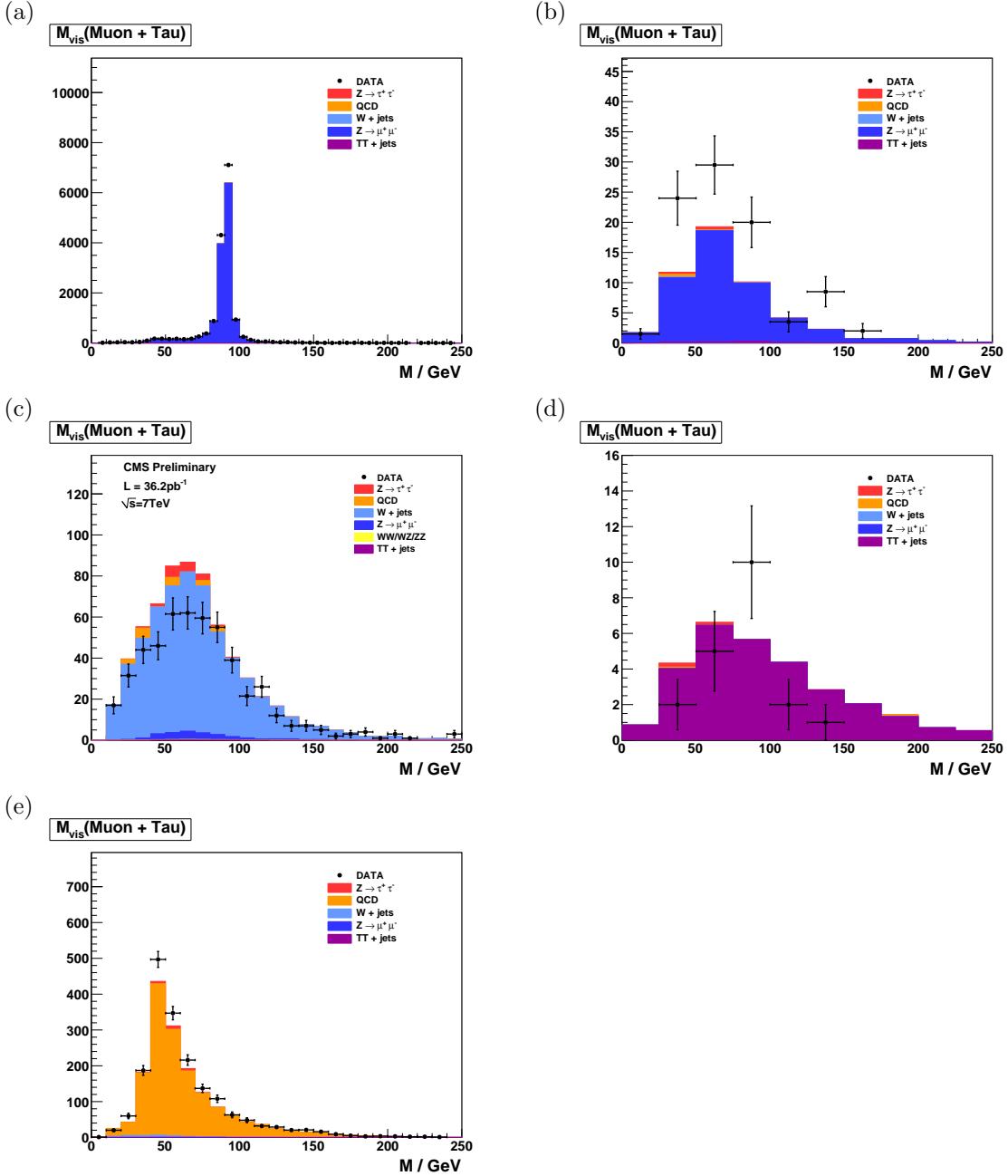


Figure 6.1: Distribution of visible mass of muon plus the tau–jet candidate reconstructed in the background enriched control samples for $Z \rightarrow \mu^+ \mu^-$ (a) and (b), $W + \text{jets}$ (c), $t\bar{t} + \text{jets}$ (d) and QCD multi-jet (e) backgrounds. In (a) reconstructed tau–jet candidates are expected to be dominantly due to misidentified muons, while in (b) they are expected to be mostly due to misidentified misidentified quark or gluon jets.

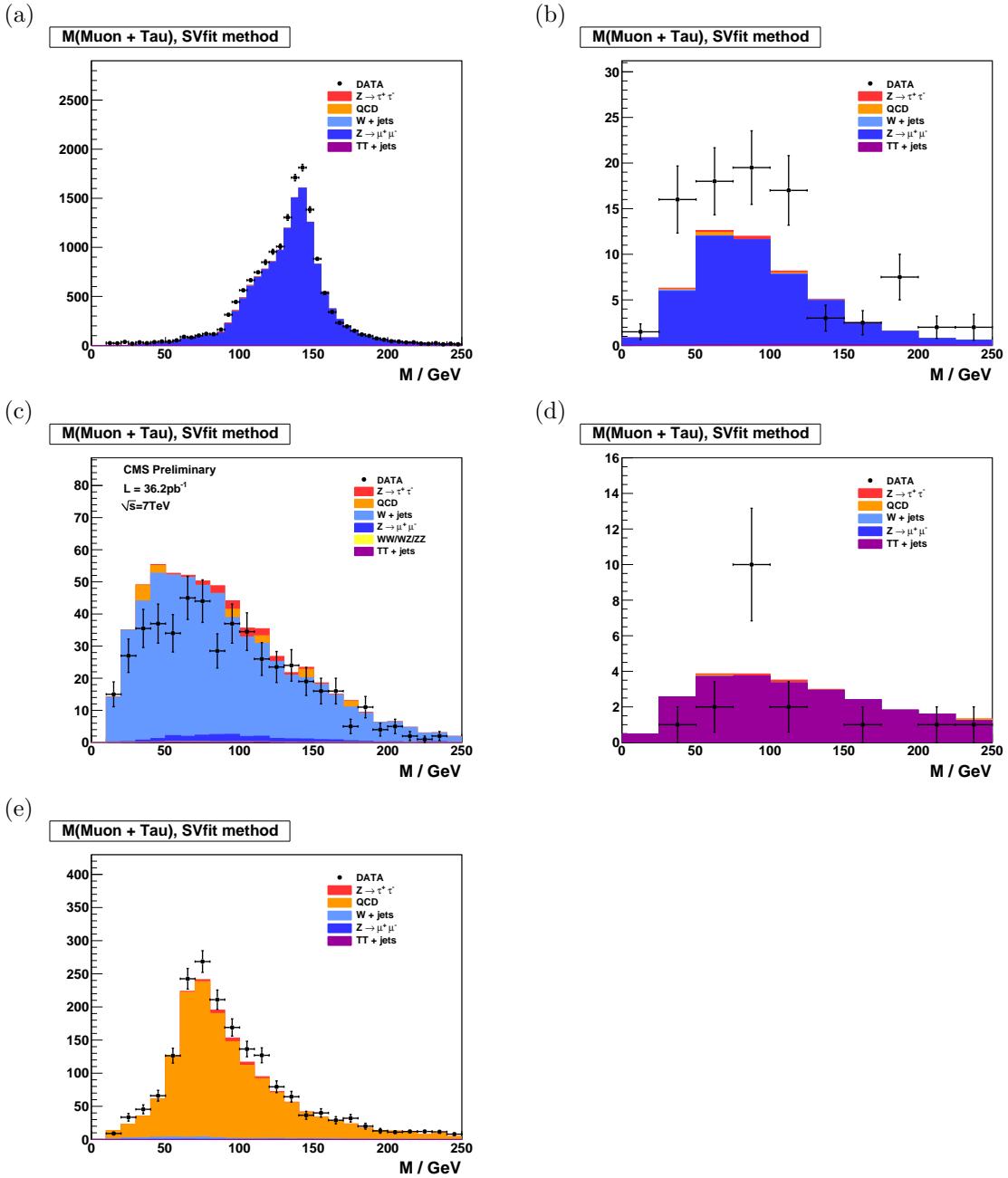


Figure 6.2: Distribution of “full” invariant mass reconstructed by the SVfit algorithm in the background enriched control samples for $Z \rightarrow \mu^+\mu^-$ (a) and (b), $W + \text{jets}$ (c), $t\bar{t} + \text{jets}$ (d) and QCD multi-jet (e) backgrounds. In (a) reconstructed tau-jet candidates are expected to be dominantly due to misidentified muons, while in (b) they are expected to be mostly due to misidentified misidentified quark or gluon jets.

1686 a simple method of computing the fake–rate, and a simpler¹ hadronic tau identification
 1687 algorithm. The closure test demonstrates that the method is self–consistent, and that the
 1688 fake–rate technique can be used to estimate the contributions of QCD, $W + \text{jets}$, $t\bar{t} + \text{jets}$
 1689 and $Z \rightarrow \mu^+\mu^-$ backgrounds. The analysis selections used in the closure test are almost
 1690 identical to the selections used in this analysis. Exact details of the selections can be found
 1691 in reference analysis [44]. The method is then extended to use fake–rates measured in data,
 1692 a multivariate method of computing the fake–rates, and the HPS + TaNC tau identification
 1693 algorithm used in this analysis.

1694 §6.2.1 Parameterization of Fake–rates

1695 Efficiencies and fake–rates of the tau identification algorithm based on requiring no tracks
 1696 of $p_T > 1$ GeV/ c and ECAL energy deposits of $p_T > 1.5$ GeV/ c reconstructed within
 1697 an “isolation cone” of size $\Delta R_{iso} = 0.5$ and outside of a “shrinking signal cone” of size
 1698 $\Delta R_{sig} = 5.0/E_T$ as it is used in the $Z \rightarrow \tau^+\tau^- \rightarrow \mu + \tau$ -jet analysis [44] are displayed
 1699 in Figure 6.3. In order to account for the visible p_T and η dependence, we parametrize
 1700 the fake–rates in bins of transverse momentum and pseudo–rapidity. As we will show in
 1701 Section 6.2.3, the parametrization of the fake–rates by p_T and η makes it possible to not
 1702 only estimate the total number of background events contributing to physics analyses, but
 1703 to model the distributions of kinematic observables with a precision that is sufficient to
 1704 extract information on the background shape.

We add a third quantity, the E_T -weighted jet–width R_{jet} , to the parametrization in
 order to account for differences between the fake–rates of quark and gluon jets, which on
 average have differing widths and different fake–rates. The jet width quantity R_{jet} is defined
 as

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

1705 where $E(\eta^2)$, $E(\phi^2)$ is the second η , ϕ moment of the jet constituents, weighted by con-
 1706 stituent transverse energy. Analyses performed by the CDF collaboration [40, 49, 46] found
 1707 that systematic uncertainties on background estimates obtained from the fake–rate method
 1708 are reduced in case differences between quark and gluon jets are accounted for in this way.

¹The closure test uses the “shrinking cone” tau identification algorithm, which is described briefly in Section 3.1. A full description can be found in [28].

1709 §6.2.2 Measurement of Fake–rates

Efficiencies and fake–rates are obtained by counting the fraction of tau–jet candidates passing all tau identification cuts and discriminators in a given bin² of p_T^{jet} , η_{jet} and R_{jet} :

$$P_{fr} \left(p_T^{jet}, \eta_{jet}, R_{jet} \right) := \frac{N_{jets} \left(p_T^{jet}, \eta_{jet}, R_{jet} | \text{tau ID passed} \right)}{N_{jets} \left(p_T^{jet}, \eta_{jet}, R_{jet} | \text{preselection passed} \right)} \quad (6.1)$$

1710 The pre–selection in the denominator of Equation 6.1 in general refers to p_T and η cuts,
 1711 which are applied with thresholds matching those applied on the final analysis level, but
 1712 may include loose tau identification criteria (which may be applied e.g. already during event
 1713 skimming). It is critical that the selection used in the denominator be identical to that of
 1714 the final analysis to ensure the fake–rates are not biased by different selections.

1715 Different sets of fake–rates are determined for the highest p_T and for the second highest
 1716 p_T jet in QCD di–jet events, for jets in a QCD event sample enriched by the contribution
 1717 of heavy quarks and gluons by requiring the presence of a muon reconstructed in the final
 1718 state, and for jets in “electroweak” events selected by requiring a W boson in the final state.

1719 §6.2.3 Application of Fake–rates

1720 Knowledge of the tau identification efficiencies and fake–rates as function of the parameters
 1721 p_T^{jet} , η_{jet} and R_{jet} as defined by Equation 6.1 is utilized to obtain an estimate for the con-
 1722 tributions of background processes to physics analyses involving tau lepton hadronic decays
 1723 in the final state. The basic idea is to replace tau identification cuts and discriminators by
 1724 appropriately chosen weights.

1725 Application of the fake–rate technique consists of two stages. The first stage consists of
 1726 loosening the tau identification cuts and discriminators and applying only the preselection
 1727 requirements defined by the denominator of Equation 6.1, in order to obtain an event
 1728 sample dominated by contributions of background processes. After disabling the selections
 1729 on hadronic tau identification, the relative contributions of the backgrounds are expected
 1730 to increase by the inverse of the (average) fake–rate, typically by a factor $\mathcal{O}(100)$. In the
 1731 second stage, weights are applied to all events in the background dominated control sample,
 1732 according to the probabilities $P_{fr} \left(p_T^{jet}, \eta_{jet}, R_{jet} \right)$ for jets to fake the signature of a hadronic

²The example presented in the closure tests bins the fake–rate calculation in bins of the parameterization variables. In Section 6.2.6 we describe a more robust multivariate method to compute the fake–rates.

1733 tau decay. After application of the weights, an estimate for the total number of background
 1734 events passing the tau identification cuts and discriminators and thus contributing to the
 1735 final analysis sample is obtained.

1736 The fake–rate technique works best if all background contributions to the analysis arise
 1737 from misidentification of quark and gluon jets as hadronic tau decays. Corrections to the
 1738 estimate obtained from the fake–rate technique are needed in case of background processes
 1739 contributing to the final analysis sample which either produce genuine tau leptons in the
 1740 final state (e.g. $t\bar{t} + \text{jets}$) or in which tau–jet candidates are due to misidentified electrons
 1741 or muons (e.g. $Z \rightarrow \mu^+\mu^-$, $Z \rightarrow e^+e^-$), as the latter may fake signatures of hadronic tau
 1742 decays with very different probabilities than quark and gluon jets.

1743 In the “simple” fake–rate method described in detail in the next section, the correc-
 1744 tions are taken from Monte Carlo simulations. Corrections based on Monte Carlo are needed
 1745 also to compensate for signal contributions to the background dominated control sample.
 1746 An alternative to Monte Carlo based corrections is to utilize additional information con-
 1747 tained in the background dominated control sample. The modified version is described in
 1748 Section 6.2.5. It has been used to estimate background contributions in searches for Higgs
 1749 boson production with subsequent decays into tau lepton pairs performed by the CDF col-
 1750 laboration in TeVatron Run II data [40, 49, 46]. We will refer to the modified version as
 1751 “CDF–type” method in the following.

1752 §6.2.4 “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:

$$w_{jet}^{simple} \left(p_T^{jet}, \eta_{jet}, R_{jet} \right) := P_{fr} \left(p_T^{jet}, \eta_{jet}, R_{jet} \right) \quad (6.2)$$

1753 These weights are applied to all jets in the background dominated control sample which
 1754 pass the preselection defined by the denominator of Equation 6.1. Note that the weights
 1755 defined by Equation 6.2 can be used to estimate the contributions of background processes
 1756 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 6.2 over all tau-jet candidates in the event which pass the preselection:

$$W_{\text{event}}^{\text{simple}} := \sum w_{\text{jet}}^{\text{simple}} \quad (6.3)$$

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 $Z \rightarrow \tau^+ \tau^- \rightarrow \mu + \tau\text{-jet}$ analysis). Per-particle weights need to be computed for such “composite particles”, depending on $p_{\text{T}}^{\text{jet}}$, η_{jet} , R_{jet} of its tau-jet candidate constituent, according to:

$$w_{\text{comp-part}}^{\text{simple}}(p_{\text{T}}^{\text{jet}}, \eta_{\text{jet}}, R_{\text{jet}}) := w_{\text{jet}}^{\text{simple}}(p_{\text{T}}^{\text{jet}}, \eta_{\text{jet}}, R_{\text{jet}}) \quad (6.4)$$

Different estimates are obtained for the fake-rate probabilities determined for the high-
1757 est and second highest p_{T} jet in QCD di-jet events, jets in a muon enriched QCD sample
1758 and jets in $W + \text{jets}$ events. The arithmetic average of the four estimates of the closure test
1759 together with the difference between the computed average and the minimum/maximum
1760 value is given in Table 6.3.

We take the average value as “best” estimate of the background contribution and
1762 the difference between the average and the minimum/maximum estimate as its systematic
1763 uncertainty. We obtain a value of $\mathcal{O}(15\%)$ for the systematic uncertainty and find that
1764 the true sum of QCD, $W + \text{jets}$, $t\bar{t} + \text{jets}$ and $Z \rightarrow \mu^+ \mu^-$ background contributions agrees
1765 well with the “best” estimate obtained by the fake-rate method within the systematic
1766 uncertainty.
1767

Note that the estimate for the sum of background contributions which one obtains
1768 in case one applies the “simple” fake-rate weights defined by Equation 6.3 to a back-
1769 ground dominated control sample selected in data is likely to overestimate the true value
1770 of background contributions by a significant amount. The reason is that contributions of
1771 $Z \rightarrow \tau^+ \tau^-$ events with true taus are non-negligible. In fact, genuine tau contributions to
1772

1773 the background dominated control sample are expected to be 14.9% and since the per-
 1774 jet weights computed by Equation 6.2 are larger on average in signal than in background
 1775 events, the signal contribution increases by the weighting and amounts to 37.1% of the sum
 1776 of event weights computed by Equation 6.3 and given in Table 6.3.

1777 The contribution of the $Z \rightarrow \tau^+ \tau^-$ signal needs to be determined by Monte Carlo
 1778 simulation and subtracted from the estimate obtained by applying the “simple” fake–rate
 1779 method to data, in order to get an unbiased estimate of the true background contributions.

1780 §6.2.5 “CDF–type” Weights

1781 Instead of subtracting from the estimate obtained for the sum of background contributions a
 1782 correction determined by Monte Carlo simulation, the genuine tau contribution contribution
 1783 to the background dominated event sample selected in data can be corrected for by adjusting
 1784 the weights, based solely on information contained in the analyzed data sample, avoiding
 1785 the need to rely on Monte Carlo based corrections.

1786 In the “CDF–type” method, additional information, namely whether or not tau–jet
 1787 candidates pass or fail the tau identification cuts and discriminators, is drawn from the data.
 1788 The desired cancellation of signal contributions is achieved by assigning negative weights to
 1789 those tau–jet candidates which pass all tau identification cuts and discriminators, i.e. to a
 1790 fair fraction of genuine hadronic tau decays, but to a small fraction of quark and gluon jets
 1791 only. The small reduction of the background estimate by negative weights assigned to quark
 1792 and gluon jets is accounted for by a small increase of the positive weights assigned to those
 1793 tau–jet candidates for which at least one of the tau identification cuts or discriminators
 1794 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:

$$w_{jet}^{CDF}(p_T^{jet}, \eta_{jet}, R_{jet}) := \begin{cases} \frac{P_{fr}(p_T^{jet}, \eta_{jet}, R_{jet}) \cdot \varepsilon(p_T^{jet}, \eta_{jet}, R_{jet})}{\varepsilon(p_T^{jet}, \eta_{jet}, R_{jet}) - P_{fr}(p_T^{jet}, \eta_{jet}, R_{jet})} & \text{all tau ID passed} \\ \frac{P_{fr}(p_T^{jet}, \eta_{jet}, R_{jet}) \cdot (1 - \varepsilon(p_T^{jet}, \eta_{jet}, R_{jet}))}{\varepsilon(p_T^{jet}, \eta_{jet}, R_{jet}) - P_{fr}(p_T^{jet}, \eta_{jet}, R_{jet})} & \text{otherwise} \end{cases} \quad (6.5)$$

1795 For the derivation of Equation 6.5 for the “CDF–type” weights assigned to tau–jet candi-
 1796 dates, we will use the following notation: Let n_τ (n_{QCD}) denote the total number of tau–jets
 1797 (quark and gluon jets) in a certain bin of transverse momentum p_T^{jet} , pseudo–rapidity η_{jet}

1798 and jet-width R_{jet} and n_τ^{sel} (n_{QCD}^{sel}) denote the number of tau-jets (quark and gluon jets)
 1799 in that bin which pass all tau identification cuts and discriminators. By definition of the
 1800 tau identification efficiency $\varepsilon := \varepsilon(p_T^{jet}, \eta_{jet}, R_{jet})$ and fake-rate $f := f(p_T^{jet}, \eta_{jet}, R_{jet})$:

$$\begin{aligned} n_\tau^{sel} &= \varepsilon \cdot n_\tau \\ n_{QCD}^{sel} &= f \cdot n_{QCD}. \end{aligned} \quad (6.6)$$

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:

$$w_{passed} \cdot f \cdot n_{QCD} + w_{failed} \cdot (1 - f) \cdot n_{QCD} \equiv n_{QCD}^{sel} = f \cdot n_{QCD} \quad (6.7)$$

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:

$$w_{passed} = -\frac{1 - \varepsilon}{\varepsilon} \cdot w_{failed}, \quad (6.8)$$

1801 associating the two types of weights. By inserting relation 6.8 into Equation 6.7 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

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

1802 which matches exactly Equation 6.5 for the ‘‘CDF-type’’ weights applied to tau-jet candidates given in Section 6.2.5.

1804 Event weights and the weights assigned to ‘‘composite particles’’ are computed in the
 1805 same way as for the ‘‘simple’’ weights, based on the weights assigned to the tau-jet candidates:

$$\begin{aligned} W_{event}^{CDF} &:= \Sigma w_{jet}^{CDF} \\ w_{comp-part}^{CDF}(p_T^{jet}, \eta_{jet}, R_{jet}) &:= w_{jet}^{CDF}(p_T^{jet}, \eta_{jet}, R_{jet}), \end{aligned} \quad (6.10)$$

1807 where the sums extend over all jets in the background dominated control sample which pass
 1808 the preselection defined by the denominator of Equation 6.1.

1809 The effect of the negative weights to compensate the positive weights in case the “CDF–
 1810 type” fake–rate method is applied to signal events containing genuine hadronic tau decays is
 1811 shown in Table 6.4 and illustrated in Figure 6.4. As expected, positive and negative weights
 1812 do indeed cancel in the statistical average.

1813 Figures 6.5, 6.6 and 6.7 demonstrate that an unbiased estimate of the background
 1814 contribution by the “CDF–type” weights is maintained. Overall, the estimates obtained
 1815 are in good agreement with the contributions expected for different background processes,
 1816 indicating that the adjustment of negative and positive weights works as expected for the
 1817 background as well.

1818 Results obtained by the “CDF–type” fake–rate method in the closure test are sum-
 1819 marized in Table 6.4, in which the total number of background events estimated by Equa-
 1820 tion 6.10 is compared to the true background contributions. The “best” estimate of the
 1821 background contribution obtained from the “CDF–type” method is again taken as the
 1822 arithmetic average of the estimates obtained by applying the fake–rate probabilities for
 1823 the highest and second highest p_T jet in QCD di–jet events, jets in a muon enriched QCD
 1824 sample and jets in $W + \text{jets}$ events. Systematic uncertainties are taken from the difference
 1825 between the computed average value and the minimum/maximum estimate. We obtain a
 1826 value of $\mathcal{O}(15 - 20\%)$ for the systematic uncertainty of the “CDF–type” method, slightly
 1827 higher than the systematic uncertainty obtained for the “simple” method. The small in-
 1828 crease of systematic uncertainties is in agreement with our expectation for fluctuations of
 1829 the jet–weights in case weights of negative and positive sign are used.

1830 §6.2.6 k–Nearest Neighbor Fake–rate Calculation

1831 For the fake–rate method to give correct results, care must be taken that the measured
 1832 fake–rate is well defined in all of the regions of phase space where it will be used. In the
 1833 closure test described above, the computation of the fake–rate was accomplished by binning
 1834 the numerator (tau ID passed) and denominator (tau ID passed and failed) distributions
 1835 in the three dimensions of the parameterizations. This method has the disadvantage that
 1836 the determination of the optimal binning is extremely difficult to determine, and that any

1837 bins with no entries in the denominator distribution caused the fake–rate to be undefined
1838 in those regions.

To overcome these problems, the fake–rate parameterization is implemented by adapting a multivariate technique known as a k –Nearest Neighbor classifier (k NN). A k NN classifier is typically used to classify events by populating (“training”) an n –dimensional space with signal and background events. The probability for a given point x in the space to be “signal–like” is determined by finding the k nearest neighbors and computing the ratio

$$p_{sig} = \frac{n_{sig}}{n_{sig} + n_{bkg}}, \quad (6.11)$$

1839 where n_{sig} , n_{bkg} are the observed number of signal and background events, respectively. By
1840 construction, $k = n_{sig} + n_{bkg}$. The principle of operation is illustrated in Figure 6.8.

1841 The classification feature of a k NN can be trivially adapted to parameterize an efficiency
1842 (fake–rate) such that it is defined everywhere. Examining the form of Equation 6.11, it is
1843 clear that by replacing n_{sig} with n_{passed} and n_{bkg} with n_{failed} , p_{sig} is equivalent to the tau–
1844 fake rate. We thus “train” the k NN with tau–candidates which pass the tau identification as
1845 signal events and those which fail as background events. The resulting classifier is a function
1846 which returns the expected fake–rate for any point in the space of the parameterization.
1847 The choice of k must be optimized. When k is low, the small number of neighbors causes
1848 large counting fluctuations in the fake rate. If k is too large, the k NN effectively averages
1849 over a large area of the space of the variables³. For the training statistics available in the
1850 2010 data, $k = 20$ is found to be the optimal choice.

1851 §6.2.7 Results of Fake–Rate Background Estimation

1852 An independent estimate of the background contributions to the analysis presented in this
1853 thesis is obtained by applying the fake–rate method in a manner analogous to the closure
1854 test. Fake–rates in QCD multi–jet events (light quark enriched sample), QCD events con-
1855 taining muons (heavy quark and gluon enriched sample) and $W +$ jets events are measured
1856 in data [34, 48] and applied to events which pass all the event selection criteria listed in
1857 Table 5.3, with the exceptions of

³In the limit $k \rightarrow \inf$, the k NN output reduces to a single number. In this extreme case, all information about the dependence of the fake–rate on the variables is lost.

- 1858 • the “loose” HPS + TaNC discriminator, and
 1859 • the requirement that the tau have unit charge.

1860 No assumption is made on the composition of $Z \rightarrow \mu^+ \mu^-$, $W + \text{jets}$, $t\bar{t} + \text{jets}$ and QCD
 1861 backgrounds contributing to the event sample selected by the analysis. Differences between
 1862 fake–rates obtained for QCD multi–jet, QCD muon enriched and $W + \text{jets}$ background
 1863 events are attributed as systematic uncertainties of the fake–rate method. Per jet and per
 1864 event weights have been computed by the “simple” and “CDF-type” weights as described in
 1865 the closure test and the results are found to be compatible within statistical and systematic
 1866 uncertainties. In the following, we present results for “CDF-type” weights. The “CDF-type”
 1867 weights have the advantage that the background estimate obtained does not change, whether
 1868 there is MSSM Higgs $\rightarrow \tau^+ \tau^-$ signal present in the data or not.

1869 Tau identification efficiencies need to be known when using “CDF-type” weights. Ded-
 1870 icated studies have checked the tau identification efficiencies in data [48]. Statistical and
 1871 systematic uncertainties of these studies are still sizeable at present, on the order of 20–30%.
 1872 No indication has been found, however, that the Monte Carlo simulation does not correctly
 1873 model hadronic tau decays in data. For the purpose of computing fake–rate weights via
 1874 the “CDF-type” method, tau identification efficiencies are taken from the Monte Carlo
 1875 simulation of hadronic tau decays in $Z \rightarrow \tau^+ \tau^-$ events. Systematic uncertainties on the
 1876 background estimate obtained by the fake–rate method are determined by varying the tau
 1877 identification efficiencies by $\pm 30\%$ relative to the value obtained from the Monte Carlo
 1878 simulation.

1879 The results of applying the fake–rate method to the mu + tau channel are summarized
 1880 in Table 6.5. The background prediction has been corrected for the expected⁴ contribution
 1881 of $13.1^{+2.8}_{-0.6}$ events from $Z \rightarrow \mu^+ \mu^-$ background events in which the reconstructed tau–jet is
 1882 due to a misidentified muon. The obtained estimate is in good agreement with the Monte
 1883 Carlo expectation.

1884 As an additional cross–check of the method, a sample of events containing a muon
 1885 plus a tau–jet of like–sign charge is selected in data and compared to the background

⁴The contribution of $Z \rightarrow \mu^+ \mu^-$ is estimated using a simulated sample.

1886 prediction obtained by applying the fake–rate method to the like–sign sample. The like–sign
 1887 sample is expected to be dominated by the contributions of $W + \text{jets}$ and QCD background
 1888 processes and allows to verify the fake–rate method in a practically signal free event sample.
 1889 The background estimate obtained by the fake–rate method is compared to the number of
 1890 events observed in the like–sign data sample in Table 6.6. The number of events expected
 1891 in the like–sign control sample from Monte Carlo simulation is indicated in the caption. All
 1892 numbers are in good agreement.

1893 The fake–rate method does not only allow to estimate the total number of background
 1894 events, but allows to model the distributions of background processes as well. The capability
 1895 to model distributions is illustrated in Figure 6.9, which shows good agreement between the
 1896 distributions observed in the like-sign data sample and the predictions obtained by the
 1897 fake–rate method for the distributions of muon plus tau–jet visible mass and of the “full”
 1898 invariant mass reconstructed by the SVfit algorithm.

Fixme: THIS
*IS FROM
 THE HPS
 NOTE!*

1899 §6.3 Template Method

1900 Shape templates for the $\mu + \tau_{had}$ visible mass M_{vis} are obtained from data, using a set of
 1901 dedicated control regions which are chosen to select a high purity sample of one particular
 1902 background process each. The number of events selected in each control region and com-
 1903 parisons to the predictions from Monte Carlo simulations are summarized in Table 6.2. The
 1904 template M_{vis} shapes obtained from data in the background enriched control regions are
 1905 compared to the signal region shapes obtained by Monte Carlo simulation in Figure 6.10.
 1906 The M_{vis} spectrum observed in the final analysis is fitted to the sum of these templates. Es-
 1907 timates for background yields are obtained from the normalization factor of each template,
 1908 determined by the fit. Further details of the method can be found in [47] and [50].

1909 The TaNC (Section 3.3, [51]) discriminators used in [50] are replaced by the correspond-
 1910 ing discriminators of the HPS + TaNC algorithm (Section 3.5, [33]). The $Z/\gamma^* \rightarrow \tau^+\tau^-$ sig-
 1911 nal shape is obtained via the $Z/\gamma^* \rightarrow \mu^+\mu^-$ embedding technique [52]. The $\mu + \tau_{had}$ visible
 1912 mass spectrum observed in the final analysis is compared to the sum of template shapes
 1913 scaled by the normalization factors determined by the fit in Figure 6.11. The corresponding
 1914 estimates for background contributions are summarized in Table 6.7.

Requirement	$Z \rightarrow \mu^+ \mu^-$			Enriched background process		
	Muon fake	Jet fake	$W + \text{jets}$	$t\bar{t} + \text{jets}$	$W + \text{jets}$	QCD
Muon rel. iso.	< 0.15	< 0.1	< 0.1	< 0.1	< 0.1	> 0.10 $\&$ < 0.30
Muon Track IP	-	-	-	-	-	-
Tau TaNC discr.	-	1	1	medium passed	1	1
Tau 1 3-Prong	-	-	-	-	-	-
Charge(Tau) = ± 1	-	-	-	-	-	-
Tau μ -Veto	inverted	applied	applied	applied	applied	applied
Charge(Muon+Tau)	applied	-	-	-	-	-
M_T (Muon-MET)	-	< 40 GeV	-	-	-	< 40 GeV
$P_\zeta - 1.5 \cdot P_\zeta^{vis}$	> -20 GeV	-	-	-	-	> -20 GeV
global Muons	< 2	-	< 2	< 2	< 2	< 2
central Jet Veto	-	-	2	-	-	-
b-Tagging	-	-	-	3	-	-

¹ vloose passed $\&$ loose failed ² no Jets of $E_T > 20$ GeV within $|\eta| < 2.1$ (other than the τ -jet candidate) ³ min. two Jets of $E_T > 40$ GeV, at least one of which was at least of which with “TrackCountingHighEff” discriminator > 2.5

Table 6.1: Criteria to select events in different background enriched control samples. Hyphens indicate event selection criteria which are not applied.

Enriched Selection	Data	Σ_{SM}	Contribution from				Purity
			$Z \rightarrow \tau^+ \tau^-$	$Z \rightarrow \mu^+ \mu^-$	$W + \text{jets}$	$t\bar{t} + \text{jets}$	
$Z \rightarrow \mu^+ \mu^-$							
Muon fake	15156	17109.8	331.6	16586.6	55.1	80.4	35.0
Jet fake	85	62.7	2.5	55.5	0.5	1.4	2.4
$W + \text{jets}$	514	642.4	17.9	22.9	581.7	0.8	16.7
$t\bar{t} + \text{jets}$	26	39.7	0.7	< 0.1	0.6	38.4	< 1.0
QCD	2510	2571.8	16.6	0.8	9.3	1.6	2543.4
							98.9%

Table 6.2: Number of events observed in the different background enriched control samples compared to Monte Carlo expectations. Σ_{SM} denotes the sum of $Z \rightarrow \tau^+ \tau^-$, $Z \rightarrow \mu^+ \mu^-$, $W + \text{jets}$ and QCD processes. The expected purity of each control sample is computed as the ratio of contribution of the enriched process to Σ_{SM} .

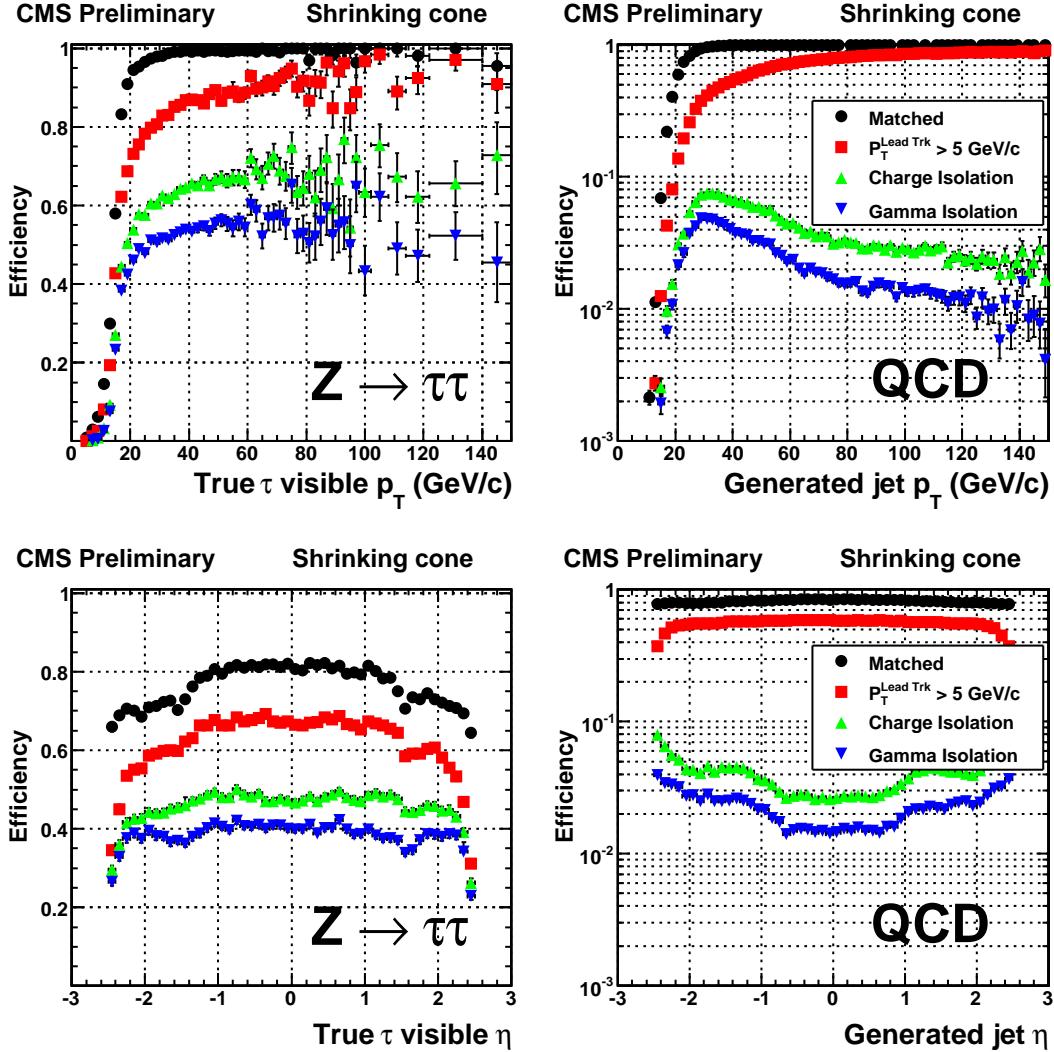


Figure 6.3: Cumulative efficiencies (left) and fake-rates (right) of successively applied tau identification cuts of the “shrinking signal cone” particle-flow based tau identification algorithm described in [28] as function of p_T^{jet} (top) and η^{jet} (bottom) of tau-jet candidates. The efficiencies/fake-rates for the complete set of tau identification criteria are represented by the blue (downwards facing) triangles.

Background Process	Expectation	Estimate obtained by applying weights of type:			Average fake-rate estimate
		QCD lead jet	QCD second jet	QCD μ -enriched	
W+jets	163.0 ± 7.1	157.2 ± 2.8	140.9 ± 2.7	129.9 ± 2.5	177.9 ± 3.2
QCD	246.4 ± 31.8	269.2 ± 14.0	246.5 ± 14.3	219.7 ± 11.8	300.8 ± 15.2
t <bar>t>+jets</bar>	12.2 ± 0.6	14.3 ± 0.3	12.6 ± 0.3	11.6 ± 0.3	16.5 ± 0.3
Z → $\mu^+ \mu^-$	68.6 ± 2.9	58.2 ± 1.3	51.2 ± 1.2	48.5 ± 1.1	65.8 ± 1.4
Σ Background	490.4 ± 32.7	499.9 ± 14.4	451.2 ± 14.6	409.7 ± 12.1	561.1 ± 15.6
Z → $\tau^+ \tau^-$	—	284.3 ± 3.7	269.0 ± 3.9	256.5 ± 3.3	325.3 ± 4.2

Table 6.3: Number of events from W+jets, QCD, tt>+jets and Z → $\mu^+ \mu^-$ background processes expected to pass all selection criteria of the Z → $\tau^+ \tau^- \rightarrow \mu^+ \tau^-$ cross-section analysis compared to the estimates obtained by weighting events in the background dominated control sample with the “simple” fake-rate weights defined by Equation 6.3.

Background Process	Expectation	Estimate obtained by applying weights of type:				Average fake-rate estimate
		QCD lead jet	QCD second jet	QCD μ -enriched	$W + \text{jets}$	
$W + \text{jets}$	163.0 ± 7.1	163.2 ± 3.8	140.6 ± 3.4	128.0 ± 3.1	188.3 ± 4.2	$155.0^{+33.6}_{-27.3}$
QCD	246.4 ± 31.8	300.5 ± 19.5	266.1 ± 19.0	236.0 ± 16.4	335.1 ± 20.4	$284.4^{+55.5}_{-52.0}$
$t\bar{t} + \text{jets}$	12.2 ± 0.6	13.1 ± 0.3	11.5 ± 0.3	10.2 ± 0.3	15.4 ± 0.4	$12.6^{+2.8}_{-2.4}$
$Z \rightarrow \mu^+\mu^-$	68.6 ± 2.9	52.7 ± 1.4	46.7 ± 1.4	41.9 ± 1.2	60.3 ± 1.6	$50.4^{+10.1}_{-8.6}$
Σ Background	490.4 ± 32.7	529.5 ± 19.9	464.9 ± 19.3	416.1 ± 16.8	599.1 ± 20.9	$502.4^{+99.4}_{-88.4}$
$Z \rightarrow \tau^+\tau^-$	–	0.3 ± 2.4	-10.6 ± 2.5	3.8 ± 2.0	-10.8 ± 2.8	$-4.3^{+8.4}_{-7.2}$

Table 6.4: Number of events from $W + \text{jets}$, QCD, $t\bar{t} + \text{jets}$ and $Z \rightarrow \mu^+\mu^-$ background processes expected to pass all selection criteria of the closure test compared to the estimates obtained by weighting events in the background dominated control sample with the “CDF-type” fake-rate weights defined by Equation 6.10.

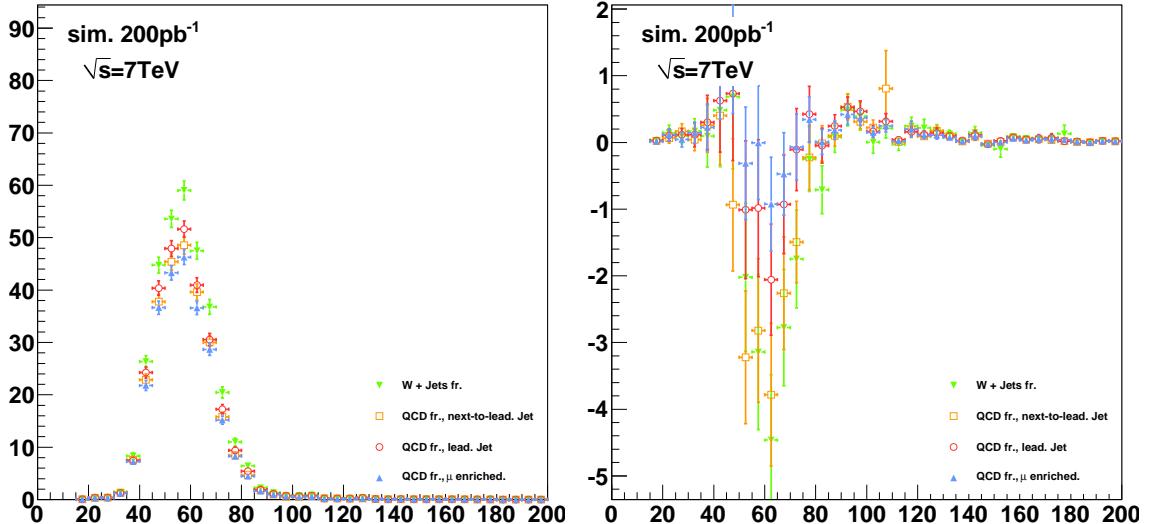


Figure 6.4: Distributions of visible invariant mass of muon plus tau-jet in $Z \rightarrow \tau^+\tau^-$ signal events weighted by “simple” weights computed according to Equation 6.4 (left) and “CDF-type” weights computed according to Equation 6.10 (right). The signal contribution to the background estimate computed by the “simple” method is non-negligible and needs to be corrected for. The “CDF-type” weights achieve a statistical cancellation of positive and negative weights, such that the total signal contribution averages to zero, avoiding the need for Monte Carlo based corrections.

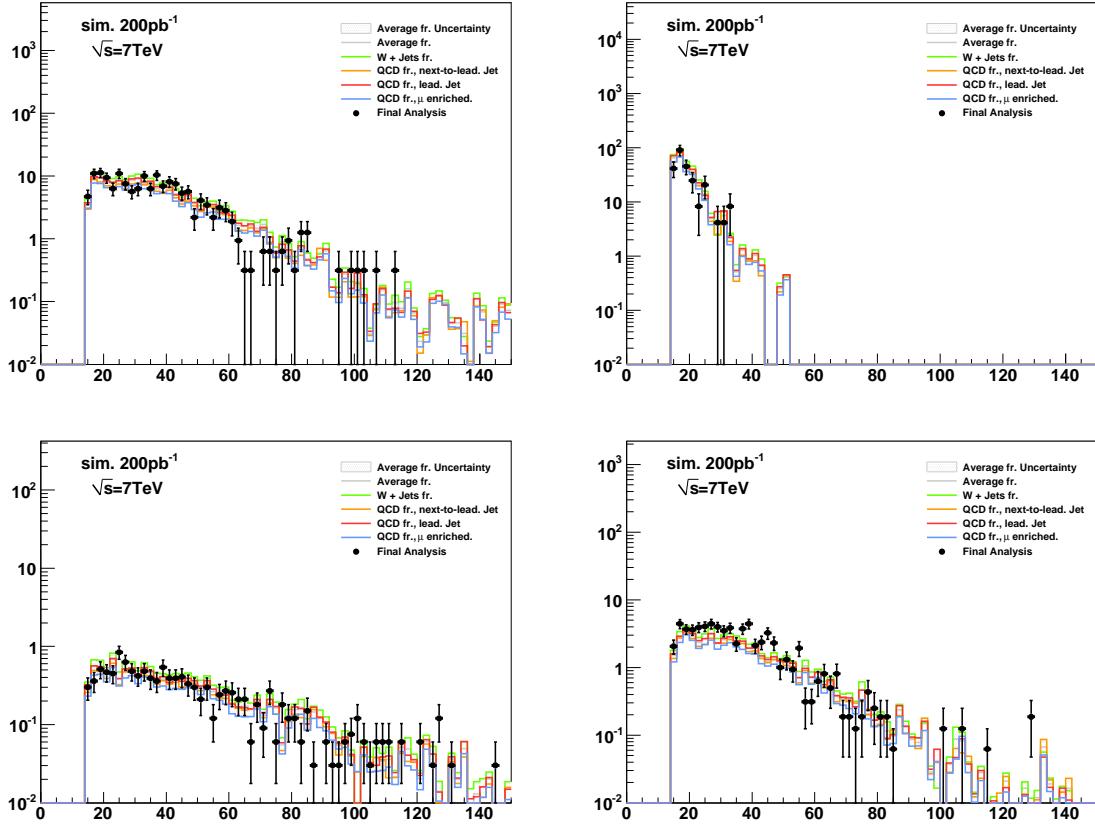


Figure 6.5: Distributions of muon transverse momentum in $W + \text{jets}$ (top left), QCD (top right), $t\bar{t} + \text{jets}$ (bottom left) and $Z \rightarrow \mu^+\mu^- \rightarrow \mu + \tau\text{-jet}$ cross-section analysis [44] compared to the estimate obtained from the “CDF method” fake-rate technique, computed according to Equation 6.10. The expected contribution of background processes is indicated by points. Lines of different colors represent the estimates obtained by applying fake-rate weights determined for different compositions of light quark, heavy quark and gluon jets, as described in Section 6.2.1. The maximum (minimum) estimate is interpreted as upper (lower) bound. The difference between the bounds is taken as systematic uncertainty on the estimate obtained from the “CDF-type” fake-rate method and is represented by the gray shaded area.

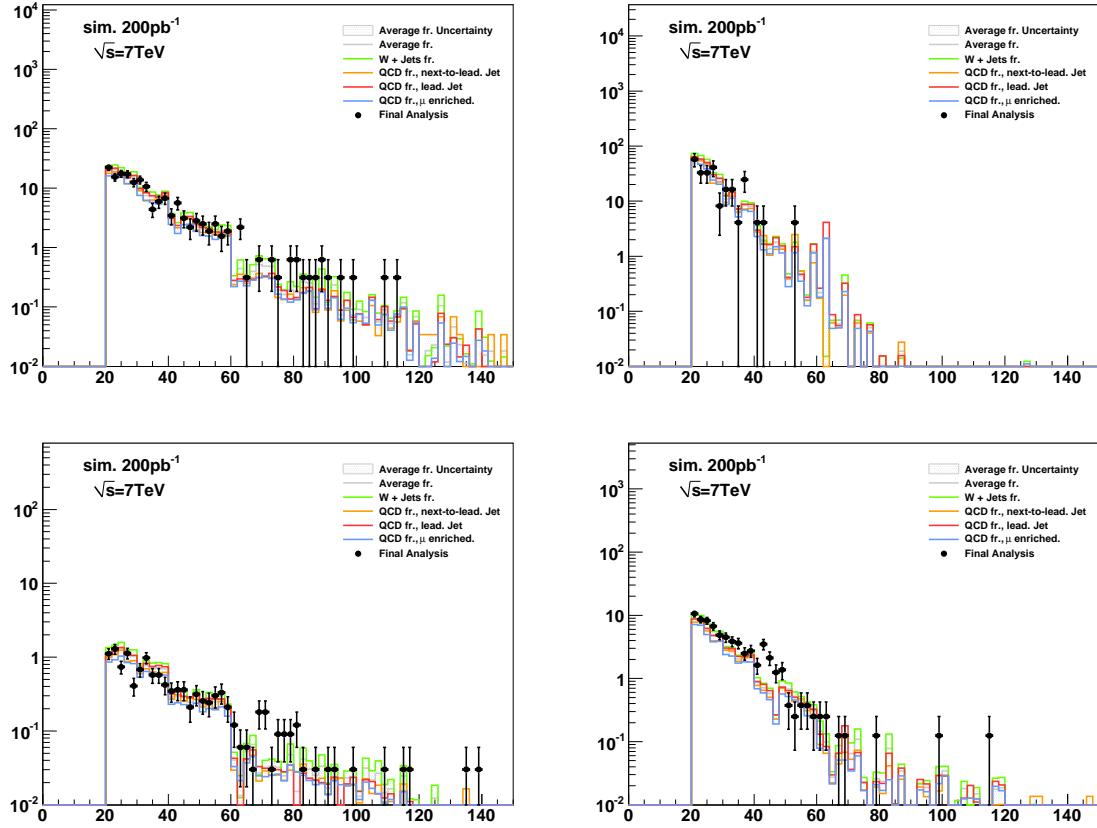


Figure 6.6: Distributions of transverse momenta of the tau–jet candidates in $W + \text{jets}$ (top left), QCD (top right), $t\bar{t} + \text{jets}$ (bottom left) and $Z \rightarrow \mu^+ \mu^-$ (bottom right) background events which pass all selection criteria of the $Z \rightarrow \tau^+ \tau^- \rightarrow \mu + \tau$ -jet cross-section analysis compared to the estimate obtained from the fake–rate technique, computed according to Equation 6.5. The expected contribution of background processes is indicated by points. Lines of different colors represent the estimates obtained by applying fake–rate weights determined for different compositions of light quark, heavy quark and gluon jets, as described in Section 6.2.1. The maximum (minimum) estimate is interpreted as upper (lower) bound. The difference between the bounds is taken as systematic uncertainty on the estimate obtained from the “CDF–type” fake–rate method and is represented by the gray shaded area.

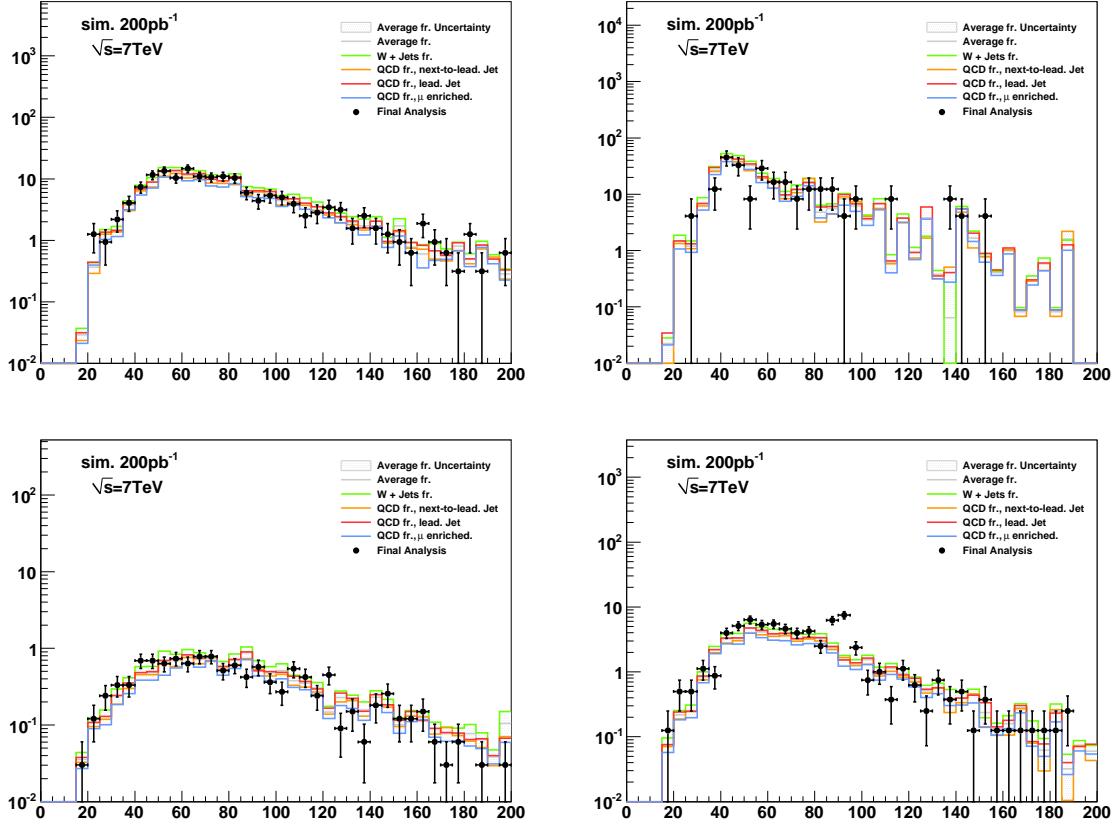


Figure 6.7: Distributions of the visible invariant mass of muon plus tau-jet in $W + \text{jets}$ (top left), QCD (top right), $t\bar{t} + \text{jets}$ (bottom left) and $Z \rightarrow \mu^+ \mu^-$ (bottom right) background events which pass all selection criteria of the closure test analysis compared to the estimate obtained from the fake-rate technique, computed according to Equation 6.10. The expected contribution of background processes is indicated by points. Lines of different colors represent the estimates obtained by applying fake-rate weights determined for different compositions of light quark, heavy quark and gluon jets, as described in Section 6.2.1. The maximum (minimum) estimate is interpreted as upper (lower) bound. The difference between the bounds is taken as systematic uncertainty on the estimate obtained from the “CDF-type” fake-rate method and is represented by the gray shaded area.

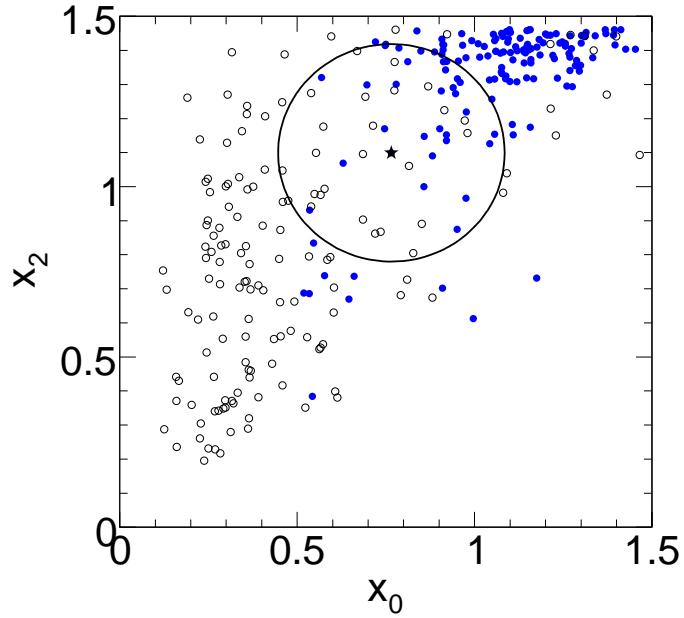


Figure 6.8: Example of the operation of a k NN classifier. The closest $k = 50$ neighbors (those inside the circle) to a test point (indicated by the star marker) are selected. The probability that the star marker is a signal event is given the number of signal neighbors (blue markers) in the circle divided by k . Image credit: [30]

Events weighted by:	Estimate
QCD lead jet	$202.1^{+14.9}_{-74.8}$
QCD second jet	$198.0^{+22.8}_{-79.3}$
QCD μ -enriched	$213.3^{+17.7}_{-82.6}$
$W + \text{jets}$	$232.8^{+21.1}_{-95.0}$
N_{bgr} estimate	$236.1^{+24.1}_{-65.9}$

Table 6.5: Estimate for background contributions obtained by weighting events passing all selection criteria listed in Table 5.3 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 $W + \text{jets}$ data samples.

Events weighted by:	Estimate
QCD lead jet	$191.7^{+2.3}_{-17.9}$
QCD second jet	$185.1^{+6.0}_{-21.1}$
QCD μ -enriched	$194.7^{+2.0}_{-20.5}$
$W + \text{jets}$	$208.9^{+0.5}_{-14.4}$
Fake-rate estimate	$201.8^{+14.2}_{-18.9}$
Observed	223

Table 6.6: Number of events observed in like-sign control region compared to the estimate obtained by fake-rate method.

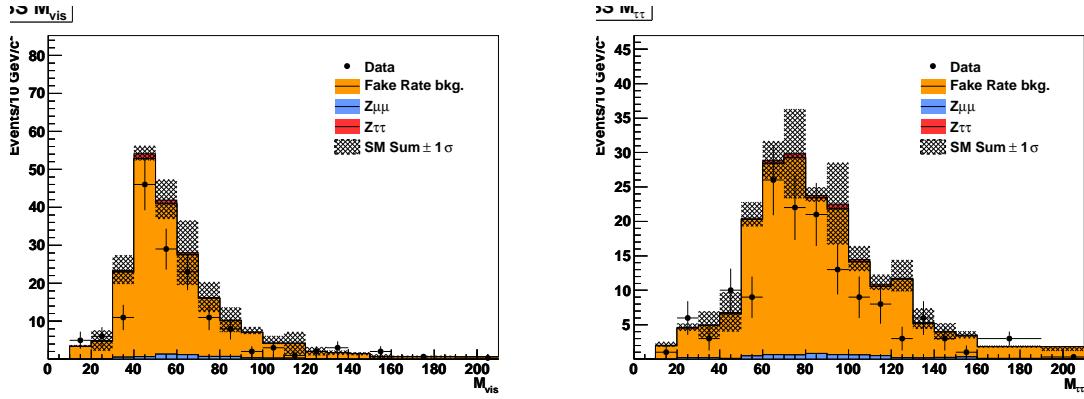


Figure 6.9: Distribution of visible mass (left) and “full” invariant mass reconstructed by the SVfit algorithm (right) observed in the like-sign charge control region compared to the background estimate obtained by the fake-rate method.

Process	Estimate
$Z \rightarrow \mu^+ \mu^-$	
Muon fake	5.7 ± 6.0
Jet fake	< 14.5
$W + \text{jets } t\bar{t} + \text{jets}$	7.6 ± 6.9
QCD	141.3 ± 40.4
N_{bgr} estimate	226.5 ± 33.1

Table 6.7: Estimated contributions of individual background processes to the signal region, obtained via the template method. As the shapes are very similar, the normalization factors for QCD and $W + \text{jets}$ background processes are anti-correlated. As a consequence, the sum of background contributions is determined by the fit more precisely than the individual contributions.

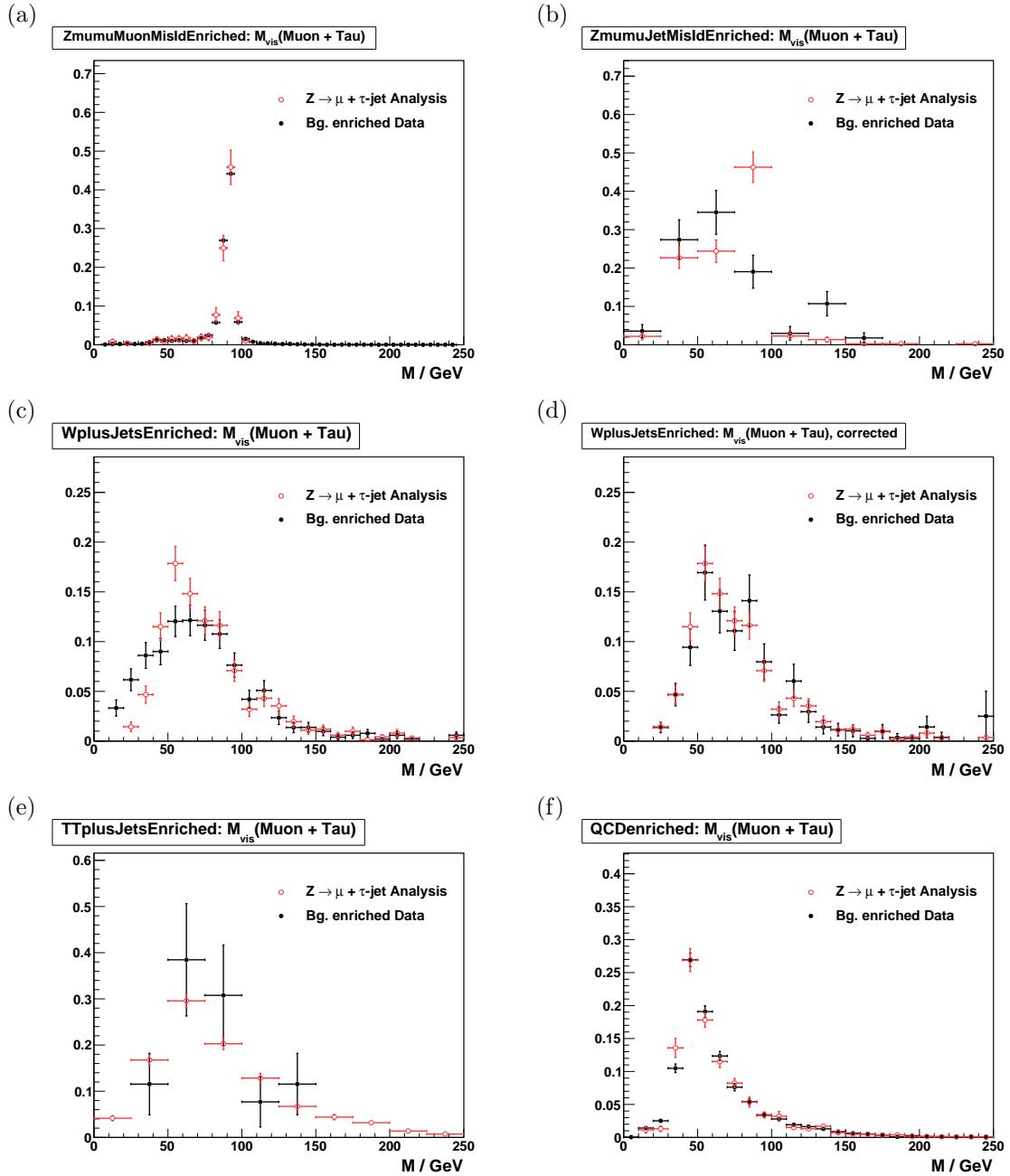


Figure 6.10: $\mu + \tau_{had}$ shape templates obtained from $Z \rightarrow \mu^+\mu^-$ (a) and (b), $W + \text{jets}$ before (c) and after (d) the bias correction explained in Section 6.3, $t\bar{t} + \text{jets}$ (e) and QCD multi-jet (f) backgrounds enriched control regions compared to the expected distribution of the enriched background process to the signal region, predicted by Monte Carlo simulations. In (a) reconstructed tau-jet candidates are expected to be dominantly due to misidentified muons, while in (b) they are expected to be mostly due to misidentified quark or gluon jets.

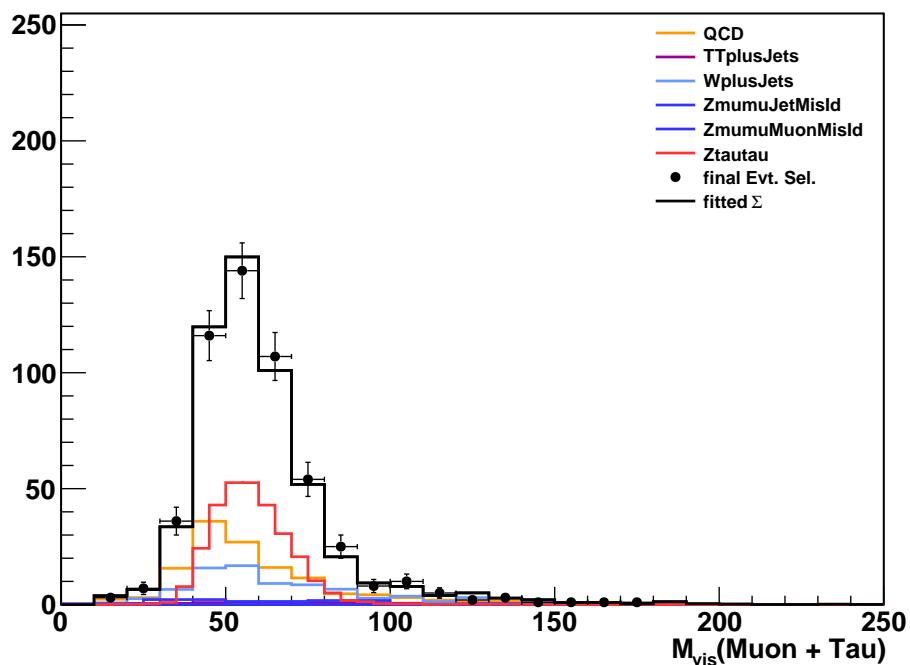


Figure 6.11: M_{vis} distribution of events selected by the $Z/\gamma^* \rightarrow \tau^+\tau^- \rightarrow \mu + \tau_{had}$ cross-section analysis compared to the sum of shape templates for signal and background processes scaled by the normalization factors determined by the fit.

1915

Chapter 7

1916

Monte Carlo Corrections

1917 One of the most important goals of the analysis is to minimize the effect of potentially
 1918 incorrect simulation effects on the final result. While the simulated CMS events have been
 1919 observed to match the 2010 data with surprising results, it is nonetheless critical to measure
 1920 in real data phenomenon which can have significant effects on the analysis whenever possible.
 1921 In practice, these measurements are used to apply a correction factor to the corresponding
 1922 measurement obtained from Monte Carlo. This measured correction factor has an associ-
 1923 ated uncertainty, and is taken into account as a systematic uncertainty. The application of
 1924 systematic uncertainties is described in the next chapter.

1925 The corrections measured and used in this analysis can be divided into two categories,
 1926 efficiency corrections and scale corrections. Identification efficiency corrections scale the
 1927 expected yield (due to a given identification selection) up or down. Scale corrections sys-
 1928 tematically scale the energy of a particle (or E_T^{miss}) up or down. In this analysis we apply
 1929 efficiency corrections for the High Level Trigger muon requirement, all stages of muon iden-
 1930 tification, and the hadronic tau identification. We apply a momentum scale correction to
 1931 the muon and tau legs, and to the resolution of the E_T^{miss} . Finally, events are simulated with
 1932 overlapping “pile-up”¹ events. The simulated events are weighted such that the number of
 1933 pile-up events in the simulation matches that observed in the data.

1934 §7.1 Muon Identification Efficiency

1935 The identification efficiencies associated with the muon are measured in $Z \rightarrow \mu^+ \mu^-$ events
 1936 using the “tag and probe” technique [44]. $Z \rightarrow \mu^+ \mu^-$ events are selected from the Muon

¹A pile-up event occurs when there are multiple interactions in one bunch proton bunch crossing. Pile-up increases with the instantaneous luminosity provided by the collider.

1937 7 TeV CMS 2010 datasets² by requiring that the events pass the “loose” Vector Boson Task
 1938 Force (VBTF) event selections [44]. In the selected events, we define the “tag” muons as
 1939 those that have transverse momentum greater than 15 GeV/c and pass the VBTF muon
 1940 selection. The tag muons are further required to pass the “combined relative isolation” de-
 1941 scribed in the VBTF paper. We finally require that the tag muon be matched to an HLT
 1942 object corresponding to the run-dependent requirements listed in Table 5.1. The trigger
 1943 match requirement ensures that the event would be recorded independently of the probe
 1944 muon. After the tag and probe muon pairs have been collected, we compare the muon identi-
 1945 fication performance in the probe collection in events selected in data to the performance in
 1946 simulated $Z \rightarrow \mu^+ \mu^-$ events. The selection of events and tag muon in the simulated sample
 1947 is the same as the data sample, with the notable exception that the only HLT requirement
 1948 applied in MC is that the tag muon is matched to an HLT_Mu9 object. Any difference in ef-
 1949 ficiency between the HLT_Mu9 path and the paths used to select the data (in the tag–probe
 1950 measurement and in the analysis) will be considered implicitly in the correction factor.

1951 The efficiencies for the muon selections applied in this analysis are measured using
 1952 the “probe” objects. We measure the following marginal efficiencies, each relative to the
 1953 previous requirement:

- 1954 • Efficiency of global probe muons to satisfy VBTF muon identification selections.
- 1955 • Efficiency of global probe muons passing the VBTF muon identification selection to
1956 satisfy the isolation criteria described in Section 5.2.1.
- 1957 • Efficiency of probe muons passing the offline analysis selection defined in Chapter 5
1958 to pass the HLT selection.

1959 In each case, the invariant mass spectrum of the tag–probe pair is fitted with a Crystal
 1960 Ball function for the signal ($Z \rightarrow \mu^+ \mu^-$) events and an exponential for the background.
 1961 The fit is done for two cases; where the probe fails the selection and the where it passes.
 1962 The method is illustrated in Figure 7.1. The signal yield N is extracted from each fit and
 1963 the efficiency is computed as $N_{\text{pass}}/(N_{\text{pass}} + N_{\text{fail}})$. Each efficiency is measured in both the

²/Mu/Run2010A-Sep17ReReco²/RECO and /Mu/Run2010B-PromptReco-v2/RECO

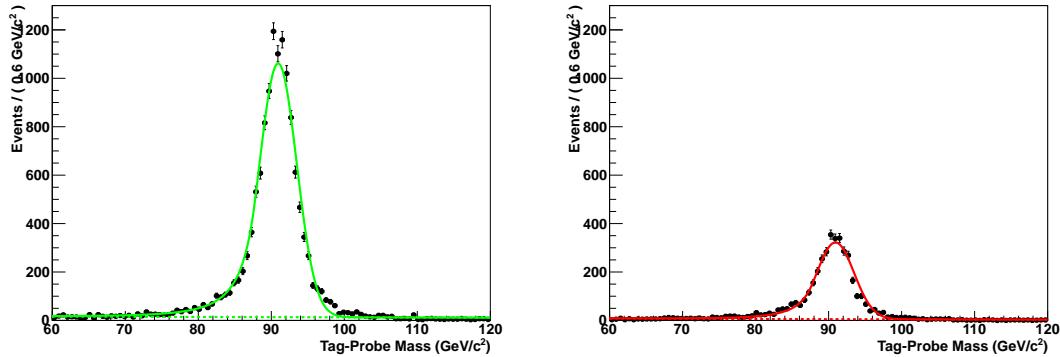


Figure 7.1: The tag–probe dimuon invariant mass spectrum in events in which the probe muon passed (left) and failed (right) the muon isolation requirement. The solid line gives the result of a simultaneous fit of the signal ($Z \rightarrow \mu^+ \mu^-$ events) and background. The fitted background contribution is shown as the dotted line. The muon isolation efficiency is then extracted from the number of signal events in the passing and failing bins.

1964 data and the simulation. The results of the measurements are shown in Table 7.1. In the
 1965 final analysis, the simulated events are weighted by the fractional difference to the measured
 1966 values; the statistical uncertainty on the weight is taken as the sum in quadrature of the
 1967 statistical uncertainties for the data and simulation efficiency measurements. The uncer-
 1968 tainty on this measurement is taken as systematic uncertainty in the final measurement.
 1969

1970 The correction for the trigger efficiency needs to take into account the differences in
 1971 the HLT selections applied during different operating periods (see Table 5.1). To determine
 1972 the overall correction factor, we measure the trigger efficiency in data for each of the op-
 1973 erating periods and compare it to the simulated efficiency of the HLT_Mu9 selection. The
 1974 overall efficiency in data is taken as the average of the three periods, weighted by integrated
 1975 luminosity.

1976 The efficiency of the “cross–triggers” used in the run–range period 148822 – 149182
 1977 (period C) cannot be measured in $Z \rightarrow \mu^+ \mu^-$ events as they require a reconstructed PFTau
 1978 object at the trigger level. A single muon trigger (HLT_Mu15) is also used in period C.
 1979 The contribution of the cross–triggers is taken as a correction to the single muon trigger
 1980 period C efficiency. The “muon leg” of the cross–triggers have the same requirements as
 1981 the single muon triggers used in the run–range 147196 – 148058 (period B). The “cross–

Muon selection	Efficiency		Ratio	Corection
	Data	Simulation		
VBTF identification	$99.2^{+0.1\%}_{-0.1\%}$	$99.1^{+0.1\%}_{-0.1\%}$	$1.001^{+0.001}_{-0.001}$	1.0
Particle Isolation	$76.8^{+0.4\%}_{-0.4\%}$	$78.3^{+0.3\%}_{-0.3\%}$	$0.981^{+0.006}_{-0.006}$	0.98
Trigger	$95.0^{+0.5\%}_{-0.5\%}$	$96.5^{+0.1\%}_{-0.2\%}$	$0.984^{+0.006}_{-0.006}$	0.98

Table 7.1: Efficiency of the various global muon selections applied in the analysis measured in data and simulated $Z \rightarrow \mu^+\mu^-$ events. The “correction” column gives the event weight correction applied to the simulated events in the final analysis. The efficiency for each selection is the marginal efficiency with respect to the selection in the row above it.

trigger” contribution is estimated as the difference between the efficiency in period B and the single-muon period C efficiency multiplied by a correction factor of $0.9 \pm 10\%$ to account for the τ leg efficiency. In the case that the measured single-muon period C efficiency is larger than the period B efficiency (due to statistical fluctuations and improvements in the trigger system), the period B efficiency is increased by 2%.

§7.2 Hadronic Tau Identification Efficiency

The hadronic tau identification efficiency has been measured in 2010 7 TeV CMS data. The most straight forward to measure the tau ID efficiency would be to use a resonance which decays to taus and has a known cross section. One could then measure the tau ID efficiency in by comparing the observed yield N_{obs} in data with that expected from the known cross section, according to the cross section equation,

$$\varepsilon = \frac{N_{\text{obs}} - N_{\text{bkg}}}{\mathcal{L} \times \mathcal{A} \times \sigma \times \text{BR}_\tau}.$$

The only suitable resonance for this method is $Z \rightarrow \tau^+\tau^-$. This method has been applied³ in the CMS $Z \rightarrow \tau^+\tau^-$ cross section analysis ??, and measured a tau identification simulation to data correction factor of 0.960 ± 0.067 .

³Actually, a slightly more complicated method is used. The analysis uses three decay channels, and the $Z \rightarrow \tau^+\tau^-$ cross section and tau identification correction factors are fitted simultaneously. The central value of the $Z \rightarrow \tau^+\tau^-$ cross section is driven by the $Z \rightarrow \tau^+\tau^- \rightarrow e\mu$ channel, which is independent of the hadronic tau identification.

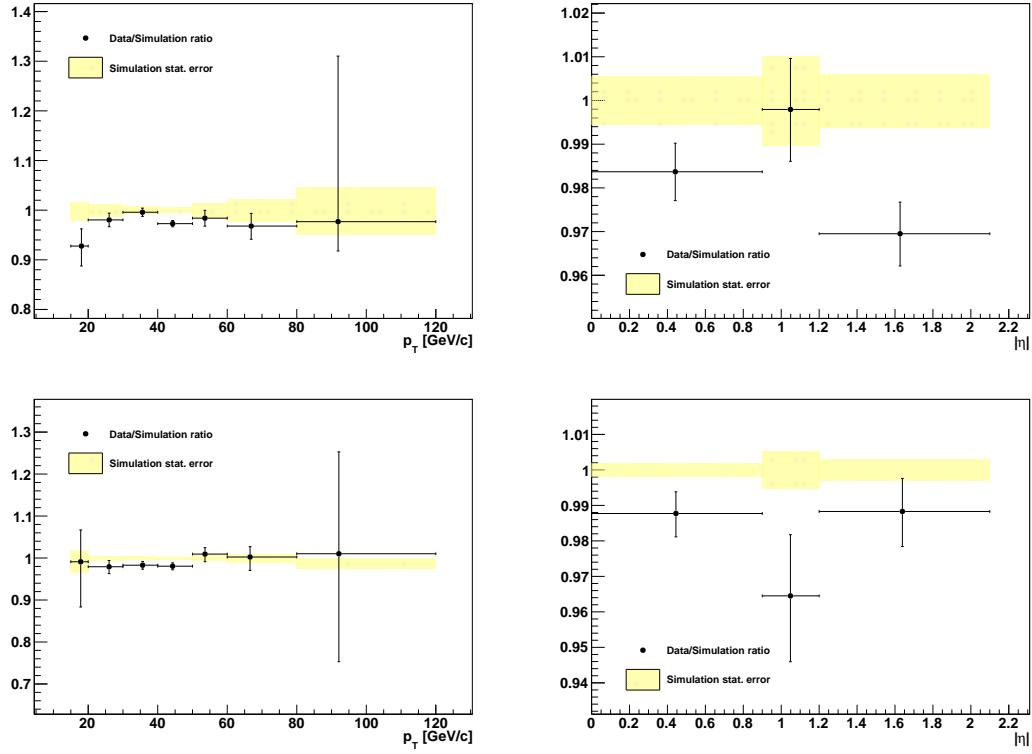


Figure 7.2: Ratio of muon isolation efficiency (top) and HLT trigger efficiency (bottom) measured in data compared to simulated $Z \rightarrow \mu^+\mu^-$ events. The correction factors are measured in bins of p_T (left) and η . In the analysis, the correction factors are taken to be flat.

Unfortunately, this method cannot be used in this analysis. The measurement using the Z resonance operates on the assumption there is no new physics contribution to the events in the Z bump. In the case that there was a Higgs signal at $m_{A^0} = 90$ GeV/ c^2 , it would be indistinguishable from the Z and would appear as an increase of N_H in the observed yield. The analysis would be completely insensitive to a Higgs boson on the Z peak, and cause the efficiency to be overestimated by a factor

$$\delta\varepsilon = \frac{N_H}{\mathcal{L} \times \mathcal{A} \times \sigma \times \text{BR}_\tau}.$$

The solution to this problem is to use a “tag and probe” approach analogous to the 1991 muon efficiency measurement of Section 7.1. The tag and probe method is only sensitive to 1992 the shapes of the distributions, and is independent of a Higgs contribution to the Z peak. 1993 This measurement has been performed by the CMS Tau Physics Object Group [48]. A 1994 loose hadronic tau preselection is applied to events which pass the selections (excluding the 1995

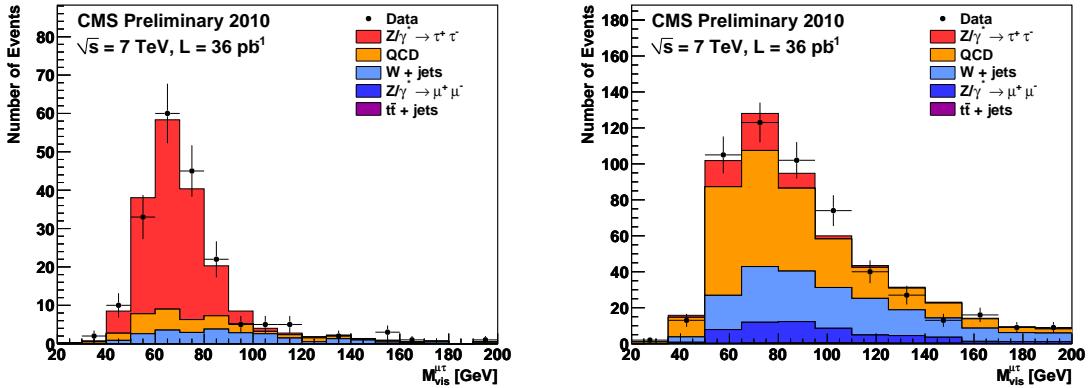


Figure 7.3: Visible mass spectrum of preselected events used to measure the hadronic tau identification efficiency in 2010 CMS 7 TeV data. The figure on the left (right) shows the preselected events that pass (fail) the hadronic tau identification. The different colors indicate the fitted yields of the different signal and background contributions. Reference: [48].

hadronic tau identification) of the CMS EWK $Z \rightarrow \tau^+\tau^-$ cross section measurement [53].
 The preselected sample is then split into two categories, those that pass the hadronic tau identification and those that fail. The signal and background yields in each category are fitted using the Template Method described in Section 6.3. An illustrative example of the fits for the yields is shown in Figure 7.3. The hadronic tau identification efficiency can then be computed using the relative size of the true tau yields in the passing and failing categories. The efficiency is measured [48] for the loose HPS + TaNC tau identification in the 2010 CMS dataset and is found to be 1.06 ± 0.30 .

§7.3 Muon and Tau Momentum Scale

Muons are one of the best measured objects at CMS. The momentum scale of CMS muons has been measured [54] using the J/ψ , $\psi(2S)$ and Υ di-muon resonant decays. The muon momentum resolution is found to be 3% or better for muons with $p_T < 100$ GeV/ c . We apply the muon momentum correction using the ‘‘MuSICleFit’’ algorithm described in [54]. The muon momentum correction and uncertainty varies as a function of muon p_T and η . The effect of the muon momentum correction uncertainty is a small effect in this analysis compared to the τ and E_T^{miss} scale uncertainties.

2012 The uncertainty on the jet energy scale is determined from an analysis of the p_T balance
 2013 between photons and jets in $\gamma + \text{jets}$ events [55]. The jet energy scale uncertainties are
 2014 determined by the JetMET group are applied to tau-jets as well as other jets in the event.
 2015 The tau energy scale correction factor is taken to be 1.0 with an uncertainty of 3%. The QCD
 2016 jet energy scale has been measured to within 3% uncertainty. In the future, the energy scale
 2017 of the tau is expected to be determined to a much better precision, as the neutral hadronic
 2018 activity of a hadronic tau decay is expected to be zero. The jet energy scale of 3% can be
 2019 confidently considered [48] an upper limit⁴, and is used in this analysis as the tau energy
 2020 scale uncertainty.

2021 §7.4 Missing Transverse Energy Correction

2022 In practice, the resolution of the reconstructed missing transverse energy is poor as it is
 2023 sensitive to the mis-measurement of any object in the event. Furthermore, a fraction of the
 2024 particles produced in the hard collision can be produced in the very forward region, outside
 2025 of the fiducial region of the calorimeters. The resolution of the E_T^{miss} reconstruction can be
 2026 measured in $Z \rightarrow \mu^+ \mu^-$ events. The true E_T^{miss} in such events is expected to be zero. The
 2027 E_T^{miss} resolution in simulated $Z \rightarrow \mu^+ \mu^-$ events is found to be smaller (better) than in the
 2028 data.

2029 The E_T^{miss} resolution depends on the “recoil” of the Z boson. The reason for this effect
 2030 is that for events where the Z is produced nearly at rest, the associated recoil products have
 2031 very small transverse momentum and are produced at very high pseudorapidity. The E_T^{miss}
 2032 is corrected using a procedure called a “Z-recoil” correction, as described in [56]. The reso-
 2033 lution of the E_T^{miss} is measured in $Z \rightarrow \mu^+ \mu^-$ events in simulation and data. The difference
 2034 in the reconstructed E_T^{miss} resolution in both samples is parameterized by the magnitude of
 2035 the transverse momenta of the particles recoiling against the Z .⁵ The reconstructed E_T^{miss}
 2036 in the simulated $Z \rightarrow \tau^+ \tau^-$, $Z \rightarrow \mu^+ \mu^-$, and $W + \text{jets}$ samples is “smeared” by a random

⁴The tau energy scale was roughly measured using the invariant mass of the hadronic decay products and shown to be compatible with 1.0, within 3%.

⁵The “recoil” particles are defined as all those not identified as Z decay products. This definition is equivalent to the total decay product transverse momentum q_T added reconstructed E_T^{miss} .

amount in each event such that the final resolution matches the observed resolution in the data.

Z -recoil corrections are determined as described in [56] and applied to simulated $Z \rightarrow \tau^+\tau^-$, $Z \rightarrow \mu^+\mu^-$ and $W + \text{jets}$ events, in order to correct for residual differences in E_T^{miss} response and resolution between data and Monte Carlo simulation [57]. The corrections are obtained by an unbinned maximum likelihood fit (in data and simulation) of the transverse recoil vector $\vec{u}_T = -(\vec{q}_T + E_T^{\text{miss}})$ as function of the transverse momentum \vec{q}_T of the Z -boson in directions parallel and perpendicular to the Z -boson transverse momentum vector.

Two additionally types of corrections are applied to the simulated E_T^{miss} , “clustered” and “unclustered.” The overall correction factor from these two effects is taken to be unity. However, the uncertainty on these two correction factors is used as a systematic uncertainty when setting the final limit. The clustered E_T^{miss} correction varies the scale of all jets with $p_T > 10$ GeV/ c in the event and recomputes the shifted E_T^{miss} . The unclustered E_T^{miss} correction varies the energy scale of all jets/particles in the event with $p_T < 10$ GeV/ c . The scale uncertainty is 3% for the clustered correction and 10% for the unclustered correction.

§7.5 Pile-up Event Weighting

The average number of pile-up interactions in the event can effect almost all aspects of the analysis. In general, increasing pile-up lowers particle identification efficiencies and lowers E_T^{miss} resolution. It is therefore important that the distribution of pile-up events in the simulation matches the distribution found in the data. Differences in the number of pile-up interactions between the data (averaged over the analyzed run-range) and pile-up Monte Carlo samples produced for “BX156⁶” pile-up conditions are corrected for by reweighting Monte Carlo simulated events according to the number of reconstructed event vertices, in order to match the distribution measured in a $W \rightarrow \mu\nu$ dataset triggered by the HLT_Mu15 High Level Trigger path. Vertices considered for this purpose are required to pass $-24 < z_{vtx} < +24$ cm, $|\rho| < 2$ cm, nDoF > 4 . In addition, the total transverse momenta of all tracks fitted to the vertex is required to exceed 10 GeV/ c , assuming that

⁶The BX156 name comes from the fact that the pile-up scenario used in this simulation corresponds to an LHC configuration with 156 bunches.

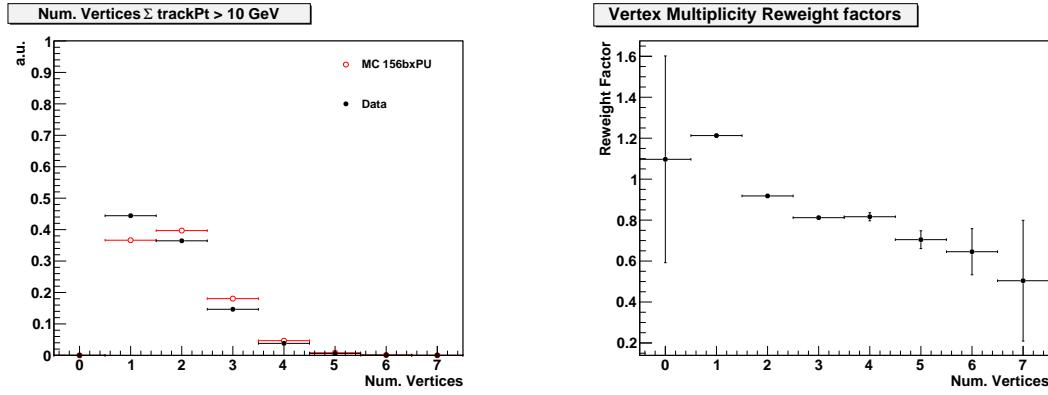


Figure 7.4: Vertex multiplicity distribution measured in the analyzed data-taking period compared to Monte Carlo simulation with “BX156” pile-up conditions (left) and resulting Monte Carlo reweighting factors (right).

2064 “softer” vertices have little or no effect on the “hard” event to pass event selection criteria.
 2065 The average vertex multiplicity distribution measured in data is compared to Monte Carlo
 2066 simulation with “BX156” pile-up conditions in Figure 7.4. Both distributions are similar,
 2067 resulting in Monte Carlo reweighting factors close to unity.

2068

Chapter 8

2069

Systematics and Limit Extraction

2070 In this chapter we discuss the systematic uncertainties affecting the search for the Higgs
 2071 boson and the statistical techniques used to establish an upper limit on the Higgs $\rightarrow \tau^+\tau^-$
 2072 branching ratio times cross section ($\sigma \times \text{BR}_\tau$). The limit can be interpreted as the largest¹
 2073 signal presence that could exist in the data and still be consistent with the null hypothesis.
 2074 The limit on $\sigma \times \text{BR}_\tau$ is roughly independent of the theoretical model². In the conclusion,
 2075 we will interpret the $\sigma \times \text{BR}_\tau$ limit result in the context of the MSSM theory.

Proper determination of systematic uncertainties is one of the most challenging and important components in performing a measurement correctly. A systematic uncertainty is the effect of the uncertainty of some ancillary measurement (or assumption) that is used in the computation of the final result. An instructive example of how a systematic uncertainty can affect the final result is a counting experiment measuring the cross section of some signal particle in the presence of background. The formula for the cross section times the branching fraction is

$$\sigma \times BR = \frac{N_{sig}}{\mathcal{L} \cdot \mathcal{A} \cdot \epsilon} = \frac{N_{obs} - N_{bkg}}{\mathcal{L} \cdot \mathcal{A} \cdot \epsilon}, \quad (8.1)$$

where N_{obs} is the number of events observed in data, N_{bkg} is the estimated number of background events in the observed data sample, \mathcal{L} is the integrated luminosity, and $\mathcal{A} \cdot \epsilon$ is the acceptance times efficiency of the signal. All of the quantities in Equation 8.1 (with the exception of the observed count N_{obs}) have some uncertainty which will effect the final measurement. Consider a situation where the expected number of background events is determined by fitting some sideband spectrum, and the fitted result has some error δN_{bkg} .

¹At some stated level of statistical confidence; the convention for limits in experimental high energy physics is 95%.

²Provided that the width of the Higgs bosons in the given model is smaller than the resolution of the SVfit mass resolution.

The total relative effect of this error can be obtained by error propagation

$$\frac{\delta(\sigma \times BR)}{\sigma \times BR} = \frac{\partial(\sigma \times BR)}{\partial N_{bkg}} \frac{1}{\sigma \times BR} \delta N_{bkg} = \frac{-\delta N_{bkg}}{N_{obs} - N_{bkg}}. \quad (8.2)$$

It is interesting to examine Equation 8.2 in two scenarios. In the limit that N_{obs} is large compared to N_{bkg} , the effect of the error on the background estimate δN_{bkg} does not affect the final result. In contrast, in a scenario when the data is dominated by background events, the relative error on the signal measurement due to the background estimation approaches infinity. The sensitivity of a measurement to a systematic uncertainty on a parameter depends on the context in which that parameter is used.

Experimental systematic uncertainties relevant for MSSM Higgs $\rightarrow \tau^+ \tau^-$ search presented in this thesis are classified in three categories: normalization uncertainties on the signal and $Z \rightarrow \tau^+ \tau^-$ background (events with true taus), normalization uncertainties on contributions from background with fake taus, and shape uncertainties. Normalization uncertainties on events with true taus are due to lepton reconstruction, identification, isolation and trigger efficiencies. These terms are equivalent to the efficiency ϵ and acceptance terms \mathcal{A} of Equation 8.2 and affect the expected yield of MSSM Higgs $\rightarrow \tau^+ \tau^-$ signal and $Z \rightarrow \tau^+ \tau^-$ background events. The uncertainties on these effects are obtained by measuring the effect in data and simulation, according to the procedures of Chapter 7, and calculating a correction factor. The uncertainty associated with the measurement of the correction factor is the systematic uncertainty. The normalization uncertainties are assumed to be uncorrelated with the shapes of visible and SVfit mass distributions which are used to extract the MSSM Higgs $\rightarrow \tau^+ \tau^-$ signal contribution in the analyzed dataset. Uncertainties on the shapes of the distributions are described by “morphing” systematics. These are due to uncertainties on the momentum/energy scale of identified electrons, muons, tau and other jets in the event. As the SVfit mass reconstruction algorithm uses the missing transverse energy, the shape of the SVfit distribution is sensitive to systematic uncertainties on the overall scale E_T^{miss} measurement. The “morphing” systematics affect the shapes of signal as well as background contributions. Normalization uncertainties on background contributions are estimated from the level of agreement between data and Monte Carlo simulation in background dominated control regions.

2103 §8.1 Signal Normalization Uncertainties

2104 The signal normalization uncertainties are due to imperfect knowledge of how improperly
 2105 modeled effects in the simulation could affect the acceptance model, the probability that
 2106 a given signal event will pass one of the selections (detailed in Chapter 5). The general
 2107 procedure to quantify these uncertainties is to measure the effect in some control region
 2108 in both the data and Monte Carlo. The ratio of data to Monte Carlo then gives a correc-
 2109 tion factor which is applied to the simulation. An uncertainty on the measurement of the
 2110 effect in control region (in the data, simulation, or both) is then taken as the systematic
 2111 uncertainties. The signal normalization uncertainties affecting this analysis on muon trig-
 2112 ger, reconstruction, identification and isolation efficiencies are taken from the tag and probe
 2113 analysis of $Z \rightarrow \mu^+ \mu^-$ events presented in Section 7.1. A very conservative estimate of 30%
 2114 is taken for the uncertainty on the tau reconstruction and identification efficiency. The tau
 2115 identification uncertainty measurement is discussed briefly in 7.2. The dependency of the
 2116 Higgs signal extraction on the tau identification efficiency has been studied, the result being
 2117 that uncertainties on the tau identification efficiency affect the limit on cross-section times
 2118 branching ratio for MSSM Higgs $\rightarrow \tau^+ \tau^-$ production by a few percent only. An uncertainty
 2119 of 11% is attributed to the luminosity measurement [58].

2120 §8.2 Background Normalization Uncertainties

2121 Uncertainties on the normalization of background processes are obtained from the study
 2122 of background enriched control regions presented in Chapter 6. The main fake tau back-
 2123 grounds in this analysis are QCD multi-jet and $W + \text{jets}$ events. These backgrounds are
 2124 produced copiously enough for the backgrounds to be studied in control regions dominated
 2125 by a single background process with a purity exceeding 90% and an event statistics exceed-
 2126 ing the expected contribution of that background to the analysis by more than one order of
 2127 magnitude. Both backgrounds are found to be well modeled by the Monte Carlo simulation.
 2128 The overall background yields used in the final fit are measured in the background enriched
 2129 control regions, and cross-checked using the Fake-rate and Template methods. For a de-
 2130 tailed discussion of the measurements, see Chapter 6. An uncertainty of 10% is attributed

2131 to the contribution of QCD and $W + \text{jets}$ backgrounds to the analysis. The cross-section
 2132 for $t\bar{t} + \text{jets}$ production makes it difficult to select a high purity sample of $t\bar{t} + \text{jets}$ events of
 2133 high event statistics. From the study of the 19 events selected in the $t\bar{t} + \text{jets}$ background
 2134 enriched control sample we assume an uncertainty on the $t\bar{t} + \text{jets}$ background contribution
 2135 in the analysis of 30%. The $Z \rightarrow \mu^+ \mu^-$ background has been studied with large statistical
 2136 precision in two separate control regions, and is dominated by events in which the recon-
 2137 structed tau-jet candidate is either due to a misidentified quark or gluon jet or due to a
 2138 misidentified muon. Good agreement between data and Monte Carlo simulation is found
 2139 in both cases. Sizeable uncertainties on the $Z \rightarrow \mu^+ \mu^-$ background contribution arise due
 2140 to the extrapolation from the background enriched control regions to the data sample con-
 2141 sidered in the analysis, however: the contribution of $Z \rightarrow \mu^+ \mu^-$ background events to the
 2142 analysis is due to events in which one of the two muons produced in the Z decay either
 2143 escapes detection or fakes the signature of a hadronic tau decay. Both cases may be difficult
 2144 to model precisely in the Monte Carlo simulation. The non-observation of a Z mass peak in
 2145 the mu + tau visible mass distribution studied with the fake-rate method on the other hand
 2146 sets a limit on possible contributions from $Z \rightarrow \mu^+ \mu^-$ background events. Conservatively,
 2147 we assume an uncertainty of 100% on both types of $Z \rightarrow \mu^+ \mu^-$ background contributions.

2148 §8.3 Shape Uncertainties

2149 Shape uncertainties on the distributions of visible and “full” invariant mass reconstructed by
 2150 the SVfit algorithm are estimated by varying the electron energy and muon momentum scale,
 2151 the energy scale of tau-jets and other jets in the event and varying the missing transverse
 2152 energy in Monte Carlo simulated events. After each variation the complete event is re-
 2153 reconstructed and passed through the event selection. Shifted visible and “full” invariant
 2154 mass shapes are obtained for each variation from the events passing all event selection
 2155 criteria. The difference between shifted shapes and the “nominal” shapes obtained from
 2156 Monte Carlo simulated events with no variation of energy or momentum scale or of the
 2157 missing transverse energy applied is then taken as shape uncertainty.

2158 The systematic uncertainties on the muon and tau energy scales have been provided
 2159 by the muon and tau CMS Physics Object Groups and are described in Section 7.3. The

modelling of missing transverse energy in different types of background events has been studied in the background enriched control regions described in Chapter 6. No significant deviations between data and Monte Carlo simulation have been found. Uncertainties due to missing transverse energy are estimated by varying the “clustered” and “unclustered” energy scales, described in Section 7.4 and recomputing the total E_T^{miss} .

§8.4 Theory Uncertainties

The signal and background normalization as well as the shape uncertainties are all experimental uncertainties in nature. Additional theoretical uncertainties arise from imprecise knowledge of parton-distribution functions (PDFs) and of the exact dependency of signal cross-sections and branching ratios on $\tan \beta$ and m_{A^0} . The PDFs describe how the energy of the protons is shared between the quarks and gluons. Since the longitudinal boost of the hard collisions depends on the PDFs, the signal acceptance is sensitive to errors in the PDFs. The uncertainties on the signal acceptance due to PDF uncertainties are estimated using tools developed by the CMS Electroweak working group [59]. The acceptance is computed with respect to MSSM Higgs $\rightarrow \tau^+ \tau^-$ decays that have muons of $p_T^\mu > 15$ GeV/c and $|\eta_\mu| < 2.1$, jets produced in hadronic tau decays with visible $p_T^{\text{vis}} > 20$ GeV/c and $|\eta_{\text{vis}}| < 2.3$ on generator level. Acceptance values are computed for the central value and 44 eigenvectors of the CTEQ66 PDF set [60]. The systematic uncertainty on the signal acceptance is computed following the PDF4LHC recommendations [61, 62].

The effect of Monte Carlo normalization, shape and theory uncertainties on the signal efficiency times acceptance is summarized in Table 8.1.

§8.5 Limit Extraction Method

The search for a new signal is performed by examining the observed distribution of the reconstructed di-tau mass $m_{\tau\tau}$ as reconstructed by the SVfit algorithm. An “bump” in this spectrum would indicate the presence of a new particle. To make a statement about the presence of a bump with confidence, the shape under a potential bump must be well described. The background shape is decomposed into the combination of shapes from the different background sources, which we refer to here as “templates.” The data and the

Source	Effect
Normalization uncertainties	
Trigger	0.981 ± 0.006
Muon identification	1.001 ± 0.001
Muon isolation	0.984 ± 0.006
Tau–jet identification	1.00 ± 0.30
Shape uncertainties	
Muon momentum scale	$\ll 1\%$
Tau–jet energy scale	$1 - 4\%^1$
Jet energy scale (JES)	$< 1\%^2$
E_T^{miss} (Z –recoil correction)	1%
Theory uncertainties	
PDF	2% ³

¹ decreasing with m_A

² number quoted for $gg \rightarrow A/H$ and $b\bar{b} \rightarrow A/H$ sample as a whole;
in the subsample of events with b-tagged jets the effect of the JES uncertainty is 4%

³ with small dependence on m_A

Table 8.1: Effect of normalization uncertainties on the $gg \rightarrow A/H$ and $b\bar{b} \rightarrow A/H$ signal efficiency times acceptance.

2188 templates for each background distribution is binned in the observable $m_{\tau\tau}$ variable. The
 2189 normalization of each template represents the total yield expected for that source.

We can then define a likelihood for any configuration of our templates given the observed data. The likelihood is a “binned Poisson likelihood,” which is defined as the product of the Poisson probability in each bin. The Poisson probability $P(n|\mu)$ is the probability to observe n events given an expectation of μ events. The Poisson probability is given by the expression

$$P(k|\mu) = \frac{\mu^k e^{-\mu}}{k!}. \quad (8.3)$$

2190 The total likelihood for observed data given some configuration of templates is then simply

2191 the product of the Poisson probabilities (Equation 8.3) in each of the N_{bin} bins:

$$\mathcal{L} = \prod_{i=1}^{N_{bin}} \frac{\mu_i^{n_i} e^{-\mu_i}}{n_i!}, \quad (8.4)$$

2192 where the expected number of events μ_i in the bin i is the sum of the number of events

2193 from all sources

$$\mu_i = \sum_{j=1}^{N_{source}} \mu_{ji}.$$

2194 The number of expected events in a source, in turn, can be written

$$\mu_{ji} = L\sigma_j \epsilon_{ji} \quad (8.5)$$

2195 where L is the integrated luminosity, σ_j is the cross section for source j , and ϵ_{ji} is the
2196 efficiency for source j in bin i .

2197 We incorporate the systematic uncertainties of the analysis by introducing a set of
2198 “nuisance parameters” $\vec{\beta}$ into the likelihood function. As the name suggests, we are not
2199 interested in the actual value of the nuisance parameters. Each nuisance parameter pa-
2200 rameterizes some phenomenon in the analysis. The shape templates (which can be defined
2201 purely in terms of σ_j and ϵ_{ji} in Equation 8.5) are now interpreted as functions of the set
2202 of nuisance parameters $\vec{\beta}$. Existing knowledge about the value of the nuisance parameter is
2203 introduced by extending the likelihood function with a constraint $\mathcal{G}(\vec{\beta})$ that expresses the
2204 knowledge about the nuisance parameters. The templates μ_{ji} can depend on the nuisance
2205 parameters in two ways. Normalization uncertainties introduce multiplicative nuisance fac-
2206 tors on the yield of some (sub)set of the sources defined by Equation 8.5. As an example,
2207 consider a simplified situation where there are only two sources: $Z \rightarrow \tau^+\tau^-$ and $W + \text{jets}$
2208 events. Both sources are sensitive to the efficiency of the muon trigger. Only the $Z \rightarrow \tau^+\tau^-$
2209 sample is sensitive to the efficiency of the hadronic tau identification algorithm, as the
2210 fake-rate in $W + \text{jets}$ is measured in data. We would then introduce two multiplicative nui-
2211 sance parameters, β_μ and β_τ , which respectively correspond to the two uncertainties. The
2212 expected number of events in the i th bin is then given by a modified form of Equation 8.5

$$\mu_i = L\beta_\mu \sigma^W \epsilon_i^W + L\beta_\mu \beta_\tau \sigma^Z \epsilon_i^Z. \quad (8.6)$$

2213 From Equation 8.6, we can see that β_μ affects both sources but β_τ only affects the $Z \rightarrow \tau^+\tau^-$
2214 source.

The shape uncertainties discussed in Section 8.3 are incorporated using a technique

$(1 - \alpha)\%$	$m = 1$	$m = 2$	$m = 3$
68.27	1.00	2.30	3.53
90.00	2.71	4.61	6.25
95.00	3.84	5.99	7.82
99.00	6.63	9.21	11.34

Table 8.2: Correspondence between a confidence level defined by $(1 - \alpha)\%$ and the corresponding interval in $2\Delta \ln \mathcal{L}$. The $2\Delta \ln \mathcal{L}$ interval is different depending the number m of parameters which are being simultaneously estimated.

called “vertical template morphing.” For each source, different templates are created for three different values of the morphing parameter, corresponding to -1, 0, and +1 standard deviation shifts in the nuisance parameter. To determine the number of expected events in the i th bin as a function of the morphing parameter, we interpret quadratically between the i th bin values of the three templates, and extrapolate linearly beyond them. The overall likelihood then, including nuisance parameters, can be written as

$$\mathcal{L} = \prod_{i=1}^{N_{bin}} \frac{[\mu_i(\vec{\beta})]^{n_i} e^{-\mu_i(\vec{\beta})}}{n_i!} \times \prod_{m=1}^{N_\beta} \mathcal{G}(\beta_m). \quad (8.7)$$

2215 In Equation 8.7, we have introduced the assumption that the nuisance parameters are
2216 uncorrelated.

2217 We wish to determine if the data is compatible with a new signal. To test for the
2218 presence of a bump, we examine the likelihood as a function of the signal cross-section. If the
2219 presence of a signal is unambiguous, one can simply determine the likelihood as a function
2220 of the cross-section. If the signal is known to be non-zero, confidence intervals about the
2221 maximum likelihood can be constructed by examining the change in the logarithm $2\Delta \ln \mathcal{L}$
2222 of the likelihood. The Frequentist interpretation of the confidence level $(1 - \alpha)$ is that if
2223 the experiments were repeated N times, the interval corresponding to $(1 - \alpha)$ would fail
2224 “cover” the *true* value of the observable in $\alpha \times N$ of the experiments. The correspondence
2225 between confidence levels and intervals in $2\Delta \ln \mathcal{L}$ is given in Table 8.2. Since the mass of
2226 a potential Higgs boson is unknown, we repeat this process for different signal masses. If
2227 there is not a significant signal, we can set upper bounds on the signal cross-section using
2228 one of several methods, which will be discussed below.

Fixme: Did I
mix this up?

At this point, the likelihood still depends on the nuisance parameters. There are two methods of removing the dependence on the nuisance parameters, “marginalization” and “profiling.” Marginalization is the process of integrating the likelihood of Equation 8.7 over the entire range of all nuisance parameters. This effectively averages the effect of the different uncertainties into the marginalized likelihood function. The profiling method maximizes the likelihood in terms of the nuisance parameters. In the profiling method, the “profile likelihood” is created by maximizing the likelihood with respect to all of the nuisance parameters. One way to interpret the profiling method is that the values of nuisance parameters are being measured *in situ*, constrained by the observed data. While not a strictly Bayesian treatment, the profile likelihood method has been shown to give almost identical results to marginalization.

An interesting situation relating to the profiling of the nuisance parameters arises in this analysis in the context of the hadronic tau identification uncertainty. As discussed in Chapter 7, the Higgs–insensitive measurement of the tau identification efficiency has a high uncertainty of 30%. If the bump at the $Z \rightarrow \tau^+\tau^-$ resonance can be considered to Higgs free, the tau identification efficiency can be measured to a much greater precision, approximately 7% [53, 48]. When profiling the likelihood (as a function of cross section) for a Higgs with a mass $m_{A^0} > m_Z$, the likelihood contains the information that there is no Higgs contribution to the Z resonance. This fact enables the profiling process to constrain the systematic uncertainty on the tau identification efficiency to the 7% level using the size of the Z resonance. Conversely, when the likelihood is evaluating the likelihood of the presence of a Higgs signal with $m_{A^0} = m_Z$, the likelihood cannot distinguish between a potential presence of Higgs events in the Z bump or a systematic undershoot of the tau identification efficiency in the simulation. In this case the profiling processing has no power to constrain the systematic to a value lower than the initial 30% value.

In the absence of the signal, or even in the presence of one, we can determine a upper 95% CL bound on the cross-section of the signal using the profile likelihood. In one method we simply use Bayes’ Theorem to convert the likelihood to a posterior density in the signal cross-section, and integrate to find the point below which 95% of the probability lies. The

Bayesian posterior PDF is computed as

$$\mathcal{P}(\sigma_H | \bar{x}, m_H) = \frac{\mathcal{L}(\bar{x} | \sigma_H m_H) \mathcal{P}(\sigma_H)}{\int \mathcal{L}(\bar{x} | \sigma'_H m_H) \mathcal{P}(\sigma'_H) d\sigma'_H}. \quad (8.8)$$

2254 The left hand side of Equation 8.8 represents the probability density for a given signal cross
 2255 section, given the observed data \bar{x} and the assumed value m_H for the Higgs mass. We refer
 2256 to this method as the “Bayesian” method of setting limits.

2257 In the other method, which is referred to as the “Delta Log Likelihood” (DLL) method,
 2258 we compute two likelihoods. The first likelihood is computed for the “null hypothesis” case.
 2259 The likelihood is profiled (maximized) assuming that no signal is present. We then construct
 2260 the profile likelihood for increasing values of the signal yield. The upper limit is achieved
 2261 when the logarithm of the profile likelihood is 1.92 units less than the value of the null
 2262 hypothesis profile likelihood.

2263 In general the limits computed by the Bayesian and DLL methods are similar. However,
 2264 the effect of upward or downward statistical fluctuations in the observed data affect the two
 2265 methods in different ways. When the data fluctuates low, the DLL method will produce
 2266 a more stringent limit than the DLL method. When the data fluctuates high, creating an
 2267 apparent signal, the Bayesian method will (correctly) set a more stringent limit.

2268 In the results presented below we use nuisance parameters corresponding to the sys-
 2269 tematic errors summarized in Table 8.3.

Fixme: Get
MET scale,
how to split
between type1
and type2?

Source	Method	Magnitude
Muon ID/trigger	Multiplicative	20%
Z cross section	Multiplicative	5%
Jet to τ fake rate	Multiplicative	20%
$\mu \rightarrow \tau$ fake rate	Multiplicative	100%
W +jets cross section	Multiplicative	10%
$t\bar{t}$ cross section	Multiplicative	40%
integrated luminosity	Multiplicative	11%
Tau energy scale	Morphing	3%
Missing E_T scale	Morphing	3%
Muon p_T scale	Morphing	neg.
EM energy scale	Morphing	neg.

Table 8.3: Summary of systematic uncertainties represented by nuisance parameters in the likelihood, their representation method and magnitudes.

2270

Chapter 9

2271

Results

2272

§9.1 Selected Events

2273

In the 36 pb^{-1} of 7 TeV data collected by CMS in 2010, the analysis selection criteria described in Chapter 5 selects a total of 573 events. The expected yields from each of the background sources, computed by the Template method (Section 6.3) and verified by the Fake-rate method (Section 6.2) are shown in Table 9.1. The total expectation from the standard model background is 577. The data agrees extremely well with the SM background expectation considering the expected statistical variance of the observed data (24 events), as well as the uncertainty on the integrated luminosity¹ and the tau identification uncertainty.

2280

2281

The distributions of the visible mass and SVfit mass in the final selected events are shown in Figure 9.3. Excellent agreement is seen between the shapes of the distributions. The p_T spectrum of the transverse momentum are sensitive to mis-modeling of the kinematics and composition of the various background sources. Control plots showing the transverse momenta of the muon and visible hadronic tau in the final analysis selection are shown in Figure 9.1. The η and ϕ distributions of the muon and tau objects are sensitive to detector effects, and the presence of cosmic muons. For example, muons from cosmic events will preferentially be produced in the $\phi = 0$ direction. Spurious candidates resulting from poorly model noise in one of the CMS subdetectors will in general be localized in $\eta - \phi$. The η and ϕ distributions of the muon and tau candidates are shown in Figure 9.2 and show excellent agreement.

¹The uncertainty on the CMS luminosity measurement was 11% at the time this analysis was performed. The measurement was later improved, and at the time of this writing the uncertainty on CMS 2010 integrated luminosity is 4% [58]. The improved luminosity measurement is not expected to change the results of this analysis significantly.

Process	Events in 36 pb ⁻¹
$t\bar{t} + \text{jets}$	6.6
$W + \text{jets}$	25.6
$Z \rightarrow \mu^+ \mu^-$	10.6
QCD	166.2
$Z \rightarrow \tau^+ \tau^-$	368.1
Standard model sum	577.1
Data	573

Table 9.1: Number of Higgs $\rightarrow \tau^+ \tau^- \rightarrow \mu + \tau_{had}$ candidate events passing the selection criteria described in Chapter 5.

2292 The expected yields from an MSSM Higgs boson signal for $\tan \beta = 30^2$ are summarized
 2293 in . The contributions from the gluon fusion production mode and the associated b –quark
 2294 production modes are given separately. The yields are divided into the exclusive categories of
 2295 events containing a b –tagged jet and those without. For a Higgs mass of $m_{A^0} = 160$ GeV/ c^2 ,
 2296 a total of 17 events are expected at $\tan \beta = 30$.

Fixme:

Get cite

2297 §9.2 Limits on Higgs Production

2298 We compute upper limits on the cross section times the branching ratio using the Bayesian
 2299 method described in Section 8.5. We compute an expected limit in the same manner as an
 2300 observed limit, but with simulated data generated in “toy” experiments. A large number
 2301 of pseudo–data sets are generated using the null hypothesis templates using Monte Carlo
 2302 techniques. The nuisance parameters are sampled within their constraints in the generation
 2303 of the pseudo–date. The pseudo–data sets are expected to have the same statistical sen-
 2304 sitivity as the observed dataset. Upper limits are then computed using the pseudo–data.
 2305 The process is repeated many times, and the spread of the obtained upper limits deter-
 2306 mines the expected upper limit band. The expected nominal upper limit, and the ± 1 , and
 2307 ± 2 confidence limits are shown in Table 9.3. The observed limit on the MSSM computed

²Details of the relationship between the MSSM Higgs cross section and $\tan \beta$ are discussed in detail in Section 9.3.

from the 573 events selected in this analysis is given in the right column of Table 9.3. The observed limit is compatible with the expected limit, within 1.5 standard deviations. The trend of the expected and observed limits versus the Higgs mass using both observables are shown in Figure 9.4. The use of the SVfit reconstructed mass as the observable increases the power of the limit significantly. The limit trend has some interesting features. When the Higgs mass is close to the mass of the Z resonance, the analysis have little power to set a limit on the presence of the Higgs. This is due to the large uncertainty on the tau identification efficiency. Essentially, when $m_{A^0} = m_Z$, the Higgs yield in the Z bump would have be larger than 30% of the $Z \rightarrow \tau^+\tau^-$ yield for the profile likelihood to be able to recognize an excess of events. Below this value, the profile likelihood can simple shift the tau identification efficiency scale factor up by 30% and “eat” any potential excess of signal.

§9.3 Interpretation in the MSSM

The limits on the cross section times branching ratio are roughly model independent,³ and could be applied to set limits on the parameter space of a number of models. In this thesis, we interpret the results in the context of the MSSM. Specifically, we exclude a region in the $\tan\beta - m_{A^0}$ parameters space of the MSSM. To find the upper limit band on $\tan\beta$, we find the minimum value of $\tan\beta$ which provides the cross section and branching ratio product found in the corresponding row in Table 9.3.

The mapping between m_{A^0} and $\tan\beta$ and the Higgs cross section is provided by the LHC Higgs Cross Section working group [63]. The cross sections and branching ratios have been computed for the h^0 , H^0 , and A^0 MSSM Higgs states in both the ggA and qqA production modes, for a grid of points in $\tan\beta - m_{A^0}$ space. In order to combine the ggA and qqA production modes, what we call our signal cross-section is the sum of the cross-section times branching ratio for both modes, assuming $\tan\beta = 30$. Additionally, as discussed in Section 1.2.3, the MSSM Higgs sector consists of two Higgs doublets, yielding five physical Higgs bosons. This search is sensitive to the three neutral Higgs particles the

³This assumption is only valid if the shape of the sum of all new physics contributions are also model independent, on the scale of the experimental resolution. For the values of $\tan\beta$ this analysis is sensitive to, this is a valid approximation in the MSSM. In a model where the width of the Higgs boson resonance was larger than the resolution of the SVfit method, the limits of Table 9.3 would not be valid.

2334 h^0, H^0 , and A^0 . The relative contributions of the three Higgs types depends on the mass
 2335 m_{A^0} of the CP-odd Higgs. An observed signal will have contributions from at least two
 2336 Higgs states. For $m_{A^0} \leq 130$ GeV/c², the A^0 and h^0 are approximately degenerate in mass
 2337 and width. In this region the H^0 has a very small relative cross section and a constant
 2338 mass of $m_{H^0} \approx 130$ GeV/c². For $m_{A^0} \geq 130$ GeV/c², the h_0 reaches a limiting mass of
 2339 ≈ 130 GeV/c², and the H^0 and A^0 become mass degenerate.

2340 The region in $\tan\beta - m_{A^0}$ MSSM parameter space excluded by this analysis at 95%
 2341 CL is shown in Figure 9.3. The limit is compared to the combined result from Run II of
 2342 the Tevatron (this result is discussed in detail in Section 1.3.3). The result of this analysis
 2343 sets a stronger limit than the Tevatron for large values of m_{A^0} . In the low m_{A^0} region, the
 2344 analysis suffers due to the large tau identification efficiency uncertainty. This effect can be
 2345 mitigated by using the $e - \mu$ channel. The combined CMS result uses this approach, and
 2346 will be discussed briefly in the conclusion.

Fixme:

check this!

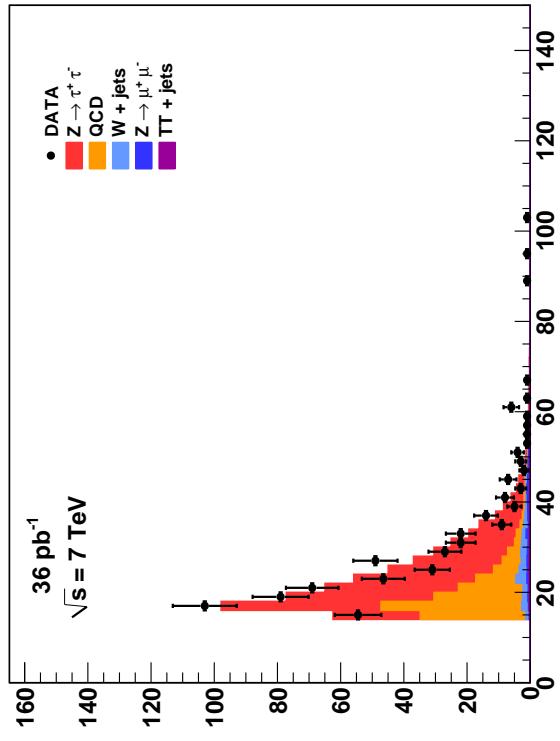
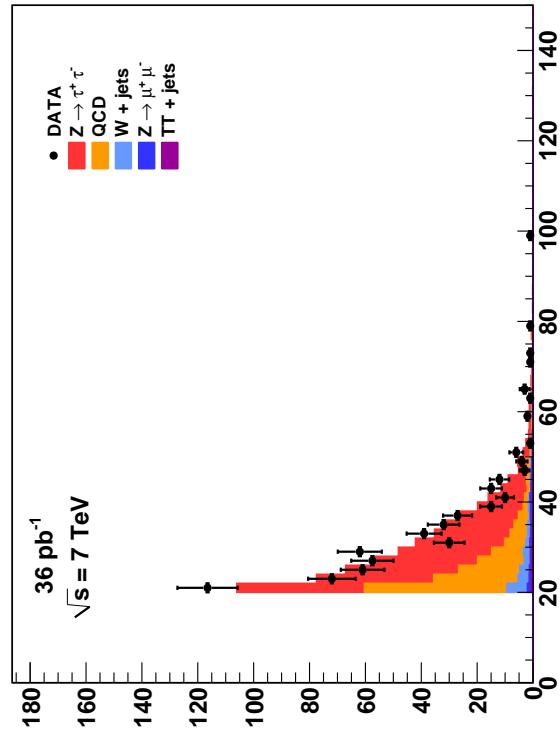
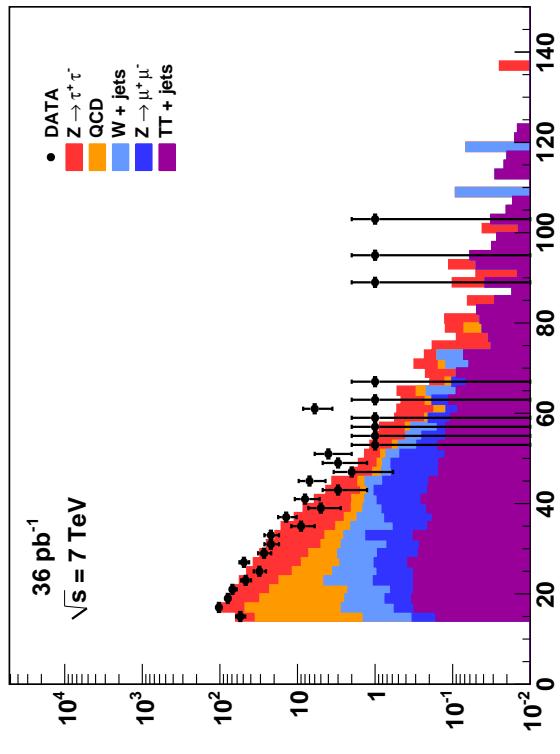
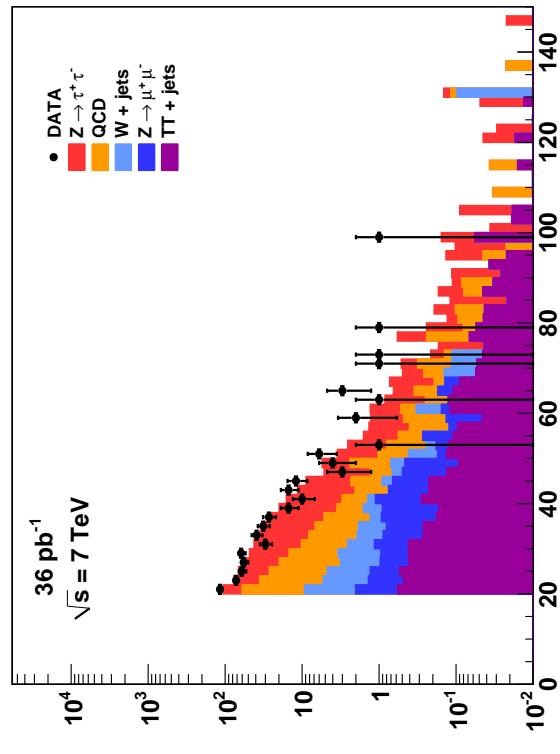
Muon (final event sample)**Tau (final event sample)****Muon (final event sample)****Tau (final event sample)**

Figure 9.1: Distribution of the transverse momentum of the muon (top) and hadronic tau in $Higgs \rightarrow \tau^+ \tau^- \rightarrow \mu + \tau_{had}$ candidate events passing the selection criteria described in Chapter 5. The distributions are shown in linear (logarithmic scale on the left (right)).

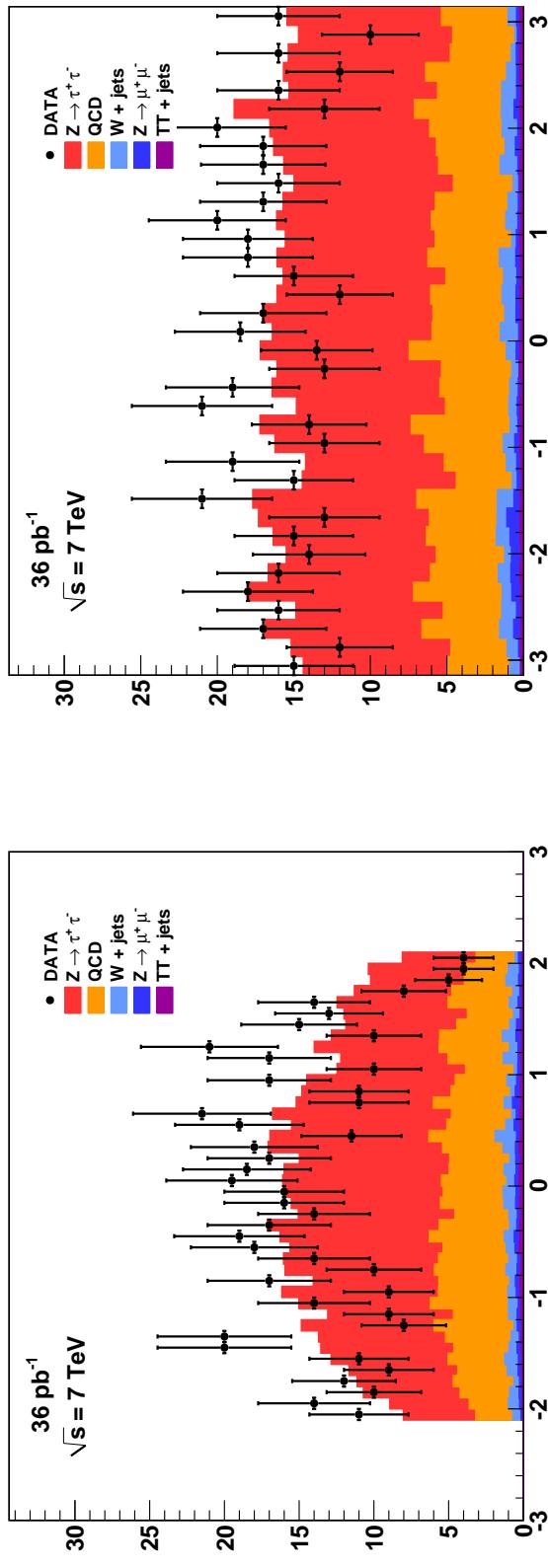
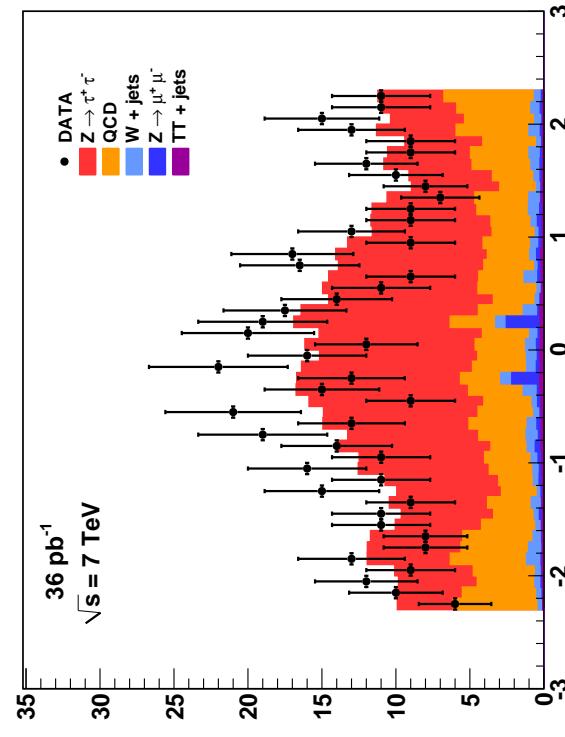
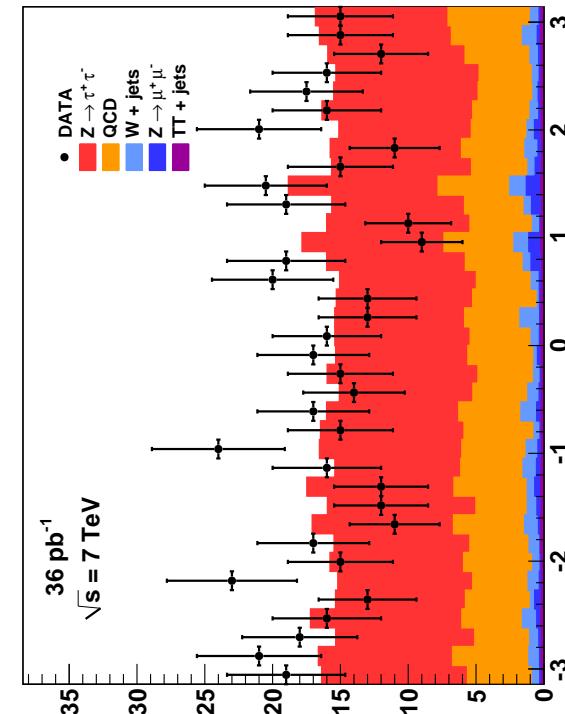
Muon (final event sample)**Tau (final event sample)****Tau (final event sample)****Muon (final event sample)**

Figure 9.2: Distribution of the η (left) and ϕ (right) of the muon (top) and hadronic tau (bottom) in Higgs $\rightarrow \tau^+\tau^- \rightarrow \mu + \tau_{had}$ candidate events passing the selection criteria described in Chapter 5.

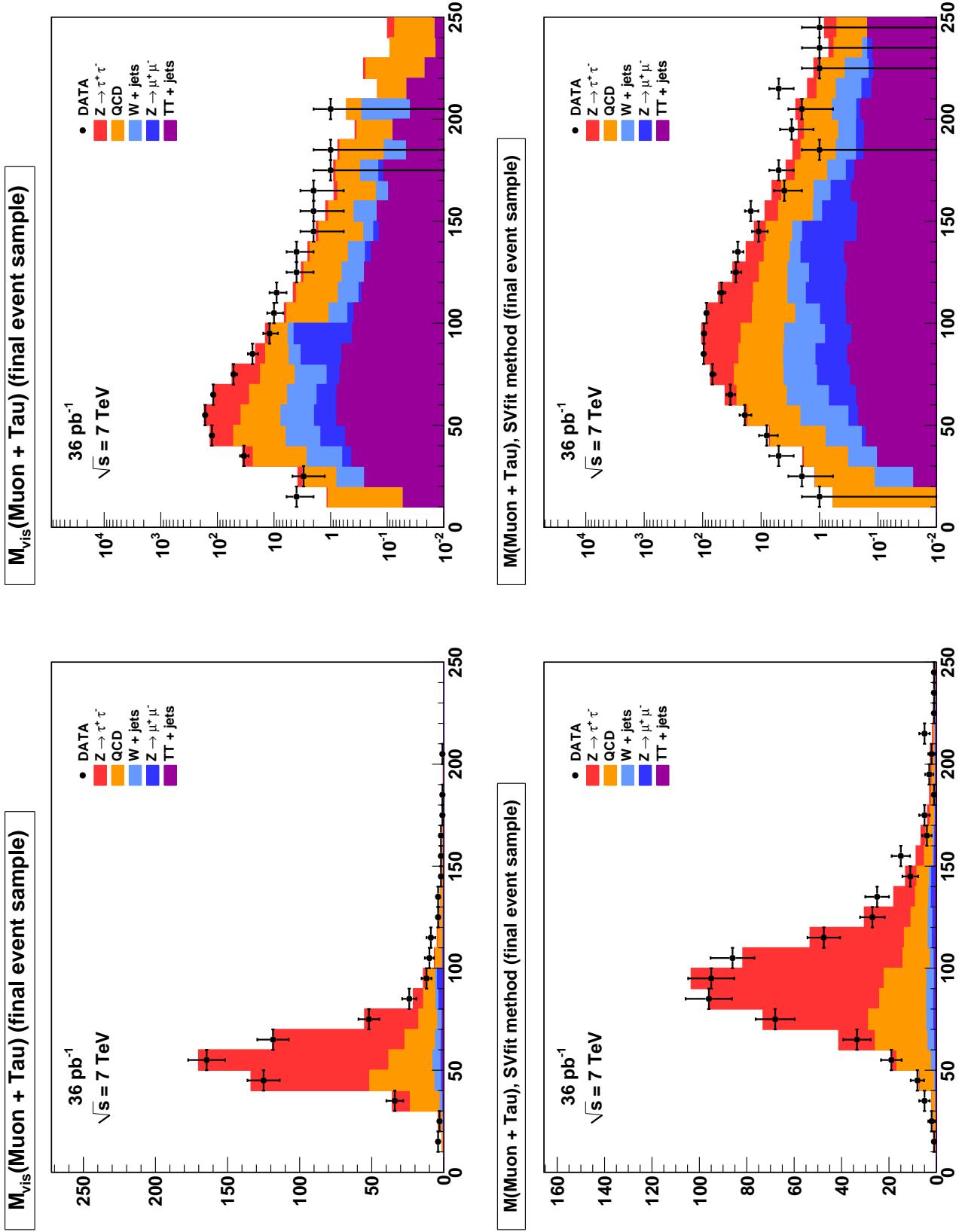


Figure 9.3: Distribution of visible (top) and “full” $\tau^+\tau^-$ invariant mass reconstructed by the SVfit algorithm (bottom) in $Higgs \rightarrow \tau^+\tau^- \rightarrow \mu + \tau_{had}$ candidate events passing the selection criteria described in Chapter 5. The distributions are shown in linear (logarithmic) scale on the left (right).

Process	Events without b -tag	Events with b -tag
Gluon fusion production		
A90	37.21	0.86
A100	27.40	0.40
A120	14.39	0.14
A130	11.81	0.18
A160	4.46	0.09
A200	1.51	0.03
A250	0.47	0.01
A300	0.15	0.0
A350	0.06	0.44
Associated b -quark production		
bbA90	33.07	5.50
bbA100	30.18	4.77
bbA120	21.91	4.02
bbA130	18.34	3.35
bbA160	10.35	2.10
bbA200	4.85	1.29
bbA250	2.11	0.55
bbA300	0.97	0.26
bbA350	0.41	0.13

Table 9.2: Number of Higgs signal event expected to pass the selection criteria described in Section 5. The expected signal yield is given for MSSM parameter $\tan \beta = 30$, using the cross sections provided by the LHC Higgs Cross Section working group.

Secondary Vertex Fit 95% CL Upper Limit (pb)						
Mass (GeV/c ²)	Expected $\sigma_H \times B_\tau$ (pb)					Observed $\sigma_H \times B_\tau$ (pb)
	-2 σ	-1 σ	Median	+1 σ	+2 σ	
90	329.2	429.2	621.9	862.9	999.1	394.7
120	30.1	41.6	59.8	82.0	116.6	86.5
130	20.7	27.6	40.5	55.6	79.4	59.9
160	10.3	13.2	19.0	26.2	35.8	28.3
200	6.3	8.3	11.2	15.8	20.2	16.4
250	4.0	5.6	7.6	10.6	14.5	12.9
300	2.9	4.0	5.7	7.8	11.1	9.4
Visible Mass 95% CL Upper Limit (pb)						
	90	120	130	160	200	250
90	376.2	523.3	688.2	980.9	998.8	573.8
120	37.0	52.1	75.4	109.2	164.1	82.6
130	26.2	35.9	52.2	74.6	117.5	64.2
160	14.3	18.3	25.1	35.2	55.1	41.2
200	8.9	11.9	16.6	22.4	32.8	31.1
250	5.9	8.1	11.5	15.9	22.3	18.1
300	4.2	5.8	8.4	11.7	15.9	10.8

Table 9.3: Expected 95% CL upper limit bands and the observed limit using the Bayesian prescription. The limit is computed using both the SVfit mass (top) as well as the visible mass (bottom) as the search observable. Use of the SVfit mass significantly improves the strength of the limit considerably.

Higgs State	Included when		
	$m_{A^0} < 130$ GeV/c ²	$m_{A^0} = 130$ GeV/c ²	$m_{A^0} > 130$ GeV/c ²
A^0	yes	yes	yes
H^0	yes	yes	no
h^0	no	yes	yes

Table 9.4: Logic for determining the MSSM Higgs cross section for a given mass of the CP-odd A^0 Higgs. In some regions of parameter space, the contributions of one of the CP-event Higgs particles is ignored.

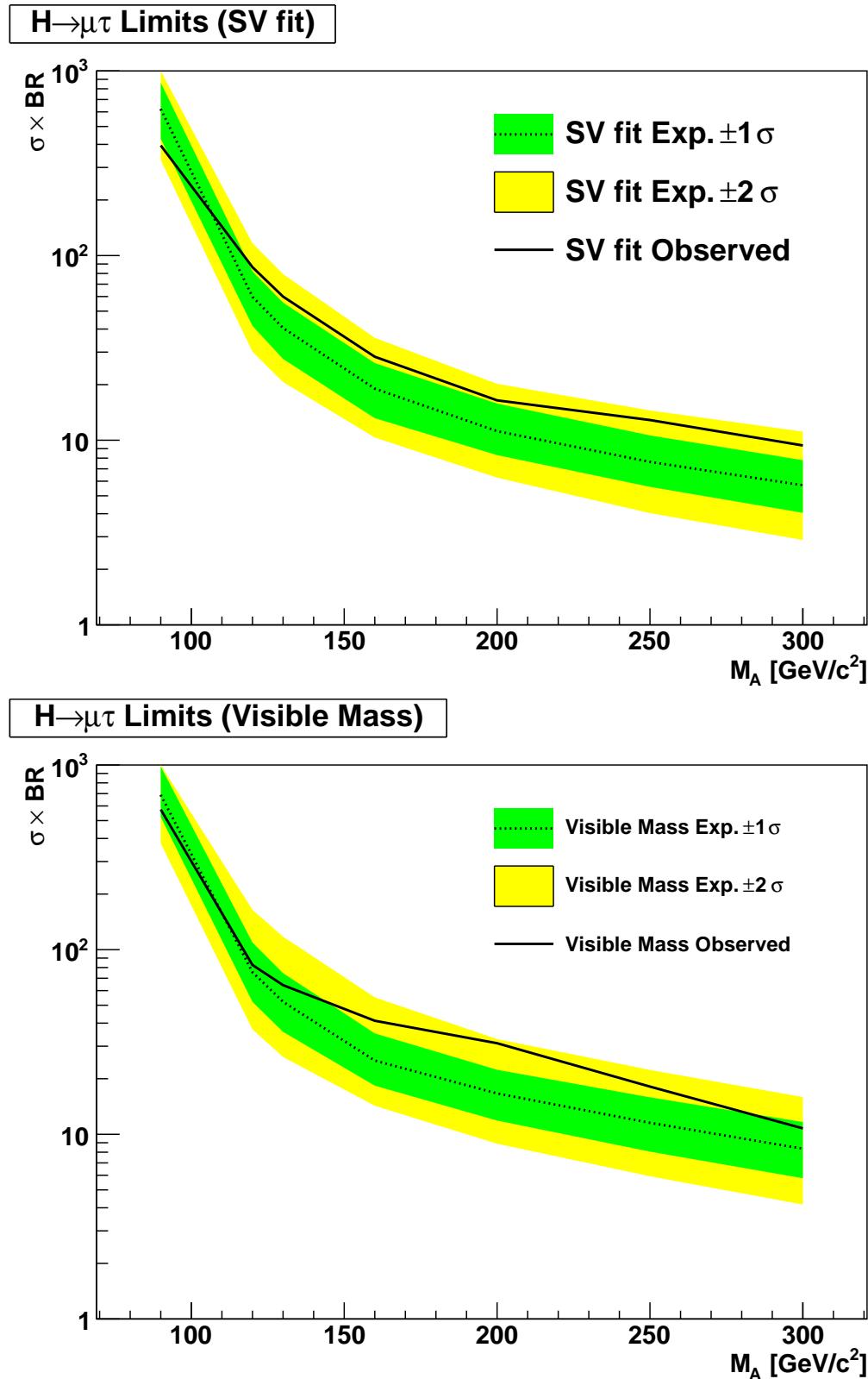


Figure 9.4: Observed and expected limits on the cross section times branching ratio of a Higgs boson versus Higgs mass. The top plot gives the limit computed using the SVfit mass as the observable, the bottom plot gives the limit computed using the visible mass. The dashed line gives the nominal expected limit. The green and yellow bands give the $+1$ and $+2$ standard deviations on the expected limit.

Figure 9.5: Region of MSSM $\tan \beta - m_{A^0}$ parameter space excluded by this analysis. THIS PLOT IS NOT DONE YET

2347 Conclusions

2348 This analysis has presented a search for MSSM Higgs bosons in the 2010 7 TeV CMS data
 2349 set. Two new experimental methods, the TaNC tau identification algorithm, and the SVfit
 2350 mass reconstruction method have been introduced in this thesis. Both methods increased
 2351 the sensitive of the Higgs search. The search was performed using 36 pb^{-1} of data. The
 2352 expected event yield from standard model sources is 577 events. In total, 573 events were
 2353 selected; the observed is compatible with the standard model. No signal-like excess of events
 2354 is observed. We set an upper limit on the production of Higgs bosons, and interpret this
 2355 limit in the context of the MSSM.

2356 The analysis presented in this thesis was part of a larger study [1] performed by the
 2357 CMS collaboration searching for the MSSM Higgs boson decaying to tau leptons. The CMS
 2358 analysis used three channels, the $H \rightarrow \tau\tau \rightarrow e - \tau_h$, $H \rightarrow \tau\tau \rightarrow e - \mu$, and the $\mu - \tau_h$
 2359 channel. The $\mu - \tau_h$ channel search presented in this thesis is very similar to the CMS result.
 2360 While not as pure as the $\mu - \tau$ channel, the inclusion of the high-statistics $e - \tau$ channel
 2361 increases the sensitivity of the CMS analysis. The $e - \mu$ channel has low statistics, but is
 2362 not sensitive to the systematic uncertainty on the hadronic tau identification. The region of
 2363 the MSSM parameter space excluded by combined CMS result is illustrated in Figure 9.3.
 2364 At the time of this writing, the CMS result described in [1] sets the most stringent limits
 2365 on the MSSM using a direct search.

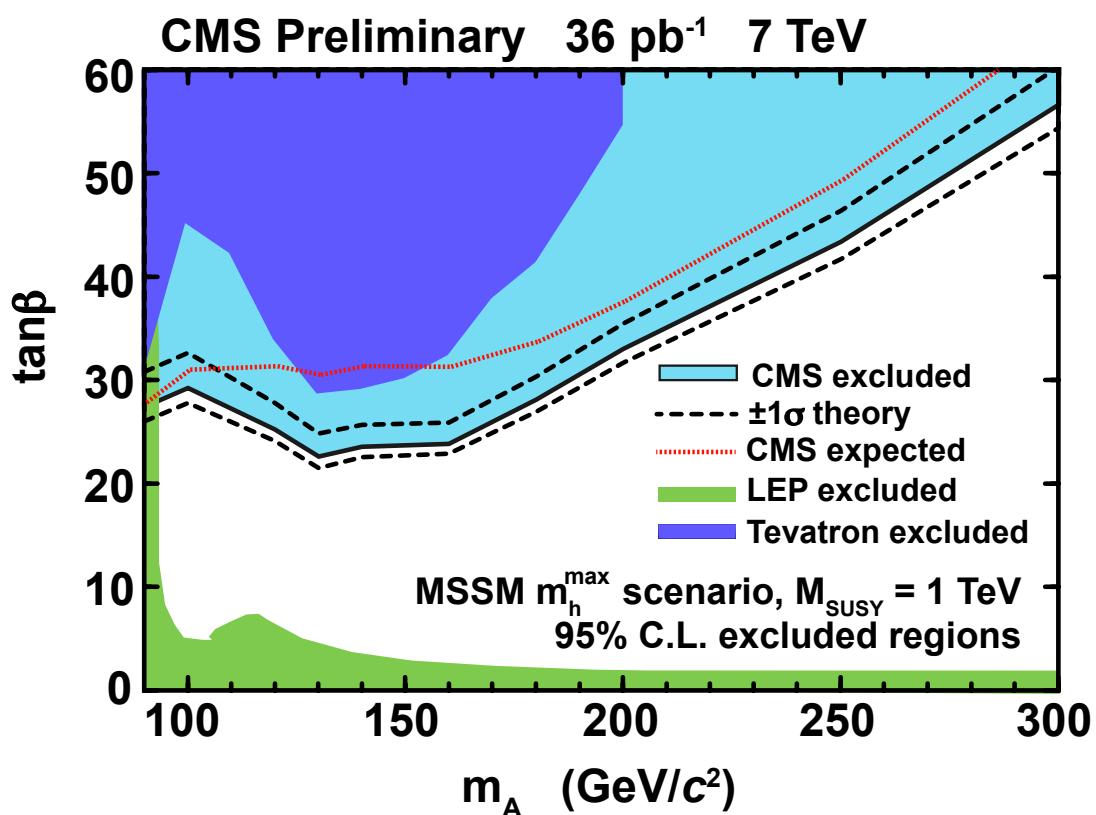


Figure 9.6: Region of MSSM $\tan \beta - m_{A^0}$ parameter space excluded by the CMS combined analysis [1].

2366 Bibliography

- 2367 [1] CMS Collaboration, “Search For Neutral MSSM Higgs Boson Production via Decays
2368 to Tau Pairs in pp Collisions at $\sqrt{s} = 7$ TeV”, *to be published* (2011).
- 2369 [2] J. Goldstone, “Field Theories with Superconductor Solutions”, *Nuovo Cim.* **19**
2370 (1961) 154–164. doi:10.1007/BF02812722.
- 2371 [3] J. Goldstone, A. Salam, and S. Weinberg, “Broken Symmetries”, *Phys. Rev.* **127**
2372 (Aug, 1962) 965–970. doi:10.1103/PhysRev.127.965.
- 2373 [4] F. Englert and R. Brout, “Broken Symmetry and the Mass of
2374 Gauge Vector Mesons”, *Phys. Rev. Lett.* **13** (Aug, 1964) 321–323.
2375 doi:10.1103/PhysRevLett.13.321.
- 2376 [5] P. W. Higgs, “Broken Symmetries and the Masses of Gauge Bosons”, *Phys. Rev.*
2377 *Lett.* **13** (Oct, 1964) 508–509. doi:10.1103/PhysRevLett.13.508.
- 2378 [6] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, “Global Conservation
2379 Laws and Massless Particles”, *Phys. Rev. Lett.* **13** (Nov, 1964) 585–587.
2380 doi:10.1103/PhysRevLett.13.585.
- 2381 [7] S. Glashow, “Partial Symmetries of Weak Interactions”, *Nucl.Phys.* **22** (1961) 579–
2382 588. doi:10.1016/0029-5582(61)90469-2.
- 2383 [8] S. Weinberg, “A Model of Leptons”, *Phys.Rev.Lett.* **19** (1967) 1264–1266.
2384 doi:10.1103/PhysRevLett.19.1264.
- 2385 [9] A. Salam, “Weak and Electromagnetic Interactions”, Originally printed in
2386 *Svartholm: Elementary Particle Theory, Proceedings Of The Nobel Symposium
2387 Held 1968 At Lerum, Sweden*, Stockholm 1968, 367–377.
- 2388 [10] D. Griffiths, “Introduction to Elementary Particles”. Wiley-VCH, 2004.
- 2389 [11] E. Fermi, “An attempt of a theory of beta radiation. 1”, *Z. Phys.* **88** (1934) 161–
2390 177. doi:10.1007/BF01351864.
- 2391 [12] S. M. T. Morii, C.S. Lim, “The Physics of the Standard Model and Beyond”. World
2392 Scientific, 2004.
- 2393 [13] UA1 Collaboration, “Experimental observation of isolated large transverse energy
2394 electrons with associated missing energy at $\sqrt{s} = 540$ GeV”, *Phys. Lett.* **B122**
2395 (1983) 103–116.
- 2396 [14] UA2 Collaboration, “Observation of single isolated electrons of high transverse mo-
2397 mentum in events with missing transverse energy at the CERN $\bar{p}p$ collider”, *Phys.*
2398 *Lett.* **B122** (1983) 476–485. doi:10.1016/0370-2693(83)91605-2.

- 2399 [15] UA1 Collaboration, “Experimental observation of lepton pairs of invariant mass
 2400 around $95 \text{ GeV}/c^2$ at the CERN SPS collider”, *Phys. Lett.* **B126** (1983) 398–410.
 2401 doi:10.1016/0370-2693(83)90188-0.
- 2402 [16] UA2 Collaboration, “Evidence for $Z^0 \rightarrow e^+e^-$ at the CERN $\bar{p}p$ collider”, *Phys. Lett.*
 2403 **B129** (1983) 130–140. doi:10.1016/0370-2693(83)90744-X.
- 2404 [17] S. P. Martin, “A Supersymmetry Primer”, *arXiv hep-ph* (sep, 1997). 128 pages.
 2405 Version 5 (December 2008) contains a change in convention that flips the signs of
 2406 sigma and sigmabar matrices. It also contains a total of about 2 pages of updates,
 2407 mostly on supersymmetry breaking issues. Errata and a version with larger type (12
 2408 pt, 142 pages) can be found at <http://zippy.physics.niu.edu/primer.html>.
- 2409 [18] CERN Computing Newsletter.
- 2410 [19] D. L. Rainwater, D. Zeppenfeld, and K. Hagiwara, “Searching for $H \rightarrow \tau^+\tau^-$
 2411 in weak boson fusion at the LHC”, *Phys. Rev.* **D59** (1999) 014037, arXiv:hep-
 2412 ph/9808468. doi:10.1103/PhysRevD.59.014037.
- 2413 [20] Particle Data Group Collaboration, “Review of particle physics”, *J. Phys.* **G37**
 2414 (2010) 075021. doi:10.1088/0954-3899/37/7A/075021.
- 2415 [21]
- 2416 [22] LHC Higgs Cross Section Working Group Collaboration, “Handbook of LHC Higgs
 2417 Cross Sections: 1. Inclusive Observables”, arXiv:1101.0593.
- 2418 [23] M. Carena, S. Heinemeyer, C. Wagner et al., “MSSM Higgs boson searches at the
 2419 Tevatron and the LHC: Impact of different benchmark scenarios”, *The European
 2420 Physical Journal C - Particles and Fields* **45** (2006) 797–814. 10.1140/epjc/s2005-
 2421 02470-y.
- 2422 [24] CMS Collaboration, “The CMS experiment at the CERN LHC”, *JINST* **3** (2008)
 2423 S08004.
- 2424 [25] CMS Collaboration, “Measurement of Momentum Scale and Resolution using Low-
 2425 mass Resonances and Cosmic-Ray Muons”, *CMS PAS CMS-PAS-TRK-10-004*
 2426 (2010).
- 2427 [26] CMS Collaboration, “Tracking and Primary Vertex Results in First 7 TeV Collisions”,
 2428 *CMS PAS CMS-PAS-TRK-10-005* (2010).
- 2429 [27] G. L. Bayatian et al., “CMS Physics Technical Design Report Volume I: Detector
 2430 Performance and Software”. Technical Design Report CMS. CERN, Geneva, 2006.
- 2431 [28] CMS Collaboration, “CMS Strategies for tau reconstruction and identification using
 2432 particle-flow techniques”, *CMS PAS CMS-PAS-PFT-08-001* (2008).
- 2433 [29] CMS Collaboration, “Particle-Flow Event Reconstruction in CMS and Performance
 2434 for Jets, Taus, and MET”, *CMS PAS CMS-PAS-PFT-09-001* (2009).
- 2435 [30] A. Hoecker et al., “TMVA - Toolkit for Multivariate Data Analysis”, *ArXiv Physics
 2436 e-prints* (March, 2007) arXiv:arXiv:physics/0703039.

- 2437 [31] A. Kolmogorov, “On the representation of continuous functions of several variables
 2438 by superposition of continuous functions of one variable and addition”, *Doklady
 2439 Akademii Nauk SSSR* **114** (1957).
- 2440 [32] J. C. *et al.*, “Size of signal cones and isolation rings in the CMS tau identification
 2441 algorithm”, *CMS Note* **2008/026** (2008).
- 2442 [33] M. Bachtis, S. Dasu, and A. Savin, “Prospects for measurement of $\sigma(pp \rightarrow Z) \cdot B(Z \rightarrow \tau^+\tau^-)$ with CMS in pp Collisions at $\sqrt{s} = 7$ TeV”, *CMS Note*
 2443 **2010/082** (2010).
- 2444 [34] CMS Collaboration, “Study of tau reconstruction algorithms using pp collisions data
 2445 collected at $\sqrt{s} = 7$ TeV”, *CMS PAS CMS-PAS-PFT-10-004* (2010).
- 2446 [35] G. P. S. M. Cacciari and G. Soyez, “The anti-kt jet clustering algorithm”, *JHEP*
 2447 **04** (2008) 063, arXiv:0802.1189.
- 2448 [36] CMS Collaboration, “Commissioning of the Particle–Flow reconstruction in
 2449 Minimum–Bias and Jet Events from pp Collisions at 7 TeV”, *CMS PAS PFT-10-*
 2450 **002** (2010).
- 2451 [37] S. M. T. Sjöstrand and P. Skands, “PYTHIA 6.4 Physics and Manual”, 2000.
- 2452 [38] S. Jadach, Z. Was, R. Decker *et al.*, “The Tau Decay Library Tauola: Version 2.4”,
 2453 *Comput. Phys. Commun.* **76** (1993) 361.
- 2454 [39] S. Agostinelli, J. Allison, K. Amako *et al.*, “G4—a simulation toolkit”, *Nuclear
 2455 Instruments and Methods in Physics Research Section A: Accelerators, Spec-
 2456 trometers, Detectors and Associated Equipment* **506** (2003), no. 3, 250 – 303.
 2457 doi:10.1016/S0168-9002(03)01368-8.
- 2458 [40] CDF Collaboration, “Search for MSSM Higgs decaying to τ pairs in $p\bar{p}$ collision at
 2459 $\sqrt{s} = 1.96$ TeV at CDF”, *Phys. Rev. Lett.* **96** (2006).
- 2460 [41] CMS Collaboration, “CMS technical design report, volume II: Physics performance”,
 2461 *J. Phys.* **G34** (2007) 995–1579. doi:10.1088/0954-3899/34/6/S01.
- 2462 [42] L. Bianchini, “Improved Collinear Approximation for VBF $H \rightarrow \tau\tau \rightarrow 3\nu + \ell + \tau_{had}$ ”,
 2463 *CMS Note* **2010/226** (2010).
- 2464 [43] B. K. Bullock, K. Hagiwara, and A. D. Martin, “Tau Polarization And Its Corre-
 2465 lations As A Probe Of New Physics”, *Nucl. Phys.* **B 395** (1993) 499.
- 2466 [44] CMS Collaboration, “Measurements of Inclusive W and Z Cross Sections in pp Col-
 2467 lisions at $\sqrt{s} = 7$ TeV”, *CMS PAS EWK-10-002* (2010).
- 2468 [45] CMS Collaboration, “MET Performance in Events Containing Electroweak Bosons
 2469 from pp Collisions at $\sqrt{s} = 7$ TeV”, *CMS PAS JME-10-005* (2010).
- 2470 [46] C. C. Almenar, “Search for the neutral MSSM Higgs bosons in the $\tau\tau$ decay channels
 2471 at CDF Run II”. PhD thesis, Departament de Física Atomica, Molecular i Nuclear
 2472 (Universitat de València and IFIC (CSIC - Universitat de València), 2008.
- 2473

- 2474 [47] L. Lusito and C. Veelken, “Estimation of Background contributions to Tau analyses
2475 via Template Fitting”, *CMS Note* **2010/088** (2010).
- 2476 [48] CMS Collaboration, “Performance of tau reconstruction algorithms in 2010 data col-
2477 lected with CMS”, *CMS PAS TAU-11-001* (2011).
- 2478 [49] D. Jang, “Search for MSSM Higgs decaying to τ pairs in $p\bar{p}$ collision at $\sqrt{s} =$
2479 1.96 TeV at CDF”, *Ph.D. Thesis, Rutgers University* (2006).
- 2480 [50] J. Conway, E. Friis, and C. Veelken, “Measurement of the $Z/\gamma^* \rightarrow \tau^+\tau^-$ Production
2481 Cross-section in the $\mu + \tau_{had}$ final state using the HPS+TaNC Tau id. algorithm”,
2482 *CMS Note* **2011/021** (2011).
- 2483 [51] J. Conway, E. Friis, M. Squires et al., “The Tau Neural Classifier algorithm: tau
2484 identification and decay mode reconstruction using neural networks”, *CMS Note*
2485 **2010/099** (2010).
- 2486 [52] T. Früboes and M. Zeise, “The TauAnalysis/MCEmbeddingTools Package”.
2487 <https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideTauAnalysisMCEmbeddingTools>.
- 2488 [53] CMS Collaboration, “Measurement of the Inclusive $Z \rightarrow \tau\tau$ Cross Section in pp Col-
2489 lisions at $\sqrt{s} = 7$ TeV”, *to be published* (2011).
- 2490 [54] S. Bolognesi, M. A. Borgia, R. Castello et al., “Calibration of track momentum using
2491 dimuon resonances in CMS”, *CMS Note* **2010/059** (2010).
- 2492 [55] CMS Collaboration, “Jet Energy Corrections determination at $\sqrt{s} = 7$ TeV”, *CMS
2493 PAS JME-10-010* (2010).
- 2494 [56] G. Bauer et al., “Modeling of $W \rightarrow \ell\nu$ MET with Boson Recoil”, *CMS Note*
2495 **2010/332** (2010).
- 2496 [57] G. Cerati et al., “Search for MSSM neutral Higgs $\rightarrow \tau^+\tau^-$ Production using the
2497 TaNC Tau id. algorithm”, *CMS Note* **2010/460** (2010).
- 2498 [58] CMS Collaboration, “Absolute luminosity normalization”, *CMS Detector Perfor-
2499 mance Summary* **CMS-DPS-2011-002** (2008).
- 2500 [59] J. Alcaraz. <https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideEWKUtilities>.
- 2501 [60] P. M. Nadolsky et al., “Implications of CTEQ global analysis for collider observ-
2502 ables”, *Phys. Rev.* **D 78** (2008) 013004, arXiv:0802.0007.
- 2503 [61] PDF4LHC Working Group. <http://www.hep.ucl.ac.uk/pdf4lhc/PDF4LHCrecos.pdf>.
- 2504 [62] R. C. G. D. Bourilkov and M. R. Whalley, “LHAPDF: PDF use from the Tevatron
2505 to the LHC”, arXiv:0605.0240.
- 2506 [63] LHC Higgs Cross Section Working Group Collaboration *arXiv:hep-ex/1101.0593*
2507 (2011).