

**Search for Neutral MSSM Higgs Bosons Decaying to Pairs of  $\tau$  Leptons at  $\sqrt{s} = 7$  TeV**

By

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

**DISSERTATION**

Submitted in partial satisfaction of the requirements for the degree of

**DOCTOR OF PHILOSOPHY**

in

Physics

in the

**OFFICE OF GRADUATE STUDIES**

of the

**UNIVERSITY OF CALIFORNIA**

**DAVIS**

Approved:

---

Professor John Conway (Chair)

---

Professor Robin Erbacher

---

Professor Mani Tripathi

Committee in Charge

2011

for Alvin A Klose

## 3 Abstract

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

19

## Acknowledgments

20

Hooray for everybody.

# Table of Contents

21	<b>Introduction</b>	1
22	<b>1 The Standard Model and Beyond</b>	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 . . . . .	9
27	1.1.4 The Higgs Mechanism . . . . .	11
28	1.1.5 Electroweak Unification . . . . .	13
29	1.1.6 Quantum Chromodynamics . . . . .	18
30	1.2 Beyond the Standard Model . . . . .	20
31	1.2.1 The Hierarchy Problem . . . . .	21
32	1.2.2 Supersymmetry . . . . .	22
33	1.2.3 The Minimal Supersymmetric Extension to the Standard Model . . . . .	22
34	1.3 Searches for the Higgs Boson . . . . .	25
35	1.3.1 Standard Model Higgs Boson Phenomenology . . . . .	25
36	1.3.2 MSSM Higgs Boson Phenomenology . . . . .	31
37	1.3.3 Results from LEP and Tevatron . . . . .	32
38	1.4 The Physics of the Tau Lepton . . . . .	37
39		
40	<b>2 The Compact Muon Solenoid Experiment</b>	39
41	2.1 The Large Hadron Collider . . . . .	40
42	2.2 Solenoid Magnet . . . . .	40
43	2.3 Charged Particle Tracking Systems . . . . .	42
44	2.4 Electromagnetic Calorimeter . . . . .	45
45	2.5 Hadronic Calorimeter . . . . .	48
46	2.6 Muon System . . . . .	49
47	2.7 Trigger System . . . . .	51
48		
49	<b>3 Tau Identification: The Tau Neural Classifier</b>	54
50	3.1 Geometric Tau Identification Algorithms . . . . .	55
51	3.2 Decay Mode Tau Identification: Motivation . . . . .	55
52	3.3 The Tau Neural Classifier . . . . .	56
53	3.3.1 Decay Mode Reconstruction . . . . .	57
54	3.3.2 Photon Merging . . . . .	57
55	3.3.3 Quality Requirements . . . . .	58
56	3.3.4 Performance . . . . .	59
57	3.3.5 Neural Network Classification . . . . .	63
58	3.3.6 Neural Network Training . . . . .	63

58	3.3.7 Discriminants . . . . .	67
59	3.3.8 Neural Network Performance . . . . .	68
60	3.4 Summary . . . . .	75
61	3.5 HPS+TaNC: A Hybrid Algorithm . . . . .	76
62	3.5.1 Decay Mode Reconstruction . . . . .	76
63	3.5.2 Hadronic Tau Discrimination . . . . .	80
64	3.6 Electron and Muon Rejection . . . . .	81
65	<b>4 Mass Reconstruction: The Secondary Vertex Fit</b>	<b>84</b>
66	4.1 Existing Mass Reconstruction Algorithms . . . . .	84
67	4.2 The Secondary Vertex Fit . . . . .	86
68	4.3 Parametrization of Tau Decays . . . . .	87
69	4.4 Likelihood for Tau Decays . . . . .	88
70	4.4.1 Likelihood for Reconstructed $E_T^{\text{miss}}$ . . . . .	89
71	4.4.2 Likelihood for Tau Transverse Momentum Balance . . . . .	90
72	4.4.3 Secondary Vertex Information . . . . .	92
73	4.5 Performance . . . . .	94
74	<b>5 Analysis Selections</b>	<b>97</b>
75	5.1 High Level Trigger . . . . .	97
76	5.2 Particle Identification . . . . .	98
77	5.2.1 Muons . . . . .	98
78	5.2.2 Hadronic Taus . . . . .	99
79	5.2.3 Missing Transverse Energy . . . . .	100
80	5.3 Event Selections . . . . .	100
81	<b>6 Data-Driven Background Estimation</b>	<b>105</b>
82	6.1 Background-Enriched Control Regions . . . . .	106
83	6.2 The Fake-rate Method . . . . .	107
84	6.2.1 Parameterization of Fake-rates . . . . .	110
85	6.2.2 Measurement of Fake-rates . . . . .	111
86	6.2.3 Application of Fake-rates . . . . .	111
87	6.2.4 “Simple” Weight Method . . . . .	112
88	6.2.5 “CDF-type” Weights . . . . .	114
89	6.2.6 k-Nearest Neighbor Fake-rate Calculation . . . . .	117
90	6.2.7 Results of Fake-Rate Background Estimation . . . . .	118
91	6.3 Template Method . . . . .	120
92	<b>7 Monte Carlo Corrections</b>	<b>133</b>
93	7.1 Muon Identification Efficiency . . . . .	133
94	7.2 Hadronic Tau Identification Efficiency . . . . .	136
95	7.3 Muon and Tau Momentum Scale . . . . .	138
96	7.4 Missing Transverse Energy Correction . . . . .	139
97	7.5 Pile-up Event Weighting . . . . .	140
98	<b>8 Systematics and Limit Extraction</b>	<b>142</b>
99	8.1 Signal Normalization Uncertainties . . . . .	144
100	8.2 Background Normalization Uncertainties . . . . .	144

101	8.3 Shape Uncertainties . . . . .	145
102	8.4 Theory Uncertainties . . . . .	146
103	8.5 Limit Extraction Method . . . . .	146
104	<b>9 Results</b>	<b>153</b>
105	9.1 Selected Events . . . . .	153
106	9.2 Limits on Higgs Boson Production . . . . .	154
107	9.3 Interpretation in the MSSM . . . . .	155
108	<b>Conclusions</b>	<b>164</b>
109	<b>Bibliography</b>	<b>164</b>

# 110 List of Figures

111	1.1	Fermi contact interaction diagram . . . . .	8
112	1.2	Muon decaying through intermediate gauge boson . . . . .	8
113	1.3	QCD Feynman Diagrams . . . . .	19
114	1.4	Loop corrections to Higgs boson mass . . . . .	21
115	1.5	Higgstrahlung production diagram at $e^+e^-$ colliders . . . . .	26
116	1.6	Gluon fusion Higgs boson production diagram . . . . .	27
117	1.7	Vector boson fusion Higgs boson production diagram . . . . .	27
118	1.8	Parton luminosity comparison of the LHC and Tevatron . . . . .	28
119	1.9	SM Higgs boson cross sections at the LHC . . . . .	29
120	1.10	SM Higgs boson branching fractions . . . . .	29
121	1.11	Cross sections of interest at hadron colliders . . . . .	30
122	1.12	MSSM Higgs boson production with association $b$ -quarks . . . . .	31
123	1.13	MSSM Higgs boson cross sections at the LHC . . . . .	32
124	1.14	LEP SM Higgs boson limit plot . . . . .	34
125	1.15	Tevatron low mass standard model Higgs boson limit plot . . . . .	35
126	1.16	Tevatron high mass standard model Higgs boson limit plot . . . . .	36
127	1.17	LEP MSSM exclusion limits . . . . .	36
128	1.18	Tevatron MSSM exclusion limits . . . . .	37
129	2.1	Schematic drawings of the CMS detector . . . . .	41
130	2.2	Material budget of the CMS tracker . . . . .	44
131	2.3	Momentum and impact parameter resolutions of CMS tracker . . . . .	45
132	2.4	Energy resolution of the CMS ECAL . . . . .	47
133	2.5	Muon system material budget and identification efficiency . . . . .	50
134	3.1	Visible invariant mass of $\tau$ lepton decay products . . . . .	56
135	3.2	Invariant mass photon pairs in reconstructed $\pi^0$ mesons . . . . .	58
136	3.3	Neutral energy fraction in visible $\tau$ decays . . . . .	59
137	3.4	Tau decay mode reconstruction performance . . . . .	61
138	3.5	Kinematic dependence of decay mode reconstruction . . . . .	62
139	3.6	Neural network over-training validation plots . . . . .	66
140	3.7	Kinematic weighting of training sample . . . . .	67
141	3.8	Neural network output in each decay mode . . . . .	70
142	3.9	Performance curves for the neural networks used in the TaNC . . . . .	71
143	3.10	Tau Neural Classifier performance curves for different $p_T$ ranges . . . . .	72
144	3.11	Tau Neural Classifier transformation performance . . . . .	74
145	3.12	Transformed neural network output . . . . .	75
146	3.13	Tau Neural Classifier performance comparison . . . . .	77
147	3.14	Tau Neural Classifier kinematic performance . . . . .	78

148	3.15 Invariant mass distribution of PF photon pairs . . . . .	80
149	4.1 Coordinate system of the SVfit . . . . .	89
150	4.2 Effect of $p_T$ -balance term on SVfit performance . . . . .	92
151	4.3 Effect of the visible $p_T$ requirements on muon and hadronic $\tau$ decays . . . . .	93
152	4.4 Effect of the visible $p_T$ requirements for $Z$ and Higgs boson events . . . . .	94
153	4.5 Comparison of SVfit with the collinear approximation algorithm . . . . .	95
154	4.6 Comparison of SVfit with the visible mass observable . . . . .	96
155	5.1 Distributions of $M_T$ and muon isolation discriminants . . . . .	102
156	5.2 Reconstruction and distribution of $P_\zeta$ discriminant . . . . .	103
157	6.1 Visible mass distribution of the backgrounds in the signal and control regions	108
158	6.2 SVfit mass distribution of the backgrounds in the signal and control regions	109
159	6.3 $p_T$ and $\eta$ dependency of tau ID performance . . . . .	123
160	6.4 Comparison of fake-rate contribution from genuine taus in the simple and CDF methods . . . . .	125
161	6.5 Muon transverse momentum in the fake-rate method . . . . .	126
162	6.6 Tau transverse momentum in the fake-rate method . . . . .	127
163	6.7 Visible mass in the fake-rate method . . . . .	128
164	6.8 $k$ -Nearest Neighbor classifier example . . . . .	129
165	6.9 Comparison of visible mass and SVfit mass . . . . .	130
166	6.10 Comparison of background shapes in the signal and control regions . . . . .	131
167	6.11 Visible mass distribution in the final fit of the template method . . . . .	132
168		
169	7.1 Tag-probe muon isolation method . . . . .	135
170	7.2 Muon isolation correction factors . . . . .	137
171	7.3 Measurement of hadronic tau identification efficiency . . . . .	138
172	7.4 Distribution of number of reconstructed primary vertices per event . . . . .	141
173		
174	9.1 Transverse momentum distributions of muon and tau in the final selected events . . . . .	157
175	9.2 Distributions of the $\eta$ and $\phi$ of the muon and tau candidates in the final selected events . . . . .	158
176	9.3 Distributions of final selected events . . . . .	159
177	9.4 Observed and expected limits on Higgs boson $\sigma \times \text{BR}$ . . . . .	162
178	9.5 Excluded regions of MSSM $\tan \beta - m_{A^0}$ parameter space . . . . .	163
179	9.6 CMS combined exclusion of MSSM $\tan \beta - m_{A^0}$ parameter space . . . . .	165
180		

## 181 List of Tables

182	1.1	Chiral supermultiplets in the MSSM . . . . .	23
183	1.2	Gauge supermultiplets in the MSSM . . . . .	23
184	1.3	Higgs boson search channels at LEP . . . . .	33
185	1.4	Decay modes of the $\tau$ lepton . . . . .	38
186	3.1	Decay mode performance – naive reconstruction . . . . .	60
187	3.2	Decay mode performance – TaNC reconstruction . . . . .	60
188	3.3	Neural network training event statistics . . . . .	64
189	3.4	Variables used in the different TaNC neural networks . . . . .	83
190	5.1	High Level Trigger paths used to select $\mu + \tau_h$ events . . . . .	98
191	5.2	Analysis backgrounds that include fake taus . . . . .	100
192	5.3	Event selection summary . . . . .	104
193	6.1	Criteria used to select background–enriched control regions . . . . .	121
194	6.2	Comparison of background control region yields in data and the prediction from simulation . . . . .	122
195	6.3	Fake–rate “simple” method closure test results . . . . .	124
196	6.4	Fake–rate “CDF” method closure test results . . . . .	125
197	6.5	Fake–rate method results . . . . .	129
198	6.6	Fake–rate method predicted yields in like–sign control region . . . . .	130
199	6.7	Background yields measured using the template method . . . . .	130
200	7.1	Muon trigger, identification, and isolation correction factors . . . . .	136
201	8.1	Effect of normalization uncertainties on signal efficiency times acceptance .	147
202	8.2	Correspondence of confidence levels and $2\Delta \ln \mathcal{L}$ intervals . . . . .	150
203	8.3	Summary of systematic uncertainties . . . . .	152
204	9.1	Final analysis yield and background expectations . . . . .	154
205	9.2	Expected signal yields at $\tan \beta = 30$ . . . . .	160
206	9.3	Expected and observed 95% CL $\sigma \times \text{BR}$ upper limits . . . . .	161
207	9.4	Contributions of different MSSM Higgs boson types at different $m_{A^0}$ . . . . .	161
208			

## 209 Introduction

210 This thesis describes a search for the Higgs boson, a new particle predicted by the standard  
211 model of particle physics. The search is optimized for a particular extension of the stan-  
212 dard model, a theory called the minimal supersymmetric extension to the standard model  
213 (MSSM). The analysis uses the 2010 dataset from the Compact Muon Solenoid (CMS) ex-  
214 periment, which contains  $36 \text{ pb}^{-1}$  of integrated luminosity at a center of mass energy of  
215 7 TeV. The Higgs boson is hypothesized to be the catalyst of electroweak symmetry break-  
216 ing, the phenomenon strongly believed to impart mass to particles that form our natural  
217 world.

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

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

233 The studies presented herein were part of a larger effort at CMS to search for an MSSM  
234 Higgs boson decaying to tau lepton pairs. In addition to the  $\mu - \tau_h$  channel described in  
235 this thesis, final states with an electronic and hadronic tau decay ( $e - \tau$ ) and electronic and  
236 muonic ( $e - \mu$ ) were considered. The combination of all three was used to set a limit on  
237 the MSSM [1]. This result has recently been accepted for publication in Physical Review  
238 Letters B. At the time of this writing, the CMS analysis sets the world's strongest limit on  
239 the MSSM Higgs boson using a direct search.

## Chapter 1

# The Standard Model and Beyond

(ch:theory)

### 1.1 The Standard Model

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

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

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

### 268 1.1.1 Quantum Electrodynamics and Gauge Invariance

QEDandGaugeInvariance  
269 The theory of QED is a modern extension of Maxwell's theory of electromagnetism, describ-  
 270 ing the interaction of matter with light. The development of QED is a result of efforts to  
 271 develop a quantum mechanical formulation of electromagnetism compatible with the the-  
 272 ory of special relativity. QED is a *gauge* theory, which means that the physical observables  
 273 are invariant under local gauge transformations. Requiring local gauge invariance gives rise  
 274 to a "gauge" field, which can be interpreted as particles that are exchanged during an  
 275 interaction.

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

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

$$283 p^\mu p_\mu - m^2 c^2 = 0. \quad (1.1) \quad \text{eq:EnergyPRelat}$$

283 Dirac sought to express this relationship in the framework of quantum mechanics by apply-  
 284 ing the transformation

$$285 p_\mu \rightarrow i\hbar\partial_\mu \quad (1.2) \quad \text{eq:QuantizeMom}$$

285 to equation Equation 1.1, but with the requirement that the resulting equation be first

286 order in time.<sup>1</sup> To achieve this, Dirac factorized Equation 1.1 into

$$( \gamma^\kappa p_\kappa + mc ) ( \gamma^\mu p_\mu - mc ) = 0, \quad (1.3) \quad \text{eq:DiracEquation}$$

287 where  $\gamma^\mu$  is a set of four  $4 \times 4$  matrices referred to as the Dirac matrices. The equation of  
288 motion is obtained by choosing either term (they are equivalent) from the left hand side of  
289 Equation 1.3 and making the substitution in Equation 1.2.

$$i\hbar\gamma^\mu\partial_\mu\psi - mc\psi = 0. \quad (1.4) \quad \text{eq:DiracEquation}$$

290 The solutions  $\psi$  of the Dirac equation are called “Dirac spinors,” and represent the quantum  
291 mechanical state of spin 1/2 particles.

292 The Lagrangian corresponding to the Dirac equation (1.4) is

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

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

$$\psi' \rightarrow e^{i\theta}\psi. \quad (1.6) \quad \text{eq:U1GaugeTrans}$$

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

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

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

---

<sup>1</sup>A detailed discussion of this topic is available in [10].

302 locally gauge invariant:

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

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

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

305 This term can be interpreted as the coupling between the particle and the gauge boson (force  
306 carrier) fields. The coupling is proportional to the constant  $e$ , which is associated with the  
307 electric charge. This is consistent with the experimental observation that particles with zero  
308 electric charge do not interact electromagnetically with each other. In this interpretation,  
309 the electromagnetic force between two charged particles is caused by the exchange of gauge  
310 bosons (photons). The existence of this “minimal coupling” is *required* if the Lagrangian  
311 is to satisfy local gauge invariance. The addition of a term with the gauge Field Strength  
312 Tensor to represent the kinetic term of the gauge (photon) field yields the QED Lagrangian:

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

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

### 318 1.1.2 The Weak Interactions

319 <sec:WeakInteractions> The theory of Weak Interactions was created to describe the physics of radioactive decay.  
320 The first formulation of the theory was done by Fermi [11] to explain the phenomenon of the  
321  $\beta$  decay of neutrons. The initial theory was a four-fermion “contact” theory. In a contact  
322 theory, all four fermions come involved in the  $\beta$ -decay are connected at a single vertex. The

323 Fermi theory Hamiltonian for the  $\beta$ -decay of a proton is then [12]

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

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

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

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

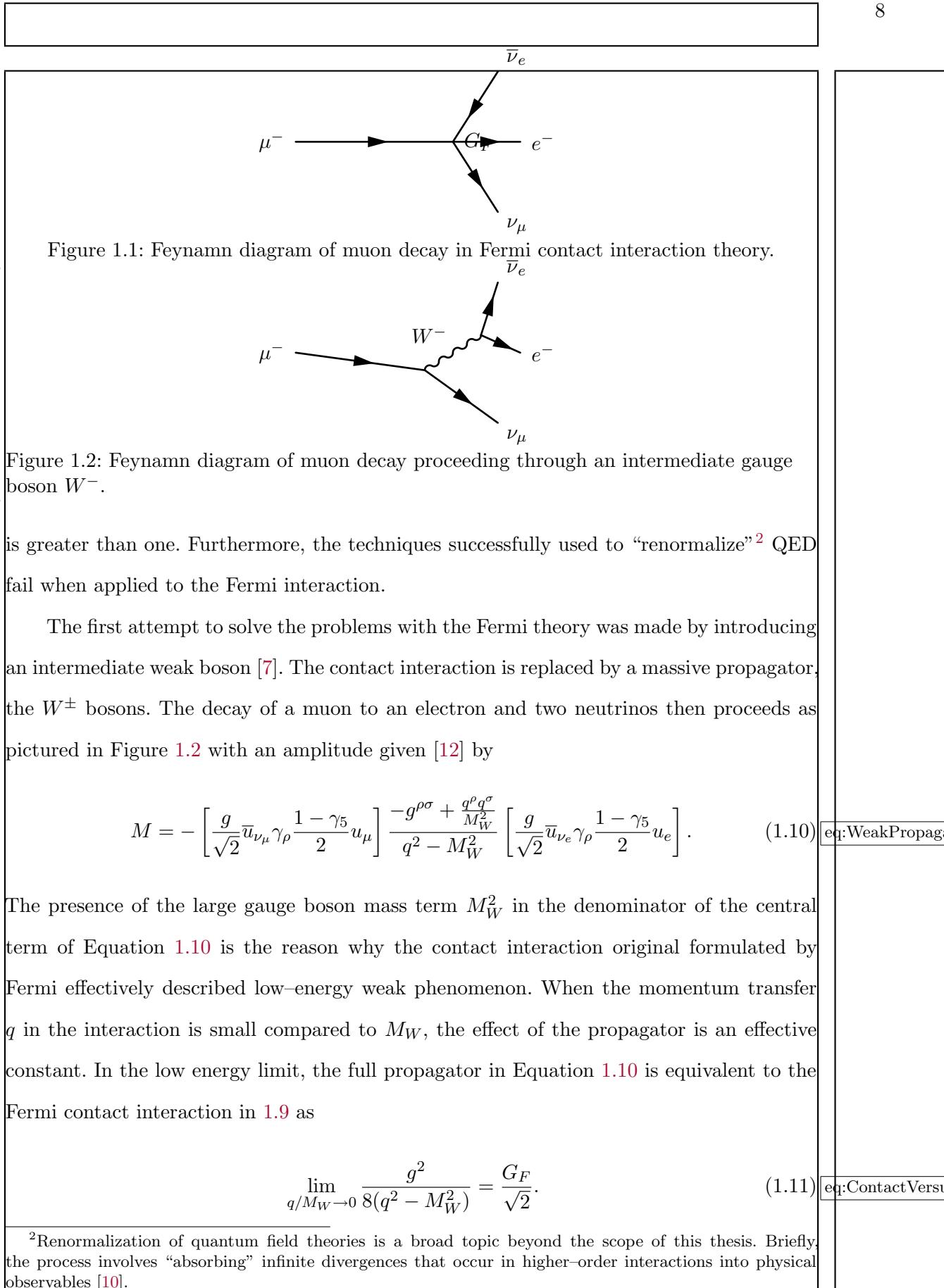
$$h\psi_L = -\frac{1}{2}\psi_L$$

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

333 The Fermi interaction can describe both nuclear  $\beta$  decay ( $p \rightarrow n + e^+ + \bar{\nu}_e$ ) as well  
 334 as the decay of a muon into an electron ( $\mu \rightarrow \nu_\mu + e + \bar{\nu}_e$ , Figure 1.1). Furthermore, the  
 335 coupling constant  $G$  is found to be a *universal* constant in weak interactions, in that it is  
 336 the same for interactions regardless of the particle species participating in the interaction.  
 337 That is,  $G_\mu = G_e = G_F$ . Using an Hamiltonian analogous to Equation 1.8 for muon decay,  
 338 the decay amplitude  $M$  is found to be

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

339 However, the contact interaction form of Fermi’s theory is not complete. When applied  
 340 to scattering processes, the interaction violates unitarity: the calculated cross section grows  
 341 with the center of mass energy, so that for some energy the probability for an interaction



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

### 361 1.1.3 Spontaneous Symmetry Breaking

362 (sec:SSB) In the early 1960s Glashow, Weinberg, and Salam published a series of papers describing  
 363 how the electromagnetic and weak forces could be unified into a common “electroweak”  
 364 force. The fact that at low energy the electromagnetic and weak forces appear to be sep-  
 365 arate phenomena is due to the fact that the symmetry of the electroweak gauge group is  
 366 “spontaneously broken.” Modern field theories (both the SM and beyond) are predicated  
 367 on the idea that the all interactions are part of a single, unified symmetry group and the  
 368 differences between various scales (electromagnetic, weak, etc.) at lower energies are due to  
 369 the unified symmetry being spontaneously broken.

370 A symmetry of a Lagrangian is spontaneously broken when the ground state, or vac-  
 371 um, is at a value about which the Lagrangian is not symmetric. In quantum field theories,  
 372 a particle is interpreted as quantized fluctuations of its corresponding field about some  
 373 constant (vacuum) ground state. The “effective” Lagrangian that we observe in the (low  
 374 energy) laboratory would be the expansion of the Lagrangian about this stable point. The  
 375 effective Lagrangian no longer obeys the original symmetry, which has been “broken.” We  
 376 give a brief example of the phenomenological effects of a spontaneously broken symmetry  
 377 in a toy model, following the treatment in [12].

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

eq:ToySSBLagran

378 The toy Lagrangian in Equation 1.12 has a global  $U(1)^3$  symmetry and consists of two  
 379 real-valued fields,  $\phi_1$  and  $\phi_2$ . The particle mass spectra of the theory is given by expanding  
 380 the field potential  $V(\phi_1, \phi_2)$  about its minimum,  $(\phi_1^{min}, \phi_2^{min})$ . The first three terms in the  
 381 series are found by

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

382 Since at the minimum the partial derivative of  $V$  is zero with respect to all fields, the  
 383 second term in Equation 1.13 is zero. The third term determines the masses of the particles  
 384 in the theory. Since a mass term for a particle corresponding to a field  $\phi_n$  in the Lagrangian  
 385 appears as  $\frac{1}{2}m^2\phi_n\phi_n$ , we can identify

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

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

390 The particle spectra of the model depends heavily on the form of the potential. An  
 391 illustrative form (that is renormalizable and bounded from below) of a possible configuration  
 392 for the potential  $V$  in Equation 1.12 is

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

393 If the parameters  $m^2$  and  $\lambda$  are both positive, then the minimum of  $V$  is at the origin ( $\phi_1 =$

<sup>3</sup>Technically, 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  $\phi_1$  and  $\phi_2$  can be seen to correspond to the real and imaginary parts of a complex field  $\phi$  that does transform according to  $U(1)$ .

394  $\phi_2 = 0$ ). In this case, the mass matrix term in Equation 1.13 takes the form  $\left(\frac{\partial^2 V}{\partial \phi_a \partial \phi_b}\right)_0 =$   
 395  $\frac{m^2}{2} \delta_{ab}$ , where  $\delta_{ab}$  is the Kronecker delta function. Therefore the mass matrix is already  
 396 diagonalized, and the  $\phi_1$  and  $\phi_2$  both correspond to particles with mass  $m$ . If the  $m^2$   
 397 parameter in Equation 1.15 is negative, the spectrum is dramatically different. After making  
 398 the replacement  $m^2 = -\mu^2 (\mu^2 > 0)$ , the extrema of  $V$  are no longer unique. The requirement  
 399 of  $\frac{\partial V}{\partial \phi_i} = 0$  for all  $i$  is satisfied in two cases:

$$(\phi_1^{min}, \phi_2^{min}) = (0, 0) \quad (1.16) \quad \text{eq:WignerPoint}$$

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

400 If the vacuum state is defined at the point in Equation 1.16, the symmetry is unbroken  
 401 and the mass spectra is unchanged. However, the system is unstable at this point, as it is  
 402 a local maximum. The true global minimum is defined as the set of points which satisfy  
 403 Equation 1.17, which form a continuous circle in  $\phi_1 - \phi_2$  space (and is therefore infinitely  
 404 degenerate). We can choose any point on the circle as the vacuum expectation value (VEV).  
 405 If the point  $(\phi_1^{min} = \nu, \phi_2^{min} = 0)$ <sup>4</sup> is chosen, evaluating Equation 1.14 yields the mass matrix

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

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

#### 410 1.1.4 The Higgs Mechanism

411 (sec:HiggsMech) As in Section 1.1.1, extending the gauge symmetry requirement to be *locally* invariant  
 412 creates interesting consequences for models that have spontaneously broken symmetry. This  
 413 gives rise to the “Higgs mechanism,” which we overview here. For simplicity we will again

<sup>4</sup>The 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=  $\nu$  and one with a VEV= 0. Therefore the physical content of the theory does not depend on the choice of VEV.

414 consider a model with  $U(1)$  symmetry. The model is identical to the one presented in  
 415 Section 1.1.3, with two exceptions. First, we express the two real fields  $\phi_1$  and  $\phi_2$  as a single  
 416 complex-valued field  $\phi$ . Second, the model is required to be locally  $U(1)$  invariant, and so  
 417 uses the gauge-covariant derivatives, minimal coupling to the gauge field, and contains the  
 418 kinetic term for the gauge field, as discussed in Section 1.1.1. The unbroken Lagrangian is

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

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

419 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  
 420 under the local  $U(1)$  gauge transformation

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

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

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

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

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

426 The kinetic term of the gauge field  $F_{\mu\nu}$  is invariant under this transformation. If the gauge  
 427 transformations of Equation 1.21 are substituted into the Lagrangian (1.18) the effective

---

<sup>5</sup>In the unitary gauge, the choice of gauge ensures that the mass matrix is diagonalized.

428 Lagrangian at the minimum of  $V$  is

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

eq:HiggsMechanism

429 The breaking of the original symmetry has dramatically altered the physical consequences  
 430 of the model. In its unbroken form, the model described by Equation 1.18 would produce  
 431 two real massive particles and one massless gauge boson mandated by local gauge invari-  
 432 ance. After symmetry breaking, the effective Lagrangian in Equation 1.22 contains a massive  
 433 scalar  $\eta$  with  $m = \sqrt{2\mu^2}$  and a *massive* gauge boson  $B_\mu$  with mass  $m = \sqrt{2}e\nu$ . By acquiring  
 434 a mass, the gauge boson  $B_\mu$  has acquired the degree of freedom (as it can now be longitudi-  
 435 nally polarized) previously associated to the second degree of freedom in the scalar  $\phi$  field.  
 436 This phenomenon, known as the Higgs mechanism, is a simplified version of the techniques  
 437 successfully used to unify the electromagnetic and weak forces that we will discuss in the  
 438 next section.

#### 439 1.1.5 Electroweak Unification

ec:ElectroweakUnification) 440 In the 1960s, the ideas of local gauge invariance in field theories, spontaneous symme-  
 441 try breaking, and the Higgs mechanism were combined by Glashow [7], Weinberg [8] and  
 442 Salam [9] to form the unified theory of electroweak interactions, the nucleus of the SM. This  
 443 model successfully unified the electromagnetic and weak interactions into a unified theory  
 444 with a larger symmetry group. The reason for the empirically observed difference in scales  
 445 between two interactions is due to the larger, unified symmetry group being broken. This  
 446 broken symmetry creates heavy gauge bosons via the Higgs mechanism, whose large mass  
 447 decreases the strength of “weak” interactions at low energy, as discussed in Section 1.1.2.  
 448 The model successfully predicted the existence and approximate masses of the weak force  
 449 carriers, the  $W^\pm$  and  $Z$  bosons. These particles were later observed [13, 14, 15, 16] with the  
 450 predicted masses at the UA1 and UA2 experiments.

451 To provide a simple introduction to the mechanisms of the model, we will start with a  
 452 model that includes only one family of leptons, the electron  $e$  and its associated neutrino  
 453  $\nu_e$ . Following once again the treatment of [12], we describe the representation of the  $e$  and  
 454  $\nu_e$  in the chosen symmetry group of the model. We then construct a locally gauge invariant  
 455 Lagrangian with spontaneously broken symmetry, and examine the particle content of the  
 456 resulting model.

457 The form of the charged current  $J_\mu(x) = \bar{u}_{\nu_e} \gamma_\rho \frac{1-\gamma_5}{2} u_e$  in the weak interaction amplitudes (1.9) indicates that the left-handed electron and neutrino (remember that the  $(1-\gamma_5)$   
 458 kills any right-handed spinors) can be combined into a doublet  $L$  of  $SU(2)$ .  
 459

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

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

$$\tau^+ = \frac{\tau^1 + i\tau^2}{2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad (1.24) \quad \text{?eq:Su2Generator}$$

$$\tau^- = \frac{\tau^1 - i\tau^2}{2} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad (1.25) \quad \text{eq:Su2Generator}$$

461 where the  $\tau^i$  are the Pauli matrices. The weak currents  $J_\mu^\pm$  can be written by combining  
 462 Equations 1.23–1.25

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

463 Since  $\tau^1$ ,  $\tau^2$ , and  $\tau^3$  are the generators of the  $SU(2)$  group, we can complete the group  
 464 by adding a neutral current to the charged currents of Equation 1.26. The  $\tau^3$  generator is

465 diagonal, so the charge of the current is zero and no mixing of the fields occur:

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

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

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

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

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

483 To accomplish the symmetry breaking, we introduce a new  $SU(2)$  doublet of complex

484 Higgs fields  $\phi$  that have hypercharge  $Y = 1$ , and contribute  $\mathcal{L}_S$  to the Lagrangian:

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

485 where  $D_\mu$  is the gauge covariant derivative containing couplings to both the  $SU(2)_L$  and  
 486  $U(1)_Y$  gauge fields, and  $V$  has a form analogous to  $V$  in Equation 1.19. At this point we  
 487 also add  $SU(2)_L \times U(1)_Y$  invariant “Yukawa” terms

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

eq:YukawaTerms

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

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

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

494 The new symmetry of the model can be confirmed by looking at the action of the different  
 495 symmetry generators on the VEV. If the generator acting on the vacuum state has a non-  
 496 zero value, then the corresponding symmetry is broken. It can then be seen that the original  
 497 symmetry generators  $T^+$ ,  $T^-$ ,  $T^3$ , and  $Y$  are all broken. The vacuum *is* invariant under  $Q$ ,  
 498 the generator of  $U(1)_{em}$ :

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

499 so the broken Lagrangian contains the correct symmetry properties.

500 The gauge boson content of the electroweak interaction is obtained by parameterizing

501 the Higgs field in the magnitude–phase notation of Equation 1.20 and using the unitary  
 502 gauge (see Section 1.1.4), where the gauge transformation is chosen so Higgs field is real.  
 503 The Higgs scalar doublet is then

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

504 The mass spectrum of the gauge bosons of the electroweak interaction (the photon,  $W^\pm$ ,  
 505 and  $Z$ ) is determined by the interaction of the gauge field terms in the covariant derivative  
 506 with the non-zero vacuum expectation value  $\nu$  of the scalar Higgs field  $\phi$

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

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

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

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

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

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

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

514 giving a massive  $Z$  boson with

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

515 and the massless photon  $A_\mu$  of QED. The mass of the  $Z$  is related to the mass of the  $W^\pm$   
516 by

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

517 where  $\theta_W$  is the “Weinberg angle,” which must be determined from experiment. As the  
518 Fermi contact interaction of Section 1.1.2 is an effective theory of the weak sector, the value  
519 of  $G_F$  obtained from  $\beta$  and muon decay experiments give clues to the masses of the  $W$  and  
520  $Z$ .

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

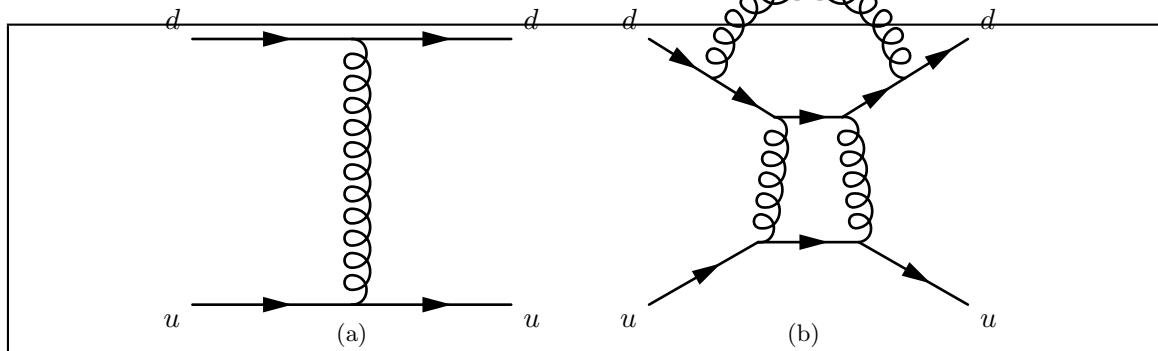
$$M_Z \approx \frac{76 \text{ GeV}}{\sin 2\theta_W} > 76 \text{ GeV}.$$

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

523 The model that is presented in this section assumes only one species of leptons, the  
524 electron and its associated neutrino. The electroweak model is trivially extended [12] to  
525 include the other species ( $\mu, \tau$ ) of leptons and the three families of quarks. The masses of  
526 the fermions are determined by the Yukawa terms in Equation 1.28. Each particle species  
527 has a Yukawa term relating the Higgs VEV to its mass that is not constrained by the theory,  
528 and must be determined by experiment.

### 529 1.1.6 Quantum Chromodynamics

530 After electroweak unification, the SM is completed by the theory of quantum chromody-  
531 namics (QCD), which describes the interactions between quarks and gluons. QCD is a broad  
532 field and only a brief introduction to its motivations and the phenomenology relevant to  
533 the analysis presented in this thesis is contained in this section. The existence of quarks as  
534 composite particles of hadrons was first proposed by Gell–Man and Zweig to explain the  
535 spectroscopy of hadrons. QCD is an  $SU(3)$  non–Abelian gauge theory which is invariant



g:QCDFeynmanDiagrams)

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.

under *color* transformations. Color is the charge of QCD and comes in three types: red, green and blue. The gauge boson that carries the force of QCD is called the gluon, which is massless as the  $SU(3)_c$  color symmetry is unbroken.

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

556 The practical consequence of color confinement to a physicist studying electroweak phe-  
 557 nomenon at a high-energy particle physics experiment is the production of quark and gluon  
 558 “jets,” which are high multiplicity sprays of particles observed in the detector. In a proton-  
 559 proton collision, quarks and gluons can be knocked off the incident protons. These quarks  
 560 and gluons immediately “hadronize,” surrounding themselves with additional hadrons, the  
 561 majority of which are charged and neutral pions. Heavier quarks, such as the charm, beauty,  
 562 and top quarks undergo a flavor-changing weak decays, which can give rise to structure  
 563 (leptons, sub-jets) within the jet. Furthermore, due to the relative strength of the strong  
 564 interaction compared that of the electroweak, collision events involving only strong inter-  
 565 actions are produced at rates many orders of magnitudes larger than that of electroweak  
 566 interactions. This makes life difficult for physicists studying the electroweak force at hadron  
 567 colliders. Sections 2.7, and Chapters 3 and 5 will discuss the techniques used to identify and  
 568 remove QCD events from the data at different stages of the analysis.

## 569 1.2 Beyond the Standard Model

?<sec:BSM>? The standard model is one of the most successful theories of the natural world ever created.  
 570 The predictions of the SM have been tested to many orders of magnitude and no experiment  
 571 to date<sup>6</sup> has found a result statistically incompatible with the SM. However, there is a  
 572 general consensus in the physics community that the SM is not complete. It is believed that  
 573 it is only an effective theory that is valid below some energy scale  $\Lambda$ . Above this energy,  
 574 there must exist some other “new physics,” which unifies the forces of the SM and correctly  
 575 describes the natural world at all scales, while maintaining equivalence to the SM at low  
 576 energy. This concept is analogous to the relationship between the effective Fermi contact  
 577 theory of Section 1.1.2 and the unified electroweak theory of Section 1.1.5. The size of  
 578 the cutoff scale  $\Lambda$  is estimated [12] to be  $\mathcal{O}(10^{15})$  GeV for a unified theory with  $SU(5)$   
 579 symmetry and even larger,  $\mathcal{O}(10^{19})$  GeV =  $M_{\text{planck}}$  if the theory is unified with gravity.  
 580

581 There are many compelling reasons that indicate that the SM is incomplete. One is  
 582 the fact that the model does not include gravity, which has still not been successfully

---

<sup>6</sup>The 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.

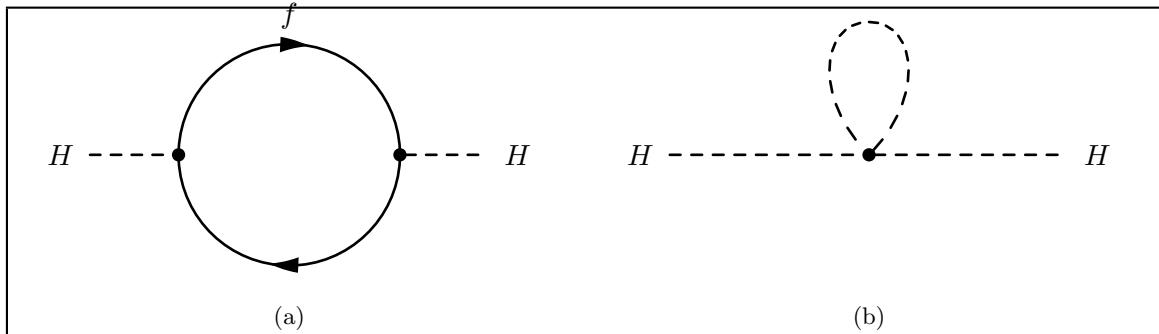


Figure 1.4: Feynman diagram of fermion (a) and scalar (b) loop corrections to Higgs boson mass.

reformulated into a quantum mechanical theory. Another is that cosmological observations indicate the presences of massive amounts of “dark matter” in the universe. Dark matter is expected to be composed of a stable massive neutral particle which interacts very weakly with other matter; no SM particle fits this description. Finally, there is the “hierarchy,” or fine-tuning problem. This problem strongly affects the Higgs sector, and motivated the development of Supersymmetry, which are the targets of the search presented in this thesis. An short overview of the hierarchy problem and Supersymmetry are presented in the next sections.

### 1.2.1 The Hierarchy Problem

The enormous size of the cutoff scale  $\Lambda$  in the SM causes a major theoretical problem in the SM. During renormalization of the SM, amplitudes with divergent integrals are cut off at  $\Lambda$ . These large constant terms are “absorbed” into the physical observables. The cutoff term appears directly in quantum corrections to the Higgs boson mass [17]. The Yukawa term  $-\lambda_f H \bar{f} f$  coupling the fermion  $f$  to the Higgs boson  $H$  produces loop corrections to Higgs boson 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 boson mass is

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

The correction scales with  $\Lambda^2$ , which is many orders of magnitude larger than the electroweak ( $M_W$ ) scale. The physical mass of the Higgs boson is expected to have the same scale as  $M_W$ .

602  $\mathcal{O}(100 \text{ GeV}/c^2)$ . The fact that each fermion contributes a loop correction (Equation 1.31)  
 603 requires that the “bare mass” of the Higgs boson to be tuned to the precision of  $(M_W/\Lambda)^2 \approx$   
 604  $10^{-26}$  for the renormalized mass to be correct! This is the so-called fine-tuning problem: it  
 605 is believed that in a natural theory there will be only one scale. The electroweak unification  
 606 analogy is in Equation 1.11, where it was noticed that the difference between the QED and  
 607 weak scale was due to the massive  $M_W$  propagator term, and that the fundamental scale  
 608  $g$  of the intermediate weak boson theory was compatible with QED. The most promising  
 609 solution to the hierarchy problem is the introduction of a new, “super” symmetry.

### 610 1.2.2 Supersymmetry

611 Supersymmetry extends the SM by positing that there exists a symmetry between the  
 612 integer-spin bosons ( $\gamma, W^\pm, Z, H$ ) and the half integer-spin fermions (quarks and leptons).  
 613 In Supersymmetry, every particle in the SM has a “superpartner” with a spin differs by  $1/2$ .  
 614 All of the other quantum numbers (including mass) of the superpartners are the same. The  
 615 introduction of this symmetry immediately solves the hierarchy problem. For every scalar  
 616 loop correction (Figure 1.4(b)) to the Higgs boson mass there is now a corresponding fermion  
 617 loop correction (Figure 1.4(a)). As the fermion and the scalar have the same quantum  
 618 numbers (except for spin) it turns out that these two diagrams have the same value, but  
 619 *opposite* sign. Thus the large  $\Lambda^2$  superpartner loop corrections to the Higgs boson mass  
 620 exactly cancel out the problematic SM corrections. It is clear that if Supersymmetry exists,  
 621 it must be broken. We have not observed a scalar charged particle with the same mass as the  
 622 electron, for example. An excellent overview of possible mechanisms that create spontaneous  
 623 symmetry breaking in supersymmetric models is given in Chapter 6 of [17].

### 624 1.2.3 The Minimal Supersymmetric Extension to the Standard Model

(sec:MSSMAndTaus)  
 625 The simplest possible supersymmetric extension to the SM is MSSM. The model groups  
 626 superpartner pairs into chiral (a left or right-handed fermion field plus a complex scalar  
 627 field) and gauge (a spin-1 vector boson and a left or right-handed *gaugino* fermion) “su-  
 628 permultiplets.” As the weak interactions of the SM fermions are chiral, they (and their  
 629 superpartners) must belong in a chiral supermultiplet. It is interesting to note that there

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

Table 1.1: Chiral supermultiplets in the MSSM. The spin-0 fields are complex scalars, and the spin-1/2 fields are left-handed two-component Weyl fermions. Source: [17]

{tab:chiral}

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

Table 1.2: Gauge supermultiplets in the MSSM. Source: [17]

is a different superpartner for the left and right-handed components of the fermions, even though the superpartners are spin-0 and cannot have any handedness. It is found that there must be two Higgs supermultiplets for the MSSM to be viable. As there are now fermionic particles in the Higgs sector (the Higgsinos), if only one supermultiplet is introduced the MSSM suffers from non-renormalizable gauge anomalies.<sup>7</sup> By introducing an additional Higgs supermultiplet with opposite hypercharge, the anomaly is canceled. The scalar portion of the MSSM Higgs sector then contains two complex doublet fields  $H_u = (H_u^+, H_u^0)$  (up-type) and  $H_d = (H_d^0, H_d^-)$  (down-type). The complete chiral and gauge supermultiplets of the MSSM are enumerated in Tables 1.1 and 1.2, respectively.

The superpotential (like the scalar potential of Section 1.1.3 but invariant under supersymmetric transformations) of the MSSM is then [17]

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

<sup>7</sup>A 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.

641 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  
 642 Yukawa  $3 \times 3$  matrices which act on the different families. It is important to note that the  
 643 up-type quarks couple to the up-type Higgs doublet  $H_u$ , while the down-type quarks and  
 644 leptons couple to the down-type Higgs doublet. This feature has large phenomenological  
 645 consequences, which are discussed in Section 1.3.2. The scalar portion of the  $W_{\text{MSSM}}$  poten-  
 646 tial defines the spontaneous symmetry breaking. Similar to the scalar potential  $V$  symmetry  
 647 breaking of Section 1.1.3, the potential of  $V$  at the minimum is found<sup>8</sup> to be

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

eq:MSSMScalarP

648 Under suitable choices<sup>9</sup> of the parameters in Equation 1.32, the up-type and down-type  
 649 neutral Higgs fields acquire a VEV,  $\nu_u$  and  $\nu_d$ , respectively. The VEVs are related to the  
 650 VEV of electroweak symmetry breaking (Equation 1.30) in the SM,

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

651 The ratio of the VEVs is expressed as

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

652 which is an important parameter of the MSSM. As there are two complex doublets, there are  
 653 a total of eight degrees of freedom in the MSSM Higgs sector. After the symmetry breaking,  
 654 three of the degrees of freedom are (like the Standard Model) eaten by the  $W^\pm$  and  $Z$  weak  
 655 gauge bosons. The remaining five degrees of freedom create five massive Higgs bosons: two  
 656 CP-even neutral scalars  $h^0$  and  $H^0$ , a CP-odd neutral scalar  $A^0$ , and two (positive and  
 657 negative) charged scalars  $H^\pm$ . The masses are of the different Higgs boson mass eigenstates

<sup>8</sup>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  $U(1)_{em}$ .

<sup>9</sup>See Chapter 7 of [17] for a detailed overview.

658 are related to each other and  $\tan \beta$  at tree level by

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

659 It can be seen that the tree level mass  $m_{h^0}$  of Equation 1.33 is bounded from above by  
 660  $m_{h^0} < m_Z |\cos(2\beta)| < 90 \text{ GeV}/c^2$ . If this is true the model would have been excluded by  
 661 the LEP experiment (see next section). However, there are important quantum corrections  
 662 to  $m_{h^0}$  from the top–quark and top–squark loop diagrams which increase  $m_{h^0}$ . The Yukawa  
 663 couplings in the MSSM depend on  $\tan \beta$ . The relationships for the most massive members  
 664 of each family are

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

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

## 669 1.3 Searches for the Higgs Boson

?<sec:PreviousSearches>?  
 670 The discovery of the Higgs boson is one of the biggest prizes in science today. Dozens of  
 671 experiments, thousands of scientists and billions of dollars (a human hierarchy problem)  
 672 have been spent in efforts to discover the Higgs boson. In this section we discuss how the  
 673 Higgs boson and the MSSM could appear in modern colliders (with an emphasis on the  
 674 LHC) and the current limits placed on the Higgs boson by the Large Electron–Positron  
 675 Collider (LEP) and Tevatron experiments.

### 676 1.3.1 Standard Model Higgs Boson Phenomenology

<sec:SMHiggsPhenom>  
 677 The phenomenology of the Higgs boson is strongly coupled to its relationship with mass.  
 678 The coupling of the Higgs boson to the fermions is determined by the Yukawa terms (Equa-  
 679 tion 1.28) in the Lagrangian. Taking the electron as an example, after symmetry breaking

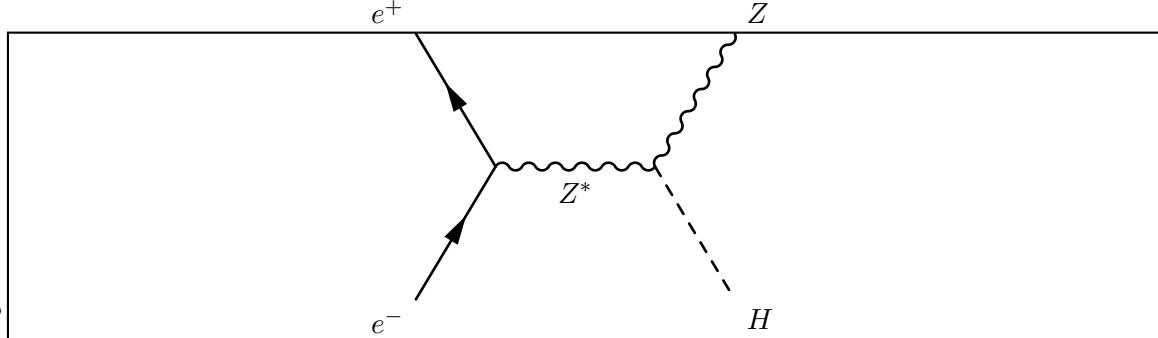


Figure 1.5: Higgstrahlung production diagram at  $e^+e^-$  colliders

680 the Yukawa term is found to be

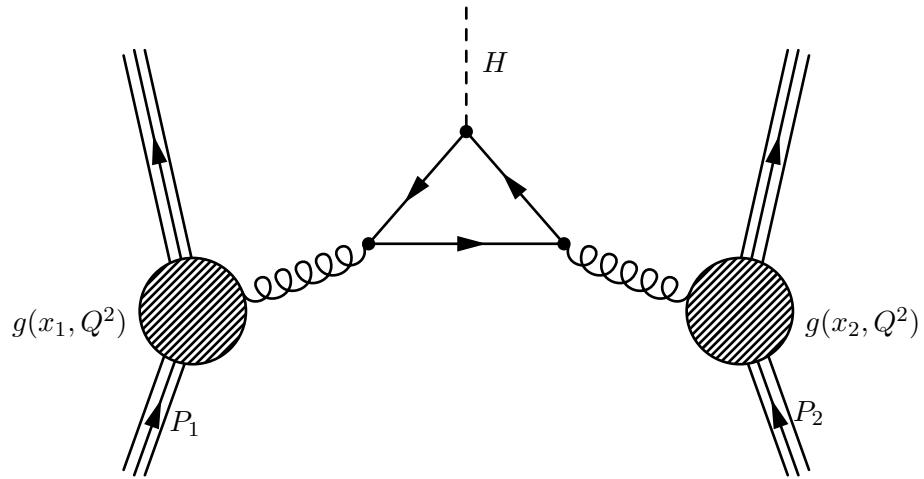
$$\mathcal{L}_e = -\frac{G_e}{\sqrt{2}}(\nu + H(x))\bar{e}e = -\frac{G_e\nu}{\sqrt{2}}\bar{e}e - \frac{G_e}{\sqrt{2}}H(x)\bar{e}e. \quad (1.34)$$

681 The value of  $G_e$  is a free parameter of the theory and is thus determined by the measurement  
682 of the electron mass and  $\nu$ , the VEV of the Higgs field

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

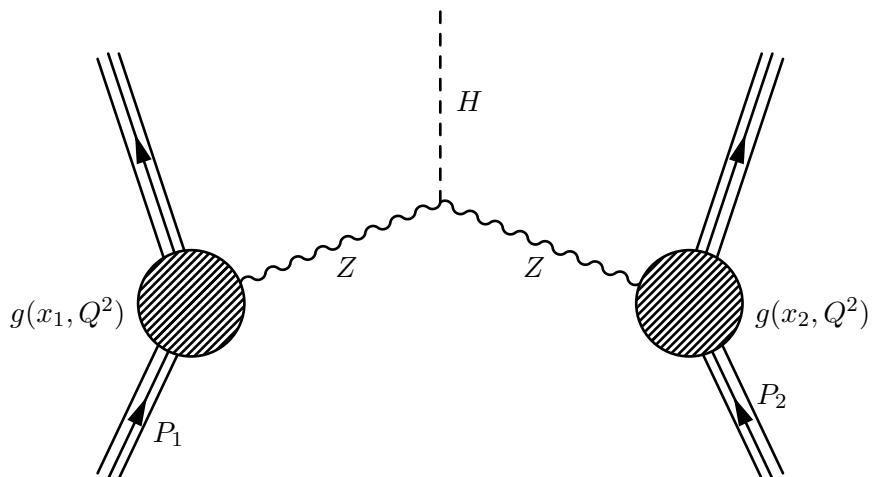
683 The left-hand side of Equation 1.35 is the same as the constant in the electron–Higgs boson  
684 coupling term ( $H(x)\bar{e}e$ ) in Equation 1.34. Therefore the coupling between the fermions  
685 and Higgs boson is proportional to their mass! This remarkable fact shapes the possible  
686 production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

687 The dominant modes of Higgs boson production depend on the type of experiment.  
688 In general, Higgs boson production is favored through high-mass intermediate states, due  
689 to the mass<sup>2</sup> proportional coupling. At the Tevatron and LEP experiments, which will be  
690 introduced in the next section, the dominant SM Higgs boson production mode is “Hig-  
691 gstrahlung,” where a virtual  $W^\pm$  or  $Z$  gauge boson is produced and then radiates a Higgs  
692 boson. Higgstrahlung is illustrated in Figure 1.3.1. At the Large Hadron Collider, higher  
693 gluon luminosities (see Figure 1.8) result in the favored cross section being “gluon fusion,”  
694 (illustrated in Figure 1.3.1) where two gluons from the incident protons combine in a quark  
695 (dominated by the massive top quark) loop which then radiates a Higgs boson. Another  
696 important channel [19] is “vector boson fusion,” (Figure 1.3.1) where weak gauge bosons  
697 ( $W^\pm$  or  $Z$ ) are radiated from the incoming quarks and fuse to produce a Higgs boson.



(fig:GluonFusion)

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



(fig:VBFProdDiagram)

Figure 1.7: Vector boson fusion (VBF) Higgs boson 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.

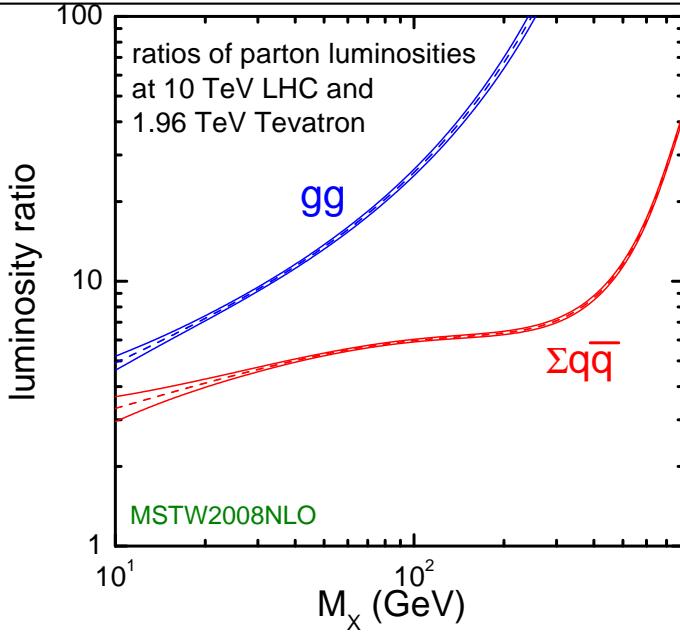


Figure 1.8: Ratio of the parton luminosity (the amount of luminosity contributed by the different species that compose the proton) of the LHC (at  $\sqrt{s} = 10$  TeV) and the Tevatron. The large increase in gluon–gluon luminosity affects the favored production mechanisms of the Higgs boson.

(fig:GluonLumiRatio)

698 This is a notable channel due to the lack of “color-flow” (gluons) between the two protons.  
 699 producing an event with low central jet activity and two “tag-jets” in the forward and  
 700 backward regions. The theoretical cross sections for the SM Higgs boson at the LHC are  
 701 shown for the various production mechanisms in Figure 1.9.

702 The branching fractions of the different decay modes of the SM Higgs boson depend  
 703 strongly on the mass of the Higgs boson. In general, the Higgs prefers (due to the Yukawa  
 704 couplings) to decay pairs of the particles with the highest mass possible. Below the threshold  
 705 to decay to pairs of weak bosons ( $M_H < 160 \text{ GeV}/c^2$ ), the Higgs boson decays predominantly  
 706 to either  $b$ -quarks ( $b\bar{b}$ , 90%) or a pair of  $\tau$  leptons ( $\tau^+\tau^-$ ,  $\approx 10\%$ ). Above the  $W^\pm W^\mp$  thresh-  
 707 old, decays to vector bosons ( $H \rightarrow W^\pm W^\mp$  and  $H \rightarrow ZZ$ ) dominate. The dependence of  
 708 branching fraction on  $M_H$  and the other rare decay modes are illustrated in Figure 1.10.  
 709 For low mass Higgs bosons, the  $\tau^+\tau^-$  decay mode plays a particularly important role. The  
 710 dominant decay mode  $H \rightarrow b\bar{b}$  suffers from enormous backgrounds from QCD jet produc-  
 711 tion. It is important to understand the magnitude of difference between expected Higgs  
 712 boson production and the rates of various backgrounds. Figure 1.11 illustrates the cross

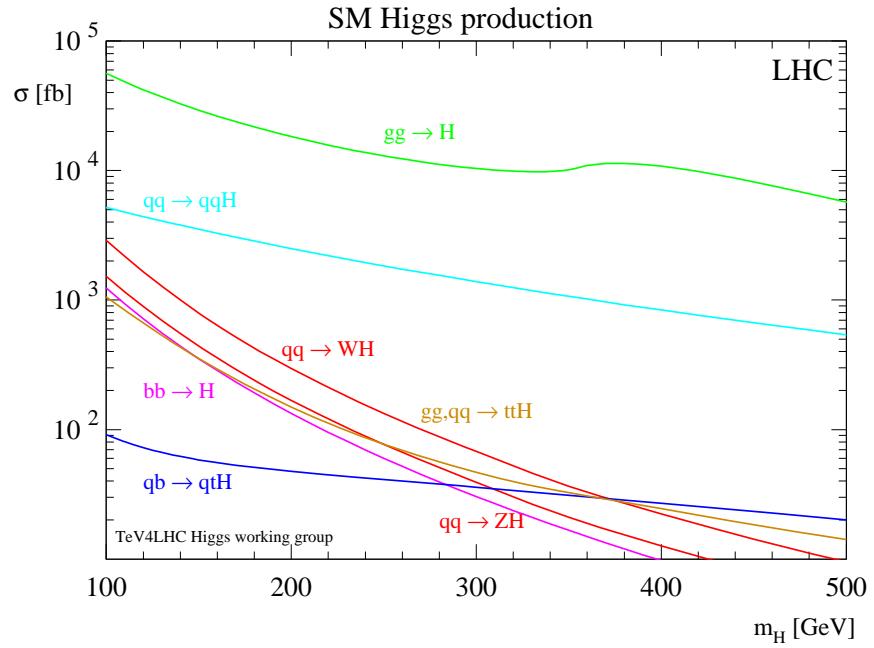


Figure 1.9: Cross section of the SM Higgs boson versus the Higgs boson mass. The different curves give the contribution to the cross section from different production mechanisms. Source: [20].

(fig:LHCSMHiggsXsec)

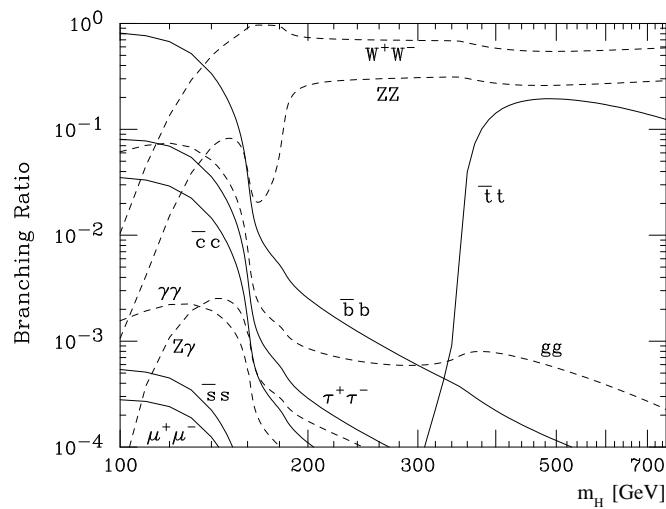


Figure 1.10: Branching fraction of the SM Higgs bosons for different values of  $M_H$ . Source: [20].

(fig:SMHiggsBR)

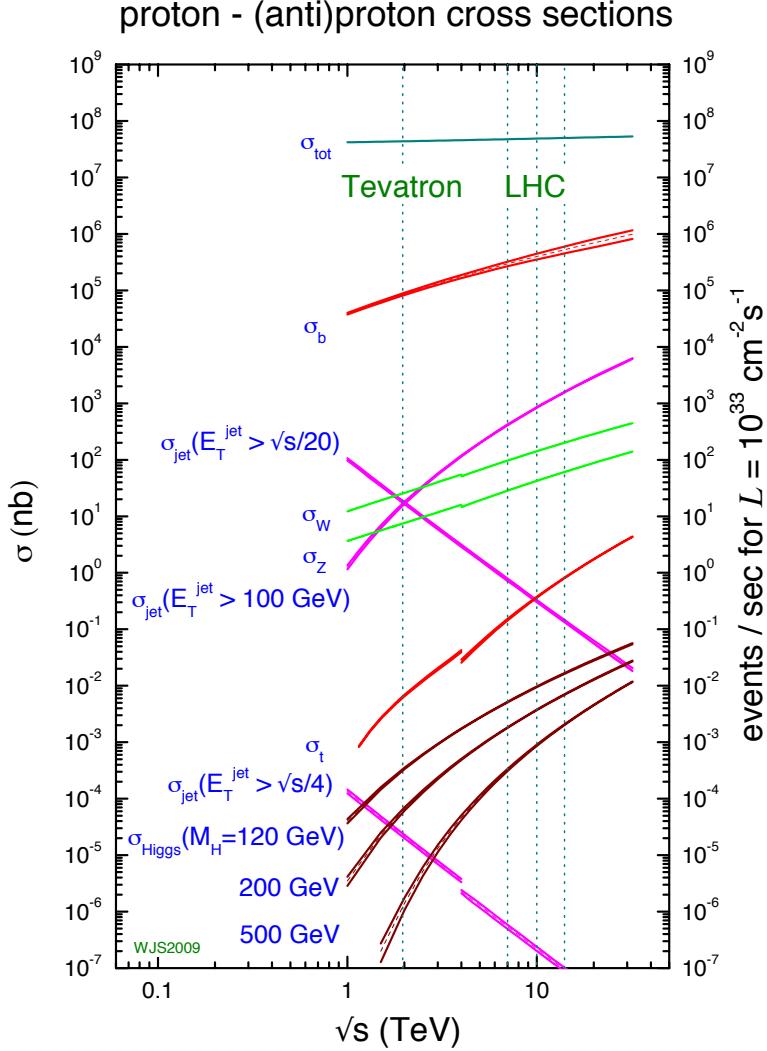
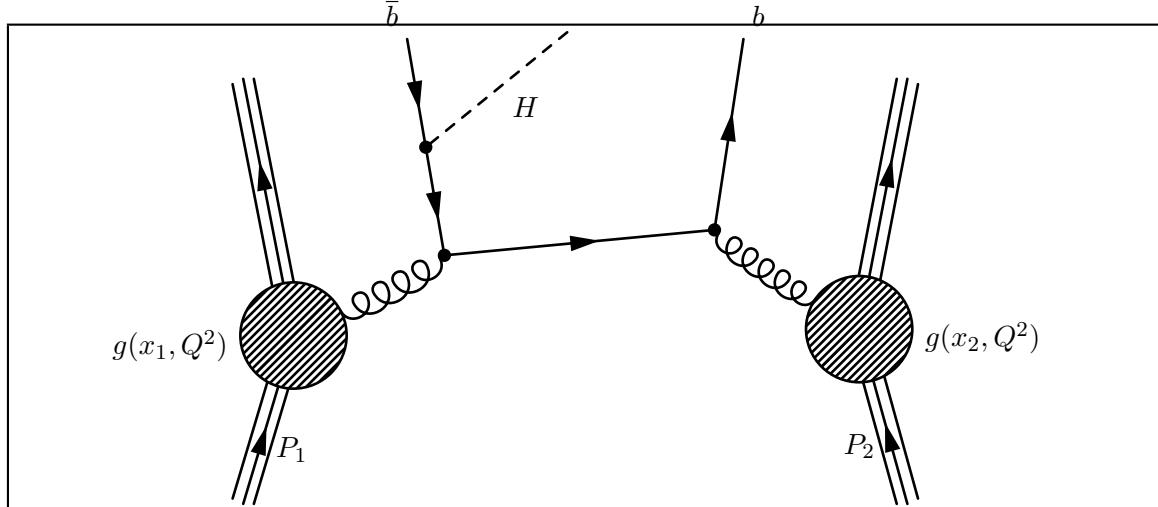


Figure 1.11: Cross sections of various processes at hadron colliders. The horizontal axis represents the center of mass energy of the collision. Of note is the vast difference in scales between Higgs boson production (maroon lines,  $\mathcal{O}(10^{-2} \text{ nb})$ ) and the QCD cross section to produce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

hadronColliderCrossSections)



g:AssociatedBProduction)

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

sections for different SM processes at hadron colliders. The rate of Higgs boson 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 Boson Phenomenology

(sec:MSSMHiggsPhenom)  
 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  $\tau$  leptons and  $b$  quarks. When the parameter  $\tan \beta$  is large, the coupling factor between the Higgs bosons and the down-type quarks and leptons (effectively the  $\tau$  and  $b$  quark) is enhanced by  $\tan \beta$ . The gluon–gluon cross section is therefore increased by  $\tan^2 \beta$ , where the top quark loop in Figure 1.3.1 is replaced by a ( $\tan \beta$  enhanced)  $b$  quark loop. Additionally, MSSM Higgs production with associated  $b$ -quarks, illustrated in Figure 1.3.2, becomes an important production mode. At tree-level, the MSSM can be defined by the mass of the CP-odd Higgs boson  $m_{A^0}$  and  $\tan \beta$ . For a reasonably high  $\tan \beta$ , there is always one CP-even Higgs boson ( $h^0$  or  $H^0$ ) which is mass-degenerate with the  $A^0$ . When  $\tan \beta$  and  $m_{A^0}$  are both large, associated  $b$  production dominates the total cross section [22]. The cross sections of the different MSSM neutral Higgs bosons are shown in Figure 1.13.

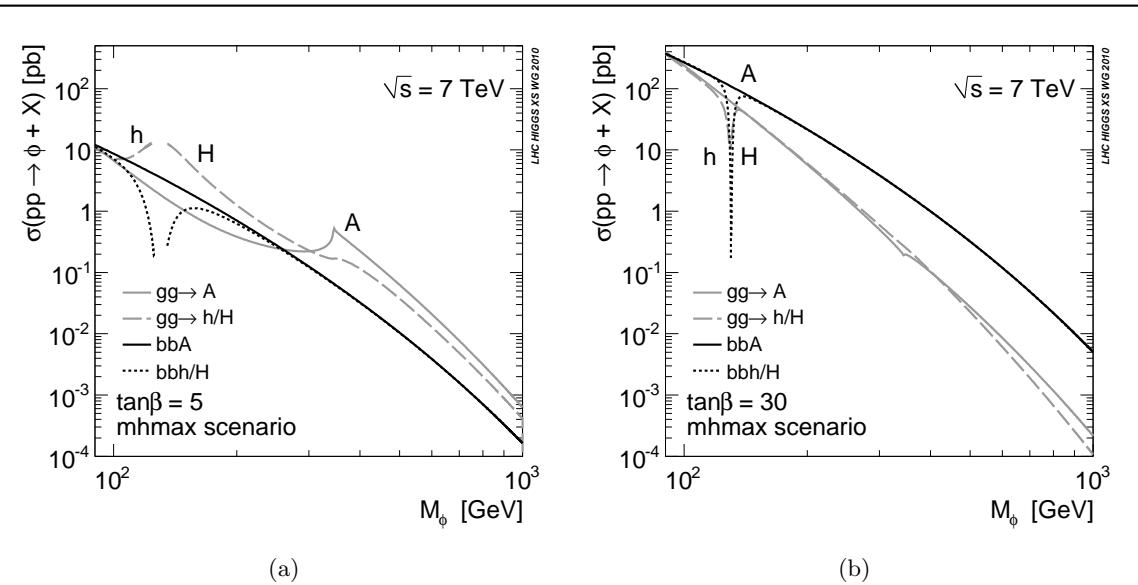


Figure 1.13: Cross sections for the different MSSM Higgs bosons versus  $m_{A^0}$  in the  $m_{h^{max}}$  benchmark scenario [23] scenario for  $\tan \beta = 5$  (a) and  $\tan \beta = 30$  (b). Source: [22]

730 The  $\tan \beta$  enhancement of the MSSM Higgs boson coupling to the  $b$  quarks and  $\tau$  leptons  
 731 causes the branching fraction of all neutral MSSM Higgs states to be  $H \rightarrow b\bar{b}$  (90%) and  
 732  $H \rightarrow \tau^+\tau^-$  (10%) across the entire range of  $m_{A^0}$ . The enhanced production rate and the  
 733 high branching fraction to  $\tau$  leptons make the MSSM Higgs bosons decaying to  $\tau$  leptons  
 734 an exciting and promising channel to search for Higgs bosons and supersymmetric physics  
 735 at colliders.

### 736 1.3.3 Results from LEP and Tevatron

The LEP and Tevatron experiments have both set limits on the existence of the SM and MSSM Higgs boson. Precision electroweak measurements give additional hints on the prospects for both models.

740 LEP was an  $e^+e^-$  collider at CERN and has effectively excluded the presence of a low  
 741 (less than 114 GeV/ $c^2$ ) mass Higgs boson. The dominant SM Higgs boson production mode  
 742 at LEP is Higgstrahlung, where the Higgs boson is produced in association with a  $Z$  boson  
 743 (see Figure 1.3.1). The search at LEP utilized a number of different decay channels [20].  
 744 The decay channels used in the LEP search are summarized in Table 1.3.3.

&lt;tab:LEPModes&gt;

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.

745 The results using all channels from the four LEP experiments<sup>10</sup> have been combined into  
 746 a single limit, shown in Figure 1.14. The analysis sets a limit on the ratio  $\xi^2 = (g_{HZZ}/g_{HZZ})^2$ .  
 747 the upper limit on the HZZ coupling divided by the predicted value of the SM. For Higgs  
 748 boson masses below 114 GeV/c<sup>2</sup>, the ratio is below unity at the 95% confidence level, ruling  
 749 out a SM Higgs boson below that mass.

750 The Tevatron is a proton–antiproton collider with a center-of-mass energy of  $\sqrt{s} =$   
 751 1.96 TeV. There are two general purpose detectors at the Tevatron, CDF and DØ . The  
 752 dominant Higgs boson production modes at the Tevatron are Higgstrahlung and gluon fusion  
 753 (see Figure 1.3.1). For low mass ( $m_H < 135$  GeV/c<sup>2</sup>) Higgs bosons the dominant channel  
 754 at the Tevatron is the Higgstrahlung production mode and  $H \rightarrow b\bar{b}$  decays. Large multi-  
 755 jet backgrounds prevent the  $H \rightarrow b\bar{b}$  decay mode from being useful for searching for Higgs  
 756 bosons produced by gluon fusion. The  $H \rightarrow \tau^+\tau^-$  and  $H \rightarrow \gamma\gamma$  decays are additionally  
 757 used in an inclusive search at low mass, but do not dominate the search sensitivity. The  
 758 combined low-mass limit on the SM Higgs boson from both Tevatron experiments is shown  
 759 in Figure 1.15. The Tevatron currently sets an upper limit on the SM Higgs boson cross  
 760 section of about 2.5 times the SM expectation.

761 When ( $m_H < 135$  GeV/c<sup>2</sup>) the  $H \rightarrow W^+W^-$  decay mode becomes significant. Low  
 762 diboson backgrounds allow this decay mode to probe both the Higgstrahlung and gluon  
 763 fusion production modes. The combined results of the CDF and DØ searches using the

<sup>10</sup>ALEPH, DELPHI, L3, and OPAL

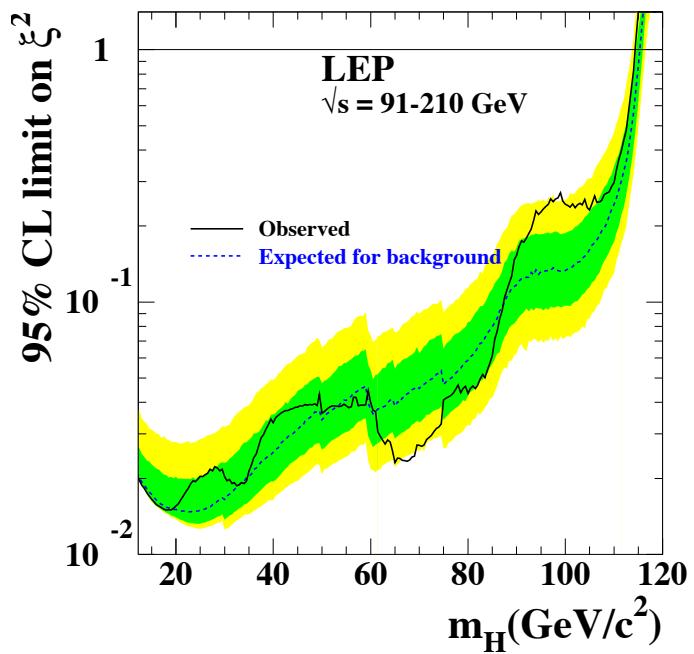


Figure 1.14: Combined LEP upper limit set on the quantity  $\xi^2 = (g_{HZZ}/g_{HZZ})^2$  at 95% confidence level. Regions where the observed ratio is less than one exclude the SM. The dashed line gives the expected limit for the null (background only) hypothesis, with the green and yellow bands representing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

(fig:LEPHiggsLimit)

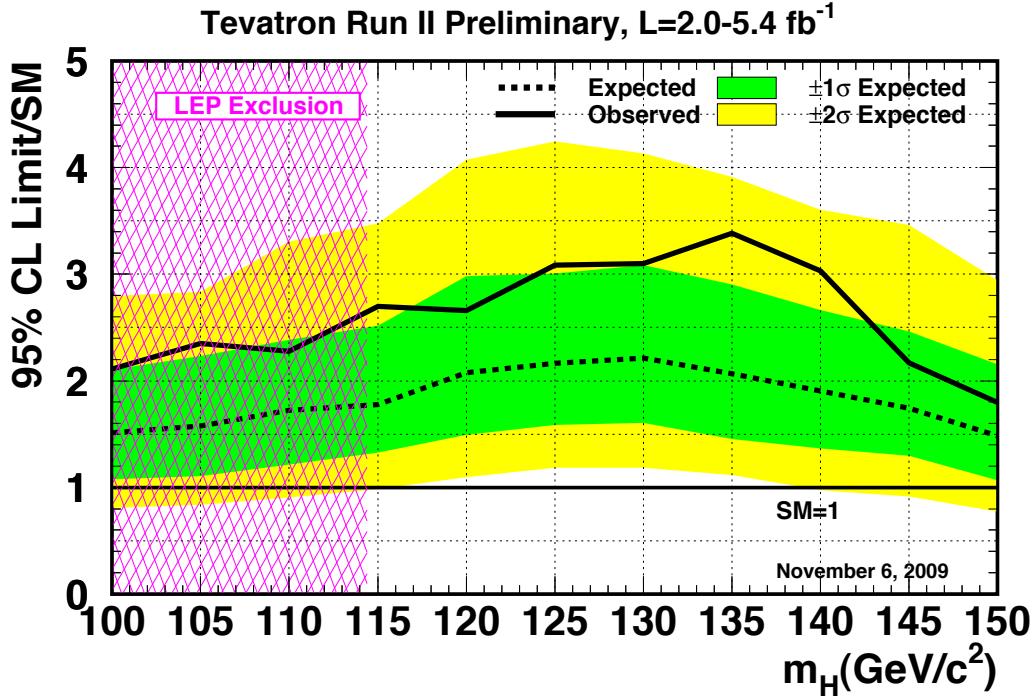


Figure 1.15: Combined CDF and D $\emptyset$  RunII upper limit on the cross section of a standard model-like Higgs boson. The LEP limit is shown in pink. Reference: [20]

764  $W^+W^-$  have decay mode recently excluded (See Figure 1.16) a SM Higgs boson with a  
 765 mass between 162 and 166  $\text{GeV}/c^2$ . This is the first exclusion in SM Higgs boson mass  
 766 parameter space since the LEP result.

767 Analyses at LEP and Tevatron have also addressed excluded regions of the MSSM. At  
 768 LEP, the dominant production modes of the MSSM Higgs bosons are Higgstrahlung and  
 769 pair production, where  $e^+e^- \rightarrow h^0A^0$  or  $H^0A^0$ . For the Higgstrahlung production mode,  
 770 the SM search can be reinterpreted in terms of the MSSM. To address the pair production  
 771 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.  
 772 Finally, LEP is also sensitive to associated MSSM Higgs boson production at low  $m_{A^0}$  and  
 773 high  $\tan\beta$  to  $e^+e^- \rightarrow \{\bar{\phi}$ , where the associated fermions  $\{\$  are  $b$ -quarks or tau leptons.  
 774 The combined limits from LEP in the  $m_{A^0} - \tan\beta$  plane are shown in Figure 1.17.

775 At the Tevatron, CDF and D $\emptyset$  have set a combined limit on the MSSM using the  
 776 inclusive  $H \rightarrow \tau^+\tau^-$  channel. The analysis presented in this thesis is very similar to the  
 777 approaches used at the Tevatron. Results from the Tevatron have excluded the MSSM for

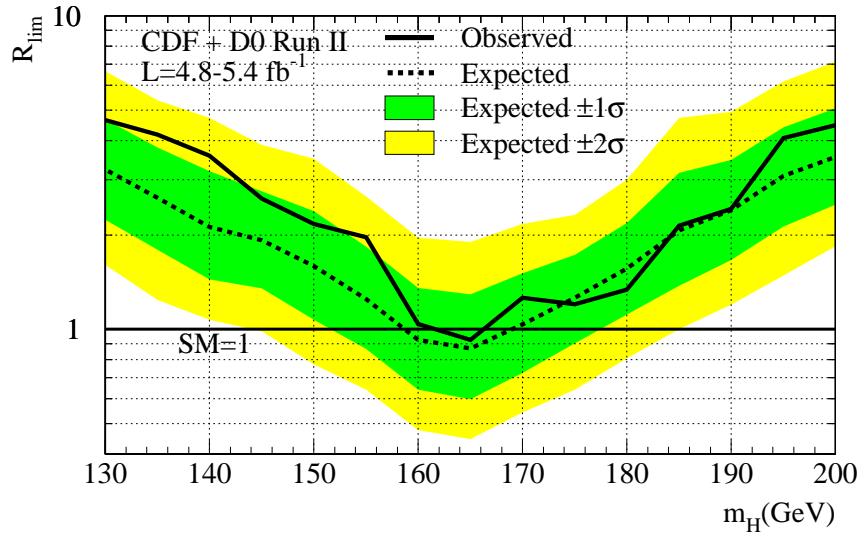


Figure 1.16: Combined CDF and D0 Run II upper limit on the cross section of a SM-like Higgs boson using the  $H \rightarrow W^+W^-$  decay mode. The SM is excluded for Higgs boson masses between 162 and 166  $\text{GeV}/c^2$ . Reference: [20]

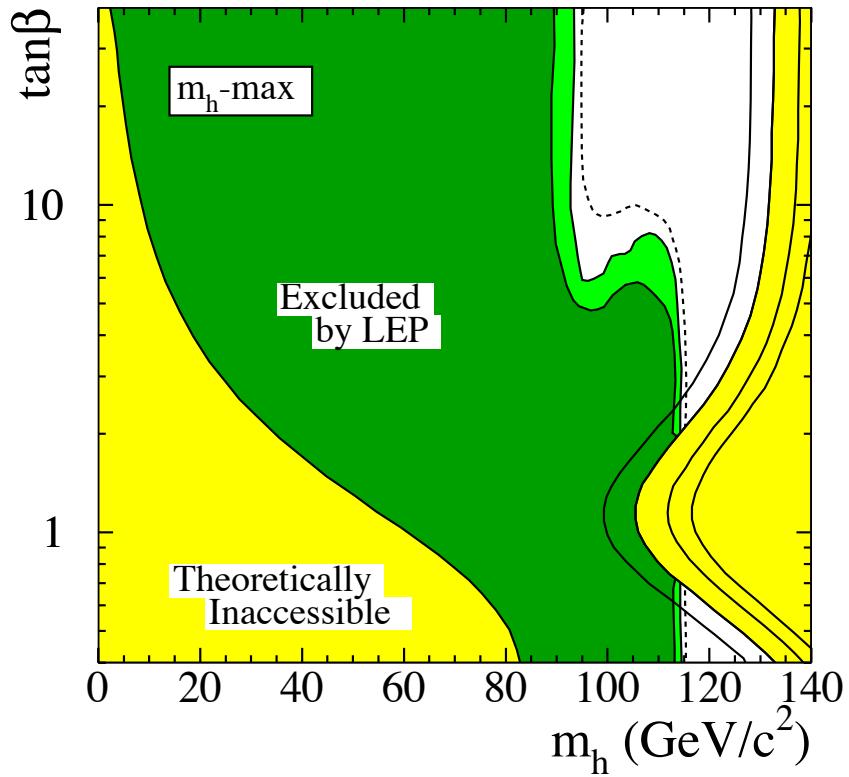


Figure 1.17: Combined LEP limits on the MSSM. The results are interpreted in the context of the  $m_h$ -max benchmark [23] scenario of the MSSM. Reference: [20]

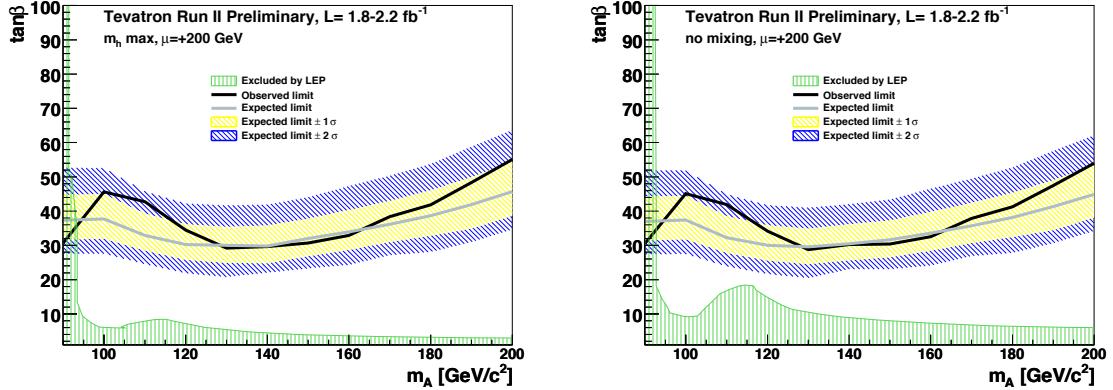


Figure 1.18: Combined Tevatron limits on the MSSM. The grey line and blue and yellow bands gives the expected limit and its one and two sigma contours. The black line is the observed limit. The results are interpreted in the context of the  $m_h$ -max benchmark (left) and “no mixing” (right) MSSM scenarios.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

Reference: [20]

(fig:TevMSSMLimits)

778  $\tan \beta$  greater than approximately 35 for MSSM Higgs boson mass  $m_{A^0} < 200$   $\text{GeV}/c^2$ . The  
 779 full exclusion plot for the  $m_h$ -max and “no mixing” MSSM benchmark scenarios are shown  
 780 in Figure 1.18.

## 781 1.4 The Physics of the Tau Lepton

782 As discussed in Sections 1.3.1 and 1.2.3, the  $\tau$  lepton is an important probe of Higgs physics.  
 783 The  $\tau$  lepton has some unusual properties which make it particularly challenging at hadron  
 784 colliders. With a mass of 1.78  $\text{GeV}/c^2$ , the  $\tau$  lepton is heaviest of the leptons. The nominal  
 785 decay distance  $c\tau$  of the  $\tau$  lepton is 87  $\mu\text{m}$ , which in practice means that the  $\tau$  will always  
 786 decay before reaching the first layer of the detector. Tau decays can be effectively classified  
 787 into two types. “Leptonic” decays consist of a  $\tau$  decaying to a light lepton ( $\ell = e, \mu$ ) and two  
 788 neutrinos  $\tau^+ \rightarrow \ell^+ \nu_\tau \bar{\nu}_\ell$ . “Hadronic” decays consist of a low-multiplicity collimated group of  
 789 hadrons, typically  $\pi^\pm$  and  $\pi^0$  mesons. The hadronic decays of the  $\tau$  lepton compose approx-  
 790 imately 65% of the  $\tau$  lepton branching fraction, with the remainder shared approximately  
 791 equally by the leptonic decays. The branching fractions for the leptonic and most common  
 792 hadronic decays are shown in Table 1.4.

793 The tau is also a challenging object in that the decay of the tau always includes neu-  
 794 trinos. The associated neutrinos are weakly interacting and do not create a signal in any

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%
$\pi^- \pi^0 \nu_\tau$	$\rho$	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%

{tab:decay`modes}

Table 1.4: Resonances and branching ratios of the dominant decay modes of the  $\tau$  lepton. The decay products listed correspond to a negatively charged  $\tau$  lepton; the table is identical under charge conjugation.

795 detector at CMS. The only sign that the neutrinos are there is an imbalance in the total  
 796 transverse<sup>11</sup> energy in the event. This thesis will describe a novel way to reconstruct the  
 797 neutrinos associated to tau decays in Chapter 4.

798 A tau with produced with energy  $E$  travels on average

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

799 before decaying in the detector. These lengths are comparable to the resolution of the CMS  
 800 tracker, therefore it is possible to reconstruct a vertex corresponding to a tau decay that is  
 801 displaced with respect to the primary vertex. This can be used as an additional discriminant  
 802 against QCD, which is expected to decay promptly. Furthermore, in Chapter 4 we will see  
 803 it may be possible to use it when reconstructing the associated neutrinos.

---

<sup>11</sup>At 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.

## 804 Chapter 2

# 805 The Compact Muon Solenoid Experiment

806 The Compact Muon Solenoid (CMS) Experiment is a general-purpose particle detector  
807 designed to measure collision events at the Large Hadron Collider (LHC), a proton-proton  
808 synchrotron located at the CERN laboratory in Geneva, Switzerland. The design goals of  
809 the CMS experiment are [24], in order of priority:

- 810 • Excellent muon identification and momentum resolution over a large solid angle. Par-  
811 ticularly important is the ability to determine the muon charge for any muon with  
812  $p < \text{TeV}/c$ .
- 813 • Charged particle tracking with excellent momentum. Triggering on  $b$ -jets requires an  
814 excellent vertex resolution, and requires pixel detectors close to the interaction region;
- 815 • An electromagnetic calorimeter with an energy resolution precise enough to measure  
816 the mass of diphoton and dielectron events with a resolution of  $\approx 1\%$  at  $100 \text{ GeV}/c^2$ .  
817 The calorimeter must have high granularity to determine isolated photons and elec-  
818 trons during high luminosity running.;
- 819 • The design goal of excellent missing-transverse-energy and dijet-mass resolution, re-  
820 quiring require hadron calorimeters large solid angle coverage and high granularity.

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

of different species of particles. The central (closest to interaction point) sub-detectors are the charged particle tracking systems (the “tracker”). The tracker is designed to be a non-destructive instrument, which means that ideally that the momentum of particles are unchanged after passing through it. Outside of the tracker is the electromagnetic and hadronic calorimeters, which are abbreviated ECAL and HCAL, respectively. The outer layers of CMS are designed to measure muons, the one<sup>1</sup> species of particle that is nearly immune to the stopping power of the calorimeter. The arrangement of destructive and non-destructive sub-detectors facilitates the identification of different types of particles. This concept is illustrated in Figure 2.1(b). In this chapter we give a brief overview of the LHC machine, and then describe the individual sub-detector systems of CMS.

## 2.1 The Large Hadron Collider

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

## 2.2 Solenoid Magnet

?<sec: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

---

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

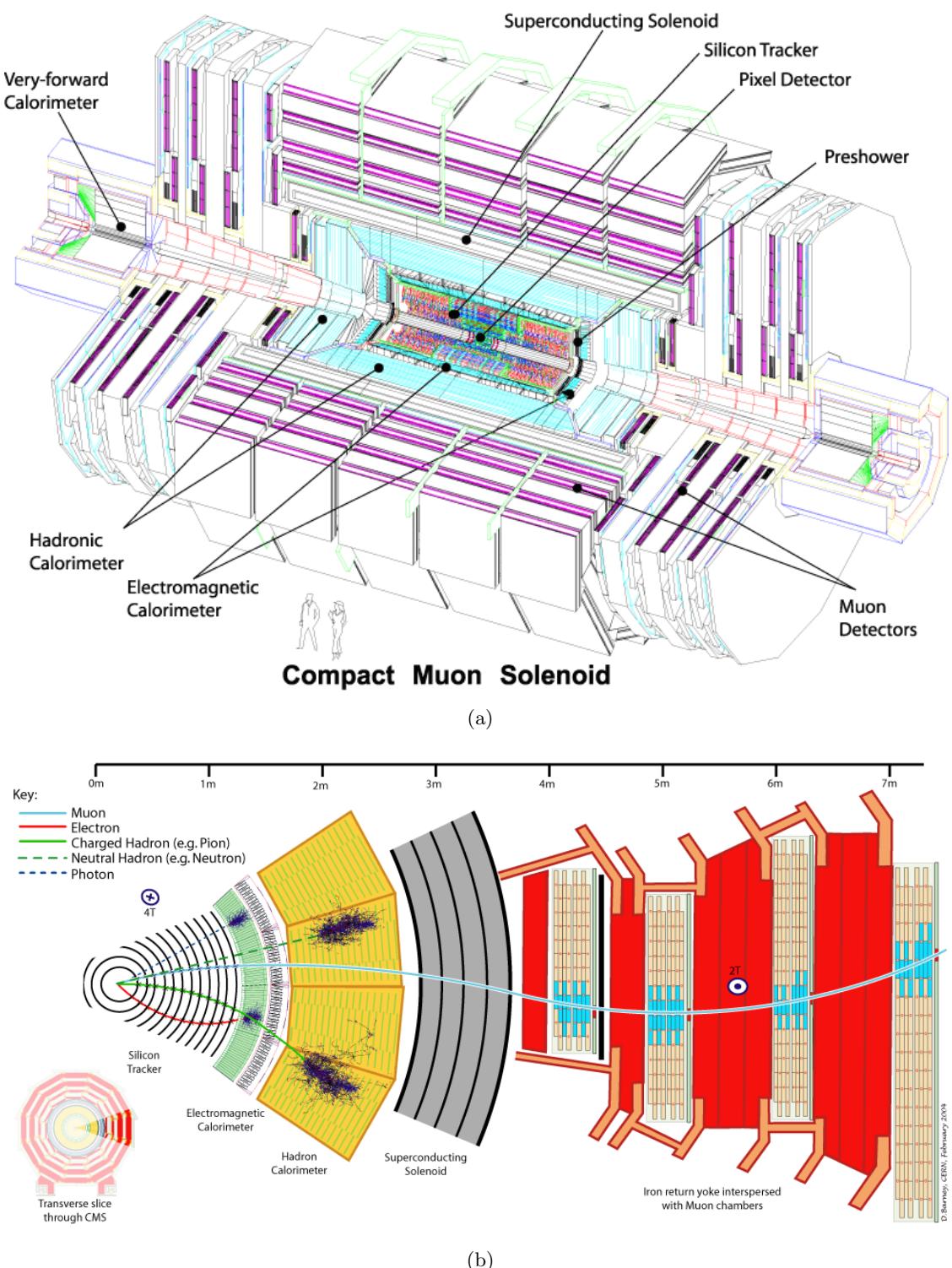


Figure 2.1: Figure (a), top, shows a schematic drawing of the CMS detector. The individual sub-detectors are labeled. Two humans are shown in the foreground for scale. Figure (b) shows a radial cross section of the detector and demonstrates how the (non-)destructiveness of different sub-detectors facilitates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

(fig:AllCMSCutaways)

854 charged particles is measured in the tracking detector by examining the curvature of the  
 855 particles path as it travels through the magnetic field. The radius of curvature  $r$  of a charged  
 856 particle in a magnetic field is given by

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

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

865 The CMS solenoid is extremely large. The radial bore of the magnet is 6.3 meters;  
 866 the magnet is 12.5 meters in length and weighs 220 tons. The large bore of the magnet  
 867 allows the tracker and calorimeter systems to be located inside the solenoid. The internal  
 868 windings of solenoid is arranged in four layers to increase the total field strength and are  
 869 cooled by liquid helium to a temperature of 4.5 Kelvin. The windings are magnetically  
 870 coupled to the support superstructure. This coupling allows the magnet to heat uniformly  
 871 during a “quench” event<sup>2</sup>, reducing localized stresses. The nominal current at full field of  
 872 the solenoid is 19.14 kA. The solenoid itself is surrounded by an iron return yoke with a  
 873 total mass of 10,000 tons. The return yoke surrounding the solenoid minimizes the fringing  
 874 field. The muon detector system is interspersed inside the yoke, and takes advantage of the  
 875 return field in the yoke to measure the momentum and charge of muons.

## 876 2.3 Charged Particle Tracking Systems

?<sec:Tracker>? The charged particle tracking system measures the trajectories of charged particles emerging  
 877 from the event. The tracker measures the trajectory of a charged particle by measuring

---

<sup>2</sup>A 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.

879 ‘‘hits’’ along the trajectory. Each hit corresponds to the global position of the trajectory  
 880 on a given surface. The trajectory can then be reconstructed by a helix to the points.  
 881 The tracker is designed to have a resolution that permits the reconstruction of ‘‘secondary  
 882 vertices’’ in  $b$ –quark and  $\tau$  lepton decays. To accomplish this, there are two types of tracking  
 883 detectors in CMS. The ‘‘pixel detector’’ comprises the inner layers (three in the barrel, two  
 884 in the endcaps). The pixel detector is situated as close as possible (4.4 cm) to the interaction  
 885 point and has a very high resolution. Outside of the pixel detector is the silicon strip tracker,  
 886 with ten layers in the barrel and 12 layers in the endcaps. A secondary vertex occurs when a  
 887 particle is semi–stable, traveling some non–negligible distance in the detector, but decaying  
 888 before the first layer of the tracking system. The pixel and strip tracking detectors have a  
 889 fiducial region which extends to a pseudorapidity of approximately  $|\eta| \approx 2.5$ .

890 Both the pixel and strip trackers are silicon–based. The principle of operation is similar  
 891 to that of a charged–coupled device (CCD) in a modern digital camera. The sensitive portion  
 892 of the detector is a silicon chip that is arranged with diode junctions formed by a  $p$ –doped  
 893 layer and an  $n$ –doped layer<sup>3</sup>. Each  $p – n$  junction is electrically isolated from adjacent layers.  
 894 The size of each junction region determines<sup>4</sup> the spatial resolution of the sensor. In the pixel  
 895 detector, each sensor region ‘‘pixel’’ is  $100 \mu\text{m} \times 150 \mu\text{m}$ . In the strip tracker, The rear side  
 896 of the chip is mounted to readout electronics. During operation, a high–voltage reverse bias  
 897 is applied to each  $p – n$  junction to achieve full depletion. When a charged particle passes  
 898 through the detector, the diode junction breaks down and the readout system registers the  
 899 hit.

900 The tracking system has been specifically designed for the high radiation environment  
 901 around the interaction point. The detector is cooled to  $-27^\circ\text{C}$  during operation to minimize  
 902 damage. Radiation exposure produced in LHC collisions can change behavior of the tracking  
 903 detector in three ways. Over time, radiation can induce positive holes in oxide layers fond  
 904 in the read–out electrons which increase the signal–to–noise ratio. In the sensor mass itself,

---

<sup>3</sup>The pixel detector actually uses a more complicated multi–layered scheme to improve radiation hardness. For details, see Section 3.2.2 of [24].

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

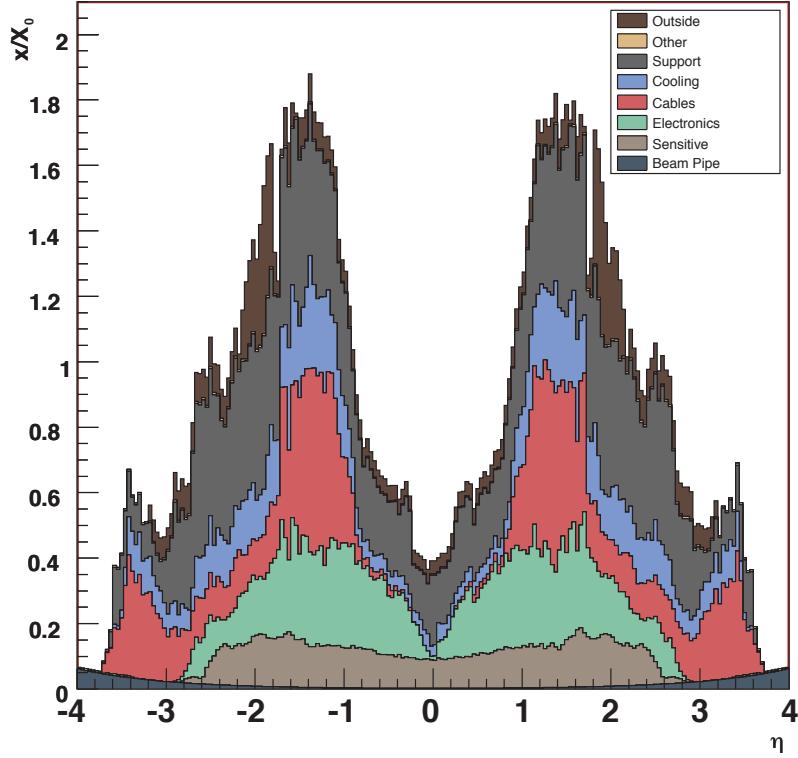


Figure 2.2: Material budget of the CMS tracker in units of radiation lengths  $X_0$  versus pseudorapidity. The material budget is broken down into the contributions from the different components of the tracker. The amount of material is largest in the “transition region” between the barrel and endcap.

`ig:TrackerMaterialBudget)`

radiation damage changes the doping from  $n$  to  $p$  over time. The required voltage to deplete the sensor will thus increase over time. The readout electronics, bias voltage supplies, and cooling systems are designed to scale with the radiation damage and maintain a signal-to-noise ratio of 10:1 or greater for 10 years of LHC operation. The final radiation effect is not an integrating effect. A “single event upset” is transient effect where an ionizing charged particle passes through the readout electronics and changes the state of the digital circuitry.

In the ideal case, the tracker would be a non-destructive instrument. However, charged particles can interact with the mass of the tracker (and its support infrastructure). These interactions limit the resolution of the tracker. The amount of matter in the tracker is referred to as the “material budget”. The material budget of the CMS tracker depends heavily on the pseudorapidity  $\eta$  and is illustrated in Figure 2.2. The relatively large ma-

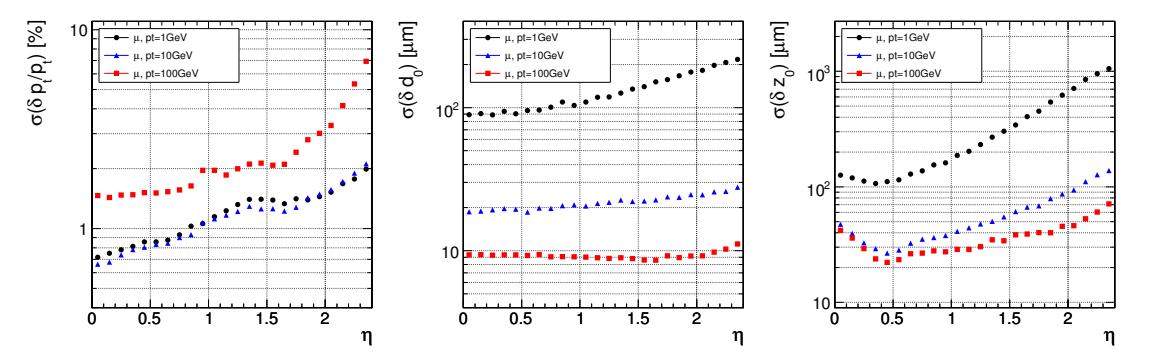


Figure 2.3: Expected resolutions of reconstructed transverse momentum (left), transverse impact parameter (center), and longitudinal impact parameter (right) versus absolute pseudorapidity  $|\eta|$ . The resolution is shown for three different cases of particle  $p_T$ , 1 GeV/c (black), 10 GeV/c (blue),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

(fig:ExpectedTrackerRes)

material budget of the CMS tracker has two effects: charged particles can undergo “multiple scattering,” interacting with material in the tracker. This can cause “kinks” in the reconstructed track. Hadronic particles (charged and neutral) can undergo nuclear interactions, 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.<sup>5</sup> The ECAL is a *scintillation*

<sup>5</sup>One 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.

933 detector, and functions by counting the number of photons produced in an electromagnetic  
934 shower inside a crystal. Upon entering the crystal, a charged particle or photon will inter-  
935 act electromagnetically with the crystal, producing a shower of electrons and photons. The  
936 shower will expand until it consists entirely of photons. The crystal is optically clear, so  
937 these photons travel to the rear face of the crystal where they are then counted by a pho-  
938 tomultiplier. The number of detected photons can then be related to the energy that was  
939 deposited in the crystal. At 18°C, about 4.5 photoelectrons will be produced per MeV of de-  
940 posited energy. The ECAL has excellent solid angle coverage, extending to a pseudorapidity  
941 of  $|\eta| = 3.0$ .

942 The ECAL uses lead tungstate ( $\text{PbWO}_4$ ) crystals as the scintillation medium. The  
943 crystals have a very large density, which allows the calorimeter to be relatively compact.  
944 To be able to correctly measure the energy of electrons and photons, an incident photon or  
945 electron must be completely stopped by interactions with the calorimeter. The quantities  
946 that determine if an electron or photon will be completely contained is the total depth of  
947 the crystal, the crystal density, and the radiation length property  $X_0$  of the crystal. The  
948 radiation length  $X_0$  is defined as the mean distance (normalized to material density) after  
949 which an electron will have lost  $(1 - \frac{1}{e})$  of its energy. The  $\text{PbWO}_4$  crystals of the CMS  
950 ECAL have a density of  $8.28 \text{ g/cm}^3$  and a depth of 230 mm. A single crystal thus has a  
951 total radiation length of  $25.8 X_0$ , and will capture on average 99.9993% of the energy of an  
952 incident electron. The front face of the crystal is  $22 \text{ mm} \times 22 \text{ mm}$ , which corresponds to an  
953  $\eta - \phi$  area of  $0.00174 \times 0.00174$ . The Molière radius of a material is the average radial profile  
954 size of an electromagnetic shower, and for  $\text{PbWO}_4$  is 2.2 cm. The fact that the Molière  
955 radius is larger than the size of the individual crystals improves the spatial resolution of  
956 the measurement. As the shower is shared between multiple crystals, the relative amounts  
957 deposited in each crystal allows the true impact point to be determined with a resolution  
958 smaller than the individual crystal size.

959 The transparency of the CMS ECAL crystals change as they are exposed to radiation.  
960 However, at the working temperature of the ECAL (18°C), the crystal transparency will  
961 naturally return to its nominal value. The transparency of the crystals thus decreases during  
962 the course of a run of collisions, then increases during the following collision-less period.

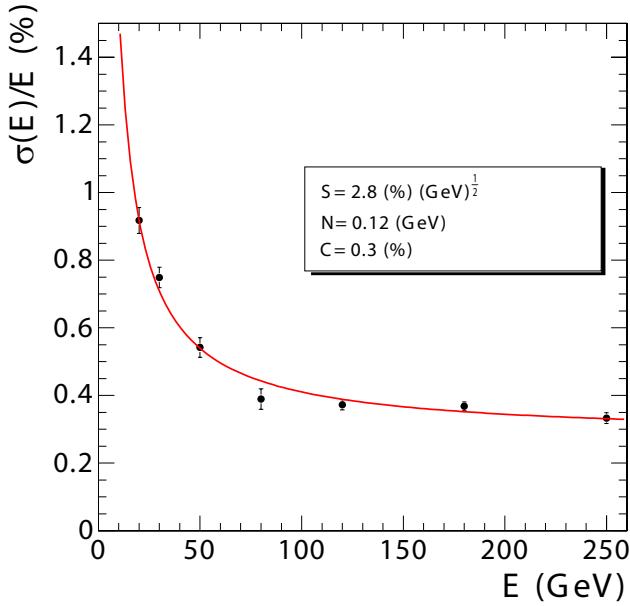


Figure 2.4: Energy resolution (in %) of the CMS ECAL measured at an electron test beam. The resolution depends on the incident energy of the electron. The points are fitted to function with the form given in Equation 2.2. The fitted parameters are given in the legend.

(fig:ECALResolution)

963 The changing transparency conditions need to be continuously monitored and corrected  
 964 for to ensure a stable detector response. The transparency of the crystals are measured  
 965 continuously using two lasers. One laser has wavelength  $\lambda = 400$  nm which corresponds to  
 966 the color of light produced in the scintillations and is sensitive to changes in transparency.  
 967 The other laser is in the near-infrared and is used to monitor the overall stability of the  
 968 crystal. The lasers are synchronized to pulse between LHC bunch trains so the transparency  
 969 can be continuously monitored while collisions are occurring.

970 The energy resolution of the ECAL is given by

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

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

## 975 2.5 Hadronic Calorimeter

?<sec:HCAL>?  
 976 The hadronic calorimeter (HCAL) surrounds the CMS ECAL and is located within the coil  
 977 of the CMS solenoid magnet. To ensure incident particles are completely contained within  
 978 the calorimeter volume, in the barrel region the HCAL employs a “tail–catcher”, an extra  
 979 layer of calorimetry outside of the magnet. The hadronic calorimeter measures the energy  
 980 of charged and neutral hadronic particles. The HCAL is a *sampling* calorimeter. Layers of  
 981 plastic scintillating tiles are interspersed between brass absorber plates. An incident hadron  
 982 produces a hadronic shower as it passes through the absorber. The particles in the shower  
 983 produce light as they pass through the scintillating tiles. Measuring the light produced in  
 984 each layer of tile allows the reconstruction of the radial profile of the shower which can be  
 985 related to the deposited energy. The response of the scintillator tiles are calibrated using a  
 986 radioactive source, either Cs<sup>137</sup> or Co<sup>60</sup>. Small stainless tubes permit the radioactive sources  
 987 to be moved into the center of the tile during calibration. The granularity of the HCAL is  
 988 0.087 × 0.087 and 0.17 × 0.17 in  $\eta - \phi$  in the barrel ( $|\eta| < 1.6$ ) and endcap ( $|\eta| > 1.6$ ).  
 989 respectively.

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

995 The HCAL includes a specially designed forward calorimeter (HF). The design of the  
 996 forward calorimeter is constrained by the extreme amount of radiation it is exposed to,  
 997 particularly at the highest rapidities. The active material of the HF are quartz fibers. The  
 998 fibers are installed inside grooves inside of a steel absorber. Charged particles created in  
 999 showers in the absorber create light in the fibers, provided they have energy greater than  
 1000 the with energy greater than the Cherenkov threshold. As Cherenkov light is created by  
 1001 the passage of charged particles through matter, the HF design is not sensitive to neutrons  
 1002 emitted by radionucleids that may be created in the absorber material durin operation.  
 1003 The fibers are grouped into two sets: one set of fibers are installed over the full depth of

1004 the detector, the other only cover half the depth. A crude form of particle identification  
1005 is possible, as showers created by electrons and photons will deposit the majority of the  
1006 energy in the front of the detector.

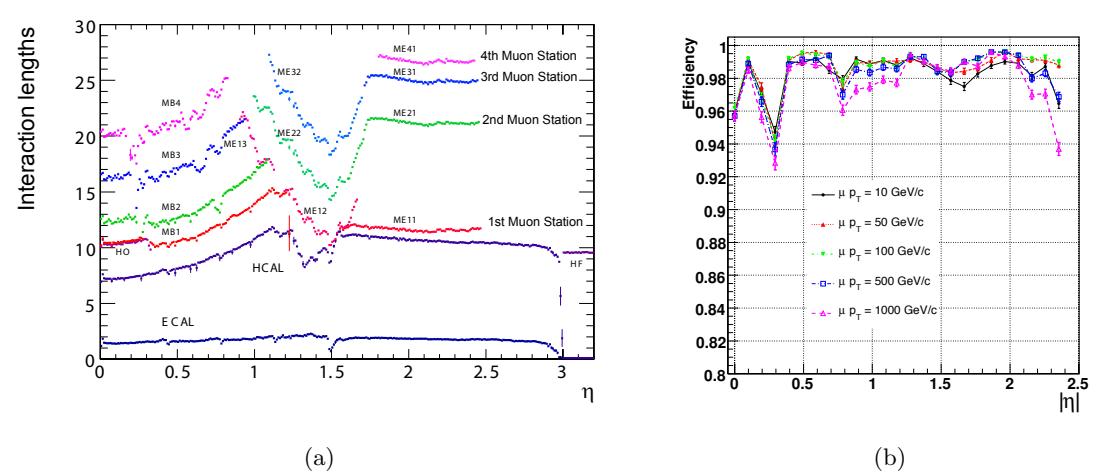
## 1007 2.6 Muon System

1008 The ability to detect and measure muons is one of the most valuable tools an experimentalist  
1009 has at a hadron collider experiment. Muons have particular properties that cause them to  
1010 leave extremely unique signatures in the detectors.

- 1011 • Muons are stable particles, for the typical energies and distances considered at a  
1012 collider.
- 1013 • Muons have non-zero charge, so their trajectories can be measured.
- 1014 • Muons are heavy enough and only interact electromagnetically so they lose very little  
1015 energy as they pass through the calorimeters.

1016 The approach to detecting muons is to build the detector to a thickness such that other  
1017 particles (electrons, photons, hadrons) will not penetrate the outermost calorimeter. Any  
1018 charged particle that is detected outside of this region can then be identified as a muon. At  
1019 CMS, the muon detection systems are built into the magnet return yoke outside of the CMS  
1020 calorimeters and magnet, giving them excellent protection (illustrated in Figure 2.5(a))  
1021 against hadronic “punch-through.” The purity of particles that reach the muon system  
1022 make it especially effective as a “trigger” of interesting physics. The CMS muon system has  
1023 the feature that it additionally can trigger on muons above a certain transverse momentum  
1024 threshold. The CMS muon system is composed of three types of detectors: drift tubes (DT),  
1025 resistive plate chambers (RPC), and cathode strip chambers (CSC).

1026 A drift tube detector is of a tube filled with a mixture of argon (85%) and carbon  
1027 dioxide (15%) gas with a positively charged ( $V = +3.6$  kV) wire running through the  
1028 middle of the tube. When a charged particle passes through the tube, it ionizes some gas.  
1029 The free electrons are then drawn to the positively charged wire inside the tube, creating a  
1030 signal when it reaches it. The speed of the detector is limited by the “drift time,” the maximum



etectorChapterMuonShit)?

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

amount of time it may take for an electron to reach a sensor wire. The precision of the spatial measurement can be increased by recording the time at which each wire records a signal and correlating the measurements across multiple tubes. The time resolution of the CMS DTs is on the order of a few nanoseconds, allowing the DT to provide a trigger on a given proton bunch crossing. The tubes in adjacent layers are offset by one half tube width to take advantage of this effect and ensure there are no gaps in the fiducial region. In CMS, the smallest unit of the DT system is the superlayer, which consists of four layers of tubes. A DT chamber consists of three or two superlayers. The tubes in the two superlayers farthest from the beam are oriented parallel to the beam and measured the bending of the muons in the magnetic field. The inner superlayer is oriented orthogonally to the beam and measures the longitudinal position of incident muons. There are four muon “stations” in the barrel which contain DT chambers. The stations correspond to available areas in the magnetic return yoke.

Cathode strip chambers (CSCs) are used in the endcap muon system, providing coverage in the pseudorapidity range  $0.9 < |\eta| < 2.4$ . A cathode strip chamber consists of a chamber filled with inert gas that with a number of internal wires held at a high voltage. A

1047 number of cathode strips are installed perpendicular<sup>6</sup> to the wires on the walls of the cham-  
 1048 ber. When a muon passes through the CSC, it ionizes some of the gas. The high voltage on  
 1049 a nearby wire causes this ionized gas to break down, forming a conductive passage in the  
 1050 gas and an “avalanche” current between the wire and a number of the cathode strips. The  
 1051 spatial position of the hit in two dimensions is found taking one coordinate from the wire  
 1052 and the other coordinate from the signal average of the cathode strips.

1053 The CSCs in the CMS endcap are positioned such that a muon in the pseudorapidity  
 1054 range  $1.2 < |\eta| < 2.4$  will cross three or four CSC detectors. The geometry of the CSC  
 1055 strips and wires is designed to provided a spatial  $r - \phi$  resolution of 2 mm at the L1 trigger  
 1056 level and a final offline reconstruction resolution of 75  $\mu\text{m}$  for the first layer and 150  $\mu\text{m}$  for  
 1057 outer layers. The RMS of the response time for a CSC layer is about 11 ns, which is too  
 1058 long to correctly associate a signal in the CSCs to an LHC bunch crossing (25 ns) with  
 1059 high efficiency. By grouping the layers into chambers, and taking the shortest response, the  
 1060 correct bunch crossing can be identified with 98–99% efficiency.

1061 The Resistive Plate Chamber (RPC) muon detectors ensure that the muon system can  
 1062 be used as a fast, first-level trigger. The RPC detector consists of two gaps filled with gas  
 1063 (up and down) with a common set of strips between the two gaps. The strips are oriented  
 1064 parallel to the beam line to permit measurement of the transverse momentum of the muons.

## 1065 2.7 Trigger System

1066 *(sec:Trigger)* At the LHC, proton bunches crossings (collisions) occur every 25 ns. This corresponds to  
 1067 an interaction rate of 40 MHz. At this high rate, and with the huge number of channels  
 1068 in the CMS detector, the front-end bandwidth readout from the detector is over 1 Pb/s.  
 1069 Due to bandwidth and storage requirements, the rate at which events are permanently  
 1070 recorded must be reduced by more than a factor of a million. This reduction is achieved  
 1071 by CMS trigger system. As only a fraction of the total events can be stored, and the rate  
 1072 of diffractive and common QCD multi-jet production is many orders of magnitude larger  
 1073 than “interesting” new physics (see Figure 1.11). The trigger must therefore be designed to

---

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

1074 select “interesting” events. A typical requirement applied at the trigger level might be the  
1075 presence of a high- $p_T$  muon, an isolated ECAL deposit, or a large deposit of energy in the  
1076 event.

1077 The CMS trigger consists of two stages: a fast Level–1 (L1) trigger and a High–Level  
1078 Trigger (HLT). The L1 trigger system is built on custom, typically reprogrammable elec-  
1079 tronics and interfaces directly to the detector subsystems. The L1 trigger has access to  
1080 information from the muon and calorimeter systems. The L1 does not have access to the  
1081 full granularity of the muon system and calorimeters but must make the decision based  
1082 on coarse segments. The design acceptance rate of the L1 trigger is 100 kHz. The trigger  
1083 typically operates at a nominal rate of 30 kHz. The maximum latency of the L1 is 3.2  $\mu$ s,  
1084 requiring that the output from detector electronics be passed through memory pipelines to  
1085 ensure that no bunch crossings go unanalyzed. The High–Level Trigger (HLT) runs on a  
1086 farm of about 1000 commercial compute nodes and processes events that are accepted by  
1087 the L1 trigger. An HLT decision (“path”) has the ability to reconstruct tracks and do a full  
1088 regional unpacking of the recorded hits in a regions of the calorimeter. Each HLT path has  
1089 a strict rate budget, as the total rate of the HLT is required to be less than 100 Hz. The  
1090 triggers used at CMS change as the conditions change. To limit the total rate to 100 Hz as  
1091 the luminosity increases, trigger paths must either increase their thresholds, or apply a  
1092 “prescale.” When a prescale is applied, a fraction of events passing the trigger are thrown  
1093 away randomly.

1094 The CMS trigger is a deep subject and a complete description is beyond the scope of  
1095 this thesis. A detailed description can be found in [27]. The triggers used in the analysis  
1096 presented in this thesis will be briefly described. Two types of trigger selections were applied  
1097 to the 2010 datasets used in this analysis. During the initial period of low luminosity running,  
1098 single muon triggers were used. As the luminosity increased, the  $p_T$  threshold of the trigger  
1099 was increased. In some cases, an “isolated muon” HLT trigger was required, in which a veto  
1100 was applied on muons with associated energy deposits in the calorimeter. In the final period  
1101 of data taking, two “cross–triggers” were used. These required the presence of both a muon  
1102 and a hadronic tau decay in the event. The triggers used in this analysis in the different  
1103 2010 run periods are enumerated in Table 5.1.

1104 The muon component of all the triggers used in this analysis is based on the “L1 seed  
1105 trigger” L1\_SingleMu7, which nominally selects event which contain a muon with  $p_T >$   
1106 7 GeV/c. The L1 muon trigger decision is determined by the Global Muon Trigger (GMT),  
1107 which combines information from the DT, CSC, and RPC sub-detectors, and is able to  
1108 trigger muons up to a pseudorapidity of  $|\eta| < 2.1$ . Each sub-detector has a “local trigger,”  
1109 which can reconstruct tracks in the muon system. For the drift tubes, the Bunch Track  
1110 Identifiers (BTI), a custom integrated circuit, searches for aligned hits in the associated  
1111 DT chamber. The CSCs and RPCs employ similar strategies to detect local muon tracks.  
1112 The sub-detectors send the GMT the charge,  $p_T$ ,  $\eta$ ,  $\phi$ , and a quality code of up to four  
1113 local muons. The measurements from the sub-detectors are combined and a final decision  
1114 is made by the GMT.

## Chapter 3

# Tau Identification: The Tau Neural Classifier

{ch:tanc}  
1118 Excellent tau identification performance is important for the discovery potential of many  
1119 possible new physics signals at CMS. The standard model background rates from true tau  
1120 leptons are typically the same order of magnitude as the expected signal rate in many  
1121 searches for new physics. The challenge of doing physics with taus is driven by the rate  
1122 at which objects are incorrectly tagged as taus. In particular, quark and gluon jets have a  
1123 significantly higher production cross-section and events where these objects are incorrectly  
1124 identified as tau leptons can dominate the backgrounds of searches for new physics using  
1125 taus. Efficient identification of hadronic tau decays and low misidentification rate for quarks  
1126 and gluons is thus essential to maximize the significance of searches for new physics at CMS.

1127 Tau leptons are unique in that they are the only type of leptons which are heavy enough  
1128 to decay to hadrons. The hadronic decays comprise approximately 65% of all tau decays, the  
1129 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  
1130 decays are typically composed of one or three charged pions and zero to two neutral pions.  
1131 The neutral pions decay almost instantaneously to pairs of photons.

1132 In this chapter, we describe a technique to identify hadronic tau decays. Tau decays  
1133 to electrons and muons are difficult to distinguish from prompt production of electrons and  
1134 muons in  $pp$  collisions. Analyses that use exclusively use the leptonic ( $e, \mu$ ) decays of taus  
1135 typically require that the decays be of opposite flavor. With the Tau Neural Classifier, we  
1136 aim to improve the discrimination of true hadronic tau decays from quark and gluon jets  
1137 using a neural network approach.

### 1138 3.1 Geometric Tau Identification Algorithms

1139 ⟨sec:GeometricTauId⟩ The tau identification strategies used in previously published CMS analyses are fully de-  
 1140 scribed in [28]. A summary of the basic methods and strategies is given here. There are  
 1141 two primary methods for selecting objects used to reconstruct tau leptons. The CaloTau  
 1142 algorithm uses tracks reconstructed by the tracker and clusters of hits in the electromag-  
 1143 netic and hadronic calorimeter. The other method (PFTau) uses objects reconstructed by  
 1144 the CMS particle flow algorithm, which is described in [29]. The particle flow algorithm  
 1145 provides a global and unique description of every particle (charged hadron, photon, elec-  
 1146 tron, etc.) in the event; measurements from subdetectors are combined according to their  
 1147 measured resolutions to improve energy and angular resolution and reduce double counting.  
 1148 All of the tau identification strategies described in this thesis use the particle flow objects.

1149 Both methods typically use an “leading object” and an isolation requirement to reject  
 1150 quark and gluon jet background. Quark and gluon jets are less collimated and have a higher  
 1151 constituent multiplicity and softer constituent  $p_T$  spectrum than a hadronic tau decay of  
 1152 the same transverse momentum. The “leading track” requirement is applied by requiring a  
 1153 relatively high momentum object near the center of the jet; typically a charged track with  
 1154 transverse momentum greater than 5 GeV/c within  $\Delta R < 0.1$  about the center of the jet  
 1155 axis. The isolation requirement exploits the collimation of true taus by defining an isolation  
 1156 annulus about the kinematic center of the jet and requiring no detector activity about a  
 1157 threshold in that annulus. This approach yields a misidentification rate of approximately 1%  
 1158 for QCD backgrounds and a hadronic tau identification efficiency of approximately 50% [28].

### 1159 3.2 Decay Mode Tau Identification: Motivation

1160 The tau identification strategy described previously can be extended by looking at the dif-  
 1161 ferent hadronic decay modes of the tau individually. The dominant hadronic decays of taus  
 1162 consist of a one or three charged  $\pi^\pm$  mesons and up to two  $\pi^0$  mesons and are enumerated  
 1163 in Table 1.4. The majority of these decays proceed through intermediate resonances and  
 1164 each of these decay modes maps directly to a tau final state multiplicity. Each intermediate  
 1165 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.

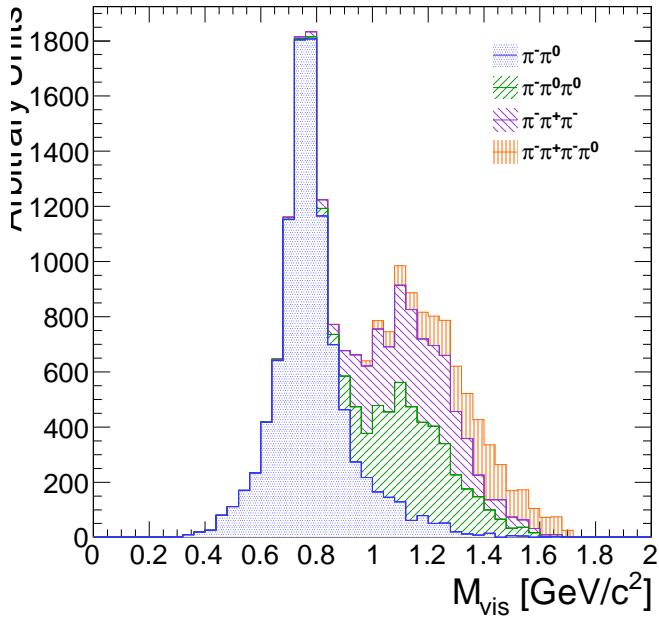


Figure 3.1: The invariant mass of the visible decay products in hadronic tau decays. The decay mode  $\tau^- \rightarrow \pi^- \nu_\tau$  is omitted. The different decay modes have different invariant masses corresponding to the intermediate resonance in the decay.

(fig:trueInvMass)

### 1172 3.3 The Tau Neural Classifier

1173 (sec:Tanc) The Tau Neural Classifier (TaNC) algorithm reconstructs the decay mode of the tau can-  
 1174 didate and then feeds the tau candidate to a discriminator associated to that decay mode  
 1175 to make the classification decision. Each discriminator therefore maps to a reconstructed  
 1176 decay mode in a one-to-one fashion. To optimize the discrimination for each of the different  
 1177 decay modes, the TaNC uses an ensemble of neural nets. Each neural net corresponds to one  
 1178 of the dominant hadronic decay modes of the tau lepton. These selected hadronic decays

1179 constitute 95% of all hadronic tau decays. Tau candidates with reconstructed decay modes  
 1180 not in the set of dominant hadronic modes are immediately tagged as background.

### 1181 3.3.1 Decay Mode Reconstruction

1182 ⟨sec:decay mode reco⟩ The major task in reconstructing the decay mode of the tau is determining the number of  
 1183  $\pi^0$  mesons produced in the decay. A  $\pi^0$  meson decays almost instantaneously to a pair of  
 1184 photons. The photon objects are reconstructed using the particle flow algorithm [29]. The  
 1185 initial collection of photon objects considered to be  $\pi^0$  candidates are the photons in the  
 1186 signal cone described by using the “shrinking-cone” tau algorithm, described elsewhere [28].

1187 The reconstruction of photons from  $\pi^0$  decays present in the signal cone is complicated  
 1188 by a number of factors. To suppress calorimeter noise and underlying event photons, all  
 1189 photons with minimum transverse energy less than 0.5 GeV are removed from the signal  
 1190 cone, which removes some signal photons. Photons produced in secondary interactions,  
 1191 pile-up events, and electromagnetic showers produced by signal photons that convert to  
 1192 electron–positron pairs can contaminate the signal cone with extra low transverse energy  
 1193 photons. Highly boosted  $\pi^0$  mesons may decay into a pair of photons with a small opening  
 1194 angle, resulting in two overlapping showers in the ECAL being reconstructed as one photon.  
 1195 The  $\pi^0$  meson content of the tau candidate is reconstructed in two stages. First, photon  
 1196 pairs are merged together into candidate  $\pi^0$  mesons. The remaining unmerged photons are  
 1197 then subjected to a quality requirement.

### 1198 3.3.2 Photon Merging

1199 Photons are merged into composite  $\pi^0$  candidates by examining the invariant mass of all  
 1200 possible pairs of photons in the signal region. Only  $\pi^0$  candidates (photon pairs) with a  
 1201 composite invariant mass less than 0.2 GeV/c are considered. The combination of the high  
 1202 granularity of the CMS ECAL and the particle flow algorithm provide excellent energy  
 1203 and angular resolution for photons; the  $\pi^0$  mass peak is readily visible in the invariant  
 1204 mass spectrum of signal photon pairs (see figure 3.3.2). The  $\pi^0$  candidates that satisfy the  
 1205 invariant mass requirement are ranked by the difference between the composite invariant  
 1206 mass of the photon pair and the invariant mass of the  $\pi^0$  meson given by the PDG [20]. The

1207 best pairs are then tagged as  $\pi^0$  mesons, removing lower-ranking candidate  $\pi^0$ s as necessary  
 1208 to ensure that no photon is included in more than one  $\pi^0$  meson.

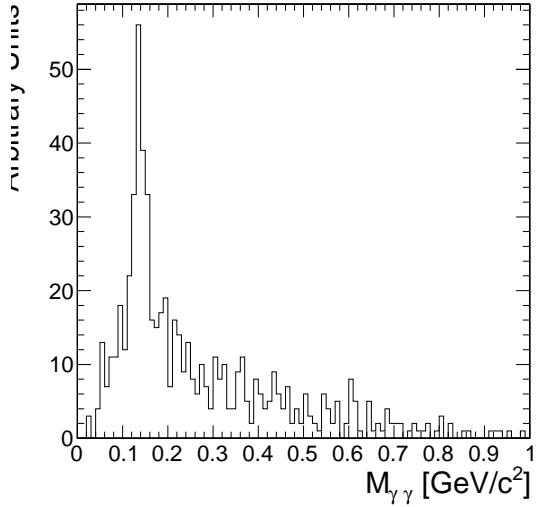


Figure 3.2: Invariant mass of the photon pair for reconstructed tau candidates with two reconstructed photons in the signal region that are matched to generator level  $\tau^- \rightarrow \pi^- \pi^0 \nu_\tau$  decays.

mDiPhotonsForTrueDM1)

### 1209 3.3.3 Quality Requirements

1210 Photons from the underlying event and other reconstruction effects cause the number of  
 1211 reconstructed photons to be greater than the true number of photons expected from a given  
 1212 hadronic tau decay. Photons that have not been merged into a  $\pi^0$  meson candidate are  
 1213 recursively filtered by requiring that the fraction of the transverse momentum carried by  
 1214 the lowest  $p_T$  photon be greater than 10% with respect to the entire (tracks,  $\pi^0$  candidates,  
 1215 and photons) tau candidate. In the case that a photon is not merged but meets the minimum  
 1216 momentum fraction requirement, it is considered a  $\pi^0$  candidate. This requirement removes  
 1217 extraneous photons, while minimizing the removal of single photons that correspond to a  
 1218 true  $\pi^0$  meson (see Figure 3.3). A mass hypothesis with the nominal [20] value of the  $\pi^0$   
 1219 is applied to all  $\pi^0$  candidates. All objects that fail the filtering requirements are moved to  
 1220 the isolation collection.

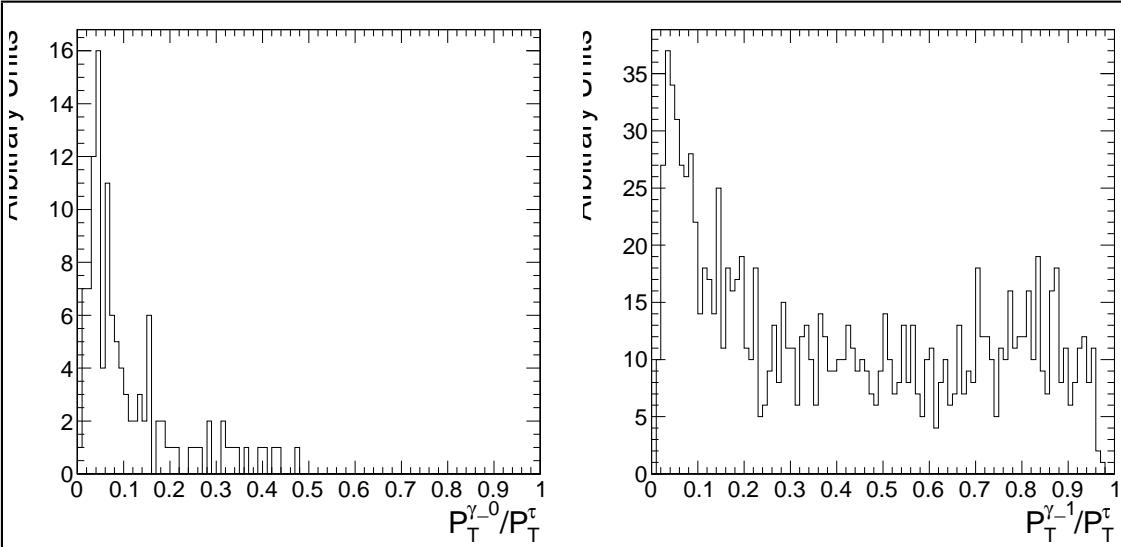


Figure 3.3: Fraction of total  $\tau$ -candidate transverse momenta carried by the photon for reconstructed taus containing a single photons for two benchmark cases. On the left, the reconstructed tau candidate is matched to generator level  $\tau^- \rightarrow \pi^- \nu_\tau$  decays, for which no photon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

(fig:photonFiltering)

### 3.3.4 Performance

The performance of the decay mode reconstruction can be measured for tau candidates that are matched to generator level hadronically decaying tau leptons by examining the correlation of the reconstructed decay mode to the true decay mode determined from the Monte Carlo generator level information. Figure 3.4 compares the decay mode reconstruction performance of a naive approach where the decay mode is determined by simply counting the number of photons to the performance of the photon merging and filtering approach described in Section 3.3.1. The correlation for the merging and filtering algorithm is much more diagonal, indicating higher performance. The performance is additionally presented for comparison in tabular form in Table 3.3.4 (merging and filtering approach) and Table 3.3.4 (naive approach).

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

1235 transverse momentum as the constituents are below the minimum  $p_T$  quality requirements  
 1236 In the forward region, nuclear interactions and conversions from the increased material  
 1237 budget enhances modes containing  $\pi^0$  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%

:dmResolutionNoNothing) 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%

:ab:dmResolutionStandard) 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.

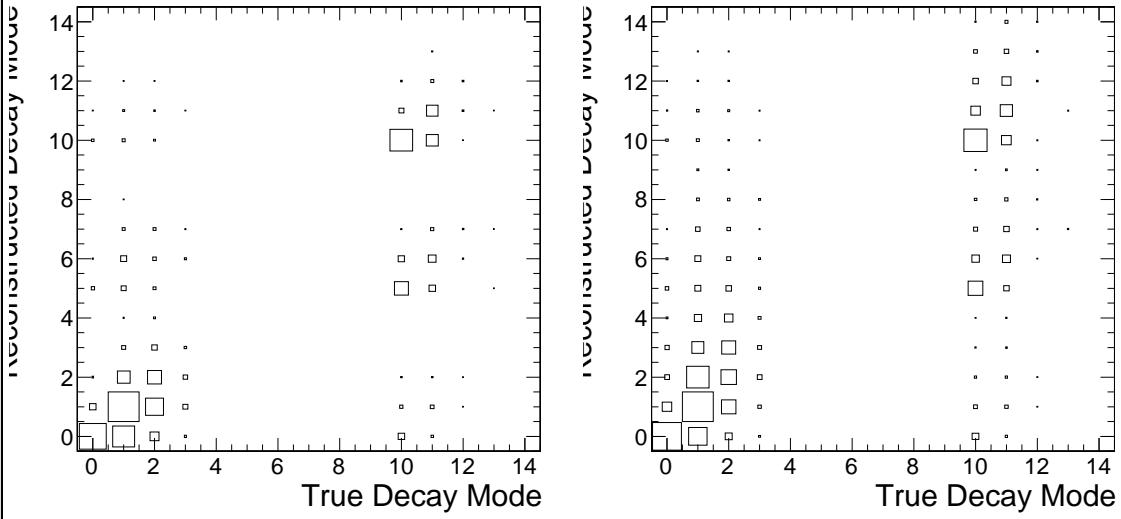


Figure 3.4: Correlations between reconstructed tau decay mode and true tau decay mode for hadronic tau decays in  $Z \rightarrow \tau^+\tau^-$  events. The correlation when no photon merging or filtering is applied is shown on the right, and the correlation for the algorithm described in Section 3.3.1 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  $\pi^0$  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.

(fig:dmResolution)

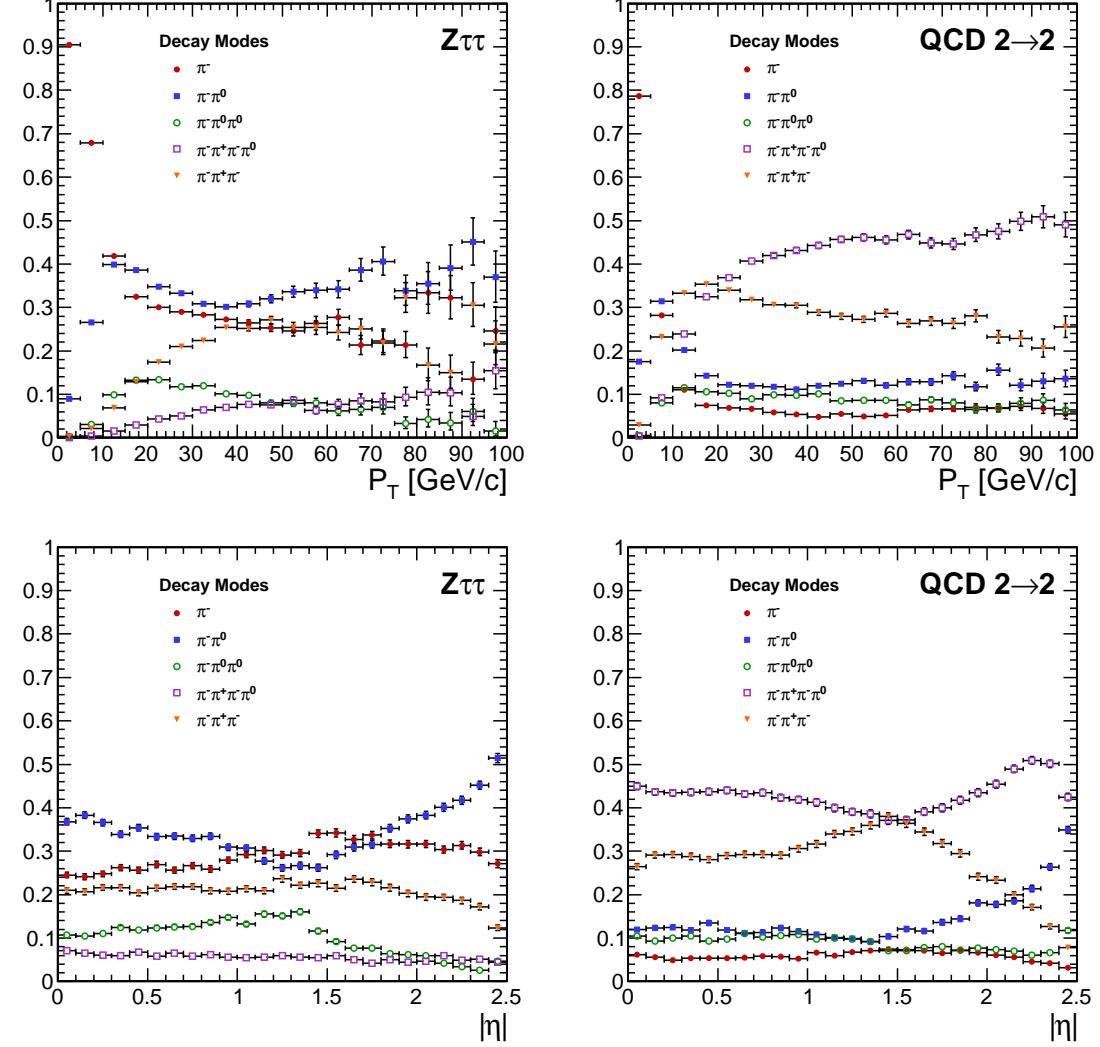


Figure 3.5: Kinematic dependence of reconstructed decay mode for tau candidates in  $Z \rightarrow \tau^+\tau^-$  (left) and QCD dijets (right) events versus transverse momentum (top) and pseudo-rapidity (bottom). Each curve is the probability for a tau candidate to be reconstructed with the associated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

(fig:dmKinematics)

### 1238 3.3.5 Neural Network Classification

(sec:tanc'nn'training)

### 1239 3.3.6 Neural Network Training

1240 The samples used to train the TaNC neural networks are typical of the signals and back-  
 1241 grounds found in common physics analyses using taus. The signal-type training sample is  
 1242 composed of reconstructed tau candidates that are matched to generator level hadronic tau  
 1243 decays coming from simulated  $Z \rightarrow \tau^+ \tau^-$  events. The background training sample consists  
 1244 of reconstructed tau candidates in simulated QCD  $2 \rightarrow 2$  hard scattering events. The QCD  
 1245  $p_T$  spectrum is steeply falling, and to obtain sufficient statistics across a broad range of  $p_T$   
 1246 the sample is split into different  $\hat{p}_T$  bins. Each binned QCD sample imposes a generator  
 1247 level cut on the transverse momentum of the hard interaction. During the evaluation of  
 1248 discrimination performance the QCD samples are weighted according to their respective  
 1249 integrated luminosities to remove any effect of the binning.

1250 The signal and background samples are split into five subsamples corresponding to  
 1251 each reconstructed decay mode. An additional selection is applied to each subsample by  
 1252 requiring a “leading pion”: either a charged hadron or gamma candidate with transverse  
 1253 momentum greater than 5 GeV/c. A large number of QCD training events is required as  
 1254 both the leading pion selection and the requirement that the decay mode match one of the  
 1255 dominant modes given in Table 1.4 are effective discriminants. For each subsample, 80% of  
 1256 the signal and background tau candidates are used for training the neural networks, with  
 1257 half (40%) used as a validation sample used to ensure the neural network is not over-trained.  
 1258 The number of signal and background entries used for training and validation in each decay  
 1259 mode subsample is given in Table 3.3.6.

1260 The remaining 20% of the signal and background samples are reserved as a statisti-  
 1261 cally independent sample to evaluate the performance of the neural nets after the training  
 1262 is completed. The TaNC uses the Multi-layer Perceptron (MLP) neural network implemen-  
 1263 tation provided by the TMVA software package, described in [30]. The MLP classifier is a  
 1264 feed-forward artificial neural network. There are two layers of hidden nodes and a single  
 1265 node in the output layer. The hyperbolic tangent function is used for the neuron activation  
 1266 function.

	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	
$\pi^-$	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

(tab:trainingEvents)

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

1267      The neural networks used in the TaNC have two hidden layers and single node in the  
 1268      output layers. The number of nodes in the first and second hidden layers are chosen to be  
 1269       $N + 1$  and  $2N + 1$ , respectively, where  $N$  is the number of input observables for that neural  
 1270      network. According to the Kolmogorov's theorem [31], any continuous function  $g(x)$  defined  
 1271      on a vector space of dimension  $d$  spanned by  $x$  can be represented by

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

1272      for suitably chosen functions for  $\Phi_j$  and  $\phi_j$ . As the form of Equation 3.1 is similar to the  
 1273      topology of a two hidden-layer neural network, Kolmogorov's theorem suggests that *any*  
 1274      classification problem can be solved with a neural network with two hidden layers containing  
 1275      the appropriate number of nodes.

1276      The neural network is trained for 500 epochs. At ten epoch intervals, the neural network  
 1277      error is computed using the validation sample to check for over-training (see Figure 3.6).  
 1278      The neural network error  $E$  is defined [30] as

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

1279 where  $N$  is the number of training events,  $y_{ANN,i}$  is the neural network output for the  $i$ th  
 1280 training event, and  $y_i$  is the desired (-1 for background, 1 for signal) output the  $i$ th event.  
 1281 No evidence of over-training is observed.

1282 The neural networks use as input observables the transverse momentum and  $\eta$  of the  
 1283 tau candidates. These observables are included as their correlations with other observables  
 1284 can increase the separation power of the ensemble of observables. For example, the opening  
 1285 angle in  $\Delta R$  for signal tau candidates is inversely related to the transverse momentum,  
 1286 while for background events the correlation is very small [32]. In the training signal and  
 1287 background samples, there is significant discrimination power in the  $p_T$  spectrum. However,  
 1288 for a general-purpose tau identification algorithm, it is desirable to eliminate any systematic  
 1289 dependence of the neural network output on  $p_T$  and  $\eta$ , as in practice the TaNC will be  
 1290 presented with tau candidates whose  $p_T - \eta$  spectrum will be analysis dependent. The  
 1291 optimal situation would be to train the neural networks using the signal and backgrounds  
 1292 that are present in the analysis where the algorithm is applied. In practice, it is difficult to  
 1293 collect enough training samples for each analysis situation. The dependence on  $p_T$  and  $\eta$  is  
 1294 removed by applying a  $p_T$  and  $\eta$  dependent weight to the tau candidates when training the  
 1295 neural nets.

The weights are defined such that in any region in the vector space spanned by  $p_T$   
 and  $\eta$  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

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

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

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

1296 where  $p_{sig}(p_T, \eta)$  and  $p_{bkg}(p_T, \eta)$  are the probability densities of the signal and background  
 1297 samples after the “leading pion” and dominant decay mode selections. Figure 3.7 shows the  
 1298 signal and background training  $p_T$  distributions before and after the weighting is applied.

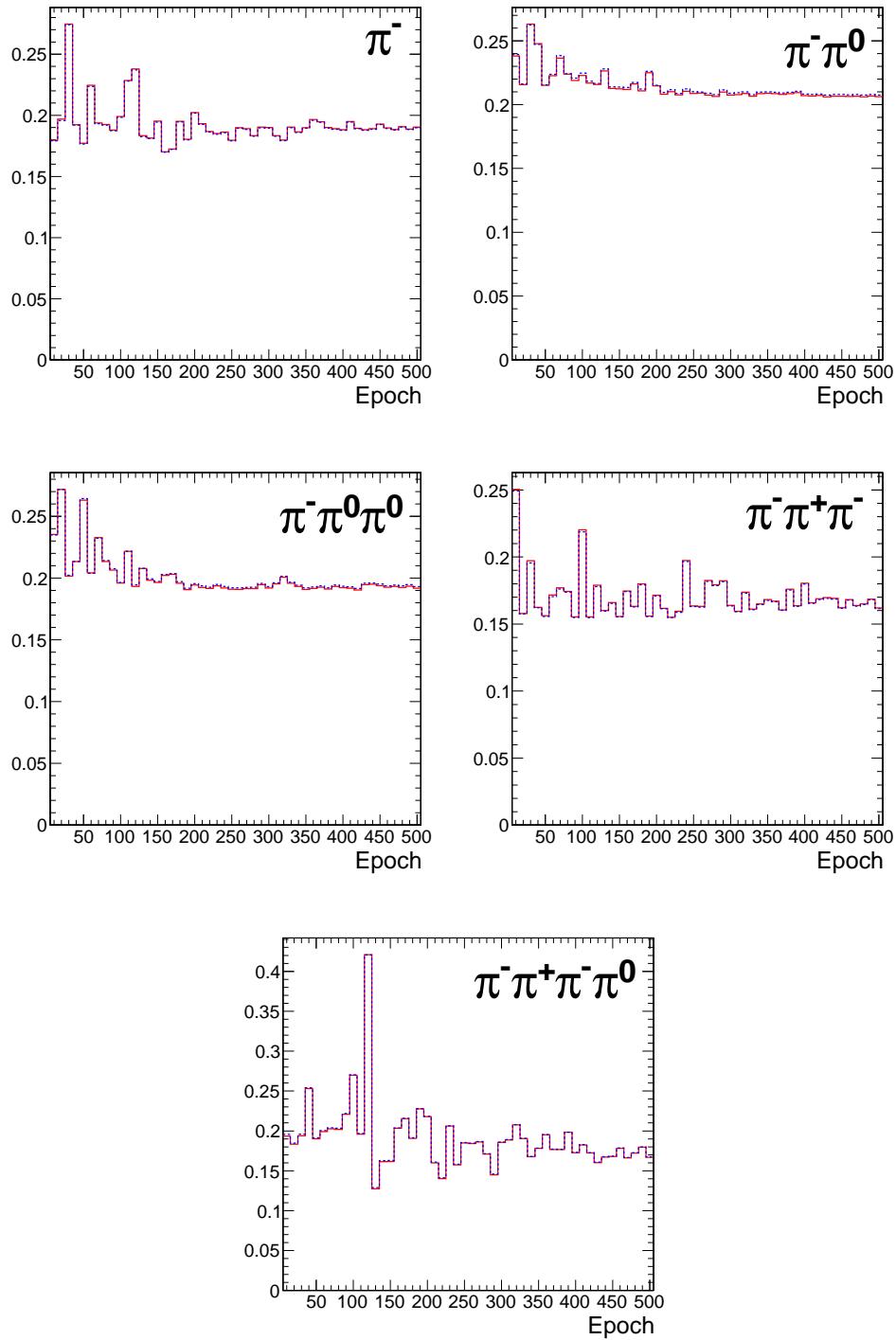


Figure 3.6: Neural network classification error for training (solid red) and testing (dashed blue) samples at ten epoch intervals over the 500 training epochs for each decay mode neural network. The vertical axis represents the classification error, defined by equation 3.2. N.B. that the choice of hyperbolic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

{fig:overTrainCheck}

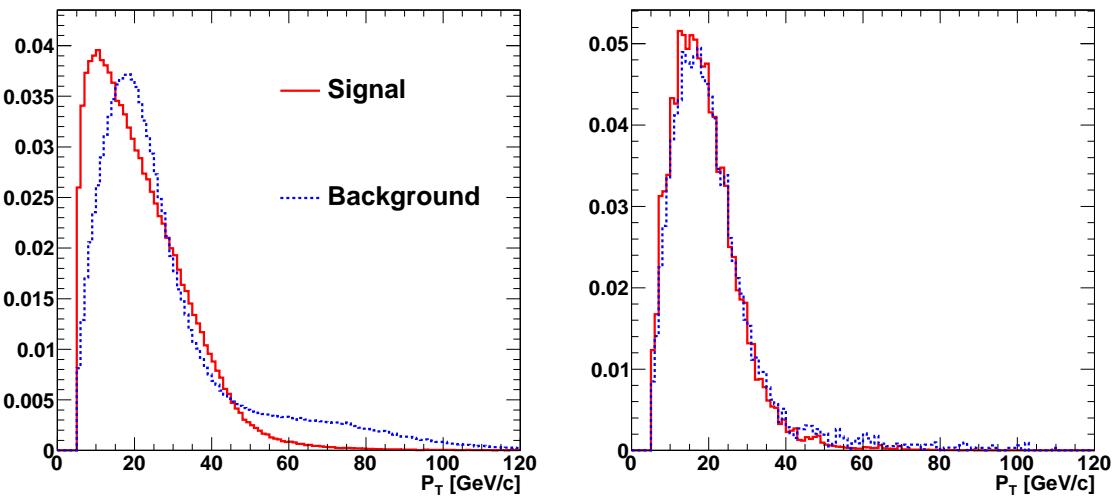


Figure 3.7: Transverse momentum spectrum of signal and background tau candidates used in neural net training before (left) and after (right) the application of  $p_T - \eta$  dependent weight function. Application of the weights lowers the training significance of tau candidates in regions of  $p_T -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

(fig:nnTrainingWeights)

### 1299 3.3.7 Discriminants

sec:tanc'nn'discriminants  
1300 Each neural network corresponds to a different decay mode topology and as such each  
1301 network uses different observables as inputs. However, many of the input observables are  
1302 used in multiple neural nets. The superset of all observables is listed and defined below.  
1303 Table 3.4 maps the input observables to their associated neural networks. In three prong  
1304 decays, the definition of the “main track” is important. The main track corresponds to the  
1305 track with charge opposite to that of the total charge of the three tracks. This distinction is  
1306 made to facilitate the use of the “Dalitz” observables, allowing identification of intermediate  
1307 resonances in three-body decays. This is motivated by the fact that the three prong decays  
1308 of the tau generally proceed through  $\tau^- \rightarrow a_1^- \nu_\tau \rightarrow \pi^- \rho^0 \nu_\tau \rightarrow \pi^- \pi^+ \pi^- \nu_\tau$ ; the oppositely  
1309 charged track can always be identified with the  $\rho^0$  decay.

1310 **ChargedOutlierAngleN**  $\Delta R$  between the Nth charged object (ordered by  $p_T$ ) in the  
1311 isolation region and the tau candidate momentum axis. If the number of isolation  
1312 region objects is less than N, the input is set at one.

1313 **ChargedOutlierPtN** Transverse momentum of the Nth charged object in the isolation  
 1314 region. If the number of isolation region objects is less than N, the input is set at zero.

1315 **DalitzN** Invariant mass of four vector sum of the “main track” and the Nth signal region  
 1316 object.

1317 **Eta** Pseudo-rapidity of the signal region objects.

1318 **InvariantMassOfSignal** Invariant mass of the composite object formed by the signal  
 1319 region constituents.

1320 **MainTrackAngle**  $\Delta R$  between the “main track” and the composite four–vector formed  
 1321 by the signal region constituents.

1322 **MainTrackPt** Transverse momentum of the “main track.”

1323 **OutlierNCharged** Number of charged objects in the isolation region.

1324 **OutlierSumPt** Sum of the transverse momentum of objects in the isolation region.

1325 **PiZeroAngleN**  $\Delta R$  between the Nth  $\pi^0$  object in the signal region (ordered by  $p_T$ ) and  
 1326 the tau candidate momentum axis.

1327 **PiZeroPtN** Transverse momentum of the Nth  $\pi^0$  object in the signal region.

1328 **TrackAngleN**  $\Delta R$  between the Nth charged object in the signal region (ordered by  $p_T$ )  
 1329 and the tau candidate momentum axis, exclusive of the main track.

1330 **TrackPtN** Transverse momentum of the Nth charged object in the signal region, exclusive  
 1331 of the main track.

1332 **3.3.8 Neural Network Performance**

neuralNetworkPerformance  
 1333 The classification power of the neural networks is unique for each of the decay modes.  
 1334 The performance is determined by the relative separation of the signal and background  
 1335 distributions in the parameter space of the observables used as neural network inputs. A  
 1336 pathological example is the case of tau candidates with the reconstructed decay mode of

1337  $\tau^- \rightarrow \pi^- \nu_\tau$ . If there is no isolation activity, the neural net has no handle with which it  
 1338 can separate the signal from the background. The neural net output for tau candidates in  
 1339 the testing sample (independent of the training and validation samples) for each of the five  
 1340 decay mode classifications is shown in Figure 3.8.

1341 When a single neural network is used for classification, choosing an operating point is  
 1342 relatively straightforward: the requirement on neural network output is tuned such that the  
 1343 desired purity is attained. However, in the case of the TaNC, multiple neural networks are  
 1344 used. Each network has a unique separation power (see Figure 3.9) and each neural network  
 1345 is associated to a reconstructed decay mode that composes different relative fractions of the  
 1346 signal and background tau candidates. Therefore, a set of five numbers is required to define  
 1347 an “operating point” (the signal efficiency and background misidentification rate) in the  
 1348 TaNC output. All points in this five dimensional cut–space map to an absolute background  
 1349 fake–rate and signal efficiency rate. Therefore there must exist a 5D “performance curve”  
 1350 which for any attainable signal efficiency gives the lowest fake–rate. A direct method to  
 1351 approximate the performance curve is possible using a Monte Carlo technique.

1352 The maximal performance curve can be approximated by iteratively sampling points in  
 1353 the five–dimensional cut space and selecting the highest performance points. The collection  
 1354 of points in the performance curve are ordered by expected fake rate. During each iteration,  
 1355 the sample point is compared to the point before the potential insertion position of the  
 1356 sample in the ordered collection. The sample point is inserted into the collection if it has  
 1357 a higher signal identification efficiency than the point before it. The sample point is then  
 1358 compared to all points in the collection after it (i.e. those with a larger fake rate); any point  
 1359 with a lower signal efficiency than the sample point is removed. After the performance curve  
 1360 has been determined, the set of cuts are evaluated on an independent validation sample  
 1361 to ensure that the measured performance curve is not influenced by favorable statistical  
 1362 fluctuations being selected by the Monte Carlo sampling. The performance curves for two  
 1363 different transverse momentum ranges are shown in Figure 3.10.

1364 The 5D performance curve can also be parameterized by using the probability for a  
 1365 tau candidate to be identified for a given decay mode. An artificial neural network maps  
 1366 a point in the space of input observables to some value of neural network output  $x$ . The

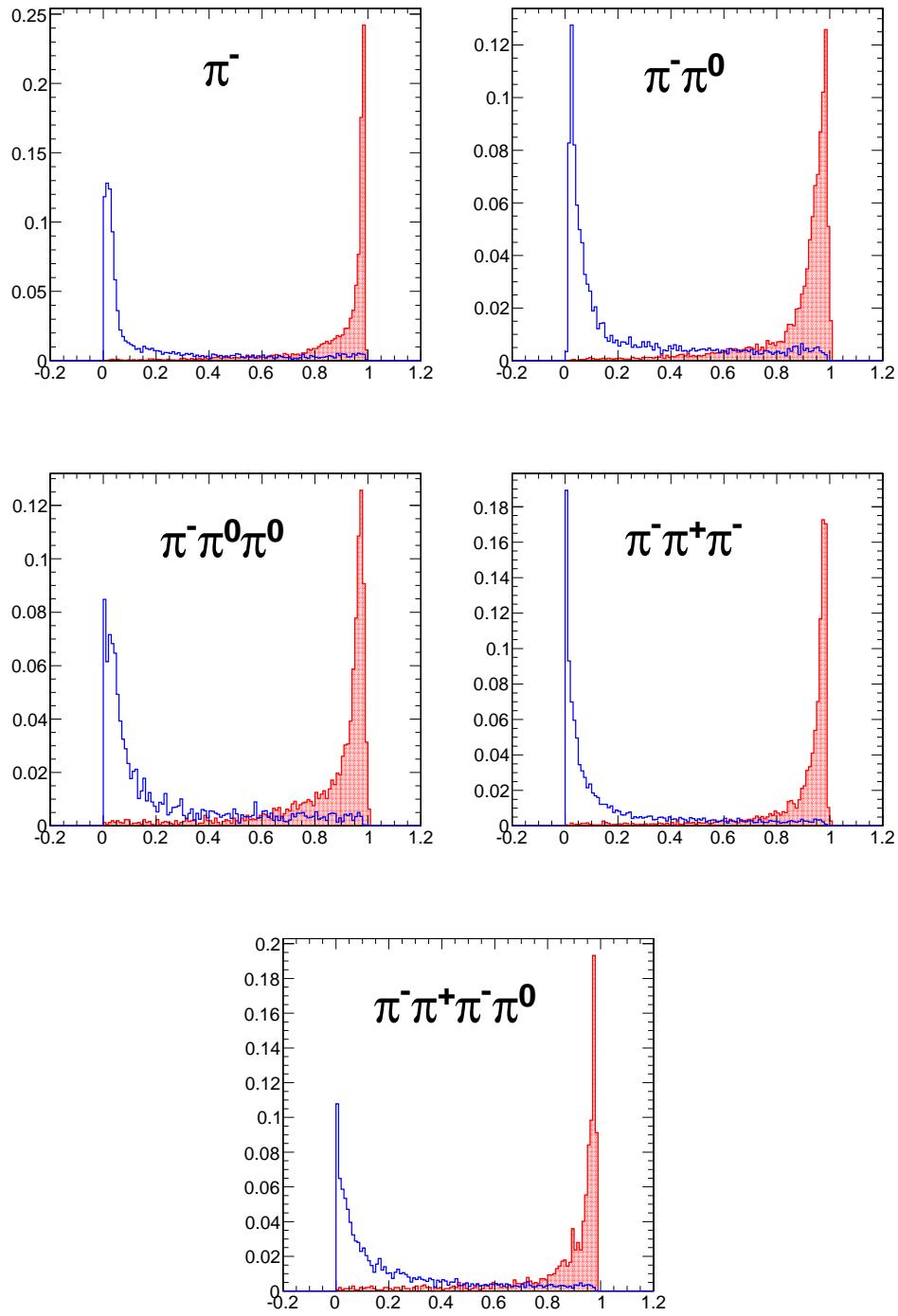


Figure 3.8: Neural network output distributions for the five reconstructed tau candidate decay modes used in the TaNC for  $Z \rightarrow \tau^+\tau^-$  events (red) and QCD dijet events (blue).

fig:NNoutputDistributions

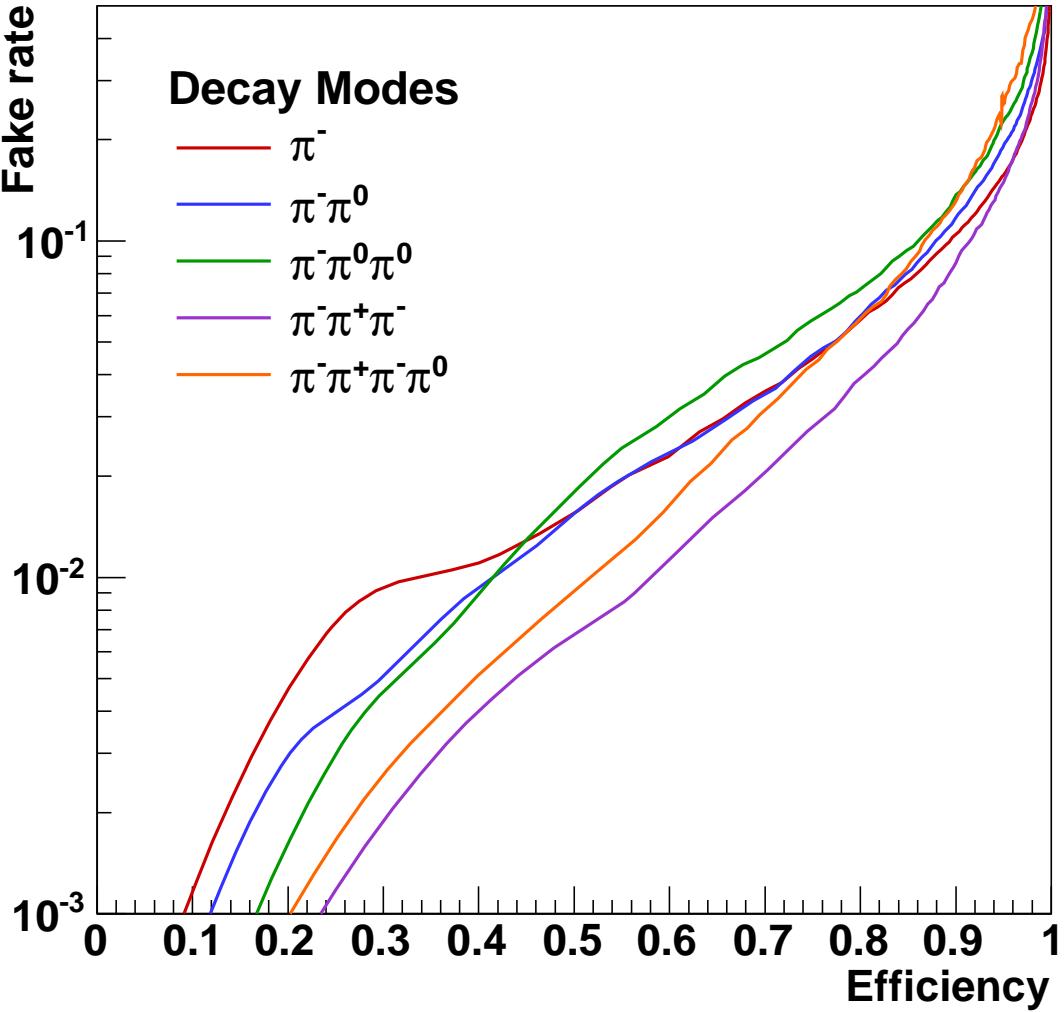


Figure 3.9: Performance curves for the five neural networks used by the TaNC for tau candidates with transverse momentum greater than 20 GeV/c. Each curve represents the signal efficiency (on the horizontal axis) and background misidentification rate (vertical axis) for a scan of the neural network selection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

(fig:nnPerfCurves)

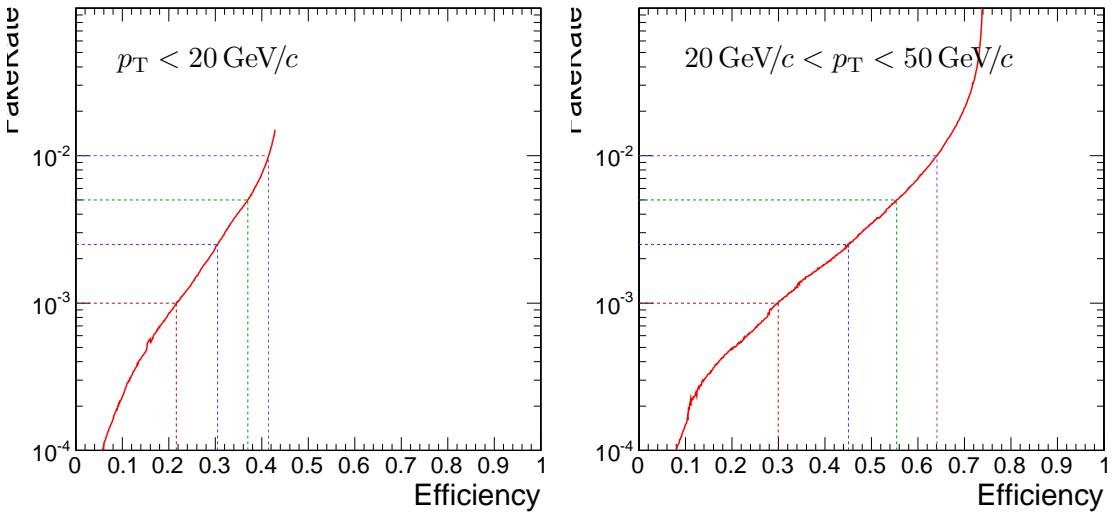


Figure 3.10: Tau Neural Classifier performance curves for tau candidates with  $p_T < 20 \text{ GeV}/c$  (left) and  $20 < p_T < 50 \text{ GeV}/c$  (right). The vertical axis represents the expected fake-rate of QCD jets and the horizontal axis the expected signal efficiency for hadronic tau decays.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

(fig:mcPerfCurves)

1367 neural network training error is given by Equation 3.2. A given point in the vector space  
 1368 spanned by the neural network input observables (denoted as “feature space”) contributes  
 1369 to the neural network training error  $E$  by

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

1370 where  $\rho^\tau(\rho^{QCD})$  denotes the training sample density of the  $\tau$  signal and QCD-jet back-  
 1371 ground 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:

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

$$\begin{aligned} \rho^\tau &= x(\rho^\tau + \rho^{QCD}) \\ \frac{\rho^{QCD}}{\rho^\tau} &= \frac{1}{x} - 1 \end{aligned} \tag{3.4} \quad \text{eq:rawTransform}$$

1372 The ratio  $\frac{\rho^{QCD}}{\rho^\tau}$  corresponds to the ratio of the normalized probability density functions of  
 1373 signal and background input observable distributions, i.e.  $\int \rho^\tau d\vec{x} = 1$ .

1374 In the case of multiple neural networks, one can derive a formula that maps the output  
 1375  $x_j$  of the neural network corresponding to decay mode  $j$  according to the “prior proba-  
 1376 bilities”  $p_j^\tau(p_j^{QCD})$  for true  $\tau$  lepton hadronic decays (quark and gluon jets) to pass the  
 1377 preselection criteria and be reconstructed with decay mode  $j$ . By substituting  $\rho^s \rightarrow \rho^s p_j^s$   
 1378 for  $s \in \{\tau, QCD\}$  in Equation 3.3, the output  $x_j$  can be related to  $p_j^\tau(p_j^{QCD})$  by

$$x'_j = \frac{\rho^\tau \cdot p_j^\tau}{\rho^\tau \cdot p_j^\tau + \rho^{QCD} \cdot p_j^{QCD}} = \frac{p_j^\tau}{p_j^\tau + \frac{\rho^{QCD}}{\rho^\tau} \cdot p_j^{QCD}} \tag{3.5} \quad \text{eq:probFracToX}$$

1379 Substituting Equation 3.4 into Equation 3.5 yields the transformation of the output  $x_j$  of the  
 1380 neural neural network corresponding to any selected decay mode  $j$  to a single discriminator  
 1381 output  $x'_j$  which for a given point on the optimal performance curve should be independent

1382 of  $j$ .

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

1383 In this manner a single number (the “transform cut”) given by Equation 3.6 can be used  
 1384 to specify any point on the performance curve. The training sample neural network output  
 1385 after the transformation has been applied is shown in Figure 3.12. The performance curve  
 1386 for the cut on the transformed output is nearly identical to the optimal performance curve  
 1387 determined by the Monte Carlo sampling technique.

1388 The discriminator output of the TaNC algorithm is a continuous quantity, enabling  
 1389 analysis specific optimization of the selection to maximize sensitivity. For the convenience  
 1390 of the user, four operating point benchmark selections are provided in addition to the  
 1391 continuous output. The four operating points are chosen such that for tau candidates with  
 1392 transverse momentum between 20 and 50 GeV/c, the expected QCD dijet fake rate will be  
 1393 0.1%, 0.25%, 0.50% and 1.0%, respectively.

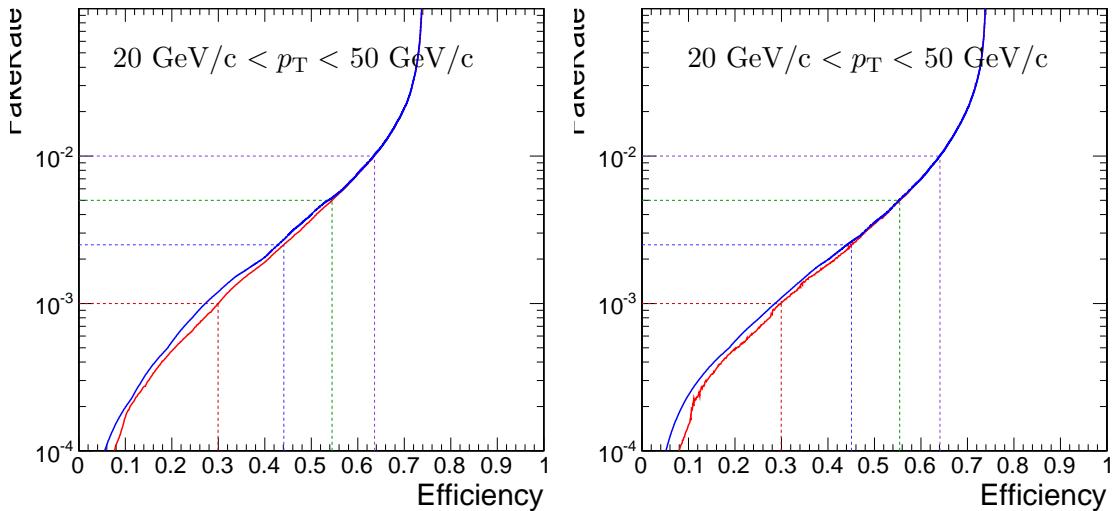


Figure 3.11: Tau Neural Classifier performance curves for tau candidates with  $20 < p_T < 50$  GeV/c. The figure on the left compares the optimal performance curve determined by the Monte Carlo sampling method (red) to the performance curve obtained by scanning the “transform cut” (blue) defined in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

ncCurvesWithTransform)?

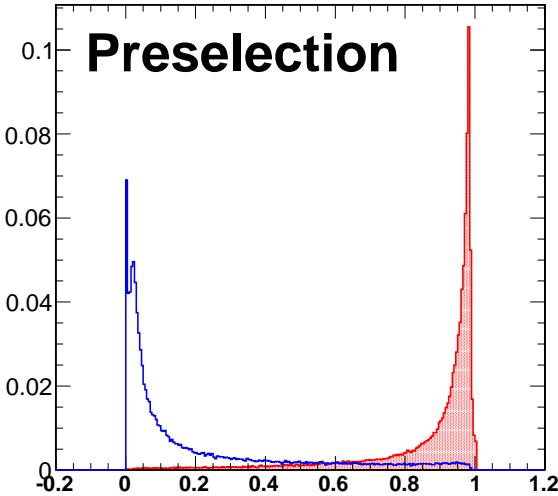


Figure 3.12: Transformed TaNC neural network output for tau candidates with transverse momentum between 20 and 50 GeV/c that pass the pre-selection criteria. The neural network output for each tau candidate has been transformation according to Equation 3.6. The decay mode probabilities  $\rho_i^{bkg}$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

`fig:transformedNNOutput`)

### 1394 3.4 Summary

1395 The Tau Neural classifier introduces two complimentary new techniques for tau lepton  
 1396 physics at CMS: reconstruction of the hadronic tau decay mode and discrimination from  
 1397 quark and gluon jets using neural networks. The decay mode reconstruction strategy pre-  
 1398 sented in Section 3.3.1 significantly improves the determination of the decay mode. This  
 1399 information has the potential to be useful in studies of tau polarization and background  
 1400 estimation.

1401 The Tau Neural classifier tau identification algorithm significantly improves tau dis-  
 1402 crimination performance compared to isolation-based approaches [28] used in previous CMS  
 1403 analyses. Figure 3.13 compares the performance of the “shrinking cone” isolation tau-  
 1404 identification algorithm [28] to the performance of the TaNC for a scan of requirements  
 1405 on the transformed neural network output. The signal efficiency and QCD dijet fake rate  
 1406 versus tau candidate transverse momentum and pseudo-rapidity for the four benchmark  
 1407 points and the isolation based tau identification are show in Figure 3.14. For tau candi-

1408 dates with transverse momentum between 20 and 50 GeV/ $c$ , the TaNC operating cut can  
 1409 be chosen such that the two methods have identical signal efficiency; at this point the TaNC  
 1410 algorithm reduces the background fake rate by an additional factor of 3.9. This reduction  
 1411 in background will directly improve the significance of searches for new physics using tau  
 1412 leptons at CMS.

### 1413 3.5 HPS+TaNC: A Hybrid Algorithm

1414 ⟨sec:TauId⟩ The techniques used in the TaNC have been hybridized with techniques used by the “Hadrons  
 1415 plus Strips” (HPS) algorithm [33]. The combined algorithm is referred to “Hadrons plus  
 1416 Strips and Tau Neural Classifier” (HPS + TaNC) identification algorithm. The algorithm  
 1417 combines the HPS methods of constructing the signal components of the tau candidate  
 1418 and the discrimination methods of the TaNC algorithm. Both algorithms are based on re-  
 1419 constructing individual tau lepton hadronic decay modes, which has been demonstrated to  
 1420 improve the tau identification performance significantly with respect to previously used cone  
 1421 isolation based algorithms [34]. The HPS + TaNC algorithm first reconstructs the hadronic  
 1422 decay mode of the tau, and applies different discriminants based on the reconstructed de-  
 1423 cay mode. Identification of hadronic tau decays by the HPS + TaNC algorithm proceeds in  
 1424 two stages: first, the hadronic decay mode of the tau is reconstructed and then different  
 1425 discriminators are applied, based on the reconstructed decay mode. In the decay mode re-  
 1426 construction particular attention is paid to the reconstruction of neutral pions, which are  
 1427 expected for the majority of hadronic decay modes.

#### 1428 3.5.1 Decay Mode Reconstruction

1429 The decay mode reconstruction algorithm is seeded by particle-flow jets reconstructed by  
 1430 the anti- $k_T$  algorithm [35]. In order to reconstruct the decay mode, the algorithm needs to  
 1431 merge photon candidates into candidate  $\pi^0$  mesons. The  $\pi^0$  candidates are reconstructed by  
 1432 two algorithms which are executed concurrently. The “combinatorial”  $\pi^0$  algorithm produces  
 1433 a  $\pi^0$  candidate for every possible pair of photons within the jet. The “strips” algorithm clus-  
 1434 ters photons strips in  $\eta - \phi$ . The results of both algorithms are combined and then “cleaned”.

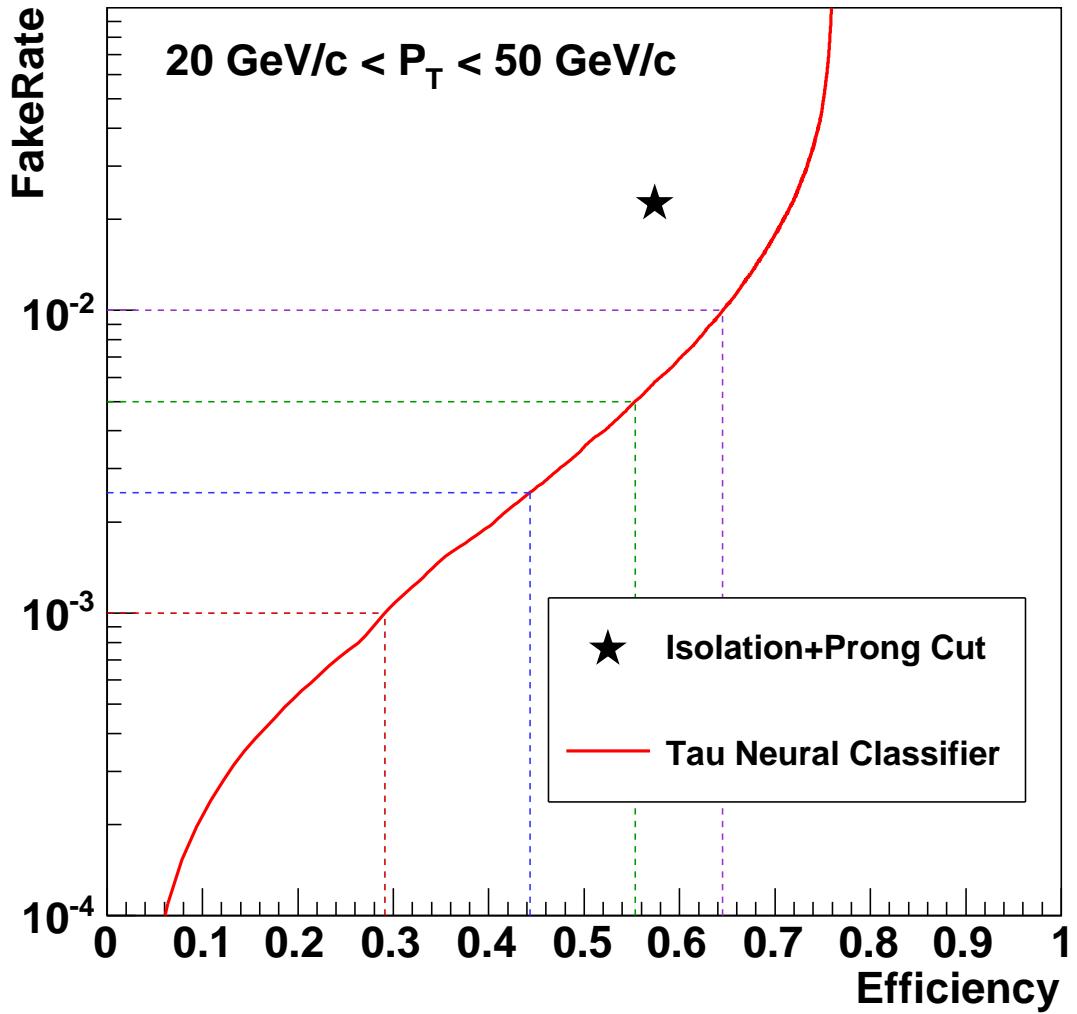


Figure 3.13: Performance curve (red) of the TaNC tau identification for various requirements on the output transformed according to Equation 3.6. The horizontal axis is the efficiency for true taus with transverse momentum between 20 and 50  $\text{GeV}/c$  to satisfy the tau identification requirements. The vertical axis gives the rate at which QCD dijets with generator-level transverse momentum between 20 and 50  $\text{GeV}/c$  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.

(fig:finalPerfCurve)

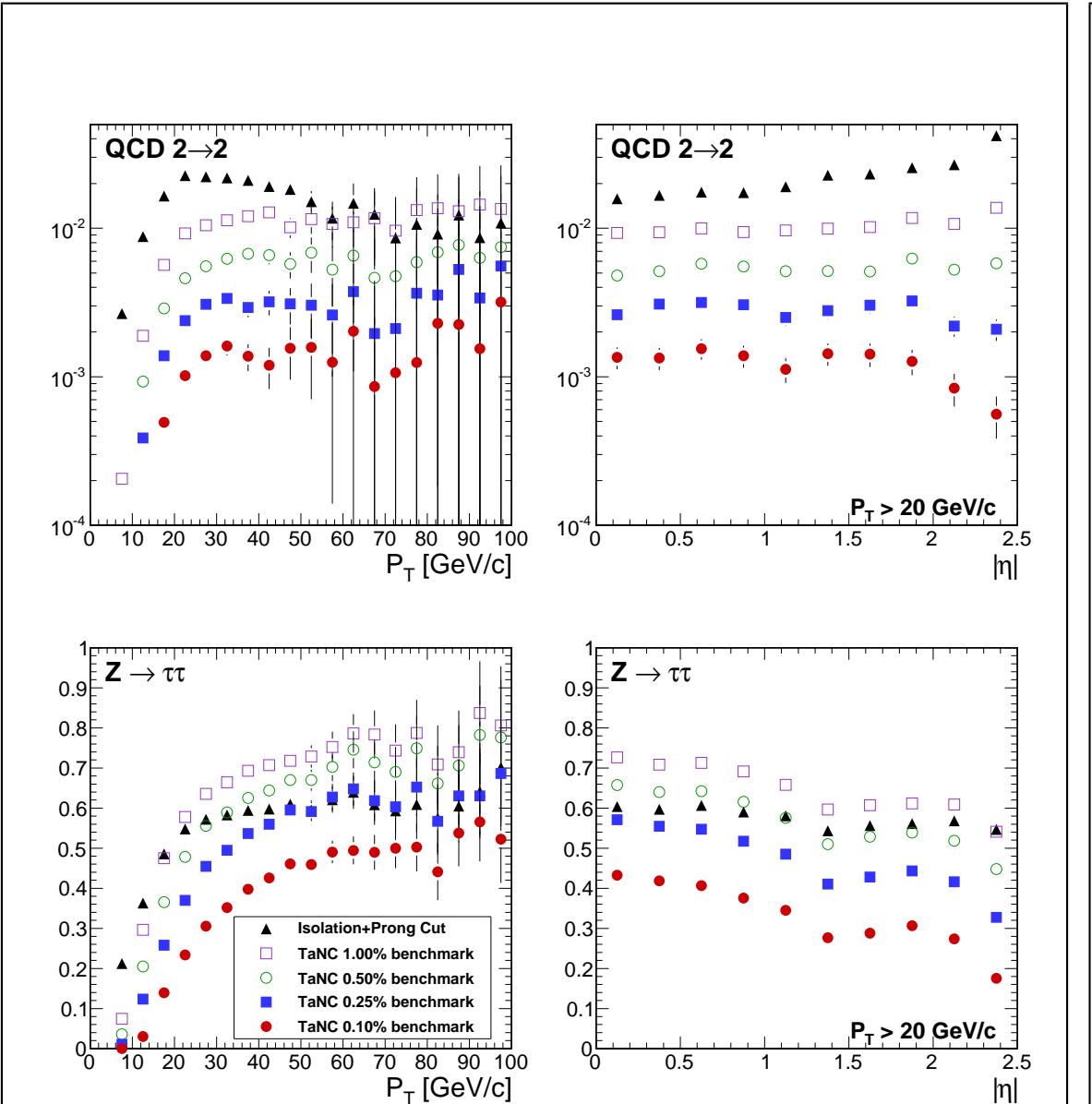


Figure 3.14: Comparison of the identification efficiency for hadronic tau decays from  $Z \rightarrow \tau^+\tau^-$  decays (bottom row) and the misidentification rate for QCD dijets (top row) versus tau candidate transverse momentum (left) and pseudo-rapidity (right) for different tau identification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

(fig:kinematicPerformance)

1435 resolving multiple hypotheses. The quality of a  $\pi^0$  candidate is determined according to the  
 1436 following categorical rankings:

- 1437 • The  $\pi^0$  candidate is in the ECAL barrel region ( $|\eta| < 1.5$ ) and has invariant mass  
 1438  $|m_{\gamma\gamma} - m_{\pi^0}| < 0.05 \text{ GeV}/c^2$ .
- 1439 • The  $\pi^0$  candidate is in the ECAL endcap region ( $|\eta| > 1.5$ ) and has invariant mass  
 1440  $m_{\gamma\gamma} < 0.2 \text{ GeV}/c^2$ .
- 1441 • The  $\pi^0$  candidate contains two or more photons within an  $\eta - \phi$  strip of size  $0.05 \times 0.20$
- 1442 • Photons not satisfying any of the other categories are considered as unresolved  $\pi^0$   
 1443 candidates in case they have  $p_T > 1.0 \text{ GeV}/c$ .

1444 The symbol  $m_{\pi^0}$  denotes the nominal neutral pion mass [20]. The size of the invariant mass  
 1445 windows in the ECAL endcap and barrel regions is motivated by the resolution of the  $\pi^0$   
 1446 mass (illustrated in Figure 3.15) measured during the commissioning of the particle-flow  
 1447 algorithm in early CMS data [36]. Multiple  $\pi^0$  candidates in the same category are ranked  
 1448 in quality according to the difference of the reconstructed photon pair mass to the nominal  
 1449  $\pi^0$  mass. After the  $\pi^0$  candidates are ranked, the highest ranked candidate is selected for  
 1450 the final collection. The photon constituents of the highest ranked candidate are removed  
 1451 from remaining  $\pi^0$  candidates not yet selected for the final collection in order to prevent  
 1452 photons from entering more than one  $\pi^0$  candidate. The rank of remaining  $\pi^0$  candidates  
 1453 is reevaluated and the  $\pi^0$  candidate with the next highest rank is selected for the output  
 1454 collection. The process is repeated until no more  $\pi^0$  candidates are remaining.

1455 Once the final collection of  $\pi^0$  candidates is determined, tau reconstruction in the  
 1456 HPS + TaNC algorithm proceeds by building tau candidates from reconstructed  $\pi^0$  candi-  
 1457 dates and charged hadrons reconstructed by the particle-flow algorithm. A combinatoric  
 1458 approach is again employed for the tau candidate building. A tau candidate hypothesis is  
 1459 built for every combination of jet constituents ( $\pi^0$  candidates plus charged hadrons) which  
 1460 has a multiplicity consistent with a hadronic tau decay. The tau candidates are ranked anal-  
 1461 ogous to the ranking utilized for the  $\pi^0$  reconstruction, but with the following categorical  
 1462 rankings:

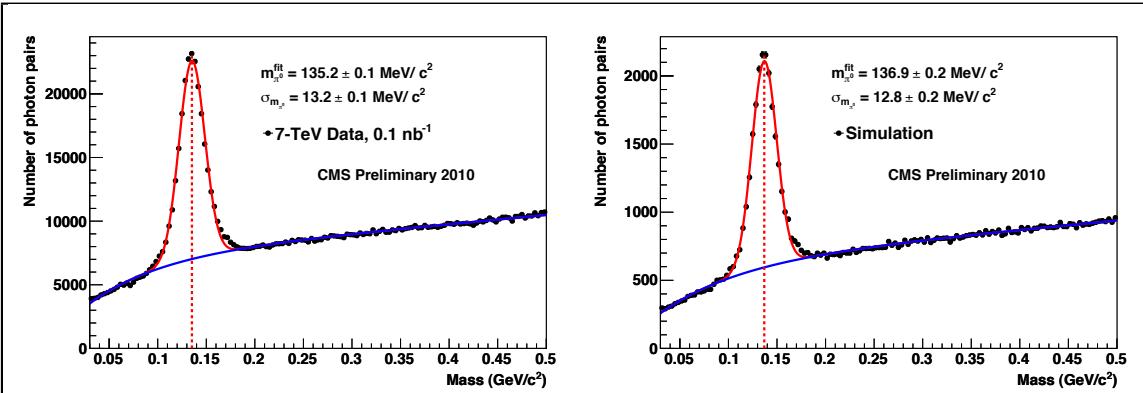


Figure 3.15: Invariant mass distribution of photon pairs reconstructed by the particle-flow in 2010 CMS minimum bias events (left), and predicted by the simulation (right). A clear resonant pick corresponding to the  $\pi_0$  meson is visible above the combinatoric background. Reference: [36]

(fig:PFPiZeroRes)

- 1463     • In each decay mode category, the tau candidate with the highest neural network
- 1464         output is selected.
- 1465     • The tau candidate has unit charge.
- 1466     • The tau candidate passes the “lead pion” criteria, requiring that there is a photon or
- 1467         charged pion candidate with  $p_T > 5 \text{ GeV}/c$ .
- 1468     • The tau candidate passes the HPS invariant mass and collimation<sup>1</sup> requirements.

1469 In case multiple tau candidates satisfy all four categorical requirements, the tau candidate  
 1470 with the highest energy sum of charged and neutral pions is selected as the highest ranking  
 1471 one.

### 1472 3.5.2 Hadronic Tau Discrimination

1473 The final level of discrimination is performed by an ensemble of neural networks, with each  
 1474 neural network corresponding to a specific decay mode, analogously to the method used  
 1475 original TaNC algorithm (Section 3.3.5). The inputs of each neural network are different and  
 1476 correspond to the observables (invariant mass, Dalitz masses) available for its associated

<sup>1</sup>The invariant mass of the signal candidates is required to be compatible with the resolution for that decay mode. The collimation selection requires the maximum  $\Delta R$  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].

decay mode. The neural networks are trained on samples simulated  $Z \rightarrow \tau^+\tau^-$  events (“signal”) and QCD dijet events selected in the 7 TeV data collected by CMS in 2010 (“background”). All of the tau hypothesis from a given jet reconstructed in data are used for training. The  $Z \rightarrow \tau^+\tau^-$  signal sample is generated by PYTHIA [37] which has been interfaced to TAUOLA [38] for the purpose of generating the decays of the tau leptons using the full matrix element information. After generation, the events passed through the “full” GEANT [39] based simulation of the CMS detector. Only tau candidates which have been reconstructed in a decay mode matching the true decay mode of the tau on generator level enter the signal training sample. The neural network implementation, network layout, and training strategies are the same as in the original TaNC algorithm described in this chapter. To account for differences in the input signal purity and separation power of the neural networks between decay modes, the outputs of each neural network are transformed according to the method described in Section 3.3.8. Multiple working-points corresponding to different purities are provided. The “loose” working point corresponds to an approximate fake-rate of 1%, and has slightly higher signal efficiency performance at high  $p_T$  than the corresponding HPS isolation-only working point.

### 1493 3.6 Electron and Muon Rejection

sec:LightLeptonRejection Additional discriminators must be applied to prevent electrons and muons from being identified as hadronic tau decays. This is especially important for removing  $Z \rightarrow e^+e^-$  and  $Z \rightarrow \mu^+\mu^-$  contributions when selecting events with two taus and requiring one of them to decay leptonically and the other hadronically. The electron and muon discrimination algorithms and performance are described in detail elsewhere [28]. A cursory overview of the techniques used are given here. Muon removal is achieved with high purity by requiring that no track in the signal collection of the tau candidate is matched to a segment<sup>2</sup> in the muon system. The rejections of true electrons is more difficult. Electrons leave no signal in the muon system and produce bremsstrahlung photons as they travel through the magnetic field. The most significant difference from a true hadronic tau is that an electron is not expected to deposit any energy in the hadronic calorimeter. Electrons are thus rejected by

<sup>2</sup>A track reconstructed in the DT or CSC sub-detectors.

1505 requiring that there is an HCAL energy deposit with a magnitude that is greater than 10%  
1506 of the momentum of the leading track in the tau.

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 3.3.7) 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.

(tab:nn'var'table)

## Chapter 4

# Mass Reconstruction: The Secondary Vertex Fit

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

### 4.1 Existing Mass Reconstruction Algorithms

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

1531 The collinear approximation is a technique previously used [19] to reconstruct the full  
 1532  $\tau^+\tau^-$  mass. In an event with two tau decays, there are a total of six<sup>1</sup> unknowns associated  
 1533 with the missing energy: the three components of the momentum of each neutrino. The  
 1534 collinear approximation makes the assumption that the neutrinos have the same direction  
 1535 as their associated visible decay products. This assumption reduces the number of unknown  
 1536 quantities to two, corresponding to the total energy of each neutrino. These two unknowns  
 1537 can be solved for by using the two components of the reconstructed missing transverse  
 1538 energy, which in the ideal case corresponds to the transverse component of the vector  
 1539 sum of the two neutrino's four momentum. The characteristic equation of the collinear  
 1540 approximation is

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

1541 where  $(E_x^{\text{miss}}, E_y^{\text{miss}})$  are the two components of the reconstructed missing transverse energy,  
 1542  $\phi_{1(2)}$  is the azimuthal angle of the visible component of the first (second) tau decay, and  
 1543  $E_{1(2)}$  is the reconstructed energy of neutrino of the first (second) tau decay.  $E_1$  and  $E_2$  can  
 1544 be extracted by inverting the matrix on the right hand side of Equation 4.1.

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

1545 The collinear approximation suffers from two problems. The approximation can fail  
 1546 (yielding unphysical negative energies for the reconstructed neutrinos) when the missing  
 1547 transverse energy is mis-measured. The events with unphysical solutions must be removed  
 1548 from the analysis, leading to a dramatic reduction in acceptance (on the order of 50% in  
 1549 this analysis). Improvements to the collinear approximation algorithm have recently been  
 1550 made which aim to recover part of the events with unphysical solutions [42]. But even with  
 1551 these improvements, no physical solution is still found for a large fraction of signal events.

---

<sup>1</sup>Technically, 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.

1552 Additionally, the method is numerically unstable when the two  $\tau$  lepton are nearly back-  
 1553 to-back in azimuth. In these cases the  $\sin(\phi_2 - \phi_1)^{-1}$  term in Equation 4.2 is very large  
 1554 and small mis-measurements of the missing transverse energy can produce a large tail on  
 1555 the reconstructed mass. This tail is particularly large for low-mass resonances. The large  
 1556 tail for high mass is predominantly due to the fact (discussed in Section 4.4.2) that the  
 1557 kinematic requirements<sup>2</sup> applied on the visible decay products preferentially selects events  
 1558 where the visible decay products carry the majority of the energy of the original  $\tau$  lepton,  
 1559 reducing the amount of true missing energy in the event.

## 1560 4.2 The Secondary Vertex Fit

1561 A novel algorithm is presented in the following, which succeeds in finding a physical solution  
 1562 for every event. As an additional benefit, the new algorithm is found to improve the ditau  
 1563 invariant mass resolution, making it easier to separate the Higgs boson signal from the  
 1564  $Z \rightarrow \tau^+\tau^-$  background.

1565 The novel Secondary Vertex fit (SVfit) algorithm for ditau invariant mass reconstruc-  
 1566 tion that we present in the following utilizes a likelihood maximization to fit a  $\tau^+\tau^-$  in-  
 1567 variant mass hypothesis for each event. The likelihood is composed of separate terms which  
 1568 represent probability densities of:

- 1569 • the tau decay kinematics,
- 1570 • the matching between the momenta of neutrinos produced in the tau decays and the  
   1571 reconstructed missing transverse momentum,
- 1572 • a regularization “ $p_T$ -balance” term which accounts for the effects on the ditau invari-  
   1573 ant mass of acceptance cuts on the visible tau decay products,
- 1574 • and the compatibility of tau decay parameters with the position of reconstructed  
   1575 tracks and the known tau lifetime of  $c\tau = 87 \mu\text{m}$  [20].

1576 The likelihood is maximized as function of a set of parameters which fully describe the tau  
 1577 decay.

---

<sup>2</sup>The 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.

### 1578 4.3 Parametrization of Tau Decays

1579 <sec:svParameterization> The decay of a tau with visible four-momentum  $p_{vis}$  measured in the CMS detector (“laboratory”)  
 1580 frame can be parametrized by three variables. The invisible (neutrino) momentum  
 1581 is fully determined by these parameters.

1582 The “opening-angle”  $\theta$  is defined as the angle between the boost direction of the tau  
 1583 lepton and the momentum vector of the visible decay products in the rest frame of the  
 1584 tau. The azimuthal angle of the tau in the lab frame is denoted as  $\bar{\phi}$  (we denote quantities  
 1585 defined in the laboratory frame by a overline). A local coordinate system is defined such  
 1586 that the  $\bar{z}$ -direction lies along the visible momentum and  $\bar{\phi} = 0$  lies in the plane spanned  
 1587 by the momentum vector of the visible decay products and the proton beam direction. The  
 1588 third parameter,  $m_{\nu\nu}$ , denotes the invariant mass of the invisible momentum system.

1589 Given  $\theta$ ,  $\bar{\phi}$  and  $m_{\nu\nu}$ , the energy and direction of the tau lepton can be computed by  
 1590 means of the following equations: The energy of the visible decay products in the rest frame  
 1591 of the tau lepton is related to the invariant mass of the neutrino system by:

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

1592 Note that for hadronic decays,  $m_{\nu\nu}$  is a constant of value zero, as only a single neutrino is  
 1593 produced. Consequently, the magnitude of  $P^{vis}$  depends on the reconstructed mass of the  
 1594 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  $\theta$  into Equations 4.3 and 4.4, the energy of the  
 tau is obtained by solving for the boost factor  $\gamma$  in the Lorentz transformation between the

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

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

1595        The energy of the tau lepton in the laboratory frame as function of the measured visible  
 1596        momentum depends on two of the three parameters only - the rest frame opening angle  $\theta$  and  
 1597        the invariant mass  $m_{\nu\nu}$  of the neutrino system. The direction of the tau lepton momentum  
 1598        vector is not fully determined by  $\theta$  and  $m_{\nu\nu}$ , but is constrained to lie on the surface of a  
 1599        cone of opening angle  $\bar{\theta}$  (given by Equation 4.4), the axis of which is given by the visible  
 1600        momentum vector. The tau lepton four–vector is fully determined by the addition of the  
 1601        third parameter  $\bar{\phi}$ , which describes the azimuthal angle of the tau lepton with respect to the  
 1602        visible momentum vector. The spatial coordinate system used is illustrated in Figure 4.1.

#### 1603 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  $\theta$  only:

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

1604        For three–body decays, the likelihood depends on the invariant mass of the neutrino system  
 1605        also:

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

eq:pdfKineLepto

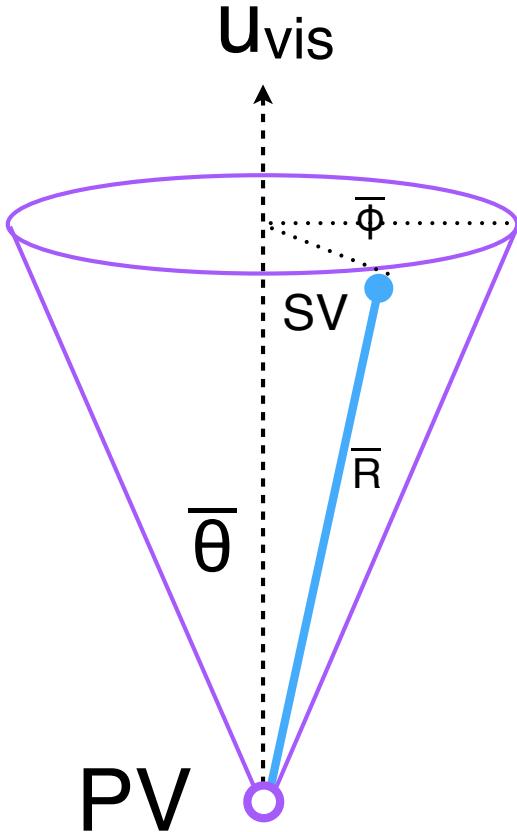


Figure 4.1: Illustration of the coordinate system used by the SVfit to describe the decays of tau leptons.

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

#### 4.4.1 Likelihood for Reconstructed $E_T^{\text{miss}}$

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

<sup>3</sup>The 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 boson and  $Z$  decays [43].

The  $E_T^{\text{miss}}$  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  $E_T^{\text{miss}}$  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  $E_T^{\text{miss}}$  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 (“ $\parallel$ ”) and perpendicular (“ $\perp$ ”) direction are:

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

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

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

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

<sup>1614</sup> where  $q_T$  denotes the transverse momentum of the tau lepton pair.

#### <sup>1615</sup> 4.4.2 Likelihood for Tau Transverse Momentum Balance

<sup>1615</sup> `(sec:ptBalance)` 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  $H \rightarrow \tau^+ \tau^-$  events as a function of the Higgs boson 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  $\theta$  to  $p_T \sim \frac{M}{2} \sin \theta$ , we obtain

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

1616 The first term of the  $p_T$ -balance likelihood is taken as the convolution of Equation 4.6 with  
 1617 a Gaussian of width  $s$ . The second term is taken to be a Gamma distribution with scale  
 1618 parameter  $\theta$  and shape parameter  $k$ , in order to account for tails in the  $p_T$  distribution of  
 1619 the tau lepton pair. The complete functional form is thus given by

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

Numerical values of the parameters  $s$ ,  $\theta$  and  $k$  are determined by fitting the PDF in Equation 4.7 to the tau lepton  $p_T$  distribution in simulated  $H \rightarrow \tau^+ \tau^-$  events. The relative weight 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 boson mass, we obtain the following numerical values for the parameters:

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

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

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

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

1620 The motivation for the  $p_T$ -balance likelihood is to add a “regularization” term which  
 1621 compensates for the effect of  $p_T$  cuts applied on the visible decay products of the two tau

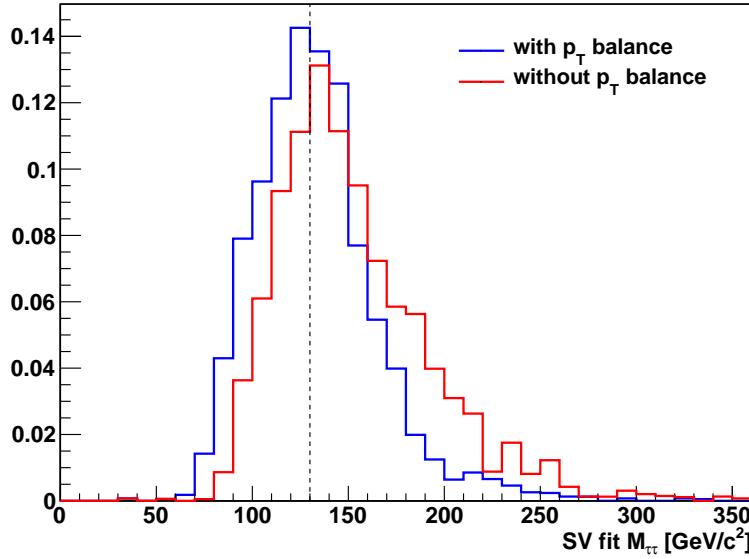


Figure 4.2: Distribution of ditau invariant mass reconstructed by the SVfit algorithm in simulated Higgs boson events with  $m_{A^0} = 130$   $\text{GeV}/c^2$ . The SVfit algorithm is run in two configurations, with (blue) and without (red) the  $p_T$ -balance likelihood term included in the fit.

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 underestimating the energy of the visible decay products (and overestimating the energy of neutrinos) produced in the tau decay. This effect produces 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  $\theta$ ,  $\bar{\phi}$  and  $m_{\nu\nu}$ , the position of the secondary (decay)

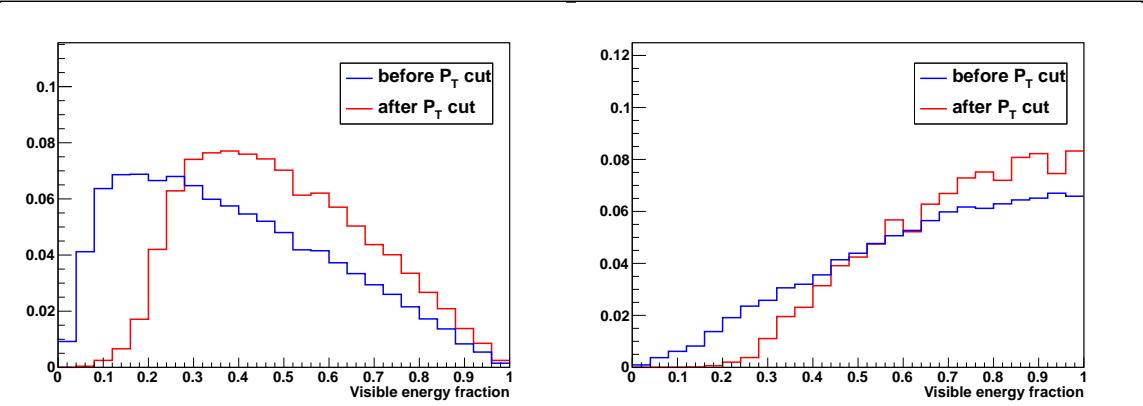


Figure 4.3: Normalized distributions of the fraction of total tau decay energy carried by the muon (left) and hadronic constituents (right) in simulated Higgs boson events with  $m_{A^0} = 130 \text{ GeV}/c^2$ . The distribution is shown before (blue) and after (red) the requirement on the  $p_T$  of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

(fig:ptBalancePtVisCuts)

1636 vertex is hence fully determined by addition of a single parameter for the flight distance.  
 1637 r. The tau lifetime  $c\tau = 87 \mu\text{m}$  is large enough to allow the displacement of the tau decay  
 1638 vertex from the primary event vertex to be resolved by the CMS tracking detector. The  
 1639 resolution provided by the CMS tracking detector is utilized to improve the resolution on  
 1640 the  $\tau^+\tau^-$  invariant mass reconstructed by the SVfit algorithm. The likelihood term based on  
 1641 the secondary vertex information represents the compatibility of the decay vertex position  
 1642 with the reconstructed tracks of charged tau decay products. Perhaps surprisingly, it turns  
 1643 out that the flight distance parameter  $R$  is sufficiently constrained even for tau decays into  
 1644 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:

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

1645 The likelihood terms for the secondary vertex fit have been implemented in the SVfit  
 1646 algorithm. In the analysis presented in this note, the decay vertex information is not used.  
 1647 however, because of systematic effects arising from tracker (mis-)alignment which are not  
 1648 yet fully understood.

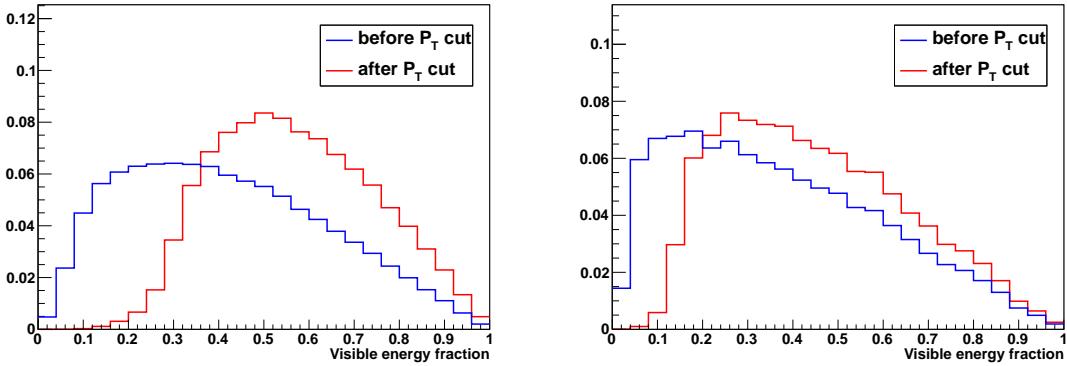


Figure 4.4: Normalized distributions of the fraction of total tau decay energy carried by the muon in simulated  $Z \rightarrow \tau^+\tau^-$  (left) and Higgs boson events with  $m_{A^0} = 200$   $\text{GeV}/c^2$  (right). The distribution is shown before (blue) and after (red) the requirement that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

PtVisCutsCompareMasses)

## 1649 4.5 Performance

1650 The tau pair invariant mass reconstructed by the Secondary Vertex fit (“SVfit mass”)  
 1651 provides the observable with the largest separation between signal Higgs boson events and  
 1652 the dominant  $Z \rightarrow \tau^+\tau^-$  background. The mean of the SVfit mass is located at the true  
 1653 mass of the ditau pair. The SVfit algorithm has a higher acceptance and better resolution  
 1654 than the collinear approximation algorithm. The SVfit always finds a physical solution,  
 1655 improving on the efficiency of the collinear approximation by a factor of two. Additionally,  
 1656 it has a much better resolution. The collinear approximation reconstructed mass distribution  
 1657 has a large tail at high mass due to events with poorly measured  $E_T^{\text{miss}}$ . The shape of the  
 1658 SVfit distribution is nearly Gaussian. The comparison is illustrated in Figure 4.5. Previous  
 1659 searches for Higgs bosons decaying to tau leptons [40] have in general used the “visible mass”  
 1660 as the observable used to search for new resonances. The SVfit method has the obvious  
 1661 difference that it reconstructs the “full” tau pair invariant mass, which is the most natural  
 1662 observable corresponding to a particle decaying to tau leptons. In addition, the relative  
 1663 resolution<sup>4</sup> of the SV fit is superior to that of the visible mass. This feature is illustrated in  
 1664 Figure 4.6. In Figure 4.6, the visible mass distribution is scaled by an arbitrary number such  
 1665 that the scaled mean of the distribution matches the true invariant mass of the tau pair

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

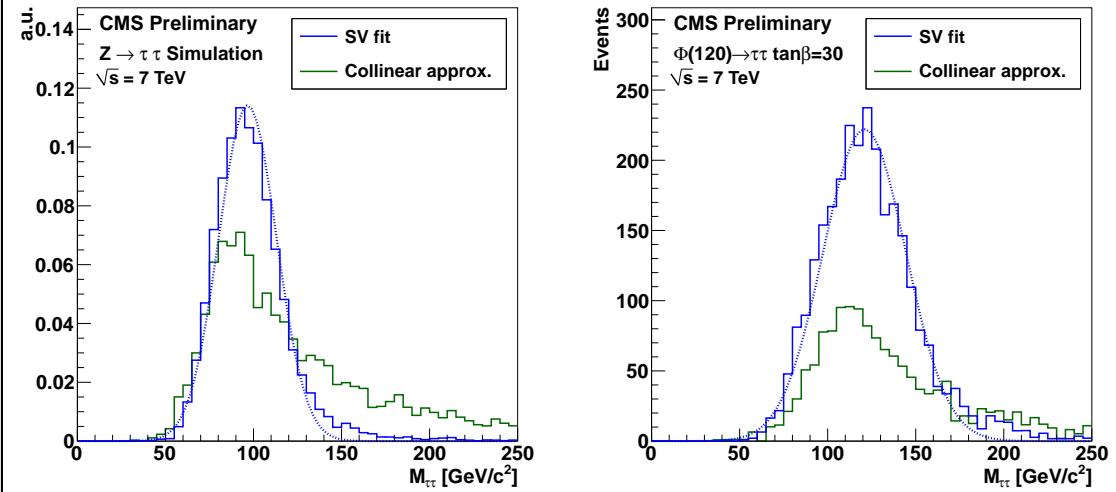


Figure 4.5: Comparison of the reconstructed tau pair invariant mass spectrum in  $Z \rightarrow \tau^+\tau^-$  (left) and MSSM  $H(120) \rightarrow \tau^+\tau^-$  (right) events after the selections described in chapter 5. The mass spectrum reconstructed by the Secondary Vertex fit is shown in blue,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

(fig:SVversusCollinear)

1666 (and the SVfit mass). The width of the SVfit distribution is smaller than that of the scaled  
 1667 visible mass distribution, indicating better performance. The increase in relative resolution  
 1668 allows an excess in the spectrum due to the presence of signal events to be more easily  
 1669 distinguished from the  $Z \rightarrow \tau^+\tau^-$  background. This increases the power of the search for  
 1670 the new signal.

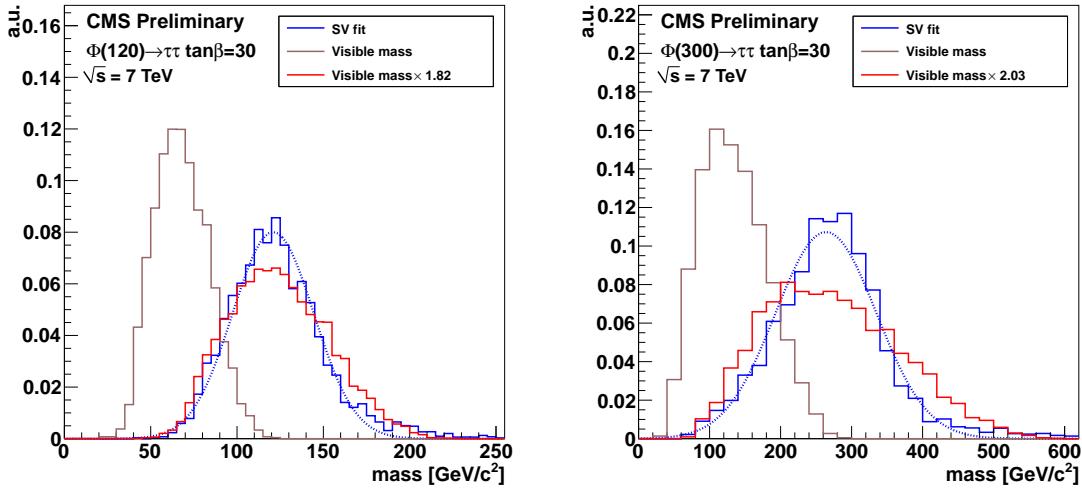


Figure 4.6: Comparison of the invariant mass of the muon and  $\tau_{\text{jet}}$  (the “visible mass”) with the full  $\tau^+\tau^-$  mass reconstructed by the SVfit. The spectrum is shown for two simulated MSSM Higgs boson samples, with  $m_{A^0} = 120 \text{ GeV}/c^2$  (left), and  $m_{A^0} = 200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

(fig:SVversusVis)

## Chapter 5

# Analysis Selections

(ch:selections) The selections applied to events in this analysis are designed to maximize the significance of Higgs boson signal events in the final set of selected events. The analysis presented in this thesis is an inclusive analysis, meaning that no preference is given to any single Higgs boson production mechanism. The analysis looks specifically at the channel in which one tau decays to a muon and the other decays to hadrons. Therefore the first step in the analysis selection is to find HLT selection that is highly efficient for the signal and is not highly prescaled<sup>1</sup>. After the trigger selection, events are required to contain at least a good muon and a good tau. Vetoos on extra leptons are applied to reduce backgrounds from dimuon events. Finally, kinematic and charge selections are applied to the event to reduce  $W + \text{jets}$  and QCD backgrounds.

## 5.1 High Level Trigger

Because only data which passes the HLT is recorded, it is critical that an appropriate trigger path is found. The events in this analysis are triggered by a combination of muon and muon + hadronic tau “cross–channel” triggers. For the muon triggers, the HLT paths with lowest  $p_T$  thresholds are used as long as the path remained unprescaled (see Table 5.1). The muon + tau-jet “cross–channel” trigger paths increase the trigger efficiency for events containing muons of transverse momenta close to the  $p_T^\mu > 15$  GeV/c cut threshold. The trigger efficiency is measured in data via the tag-and-probe technique. Details of the muon trigger efficiency measurement are given in Section 7.1. Monte Carlo simulated

<sup>1</sup>If 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 “cross–channel” trigger paths utilized to trigger events in different data–taking periods.

{tab:AHToMuTauTriggers}

1692 events are required to pass the HLT\_Mu9 trigger path. Weights are applied to simulated  
 1693 events to account for the difference between the simulated HLT\_Mu9 efficiency and the com-  
 1694 bined efficiency of the set HLT\_Mu9, HLT\_IsoMu9, HLT\_Mu11, HLT\_IsoMu13, HLT\_Mu15,  
 1695 HLT\_IsoMu9\_PFTau15 and HLT\_Mu11\_PFTau15 used to trigger the data.

## 1696 5.2 Particle Identification

### 1697 5.2.1 Muons

1698 {sec:MuonId} Muon candidates are required to be reconstructed as global and as tracker muons, meaning  
 1699 that a full track is reconstructed in the muon system and is well matched to a track in the  
 1700 silicon strip and pixel trackers. Additionally, they are required to pass the “Vector Boson  
 1701 Task Force” (VBTF) muon identification criteria developed for the  $Z \rightarrow \mu^+ \mu^-$  cross–section  
 1702 measurement [44], which consist of:

- 1703 •  $\geq 1$  pixel hits,
- 1704 •  $\geq 10$  hits in silicon pixel and strip detectors,
- 1705 •  $\geq 1$  hit(s) in muon system,
- 1706 •  $\geq 2$  matched segments,

- 1707     •  $\chi^2/DoF < 10$  for global track fit,
- 1708     • and an inner track transverse impact parameter  $d_{\text{IP}} < 2$  mm with respect to the  
1709       beamspot.

1710       In order to reduce background contributions from muons originating from heavy quark  
1711       decays in QCD multi-jet events, muons are required to be isolated. Isolation is computed  
1712       as the  $p_{\text{T}}$  sum of charged and neutral hadrons plus photons reconstructed by the CMS  
1713       particle-flow algorithm [29] within a cone of size  $\Delta R_{\text{iso}} = 0.4$  around the muon direction  
1714       divided by the muon  $p_{\text{T}}$ . The innermost region of size  $\Delta R_{\text{veto}} = 0.08$  (0.05) is excluded  
1715       from the computation of the isolation  $p_{\text{T}}$  sum with respect to neutral hadrons (photons),  
1716       in order to avoid energy deposits in the electromagnetic and hadronic calorimeters which  
1717       are due to the muon to enter the sum. In order to reduce pile-up effects, particles entering  
1718       the isolation  $p_{\text{T}}$  sum are required to have transverse momenta  $p_{\text{T}} > 1.0$  GeV/c. Charged  
1719       particles are additionally required to originate from the same vertex as the muon. The  
1720       muons are required to be isolated with respect to charged hadrons of  $p_{\text{T}} > 1.0$  GeV/c and  
1721       photons of  $p_{\text{T}} > 1.5$  GeV/c as reconstructed by the particle-flow algorithm [29] in a cone  
1722       of size  $\Delta R = 0.4$  around the direction of the muon. The distribution of the muon isolation  
1723       discriminant is shown in Figure 5.1(a).

### 1724 5.2.2 Hadronic Taus

1725 Hadronic decays of taus are identified by the HPS + TaNC hybrid algorithm described in  
1726 Section 3.5. The expected sensitivity of the search was evaluated using each of the hadronic  
1727 tau identification work points. The “loose” working point, corresponding to an expected  
1728 QCD fake-rate of about 1%, was found to have the highest performance and is used in this  
1729 analysis.  $Z \rightarrow \mu^+\mu^-$  background contributions are largely due to muons which failed to  
1730 get reconstructed as global muons (thus failing the muon identification requirement) and  
1731 are misidentified as tau candidates. These muons are typically isolated and have a large  
1732 chance to pass the hadronic tau ID discriminators. To reject these events, hadronic taus are  
1733 additionally required to pass an anti-muon veto described in Section 3.6.

Background	Cross Section ( pb)
QCD Heavy Flavor	84679 <sup>3</sup>
$W \rightarrow \mu\nu + \text{jets}$	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.

### 5.2.3 Missing Transverse Energy

The missing transverse energy  $E_T^{\text{miss}}$ , in the event is reconstructed based on the vectorial momentum sum of particle candidates reconstructed by the particle-flow algorithm [29, 45]. In the ideal case, the  $E_T^{\text{miss}}$  corresponds to the vector sum of the transverse components of all neutrinos in the event. The  $E_T^{\text{miss}}$  resolution in simulated  $Z \rightarrow \mu^+\mu^-$  events is found to be smaller (better) than in the data. The reconstructed  $E_T^{\text{miss}}$  in the simulated events is “smeared” by a correction factor such that the data and simulation are in agreement. The “Z-recoil”  $E_T^{\text{miss}}$  correction procedure is described in Section 7.4.

## 5.3 Event Selections

The selections applied to the analysis are designed to reject large fractions of the background while maintaining a high efficiency for identifying signal Higgs boson events. The backgrounds can be divided into two classifications: “fake” backgrounds, in which there is at least one misidentified hadronic tau decay, and the irreducible  $Z \rightarrow \tau^+\tau^-$  background, which cannot<sup>2</sup> be distinguished from the potential presence of a Higgs boson of the same mass. Strategies for dealing with the irreducible  $Z$  background will be discussed in the Chapter 9. The different fake backgrounds, their cross section, and the basic removal strategies are outlined in Table 5.2.

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

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

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

Additional event selection criteria are applied to reduce contributions of specific background processes. In order to reject  $Z \rightarrow \mu^+ \mu^-$  background, a dedicated discriminator against muons is applied [28]. The remaining dimuon background is suppressed by rejecting events which have a track of  $p_T > 15$  GeV/c and for which the sum of energy deposits in ECAL plus HCAL is below  $0.25 \cdot P$  within a cylinder of radius 15 cm (ECAL) and 25 cm (HCAL), respectively. Contamination from  $Z \rightarrow \tau^+ \tau^-$  events in which the reconstructed tau candidate is due to a  $\tau \rightarrow e \nu \nu$  decay is reduced by applying a dedicated 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  $P_\zeta$  variable. The transverse mass ( $M_T$ ) cut is defined as the quantity

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

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

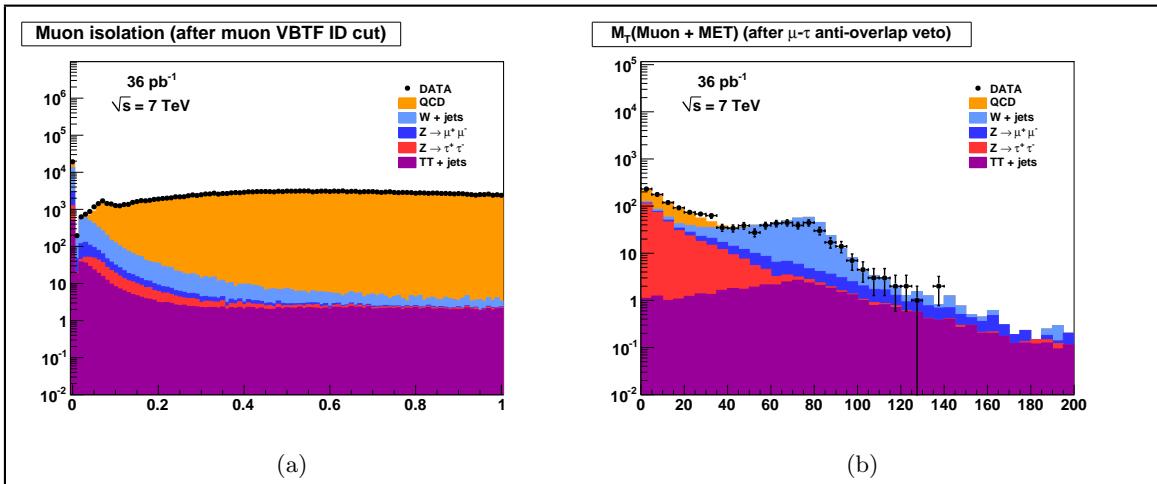


Figure 5.1: Distributions of the muon isolation (a) and  $M_T$  (b) discriminant variables. The muon isolation discriminant rejects the QCD background at a high rate. The  $M_T$  cut is designed to reject  $W + \text{jets}$  and  $t\bar{t} + \text{jets}$  backgrounds. The distributions shown are computed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

The  $P_\zeta$  variable is another quantity with discriminant power against  $W + \text{jets}$  and  $t\bar{t}$  backgrounds. The observable has been introduced in the search for  $H \rightarrow \tau^+\tau^-$  events performed by the CDF collaboration [46]. The observable is motivated by the fact that in  $\tau^+\tau^-$  signal events the neutrinos are produced nearly collinear with their associated visible decay products. It is therefore expected that the direction of the missing transverse energy vector in these events points in a direction between<sup>4</sup> the visible  $\tau$  decay products. This event topology is not preferred in  $W + \text{jets}$ ,  $t\bar{t}$  and QCD background events. The observable is computed as the 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  $\zeta$  bisecting the directions  $\vec{P}_T^{vis_1}$  and  $\vec{P}_T^{vis_2}$  of the visible  $\tau$  lepton decay products in the transverse plane (see Figure 5.2(a) for an illustration). The distribution of  $P_\zeta$  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.

<sup>4</sup>In other words, the projection of  $\vec{E}_T^{\text{miss}}$  is positive along the bisector of the muon and hadronic tau momenta.

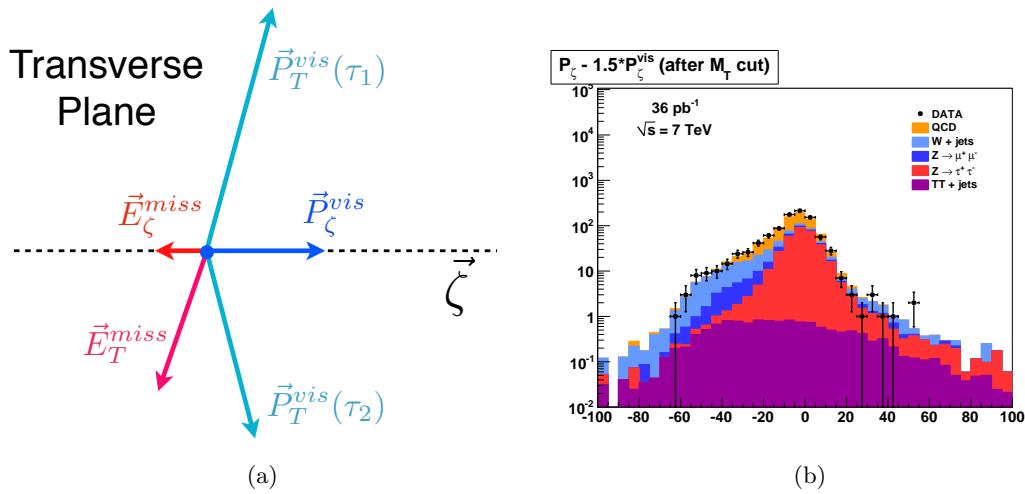


Figure 5.2: The vector quantities used in construction of the quantity  $P_\zeta - 1.5 \cdot P_\zeta^{\text{vis}}$  are illustrated in Figure (a). Image credit: [46] The distribution of the  $P_\zeta$  variable in the different background sources and 2010 data after the  $M_T$  cut has been applied is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

?<fig:PzetaPlots>?

Requirement	
Trigger	HLT Mu9 for MC <i>cf.</i> table 5.1 for Data
Vertex	reconstructed with beam-spot constraint:
Muon	$-24 < z_{vtx} < +24 \text{ cm}$ , $ \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 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	$\text{charge}(\text{Muon}) + \text{charge}(\text{Tau}) = 0$ , $\Delta R(\text{Muon}, \text{Tau}) > 0.5$
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.

## Chapter 6

# Data–Driven Background Estimation

(ch:backgrounds) For the result of this analysis to be reliable, it is of paramount importance that the backgrounds be well understood. If at all possible, all background processes should be measured in a “data–driven” way. By requiring that the background comes from data, biases due to incorrectly modeling background processes in simulation can be minimized or eliminated. In general, the data–driven methods also have the advantage that they are independent of the uncertainty on the integrated luminosity. This analysis measures the backgrounds using two complementary methods, the “template method” and the “fake–rate method.” In both cases, predictions are made about backgrounds in the signal region using measurements obtained in background–enriched control regions of the data.

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

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

1815 shapes of the backgrounds, while the fake–rate method uses only information about the  
 1816 hadronic tau fake–rate.

## 1817 6.1 Background–Enriched Control Regions

?<sec:controlregions>?  
 1818 The criteria applied to select events in the background–enriched control regions for the  
 1819 template method is based on the work described in [47]. With respect to that work, the muon  
 1820 isolation criteria applied to select  $Z \rightarrow \mu^+ \mu^-$ ,  $W + \text{jets}$ ,  $t\bar{t} + \text{jets}$  and QCD background  
 1821 enriched control samples has been changed to relative isolation with respect to charged  
 1822 hadrons and neutral electromagnetic objects reconstructed by the particle–flow algorithm.  
 1823 The selection of the enriched backgrounds is accomplished by disabling or inverting specific  
 1824 selections of Chapter 5 that were implemented to reject the given background. The selection  
 1825 of control regions used to measure the fake–rates for different types of background processes  
 1826 are very similar to the selections used for the template method. The details of the fake–rate  
 1827 measurement selections may be found in [48].

1828 All control regions are selected from the 2010 CMS muon primary datasets using single  
 1829 muon HLT trigger paths. The set of triggers and run–ranges used to select events in the  
 1830 background–enriched control samples is the same as for the analysis (see Table 5.1). The  
 1831 Monte Carlo simulated events used for comparison with the control region selections are  
 1832 required to pass the HLT\_Mu9 trigger path and are weighted according to the description  
 1833 in Chapter 7 to account for the difference in efficiency between HLT\_Mu9 and the trigger  
 1834 paths required to have passed in the data.

1835 QCD dijet events containing a muon (originating from the leptonic decay of a  $b$  or  
 1836  $c$  quark) are selected by applying an anti–isolation requirement on the jet containing a  
 1837 muon.  $W + \text{jets}$  and  $t\bar{t} + \text{jets}$  are selected by requiring an isolated muon, and inverting  
 1838 the transverse mass ( $M_T$ ) and  $P_\zeta$  selections. Tau candidates considered in the  $Z \rightarrow \mu^+ \mu^-$   
 1839 sample where the reconstructed tau candidate is faked by a misidentified muon and in  
 1840 the  $t\bar{t} + \text{jets}$  control sample are required to pass the “loose” TaNC discriminator. For the  
 1841 template method, the  $Z \rightarrow \mu^+ \mu^-$  sample where the reconstructed tau candidate is faked by  
 1842 a misidentified quark or gluon jet, the  $W + \text{jets}$  and the QCD enriched control samples have  
 1843 a loose hadronic tau “preselection” applied. The tau candidates are required to pass the

1844 “very loose”, but fail the “loose” TaNC discriminator. The criteria applied to select events  
 1845 in the different background–enriched control samples are summarized in Table 6.1. The goal  
 1846 of the background–enriched selection process is to select different background processes with  
 1847 high purity. A highly pure background control sample improves the stability of inferences  
 1848 about the signal region made using information in the enriched control region. The purity  
 1849 of the control regions (estimated using simulation) are summarized in Table 6.2.

1850 The number of events observed in the different control samples is compared to the  
 1851 Monte Carlo expectation in Table 6.2. Except for the contribution of  $Z \rightarrow \mu^+\mu^-$  events  
 1852 in which the reconstructed tau candidate is due to a misidentified quark or gluon jet,  
 1853 good agreement between data and Monte Carlo simulation is observed. Differences observed  
 1854 between data and simulation will be accounted for as systematic uncertainties.

1855 The distributions of visible and “full”  $\tau^+\tau^-$  invariant mass reconstructed by the SVfit  
 1856 algorithm (see Chapter 4) observed in the background enriched control regions is compared  
 1857 to the Monte Carlo simulation in Figures 6.1 and 6.2. The template for the  $W +$  jets back-  
 1858 ground has been corrected for the bias on the  $M_{vis}^{\mu\tau_{had}}$  shape caused by the  $M_T^{\mu E_T^{\text{miss}}} <$   
 1859  $50 \text{ GeV}/c^2$  and  $P_\zeta - 1.5 \cdot P_\zeta^{vis} > -20 \text{ GeV}$  requirements applied in the final analysis via the  
 1860 reweighting procedure described in [47]. In the  $t\bar{t} +$  jets enriched control region a peak at the  
 1861  $Z$  mass is observed in data, which is not modeled by the Monte Carlo samples considered.  
 1862 The peak could be due to  $Z \rightarrow \mu^+\mu^-$  events produced in association with  $b$  quarks. On  
 1863 the other hand, the contribution from  $t\bar{t} +$  jets events to that sample seems to be overesti-  
 1864 mated. The origin of the  $Z$  mass peak merits further investigation, but overall the  $t\bar{t} +$  jets  
 1865 is a negligible background contribution.

## 1866 6.2 The Fake–rate Method

(sec:fakerate)  
 1867 The probabilities with which quark and gluon jets get misidentified as taus may be utilized  
 1868 to obtain an estimate of background contributions in physics analyses. As an illustrative  
 1869 example and in order to demonstrate the precision achievable with the method, we introduce  
 1870 the method in the context of a “closure test,” using a simulated samples, a simple method of

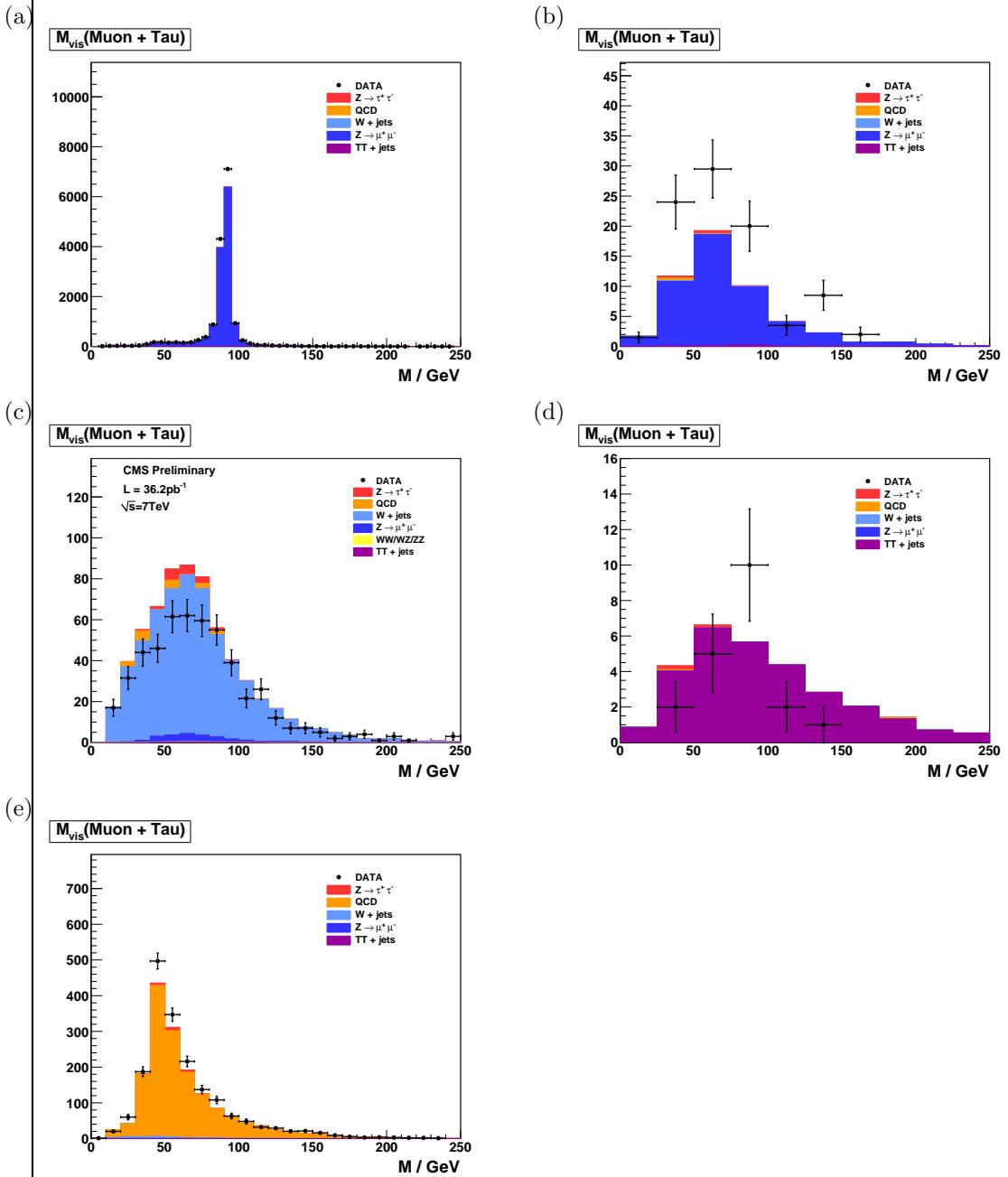
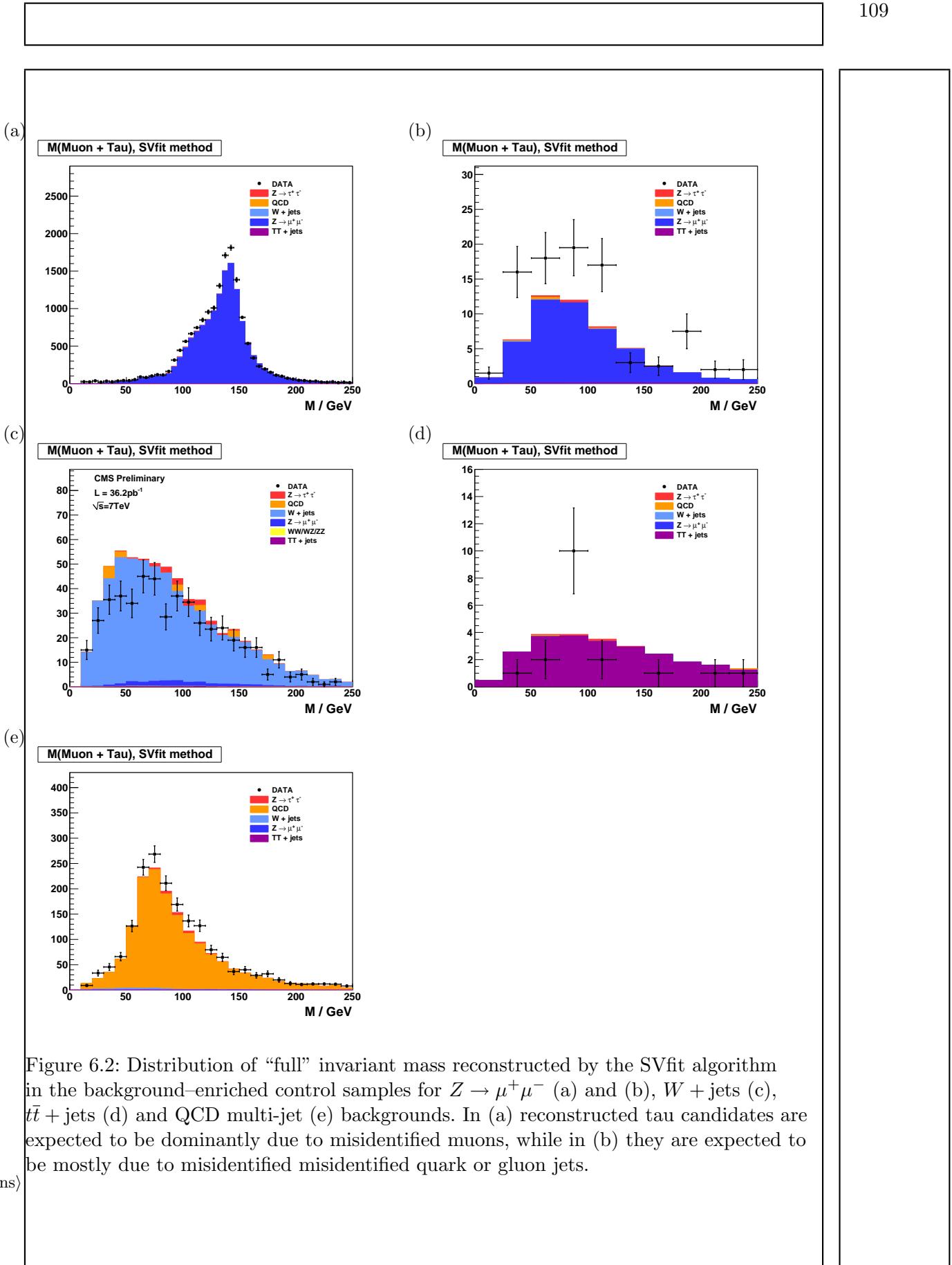


Figure 6.1: Distribution of visible mass of muon plus the tau candidate reconstructed in the background-enriched control samples for  $Z \rightarrow \mu^+ \mu^-$  (a) and (b),  $W + \text{jets}$  (c),  $t\bar{t} + \text{jets}$  (d) and QCD multi-jet (e) backgrounds. In (a) reconstructed tau candidates 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.



computing the fake–rate, and a simpler<sup>4</sup> hadronic tau identification algorithm. The closure test demonstrates that the method is self–consistent, and that the fake–rate technique can be used to estimate the contributions of QCD,  $W + \text{jets}$ ,  $t\bar{t} + \text{jets}$  and  $Z \rightarrow \mu^+ \mu^-$  backgrounds. The analysis selections used in the closure test are almost identical to the selections used in this analysis. Exact details of the selections can be found in reference analysis [44]. The method is then extended to use fake–rates measured in data, a multivariate method of computing the fake–rates, and the HPS + TaNC tau identification algorithm used in this analysis.

### 6.2.1 Parameterization of Fake–rates

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

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

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

where  $E(\eta^2)$ ,  $E(\phi^2)$  is the second  $\eta$ ,  $\phi$  moment of the jet constituents, weighted by constituent transverse energy. Analyses performed by the CDF collaboration [40, 49, 46] found

<sup>4</sup>The 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].

1896 that systematic uncertainties on background estimates obtained from the fake-rate method  
 1897 are reduced in case differences between quark and gluon jets are accounted for in this way.

### 1898 6.2.2 Measurement of Fake-rates

1899 Efficiencies and fake-rates are obtained by counting the fraction of tau candidates passing  
 1900 all tau identification cuts and discriminators in a given bin<sup>5</sup> of  $p_T^{jet}$ ,  $\eta_{jet}$  and  $R_{jet}$ :

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

1901 The pre-selection in the denominator of Equation 6.1 in general refers to  $p_T$  and  $\eta$  cuts,  
 1902 which are applied with thresholds matching those applied on the final analysis level, but  
 1903 may include loose tau identification criteria (which may be applied e.g. already during event  
 1904 skimming). It is critical that the selection used in the denominator be identical to that of  
 1905 the final analysis to ensure the fake-rates are not biased by different selections.

1906 Different sets of fake-rates are determined for the highest  $p_T$  and for the second highest  
 1907  $p_T$  jet in QCD dijet events, for jets in a QCD event sample enriched by the contribution  
 1908 of heavy quarks and gluons by requiring the presence of a muon reconstructed in the final  
 1909 state, and for jets in “electroweak” events selected by requiring a  $W$  boson in the final state.

### 1910 6.2.3 Application of Fake-rates

(sec:FakeRateApplication)  
 1911 Knowledge of the tau identification efficiencies and fake-rates as function of the parameters  
 1912  $p_T^{jet}$ ,  $\eta_{jet}$  and  $R_{jet}$  as defined by Equation 6.1 is utilized to obtain an estimate for the con-  
 1913 tributions of background processes to physics analyses involving tau lepton hadronic decays  
 1914 in the final state. The basic idea is to replace tau identification cuts and discriminators by  
 1915 appropriately chosen weights.

1916 Application of the fake-rate technique consists of two stages. The first stage consists of  
 1917 loosening the tau identification cuts and discriminators and applying only the preselection  
 1918 requirements defined by the denominator of Equation 6.1, in order to obtain an event  
 1919 sample dominated by contributions of background processes. After disabling the selections

---

<sup>5</sup>The 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.

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

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

In the “simple” fake–rate method described in detail in the next section, the corrections are taken from Monte Carlo simulations. Corrections based on Monte Carlo are needed also to compensate for signal contributions to the background dominated control sample. An alternative to Monte Carlo based corrections is to utilize additional information contained in the background dominated control sample. The modified version is described in Section 6.2.5. It has been used to estimate background contributions in searches for Higgs boson production with subsequent decays into tau lepton pairs performed by the CDF collaboration in Tevatron Run II data [40, 49, 46]. We will refer to the modified version as ‘CDF–type’ method in the following.

#### 6.2.4 “Simple” Weight Method

In the “simple” method, all tau 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

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

eqBgEstFakeRate

1947 These weights are applied to all jets in the background dominated control sample which  
 1948 pass the preselection defined by the denominator of Equation 6.1. Note that the weights  
 1949 defined by Equation 6.2 can be used to estimate the contributions of background processes  
 1950 to distributions of tau related observables. They cannot be used as event weights.

1951 In order to compare distributions of event–level quantities or per–particle quantities  
 1952 for particles of types different from tau leptons decaying hadronically, event weights need  
 1953 to be defined. Neglecting the small fraction of background events in which multiple tau  
 1954 candidates pass the complete set of all tau identification cuts and discriminators, event  
 1955 weights can be computed by summing up the per–jet weights defined by Equation 6.2 over  
 1956 all tau candidates in the event which pass the preselection:

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

1957 A bit of care is needed in case one wants to compare distributions of observables  
 1958 related to “composite particles,” the multiplicity of which depends on the multiplicity of  
 1959 tau candidates in the event (e.g. combinations of muon + tau pairs in case of the  $Z \rightarrow$   
 1960  $\tau^+\tau^- \rightarrow \mu + \tau$ -jet analysis). Per–particle weights need to be computed for such “composite  
 1961 particles”, depending on  $p_{\text{T}}^{\text{jet}}$ ,  $\eta_{\text{jet}}$ ,  $R_{\text{jet}}$  of its tau candidate constituent, according to:

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

1962 Different estimates are obtained for the fake–rate probabilities determined for the high–  
 1963 est and second highest  $p_{\text{T}}$  jet in QCD dijet events, jets in a muon enriched QCD sample  
 1964 and jets in  $W + \text{jets}$  events. The arithmetic average of the four estimates of the closure test  
 1965 together with the difference between the computed average and the minimum/maximum  
 1966 value is given in Table 6.3.

1967 We take the average value as the “best” estimate of the background contribution and  
 1968 the difference between the average and the minimum/maximum estimate as its systematic  
 1969 uncertainty. We obtain a value of  $\mathcal{O}(15\%)$  for the systematic uncertainty and find that  
 1970 the true sum of QCD,  $W + \text{jets}$ ,  $t\bar{t} + \text{jets}$  and  $Z \rightarrow \mu^+\mu^-$  background contributions agrees

1971 well with the “best” estimate obtained by the fake–rate method within the systematic  
 1972 uncertainty.

1973 Note that the estimate for the sum of background contributions which one obtains  
 1974 in case one applies the “simple” fake–rate weights defined by Equation 6.3 to a back-  
 1975 ground dominated control sample selected in data is likely to overestimate the true value  
 1976 of background contributions by a significant amount. The reason is that contributions of  
 1977  $Z \rightarrow \tau^+ \tau^-$  events with true taus are non–negligible. In fact, genuine tau contributions to  
 1978 the background dominated control sample are expected to be 14.9% and since the per-  
 1979 jet weights computed by Equation 6.2 are larger on average in signal than in background  
 1980 events, the signal contribution increases by the weighting and amounts to 37.1% of the sum  
 1981 of event weights computed by Equation 6.3 and given in Table 6.3.

1982 The contribution of the  $Z \rightarrow \tau^+ \tau^-$  signal needs to be determined by Monte Carlo  
 1983 simulation and subtracted from the estimate obtained by applying the “simple” fake–rate  
 1984 method to data, in order to get an unbiased estimate of the true background contributions.

### 1985 6.2.5 “CDF–type” Weights

1986 The “simple” method has the disadvantage that the use of Monte Carlo simulation is re-  
 1987 quired to subtract the contribution from true tau leptons. The CDF–method removes this  
 1988 dependence by adjusting the weights such that true tau contributions average out to zero.  
 1989 The information used to adjust the weights is based solely on information in the data,  
 1990 avoiding the need to rely on Monte Carlo based corrections.

1991 In the “CDF–type” method, additional information, namely whether or not tau can-  
 1992 didates pass or fail the tau identification cuts and discriminators, is drawn from the data.  
 1993 The desired cancellation of signal contributions is achieved by assigning negative weights  
 1994 to those tau candidates which pass all tau identification cuts and discriminators, i.e. to a  
 1995 fair fraction of genuine hadronic tau decays, but to a small fraction of quark and gluon jets  
 1996 only. The small reduction of the background estimate by negative weights assigned to quark  
 1997 and gluon jets is accounted for by a small increase of the positive weights assigned to those  
 1998 tau candidates for which at least one of the tau identification cuts or discriminators fails.  
 1999 In this way, an unbiased estimate of the background contribution is maintained.

2000 To be specific, the “CDF–type” weights assigned to tau candidates are computed as:

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

2001 For the derivation of Equation 6.5 for the “CDF–type” weights assigned to tau candidates,  
 2002 we will use the following notation: Let  $n_\tau$  ( $n_{QCD}$ ) denote the total number of taus (quark  
 2003 and gluon jets) in a certain bin of transverse momentum  $p_T^{jet}$ , pseudo–rapidity  $\eta_{jet}$  and  
 2004 jet–width  $R_{jet}$  and  $n_\tau^{sel}$  ( $n_{QCD}^{sel}$ ) denote the number of taus (quark and gluon jets) in that  
 2005 bin which pass all tau identification cuts and discriminators. By the definition of the tau  
 2006 identification efficiency<sup>6</sup>  $\varepsilon$  and fake–rate  $f$ :

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

2007 Depending on whether or not a given tau candidate passes all tau identification cuts and  
 2008 discriminators or not, we will assign a weight of value  $w_{passed}$  or  $w_{failed}$  to it. The values  
 2009 of the weights  $w_{passed}$  and  $w_{failed}$  shall be adjusted such that they provide an unbiased  
 2010 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:

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

2011 The latter equation yields the relation:

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

<sup>6</sup>As discussed in Section 6.2.1, the efficiency  $\varepsilon$  and fake–rate  $f$  are both functions of the  $p_T$ ,  $\eta$ , and width of the tau candidates.

2012 associating the two types of weights. By inserting relation 6.8 into Equation 6.7 we obtain:

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

2013 and

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

2014 which matches exactly Equation 6.5 for the “CDF–type” weights applied to tau candidates  
2015 given in Section 6.2.5.

2016 Event weights and the weights assigned to “composite particles” are computed in the  
2017 same way as for the “simple” weights, based on the weights assigned to the tau candidates:

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

2018 where the sums extend over all jets in the background-dominated control sample which  
2019 pass the preselection defined by the denominator of Equation 6.1.

2020 The effect of the negative weights to compensate the positive weights in case the “CDF–  
2021 type” fake–rate method is applied to signal events containing genuine hadronic tau decays is  
2022 shown in Table 6.4 and illustrated in Figure 6.4. As expected, positive and negative weights  
2023 do indeed cancel in the statistical average.

2024 Figures 6.5, 6.6 and 6.7 demonstrate that an unbiased estimate of the background  
2025 contribution by the “CDF–type” weights is maintained. Overall, the estimates obtained  
2026 are in good agreement with the contributions expected for different background processes,  
2027 indicating that the adjustment of negative and positive weights works as expected for the  
2028 background as well.

2029 Results obtained by the “CDF–type” fake–rate method in the closure test are sum-  
2030 marized in Table 6.4, in which the total number of background events estimated by Equa-

2031 tion 6.10 is compared to the true background contributions. The “best” estimate of the  
 2032 background contribution obtained from the “CDF-type” method is again taken as the  
 2033 arithmetic average of the estimates obtained by applying the fake-rate probabilities for  
 2034 the highest and second highest  $p_T$  jet in QCD dijet events, jets in a muon enriched QCD  
 2035 sample and jets in  $W + \text{jets}$  events. Systematic uncertainties are taken from the difference  
 2036 between the computed average value and the minimum/maximum estimate. We obtain a  
 2037 value of  $\mathcal{O}(15 - 20\%)$  for the systematic uncertainty of the “CDF-type” method, slightly  
 2038 higher than the systematic uncertainty obtained for the “simple” method. The small in-  
 2039 crease of systematic uncertainties is in agreement with our expectation for fluctuations of  
 2040 the jet-weights in case weights of negative and positive sign are used.

#### 2041 6.2.6 k-Nearest Neighbor Fake-rate Calculation

2042 ⟨sec:KNN⟩ For the fake-rate method to give correct results, care must be taken that the measured  
 2043 fake-rate is well-defined in all of the regions of phase space where it will be used. In the  
 2044 closure test described above, the computation of the fake rate was accomplished by binning  
 2045 the numerator (tau ID passed) and denominator (tau ID passed and failed) distributions  
 2046 in the three dimensions of the parameterizations. This method has the disadvantage that  
 2047 the determination of the optimal binning is extremely difficult to determine, and that any  
 2048 bins with no entries in the denominator distribution caused the fake-rate to be undefined  
 2049 in those regions.

2050 To overcome these problems, the fake-rate parameterization is implemented by adapt-  
 2051 ing a multivariate technique known as a  $k$ -Nearest Neighbor classifier ( $k$ NN). A  $k$ NN  
 2052 classifier is typically used to classify events by populating (“training”) an  $n$ -dimensional  
 2053 space with signal and background events. The probability for a given point  $x$  in the space to  
 2054 be “signal-like” is determined by finding the  $k$  nearest neighbors and computing the ratio

$$2055 p_{sig} = \frac{n_{sig}}{n_{sig} + n_{bkg}}, \quad (6.11) \quad \text{eq:KNNEquation}$$

2056 where  $n_{sig}$ ,  $n_{bkg}$  are the observed number of signal and background events, respectively. By  
 construction,  $k = n_{sig} + n_{bkg}$ . The principle of operation is illustrated in Figure 6.8.

2057 The classification feature of a  $k$ NN can be trivially adapted to parameterize an efficiency  
 2058 (fake–rate) such that it is defined everywhere. Examining the form of Equation 6.11, it is  
 2059 clear that by replacing  $n_{sig}$  with  $n_{passed}$  and  $n_{bkg}$  with  $n_{failed}$ ,  $p_{sig}$  is equivalent to the tau–  
 2060 fake rate. We thus “train” the  $k$ NN with tau candidates which pass the tau identification as  
 2061 signal events and those which fail as background events. The resulting classifier is a function  
 2062 which returns the expected fake–rate for any point in the space of the parameterization.  
 2063 The choice of  $k$  must be optimized. When  $k$  is low, the small number of neighbors causes  
 2064 large counting fluctuations in the fake rate. If  $k$  is too large, the  $k$ NN effectively averages  
 2065 over a large area of the space of the variables<sup>7</sup>. For the training statistics available in the  
 2066 2010 data,  $k = 20$  is found to be the optimal choice.

### 2067 6.2.7 Results of Fake–Rate Background Estimation

2068 An independent estimate of the background contributions to the analysis presented in this  
 2069 thesis is obtained by applying the fake–rate method in a manner analogous to the closure  
 2070 test. Fake–rates in QCD multi–jet events (light quark enriched sample), QCD events con-  
 2071 taining muons (heavy quark and gluon enriched sample) and  $W +$  jets events are measured  
 2072 using observed data [34, 48] and applied to events which pass all the event selection criteria  
 2073 listed in Table 5.3, with the exceptions of

- 2074 • the “loose” HPS + TaNC discriminator, and
- 2075 • the requirement that the tau have unit charge.

2076 No assumption is made on the composition of  $Z \rightarrow \mu^+ \mu^-$ ,  $W +$  jets,  $t\bar{t} +$  jets and QCD  
 2077 backgrounds contributing to the event sample selected by the analysis. Differences between  
 2078 fake–rates obtained for QCD multi–jet, QCD muon enriched and  $W +$  jets background  
 2079 events are attributed as systematic uncertainties of the fake–rate method. Per jet and per  
 2080 event weights have been computed by the “simple” and “CDF-type” weights as described in  
 2081 the closure test and the results are found to be compatible within statistical and systematic  
 2082 uncertainties. In the following, we present results for “CDF-type” weights. The “CDF-type”

---

<sup>7</sup>In 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.

2083 weights have the advantage that the background estimate obtained does not change, whether  
 2084 there is MSSM  $H \rightarrow \tau^+\tau^-$  signal present in the data or not.

2085 Tau identification efficiencies need to be known when using “CDF-type” weights. Ded-  
 2086 icated studies have checked the tau identification efficiencies in data [48]. Statistical and  
 2087 systematic uncertainties of these studies are still sizeable at present, on the order of 20–30%.  
 2088 No indication has been found, however, that the Monte Carlo simulation does not correctly  
 2089 model hadronic tau decays in data. For the purpose of computing fake-rate weights via  
 2090 the “CDF-type” method, tau identification efficiencies are taken from the Monte Carlo  
 2091 simulation of hadronic tau decays in  $Z \rightarrow \tau^+\tau^-$  events. Systematic uncertainties on the  
 2092 background estimate obtained by the fake-rate method are determined by varying the tau  
 2093 identification efficiencies by  $\pm 30\%$  relative to the value obtained from the Monte Carlo  
 2094 simulation.

2095 The results of applying the fake-rate method to the mu + tau channel are summarized  
 2096 in Table 6.5. The background prediction has been corrected for the expected<sup>8</sup> contribution  
 2097 of  $13.1^{+2.8}_{-0.6}$  events from  $Z \rightarrow \mu^+\mu^-$  background events in which the reconstructed tau is due  
 2098 to a misidentified muon. The obtained estimate is in good agreement with the Monte Carlo  
 2099 expectation.

2100 As an additional cross-check of the method, a sample of events containing a muon plus  
 2101 a tau of like-sign charge is selected in data and compared to the background prediction  
 2102 obtained by applying the fake-rate method to the like-sign sample. The like-sign sample is  
 2103 expected to be dominated by the contributions of  $W + \text{jets}$  and QCD background processes  
 2104 and allows to verify the fake-rate method in a practically signal free event sample. The  
 2105 background estimate obtained by the fake-rate method is compared to the number of events  
 2106 observed in the like-sign data sample in Table 6.6. The number of events expected in  
 2107 the like-sign control sample from Monte Carlo simulation is indicated in the caption. All  
 2108 numbers are in good agreement.

2109 The fake-rate method does not only allow estimation of the total number of background  
 2110 events, but models the distributions of background processes as well. The capability to  
 2111 model distributions is illustrated in Figure 6.9, which shows good agreement between the

---

<sup>8</sup>The contribution of  $Z \rightarrow \mu^+\mu^-$  is estimated using a simulated sample.

2112 distributions observed in the like-sign data sample and the predictions obtained by the fake–  
 2113 rate method for the distributions of muon plus tau visible mass and of the “full” invariant  
 2114 mass reconstructed by the SVfit algorithm.

### 2115 6.3 Template Method

2116 ⟨sec:template⟩ Shape templates for the  $\mu + \tau_{had}$  visible mass  $M_{vis}$  are obtained from observed data, using  
 2117 a set of dedicated control regions which are chosen to select a high purity sample of one  
 2118 particular background process each. The number of events selected in each control region  
 2119 and comparisons to the predictions from Monte Carlo simulations are summarized in Ta-  
 2120 ble 6.2. The template  $M_{vis}$  shapes obtained from data in the background–enriched control  
 2121 regions are compared to the signal region shapes obtained by Monte Carlo simulation in  
 2122 Figure 6.10. The  $M_{vis}$  spectrum observed in the final analysis is fitted to the sum of these  
 2123 templates. Estimates for background yields are obtained from the normalization factor of  
 2124 each template, determined by the fit. Further details of the method can be found in [47]  
 2125 and [50].

2126 The TaNC (Section 3.3, [51]) discriminators used in [50] are replaced by the correspond-  
 2127 ing discriminators of the HPS + TaNC algorithm (Section 3.5, [33]). The  $Z/\gamma^* \rightarrow \tau^+\tau^-$  sig-  
 2128 nal shape is obtained via the  $Z/\gamma^* \rightarrow \mu^+\mu^-$  embedding technique [52]. The  $\mu + \tau_{had}$  visible  
 2129 mass spectrum observed in the final analysis is compared to the sum of template shapes  
 2130 scaled by the normalization factors determined by the fit in Figure 6.11. The corresponding  
 2131 estimates for background contributions are summarized in Table 6.7.

**Fixme:** THIS  
 IS FROM  
 THE HPS  
 NOTE!

Requirement	$Z \rightarrow \mu^+ \mu^-$			Enriched background process		
	Muon fake	Jet fake	$W + \text{jets}$	$t\bar{t} + \text{jets}$	QCD	
Muon rel. iso.	< 0.15	< 0.1	< 0.1	< 0.1	< 0.1	> 0.10 & <sup>a</sup> < 0.30
Muon Track IP	-	-	-	-	-	-
Tau TAN <sub>C</sub> discr.	-	1	1	medium passed	1	
Tau 1  3-Prong	-	-	-	-	-	-
Charge(Tau) = ±1	-	-	-	-	-	-
Tau $\mu$ -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	-	-

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

gControlRegions

<sup>a</sup>“Very loose” working point passed and “Loose” working point failed

<sup>b</sup>No jets of  $E_T > 20$  GeV within  $|\eta| < 2.1$  (other than the  $\tau$  candidate)

<sup>c</sup>Minimum two jets of  $E_T > 40$  GeV, at least one of which has  $E_T > 60$  GeV and at least one with “TrackCountingHighEff”  $b$ -jet discriminator  $> 2.5$

Enriched Selection	Data	$\Sigma$ 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.  $\Sigma$  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  $\Sigma$  SM.

gControlRegions)

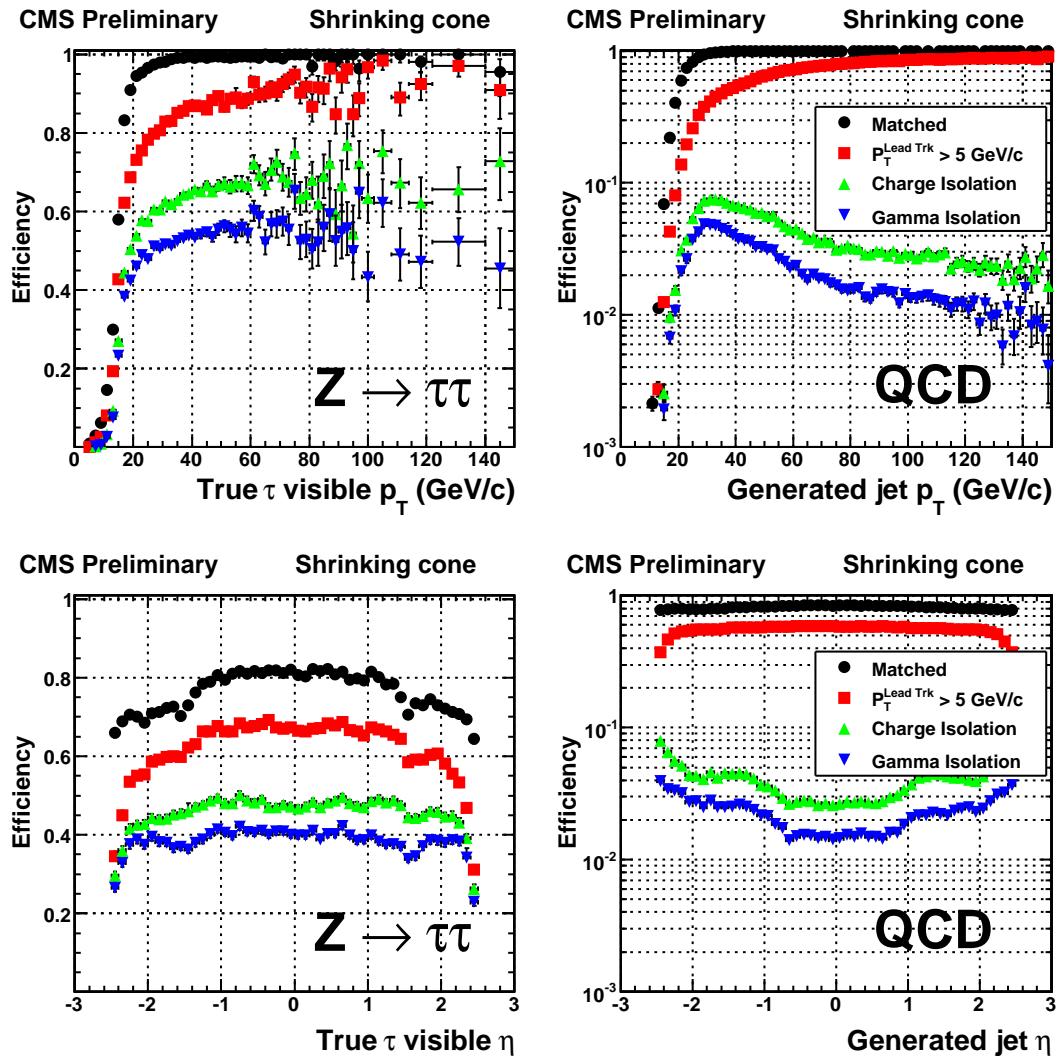


Figure 6.3: Cumulative efficiencies (left) and fake-rates (right) of successively applied tau identification cuts of the “shrinking signal cone” particle-flow based tau identification algorithm described in [28] as function of  $p_T^{\text{jet}}$  (top) and  $\eta^{\text{jet}}$  (bottom) of tau candidates.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

EfficienciesAndFakeRates)

Background Process	Expectation	Estimate obtained by applying weights of type:			Average fake-rate estimate
		QCD lead jet	QCD second jet	QCD $\mu$ -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 →  $\tau^+ \tau^- \rightarrow \mu+\tau$ -jet 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.

‘frSimpleResults’

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

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.

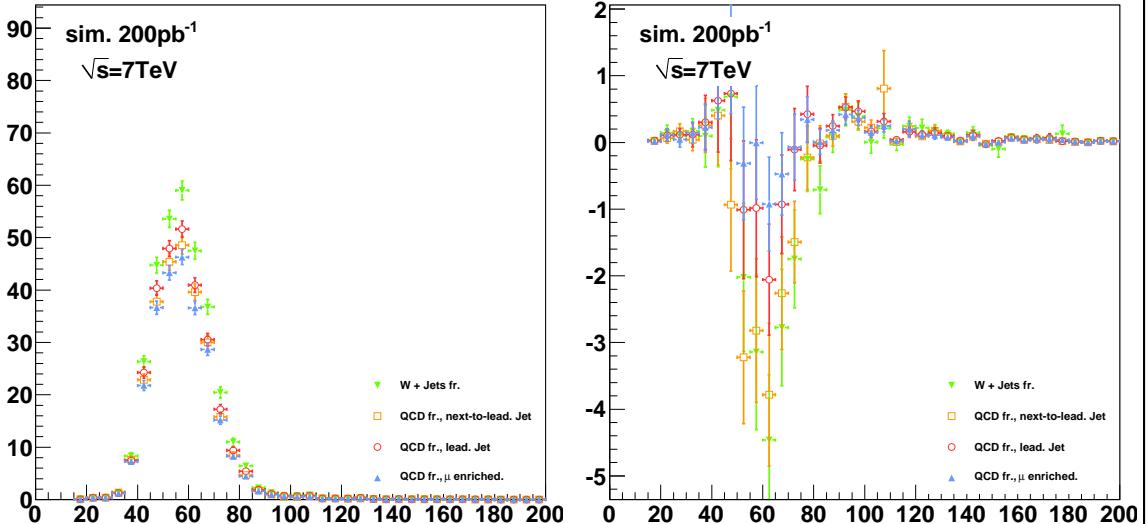


Figure 6.4: Distributions of visible invariant mass of muon plus tau in  $Z \rightarrow \tau^+\tau^-$  signal events weighted by ‘simple’ weights computed according to Equation 6.4 (left) and ‘CDF-type’ weights computed according to Equation 6.10 (right). The signal contribution to the background estimate 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.

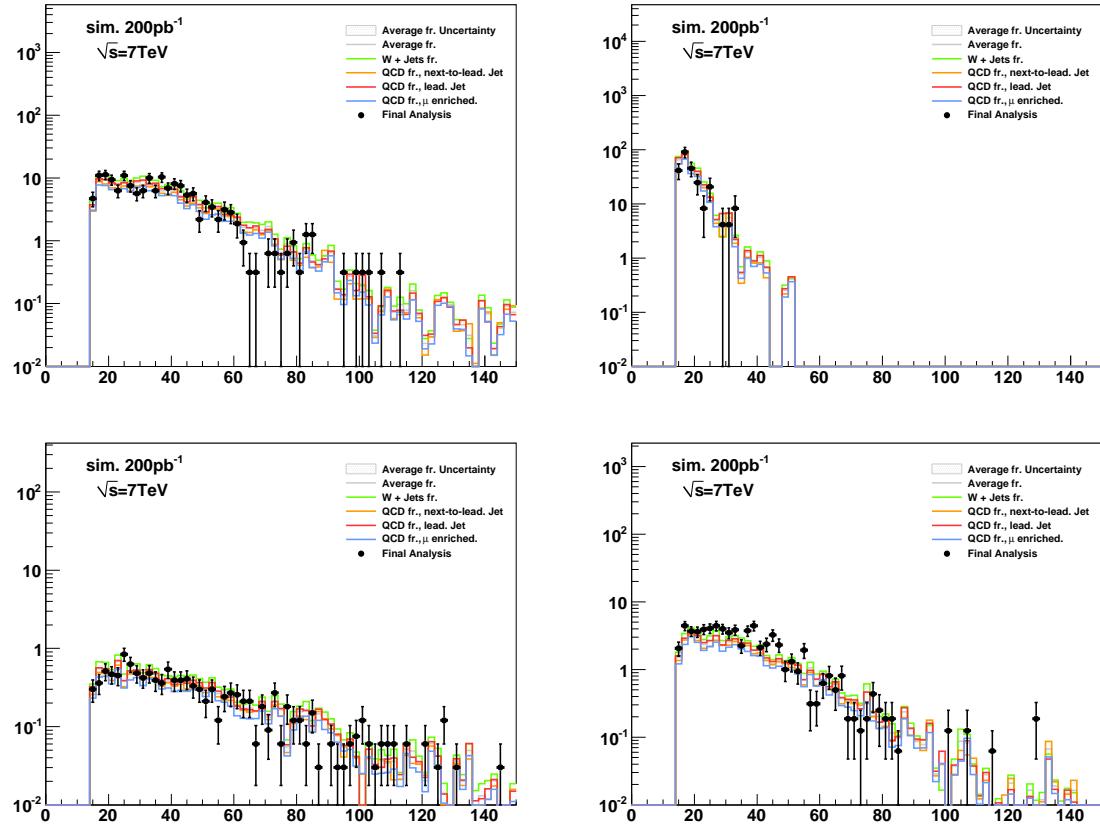


Figure 6.5: Distributions of muon transverse momentum in  $W + \text{jets}$  (top left), QCD (top right),  $t\bar{t} + \text{jets}$  (bottom left) and  $Z \rightarrow \mu^+\mu^-$  (bottom right) background events which pass all selection criteria of the  $Z \rightarrow \tau^+\tau^- \rightarrow \mu + \tau\text{-jet}$  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.

CDFtypeResults' muonPt)

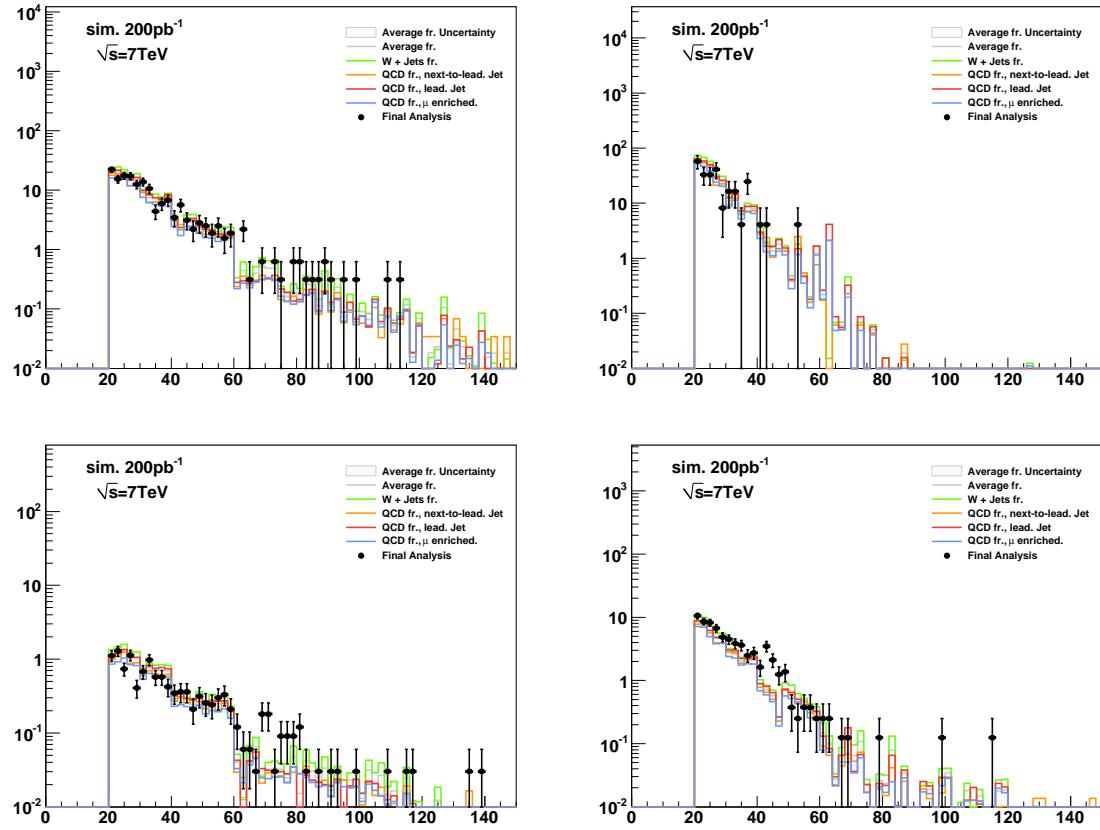


Figure 6.6: Distributions of transverse momenta of the tau candidates in  $W + \text{jets}$  (top left),  $QCD$  (top right),  $t\bar{t} + \text{jets}$  (bottom left) and  $Z \rightarrow \mu^+\mu^-$  (bottom right) background events which pass all selection criteria of the  $Z \rightarrow \tau^+\tau^- \rightarrow \mu + \tau\text{-jet}$  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.

CDFtypeResults'tauJetPt

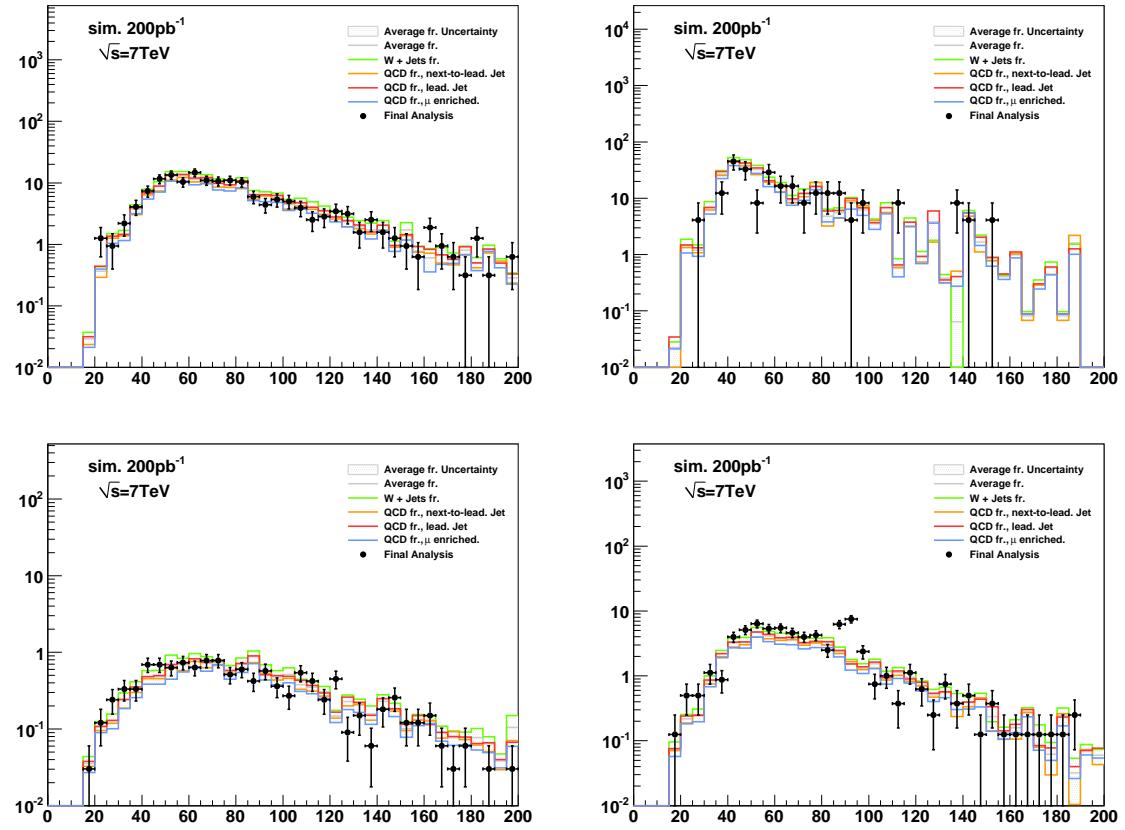


Figure 6.7: Distributions of the visible invariant mass of muon plus tau in  $W + \text{jets}$  (top left),  $\text{QCD}$  (top right),  $t\bar{t} + \text{jets}$  (bottom left) and  $Z \rightarrow \mu^+ \mu^-$  (bottom right) background events which pass all selection criteria of the closure test analysis 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.

CDFtypeResults'mVisible)

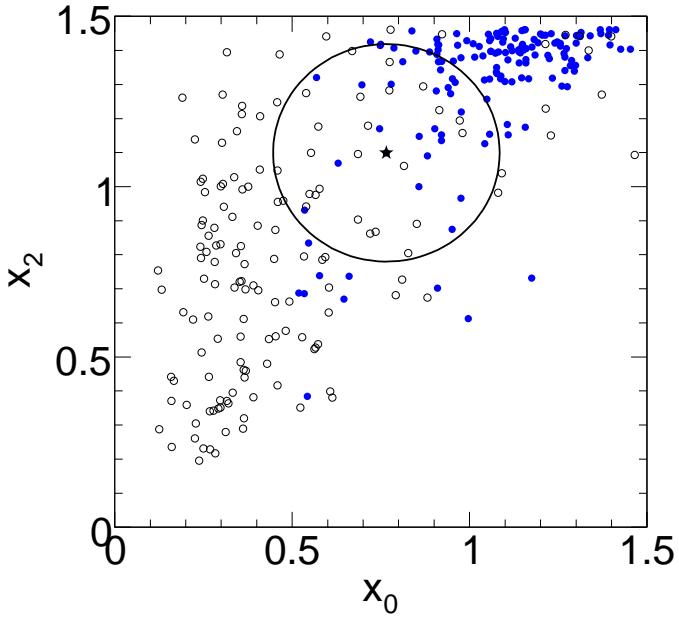


Figure 6.8: Example of the operation of a  $k$ NN classifier. The closest  $k = 50$  neighbors (those inside the circle) to a test point (indicated by the star marker) are selected. The probability that the star marker is a signal event is given the number of signal neighbors (blue markers) in the circle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

(fig:KNN)

Events weighted by:	Estimate
QCD lead jet	$202.1^{+14.9}_{-74.8}$
QCD second jet	$198.0^{+22.8}_{-79.3}$
QCD $\mu$ -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 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.

IuTauFakeRateResultsOS

Events weighted by:	Estimate
QCD lead jet	$191.7^{+2.3}_{-17.9}$
QCD second jet	$185.1^{+6.0}_{-21.1}$
QCD $\mu$ -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.

MuTauFakeRateResultsSS)

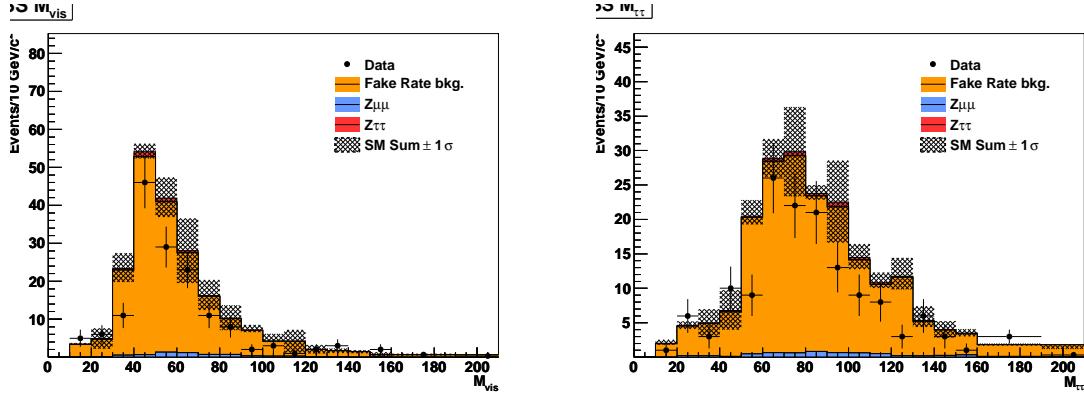


Figure 6.9: Distribution of visible mass (left) and “full” invariant mass reconstructed by the SVfit algorithm (right) observed in the like-sign charge control region compared to the background estimate obtained by the fake-rate method.

MuTauFakeRateResultsSS)

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

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.

b:BgEstTemplateMethod)

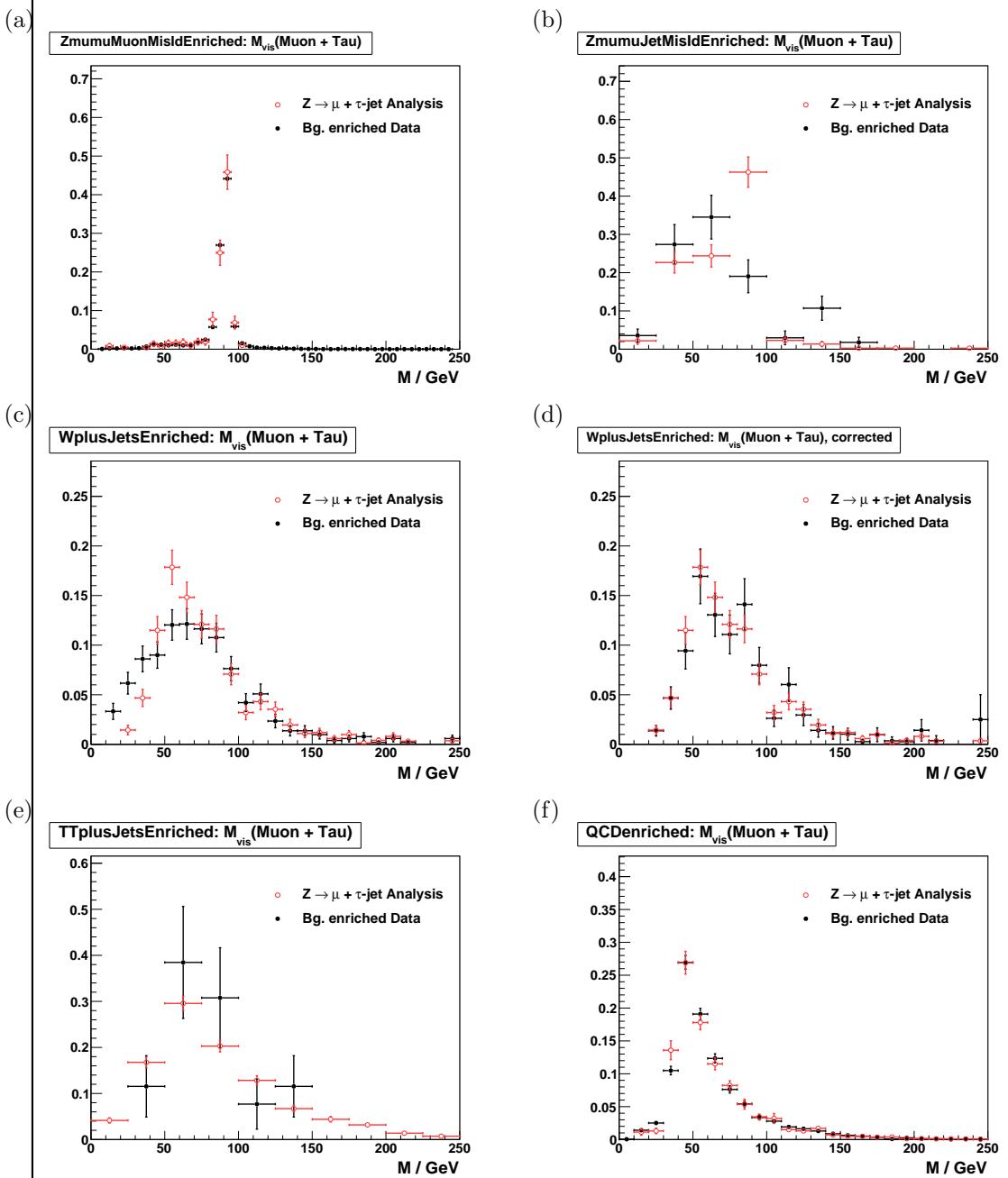


Figure 6.10:  $\mu + \tau_{had}$  shape templates obtained from  $Z \rightarrow \mu^+ \mu^-$  (a) and (b),  $W + \text{jets}$  before (c) and after (d) the bias correction explained in Section 6.3,  $t\bar{t} + \text{jets}$  (e) and QCD multi-jet (f) backgrounds enriched control regions compared to the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

{fig:VisMassTemplates}

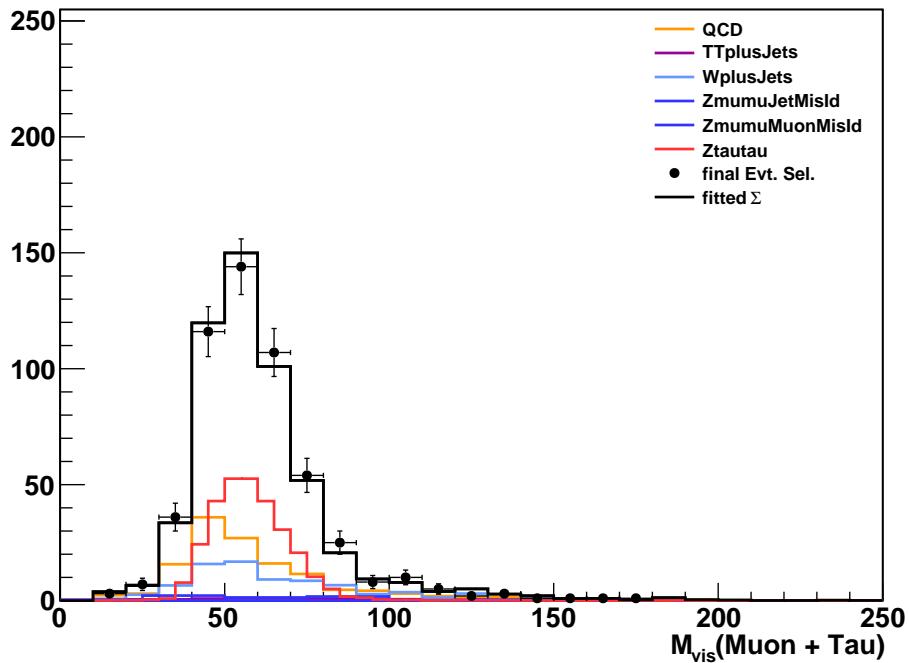


Figure 6.11:  $M_{vis}$  distribution of events selected by the  $Z/\gamma^* \rightarrow \tau^+\tau^- \rightarrow \mu + \tau_{had}$  cross-section analysis compared to the sum of shape templates for signal and background processes scaled by the normalization factors determined by the fit.

g:TemplateFitControlPlot)

## 2132 Chapter 7

# 2133 Monte Carlo Corrections

2134 One of the most important goals of the analysis is to minimize the effect of potentially  
 2135 incorrect simulation effects on the final result. While the simulated CMS events have been  
 2136 observed to match the 2010 data with surprising results, it is nonetheless critical to measure  
 2137 in real data phenomenon which can have significant effects on the analysis whenever possible.  
 2138 In practice, these measurements are used to apply a correction factor to the corresponding  
 2139 measurement obtained from Monte Carlo. This measured correction factor has an associ-  
 2140 ated uncertainty, and is taken into account as a systematic uncertainty. The application of  
 2141 systematic uncertainties is described in the next chapter.

2142 The corrections measured and used in this analysis can be divided into two categories,  
 2143 efficiency corrections and scale corrections. Identification efficiency corrections scale the  
 2144 expected yield (due to a given identification selection) up or down. Scale corrections sys-  
 2145 tematically scale the energy of a particle (or  $E_T^{\text{miss}}$ ) up or down. In this analysis we apply  
 2146 efficiency corrections for the High Level Trigger muon requirement, all stages of muon iden-  
 2147 tification, and the hadronic tau identification. We apply a momentum scale correction to  
 2148 the muon and tau legs, and to the resolution of the  $E_T^{\text{miss}}$ . Finally, events are simulated with  
 2149 overlapping “pile-up”<sup>1</sup> events. The simulated events are weighted such that the number of  
 2150 pile-up events in the simulation matches that observed in the data.

### 2151 7.1 Muon Identification Efficiency

2152 The identification efficiencies associated with the muon are measured in  $Z \rightarrow \mu^+ \mu^-$  events  
 2153 using the “tag and probe” technique [44].  $Z \rightarrow \mu^+ \mu^-$  events are selected from the Muon

---

<sup>1</sup>A 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.

2154 7 TeV CMS 2010 datasets<sup>2</sup> by requiring that the events pass the “loose” Vector Boson Task  
 2155 Force (VBTF) event selections [44]. In the selected events, we define the “tag” muons as  
 2156 those that have transverse momentum greater than 15 GeV/c and pass the VBTF muon  
 2157 selection. The tag muons are further required to pass the “combined relative isolation” de-  
 2158 scribed in the VBTF paper. We finally require that the tag muon be matched to an HLT  
 2159 object corresponding to the run-dependent requirements listed in Table 5.1. The trigger  
 2160 match requirement ensures that the event would be recorded independently of the probe  
 2161 muon. After the tag and probe muon pairs have been collected, we compare the muon identi-  
 2162 fication performance in the probe collection in events selected in data to the performance in  
 2163 simulated  $Z \rightarrow \mu^+ \mu^-$  events. The selection of events and tag muon in the simulated sample  
 2164 is the same as the data sample, with the notable exception that the only HLT requirement  
 2165 applied in MC is that the tag muon is matched to an HLT\_Mu9 object. Any difference in ef-  
 2166 ficiency between the HLT\_Mu9 path and the paths used to select the data (in the tag–probe  
 2167 measurement and in the analysis) will be considered implicitly in the correction fraction.

2168 The efficiencies for the muon selections applied in this analysis are measured using  
 2169 the “probe” objects. We measure the following marginal efficiencies, each relative to the  
 2170 previous requirement:

- 2171 • the efficiency of global probe muons to satisfy VBTF muon identification selections,
- 2172 • the efficiency of global probe muons passing the VBTF muon identification selection  
     to satisfy the isolation criteria described in Section 5.2.1,
- 2174 • and the efficiency of probe muons passing the offline analysis selection defined in  
     Chapter 5 to pass the HLT selection.

2176 In each case, the invariant mass spectrum of the tag–probe pair is fitted with a Crystal  
 2177 Ball function for the signal ( $Z \rightarrow \mu^+ \mu^-$ ) events and an exponential for the background.  
 2178 The fit is done for two cases; where the probe fails the selection and the where it passes.  
 2179 The method is illustrated in Figure 7.1. The signal yield  $N$  is extracted from each fit and  
 2180 the efficiency is computed as  $N_{\text{pass}}/(N_{\text{pass}} + N_{\text{fail}})$ . Each efficiency is measured in both the

---

<sup>2</sup>/Mu/Run2010A-Sep17ReReco<sup>2</sup>/RECO and /Mu/Run2010B-PromptReco-v2/RECO

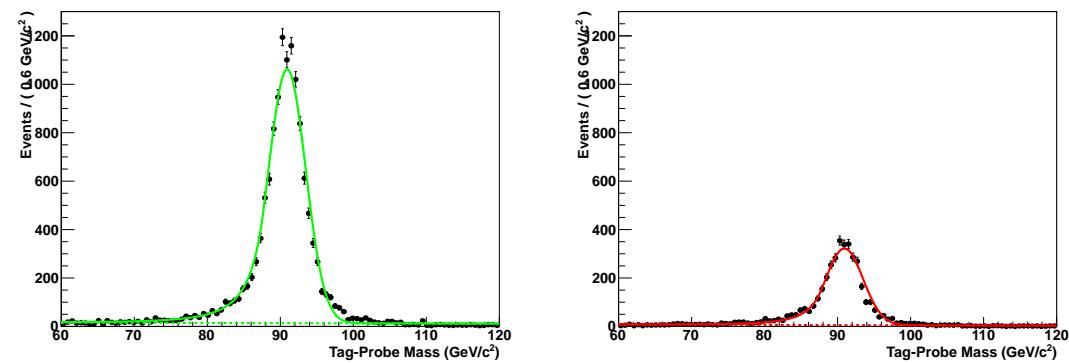


Figure 7.1: The tag–probe dimuon invariant mass spectrum in events in which the probe muon passed (left) and failed (right) the muon isolation requirement. The solid line gives the result of a simultaneous fit of the signal (real  $Z \rightarrow \mu^+\mu^-$  events) and background. The fitted background 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.

`(fig:TagAndProbeFits)`

2181 data and the simulation. The results of the measurements are shown in Table 7.1. In the  
 2182 final analysis, the simulated events are weighted by the fractional difference to the measured  
 2183 values; the statistical uncertainty on the weight is taken as the sum in quadrature of the  
 2184 statistical uncertainties for the data and simulation efficiency measurements. The uncer-  
 2185 tainty on this measurement is taken as systematic uncertainty in the final measurement.  
 2186

2187 The correction for the trigger efficiency needs to take into account the differences in  
 2188 the HLT selections applied during different operating periods (see Table 5.1). To determine  
 2189 the overall correction factor, we measure the trigger efficiency in data for each of the op-  
 2190 erating periods and compare it to the simulated efficiency of the HLT\_Mu9 selection. The  
 2191 overall efficiency in data is taken as the average of the three periods, weighted by integrated  
 2192 luminosity.

2193 The efficiency of the “cross–triggers” used in the run range 148822 – 149182 (period C)  
 2194 cannot be measured in  $Z \rightarrow \mu^+\mu^-$  events as they require a reconstructed PFTau object at  
 2195 the trigger level. A single muon trigger (HLT\_Mu15) is also used in period C. The contri-  
 2196 bution of the cross–triggers is taken as a correction to the single muon trigger period C  
 2197 efficiency. The “muon leg” of the cross–triggers have the same requirements as the single  
 2198 muon triggers used in the run range 147196 – 148058 (period B). The “cross–trigger” contri-

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.

bution 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  $\tau$  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  $N_{\text{obs}}$  in data with that expected from the known cross section, according to the cross section equation,

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

The only suitable resonance for this method is  $Z \rightarrow \tau^+\tau^-$ . This method has been applied<sup>3</sup> in the CMS  $Z \rightarrow \tau^+\tau^-$  cross section analysis [53], and measured a tau identification simulation

<sup>3</sup>Actually, 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. When measuring only the tau identification efficiency, the expected number of  $Z \rightarrow \tau^+\tau^-$  events is determined by measuring the  $Z \rightarrow \mu^+\mu^-$  rate in data, and inferring the expected number of  $Z \rightarrow \tau^+\tau^-$  events using Monte Carlo simulation. 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.

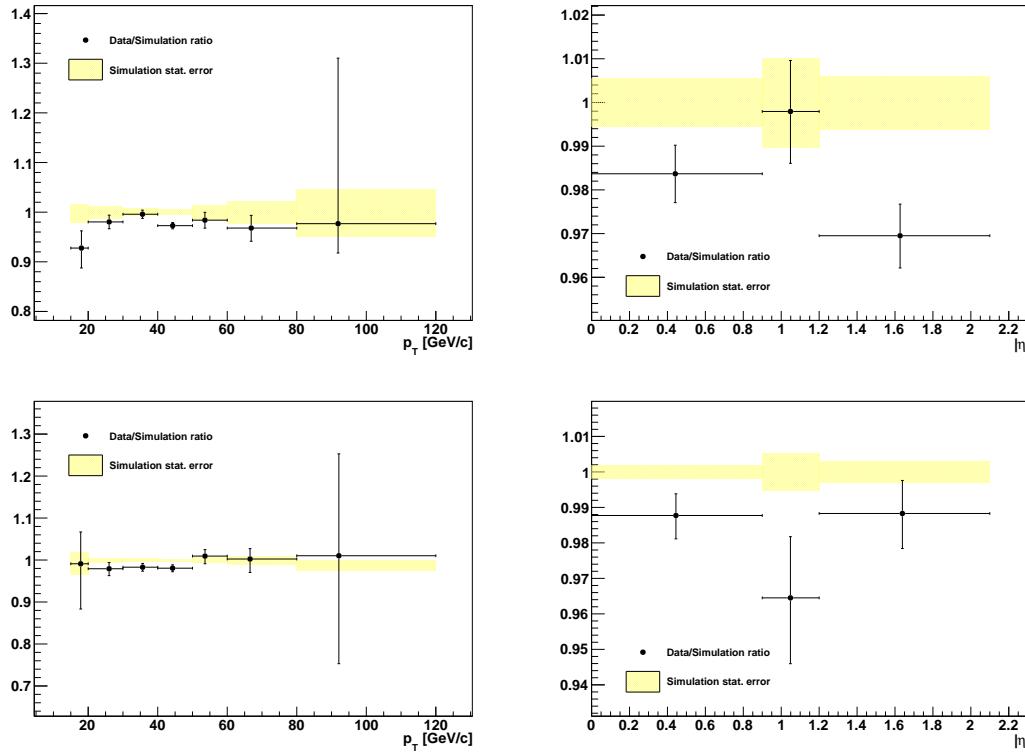


Figure 7.2: Ratio of muon isolation efficiency (top) and HLT trigger efficiency (bottom) measured in data compared to simulated  $Z \rightarrow \mu^+\mu^-$  events. The correction factors are measured in bins of  $p_T$  (left) and  $\eta$ . In the analysis, the correction factors are taken to be flat.

to data correction factor of  $0.960 \pm 0.067$ .

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 boson 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

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

The solution to this problem is to use a “tag and probe” approach analogous to the muon efficiency measurement of Section 7.1. The tag and probe method is only sensitive to

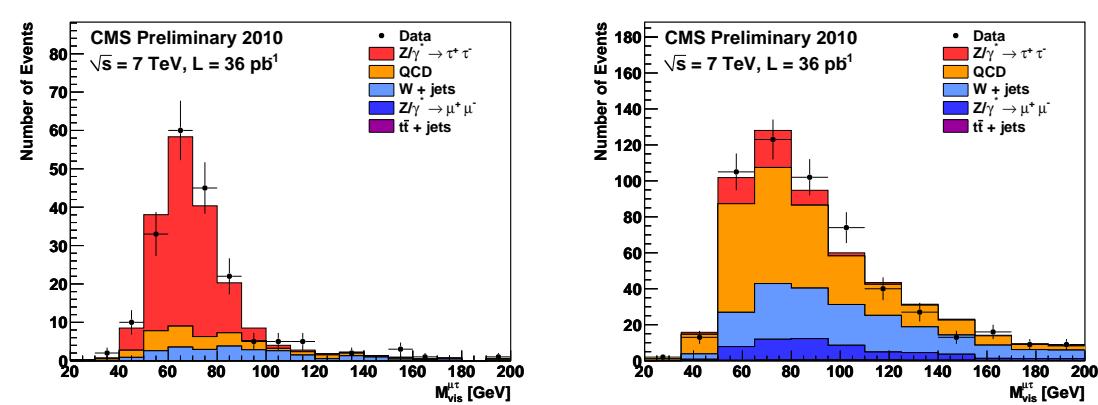


Figure 7.3: Visible mass spectrum of preselected events used to measure the hadronic tau identification efficiency in 2010 CMS 7 TeV data. The figure on the left (right) shows the preselected events that pass (fail) the hadronic tau identification. The different colors indicate the fitted yields of the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

(fig:TauIdEffFits)

the shapes of the distributions, and is independent of a Higgs boson contribution to the  $Z$  peak. This measurement has been performed by the CMS Tau Physics Object Group [48]. A loose hadronic tau preselection is applied to events which pass the selections (excluding the 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  $1.06 \pm 0.30$ .

### 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  $J/\psi$ ,  $\psi(2S)$  and  $\Upsilon$  dimuon 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 ‘‘MusCleFit’’ algorithm described in [54]. The muon momentum correction and correction and uncertainty varies as a function of

MuonTauMomentumScale)

2238 muon  $p_T$  and  $\eta$ . The effect of the muon momentum correction uncertainty is a small effect  
 2239 in this analysis compared to the  $\tau$  and  $E_T^{\text{miss}}$  scale uncertainties.

2240 The uncertainty on the jet energy scale is determined from an analysis of the  $p_T$  balance  
 2241 between photons and jets in  $\gamma + \text{jets}$  events [55]. The jet energy scale uncertainties are  
 2242 determined by the JetMET group are applied to tau-jets as well as other jets in the event.  
 2243 The tau energy scale correction factor is taken to be 1.0 with an uncertainty of 3%. The QCD  
 2244 jet energy scale has been measured to within 3% uncertainty. In the future, the energy scale  
 2245 of the tau is expected to be determined to a much better precision, as the neutral hadronic  
 2246 activity of a hadronic tau decay is expected to be zero. The jet energy scale of 3% can be  
 2247 confidently considered [48] an upper limit<sup>4</sup>, and is used in this analysis as the tau energy  
 2248 scale uncertainty.

## 2249 7.4 Missing Transverse Energy Correction

(sec:ZRecoilCorr) 2250 In practice, the resolution of the reconstructed missing transverse energy is poor as it is  
 2251 sensitive to the mis-measurement of any object in the event. Furthermore, a fraction of the  
 2252 particles produced in the hard collision can be produced in the very forward region, outside  
 2253 of the fiducial region of the calorimeters. The resolution of the  $E_T^{\text{miss}}$  reconstruction can be  
 2254 measured in  $Z \rightarrow \mu^+ \mu^-$  events. The true  $E_T^{\text{miss}}$  in such events is expected to be zero. The  
 2255  $E_T^{\text{miss}}$  resolution in simulated  $Z \rightarrow \mu^+ \mu^-$  events is found to be smaller (better) than in the  
 2256 data.

2257 The  $E_T^{\text{miss}}$  resolution depends on the “recoil” of the  $Z$  boson. The reason for this effect  
 2258 is that for events where the  $Z$  is produced nearly at rest, the associated recoil products have  
 2259 very small transverse momentum and are produced at very high pseudorapidity. The  $E_T^{\text{miss}}$   
 2260 is corrected using a procedure called a “ $Z$ -recoil” correction, as described in [56]. The reso-  
 2261 lution of the  $E_T^{\text{miss}}$  is measured in  $Z \rightarrow \mu^+ \mu^-$  events in simulation and data. The difference  
 2262 in the reconstructed  $E_T^{\text{miss}}$  resolution in both samples is parameterized by the magnitude of  
 2263 the transverse momenta of the particles recoiling against the  $Z$ .<sup>5</sup> The reconstructed  $E_T^{\text{miss}}$

<sup>4</sup>The 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%.

<sup>5</sup>The “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  $E_T^{\text{miss}}$ .

2264 in the simulated  $Z \rightarrow \tau^+ \tau^-$ ,  $Z \rightarrow \mu^+ \mu^-$ , and  $W + \text{jets}$  samples is “smeared” by a random  
 2265 amount in each event such that the final resolution matches the observed resolution in the  
 2266 data.

2267 Corrections related to the  $Z$ -recoil are determined as described in [56] and applied to  
 2268 simulated  $Z \rightarrow \tau^+ \tau^-$ ,  $Z \rightarrow \mu^+ \mu^-$  and  $W + \text{jets}$  events, in order to correct for residual  
 2269 differences in  $E_T^{\text{miss}}$  response and resolution between data and Monte Carlo simulation [57].  
 2270 The corrections are obtained by an unbinned maximum likelihood fit (in data and simu-  
 2271 lation) of the transverse recoil vector  $\vec{u}_T = -(\vec{q}_T + E_T^{\text{miss}})$  as function of the transverse  
 2272 momentum  $\vec{q}_T$  of the  $Z$ -boson in directions parallel and perpendicular to the  $Z$ -boson  
 2273 transverse momentum vector.

2274 Two additional types of corrections are applied to the simulated  $E_T^{\text{miss}}$ , “clustered”  
 2275 and “unclustered.” The overall correction factor from these two effects is taken to be unity.  
 2276 However, the uncertainty on these two correction factors is used as a systematic uncertainty  
 2277 when setting the final limit. The clustered  $E_T^{\text{miss}}$  correction varies the scale of all jets with  
 2278  $p_T > 10$  GeV/c in the event and recomputes the shifted  $E_T^{\text{miss}}$ . The unclustered  $E_T^{\text{miss}}$  cor-  
 2279 rection varies the energy scale of all jets/particles in the event with  $p_T < 10$  GeV/c. The  
 2280 scale uncertainty is 3% for the clustered correction and 10% for the unclustered correction.

## 2281 7.5 Pile-up Event Weighting

?<sec:PUweighting>  
 2282 The average number of pile-up interactions in the event can effect almost all aspects of  
 2283 the analysis. In general, increasing pile-up lowers particle identification efficiencies and  
 2284 lowers  $E_T^{\text{miss}}$  resolution. It is therefore important that the distribution of pile-up events  
 2285 in the simulation matches the distribution found in the data. Differences in the number  
 2286 of pile-up interactions between the data (averaged over the analyzed run range) and pile-  
 2287 up Monte Carlo samples produced for “BX156”<sup>6</sup> pile-up conditions are corrected for by  
 2288 reweighting Monte Carlo simulated events according to the number of reconstructed event  
 2289 vertices, in order to match the distribution measured in a  $W \rightarrow \mu\nu$  dataset triggered by  
 2290 the HLT\_Mu15 High Level Trigger path. Vertices considered for this purpose are required

---

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

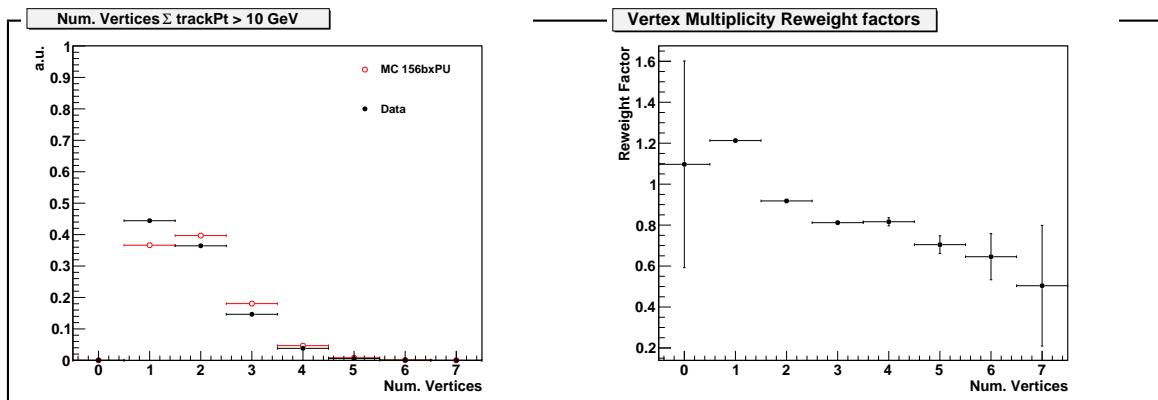


Figure 7.4: Vertex multiplicity distribution measured in the analyzed data-taking period compared to Monte Carlo simulation with “BX156” pile-up conditions (left) and resulting Monte Carlo reweighting factors (right).

{fig:pileUpReweighting}

2291 to pass  $-24 < z_{vtx} < +24$  cm,  $|\rho| < 2$  cm, nDoF  $> 4$ . In addition, the total transverse  
 2292 momenta of all tracks fitted to the vertex is required to exceed 10 GeV/ $c$ , assuming that  
 2293 “softer” vertices have little or no effect on the “hard” event to pass event selection criteria.  
 2294 The average vertex multiplicity distribution measured in data is compared to Monte Carlo  
 2295 simulation with “BX156” pile-up conditions in Figure 7.4. Both distributions are similar,  
 2296 resulting in Monte Carlo reweighting factors close to unity.

## Chapter 8

# Systematics and Limit Extraction

(ch:systematics) In this chapter we discuss the systematic uncertainties affecting the search for the Higgs boson and the statistical techniques used to establish an upper limit on the Higgs  $\rightarrow \tau^+\tau^-$  branching ratio times cross section ( $\sigma \times \mathcal{B}_{\tau\tau}$ ). The limit can be interpreted as the largest<sup>1</sup> signal presence that could exist in the data and still be consistent with the null hypothesis. The limit on  $\sigma \times \mathcal{B}_{\tau\tau}$  is roughly independent of the theoretical model<sup>2</sup>. In the conclusion, we will interpret the  $\sigma \times \mathcal{B}_{\tau\tau}$  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

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

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  $\delta N_{bkg}$ .

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

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

2318 The total relative effect of this error can be obtained by error propagation

$$\frac{\delta(\sigma \times \mathcal{B}_{\tau\tau})}{\sigma \times \mathcal{B}_{\tau\tau}} = \frac{\sigma \times \mathcal{B}_{\tau\tau}}{\partial N_{bkg}} \frac{1}{\sigma \times \mathcal{B}_{\tau\tau}} \delta N_{bkg} = \frac{-\delta N_{bkg}}{N_{obs} - N_{bkg}}. \quad (8.2)$$
eq:CrossSectionF

2319 It is interesting to examine Equation 8.2 in two scenarios. In the limit that  $N_{obs}$  is large  
 2320 compared to  $N_{bkg}$ , the effect of the error on the background estimate  $\delta N_{bkg}$  does not affect  
 2321 the final result. In contrast, in a scenario when the data is dominated by background events,  
 2322 the relative error on the signal measurement due to the background estimation approaches  
 2323 infinity. The sensitivity of a measurement to a systematic uncertainty on a parameter de-  
 2324 pends on the context in which that parameter is used.

2325 Experimental systematic uncertainties relevant for MSSM  $H \rightarrow \tau^+\tau^-$  search presented  
 2326 in this thesis are classified in three categories: normalization uncertainties on the signal and  
 2327  $Z \rightarrow \tau^+\tau^-$  background (events with true taus), normalization uncertainties on contribu-  
 2328 tions from background with fake taus, and shape uncertainties. Normalization uncertainties  
 2329 on events with true taus are due to lepton reconstruction, identification, isolation and trig-  
 2330 ger efficiencies. These terms are equivalent to the efficiency  $\epsilon$  and acceptance terms  $\mathcal{A}$  of  
 2331 Equation 8.2 and affect the expected yield of MSSM  $H \rightarrow \tau^+\tau^-$  signal and  $Z \rightarrow \tau^+\tau^-$   
 2332 background events. The uncertainties on these effects are obtained by measuring the effect  
 2333 in data and simulation, according to the procedures of Chapter 7, and calculating a cor-  
 2334 rection factor. The uncertainty associated with the measurement of the correction factor  
 2335 is the systematic uncertainty. The normalization uncertainties are assumed to be uncor-  
 2336 related with the shapes of visible and SVfit mass distributions which are used to extract  
 2337 the MSSM  $H \rightarrow \tau^+\tau^-$  signal contribution in the analyzed dataset. Uncertainties on the  
 2338 shapes of the distributions are described by “morphing” systematics. These are due to  
 2339 uncertainties on the momentum/energy scale of identified electrons, muons, tau and other  
 2340 jets in the event. As the SVfit mass reconstruction algorithm uses the missing transverse  
 2341 energy, the shape of the SVfit distribution is sensitive to systematic uncertainties on the  
 2342 overall scale  $E_T^{\text{miss}}$  measurement. The “morphing” systematics affect the shapes of signal as  
 2343 well as background contributions. Normalization uncertainties on background contributions  
 2344 are estimated from the level of agreement between data and Monte Carlo simulation in

2345 background dominated control regions.

## 2346 8.1 Signal Normalization Uncertainties

2347 The signal normalization uncertainties are due to imperfect knowledge of how improperly  
 2348 modeled effects in the simulation could affect the acceptance model, the probability that  
 2349 a given signal event will pass one of the selections (detailed in Chapter 5). The general  
 2350 procedure to quantify these uncertainties is to measure the effect in some control region in  
 2351 both the data and Monte Carlo. The ratio of data to Monte Carlo then gives a correction  
 2352 factor which is applied to the simulation. An uncertainty on the measurement of the effect in  
 2353 control region (in the data, simulation, or both) is then taken as the systematic uncertainty.  
 2354 The signal normalization uncertainties affecting this analysis on muon trigger, reconstruc-  
 2355 tion, identification and isolation efficiencies are taken from the tag and probe analysis of  
 2356  $Z \rightarrow \mu^+ \mu^-$  events presented in Section 7.1. A very conservative estimate of 30% is taken  
 2357 for the uncertainty on the tau reconstruction and identification efficiency. The tau identifi-  
 2358 cation uncertainty measurement is discussed briefly in 7.2. The dependency of the Higgs  
 2359 boson signal extraction on the tau identification efficiency has been studied, the result being  
 2360 that uncertainties on the tau identification efficiency affect the limit on cross-section times  
 2361 branching ratio for MSSM  $H \rightarrow \tau^+ \tau^-$  production by a few percent only. An uncertainty of  
 2362 11% is attributed to the luminosity measurement [58].

## 2363 8.2 Background Normalization Uncertainties

2364 Uncertainties on the normalization of background processes are obtained from the study  
 2365 of background-enriched control regions presented in Chapter 6. The main fake tau back-  
 2366 grounds in this analysis are QCD multi-jet and  $W +$  jets events. These backgrounds are  
 2367 produced copiously enough for the backgrounds to be studied in control regions dominated  
 2368 by a single background process with a purity exceeding 90% and an event statistics exceed-  
 2369 ing the expected contribution of that background to the analysis by more than one order of  
 2370 magnitude. Both backgrounds are found to be well modeled by the Monte Carlo simulation.  
 2371 The overall background yields used in the final fit are measured in the background-enriched

control regions, and cross-checked using the Fake-rate and Template methods. For a detailed discussion of the measurements, see Chapter 6. An uncertainty of 10% is attributed to the contribution of QCD and  $W + \text{jets}$  backgrounds to the analysis. The cross-section for  $t\bar{t} + \text{jets}$  production makes it difficult to select a high purity sample of  $t\bar{t} + \text{jets}$  events of high event statistics. From the study of the 19 events selected in the  $t\bar{t} + \text{jets}$  background-enriched control sample we assume an uncertainty on the  $t\bar{t} + \text{jets}$  background contribution in the analysis of 30%. The  $Z \rightarrow \mu^+ \mu^-$  background has been studied with large statistical precision in two separate control regions, and is dominated by events in which the reconstructed tau candidate is either due to a misidentified quark or gluon jet or due to a misidentified muon. Good agreement between data and Monte Carlo simulation is found in both cases. Sizeable uncertainties on the  $Z \rightarrow \mu^+ \mu^-$  background contribution arise due to the extrapolation from the background-enriched control regions to the data sample considered in the analysis, however: the contribution of  $Z \rightarrow \mu^+ \mu^-$  background events to the analysis is due to events in which one of the two muons produced in the  $Z$  decay either escapes detection or fakes the signature of a hadronic tau decay. Both cases may be difficult to model precisely in the Monte Carlo simulation. The non-observation of a  $Z$  mass peak in the mu + tau visible mass distribution studied with the fake-rate method on the other hand sets a limit on possible contributions from  $Z \rightarrow \mu^+ \mu^-$  background events. Conservatively, we assume an uncertainty of 100% on both types of  $Z \rightarrow \mu^+ \mu^-$  background contributions.

### 8.3 Shape Uncertainties

`<sec:ShapeUncertainties>` Shape uncertainties on the distributions of visible and “full” invariant mass reconstructed by the SVfit algorithm are estimated by varying the electron energy and muon momentum scale, the energy scale of taus and other jets in the event and varying the missing transverse energy in Monte Carlo simulated events. After each variation the complete event is reconstructed and passed through the event selection. Shifted visible and “full” invariant mass shapes are obtained for each variation from the events passing all event selection criteria. The difference between shifted shapes and the “nominal” shapes obtained from Monte Carlo simulated events with no variation of energy or momentum scale or of the missing transverse energy applied is then taken as shape uncertainty.

2401 The systematic uncertainties on the muon and tau energy scales have been provided  
 2402 by the muon and tau CMS Physics Object Groups and are described in Section 7.3. The  
 2403 modelling of missing transverse energy in different types of background events has been  
 2404 studied in the background–enriched control regions described in Chapter 6. No significant  
 2405 deviations between data and Monte Carlo simulation have been found. Uncertainties due  
 2406 to missing transverse energy are estimated by varying the “clustered” and “unclustered”  
 2407 energy scales, described in Section 7.4 and recomputing the total  $E_T^{\text{miss}}$ .

## 2408 8.4 Theory Uncertainties

2409 The signal and background normalization as well as the shape uncertainties are all exper-  
 2410 imental uncertainties in nature. Additional theoretical uncertainties arise from imprecise  
 2411 knowledge of parton–distribution functions (PDFs) and of the exact dependency of signal  
 2412 cross–sections and branching ratios on  $\tan\beta$  and  $m_{A^0}$ . The PDFs describe how the energy  
 2413 of the protons is shared between the quarks and gluons. Since the longitudinal boost of  
 2414 the hard collisions depends on the PDFs, the signal acceptance is sensitive to errors in the  
 2415 PDFs. The uncertainties on the signal acceptance due to PDF uncertainties are estimated  
 2416 using tools developed by the CMS Electroweak working group [59]. The acceptance is com-  
 2417 puted with respect to MSSM  $H \rightarrow \tau^+\tau^-$  decays that have muons of  $p_T^\mu > 15$  GeV/c and  
 2418  $|\eta_\mu| < 2.1$ , jets produced in hadronic tau decays with visible  $p_T^{\text{vis}} > 20$  GeV/c and  $|\eta_{\text{vis}}| < 2.3$   
 2419 on generator level. Acceptance values are computed for the central value and 44 eigenvec-  
 2420 tors of the CTEQ66 PDF set [60]. The systematic uncertainty on the signal acceptance is  
 2421 computed following the PDF4LHC recommendations [61, 62].

2422 The effect of Monte Carlo normalization, shape and theory uncertainties on the signal  
 2423 efficiency times acceptance is summarized in Table 8.1.

## 2424 8.5 Limit Extraction Method

2425 `<sec:statmethod>` The search for a new signal is performed by examining the observed distribution of the  
 2426 reconstructed ditau invariant mass  $m_{\tau\tau}$  as reconstructed by the SVfit algorithm. A localized  
 2427 excess in this spectrum would indicate the presence of a new particle. To make a statement  
 2428 about the presence of an excess with confidence, the shape of the spectrum under the

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

<sup>1</sup> decreasing with  $m_A$

<sup>2</sup> 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%

<sup>3</sup> 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.

(tab:ExpUncertainties)

2429 potential signal must be well described. The background shape is decomposed into the  
 2430 combination of shapes from the difference background sources, which we refer to here as  
 2431 “templates.” The data and the templates for each background distribution is binned in  
 2432 the observable  $m_{\tau\tau}$  variable. The normalization of each template represents the total yield  
 2433 expected for that source.

2434 We can then define a likelihood for any configuration of our templates given the ob-  
 2435 served data. The likelihood is a “binned Poisson likelihood,” which is defined as the product  
 2436 of the Poisson probability in each bin. The Poisson probability  $P(n|\mu)$  is the probability to  
 2437 observe  $n$  events given an expectation of  $\mu$  events. The Poisson probability is given by the  
 2438 expression

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

eq:PoissonProb

2439 The total likelihood for observed data given some configuration of templates is then simply  
 2440 the product of the Poisson probabilities (Equation 8.3) in each of the  $N_{bin}$  bins:

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

2441 where the expected number of events  $\mu_i$  in the bin  $i$  is the sum of the number of events  
 2442 from all sources

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

2443 The number of expected events in a source, in turn, can be written

$$\mu_{ji} = L \sigma_j \epsilon_{ji} \quad (8.5)$$
?eq:ShapeParameters

2444 where  $L$  is the integrated luminosity,  $\sigma_j$  is the cross section for source  $j$ , and  $\epsilon_{ji}$  is the  
 2445 efficiency for source  $j$  in bin  $i$ .

2446 We incorporate the systematic uncertainties of the analysis by introducing a set of  
 2447 “nuisance parameters”  $\vec{\beta}$  into the likelihood function. As the name suggests, we are not  
 2448 interested in the actual value of the nuisance parameters. Each nuisance parameter pa-  
 2449 rameterizes some phenomenon in the analysis. The shape templates (which can be defined  
 2450 purely in terms of  $\sigma_j$  and  $\epsilon_{ji}$  in Equation 8.5) are now interpreted as functions of the set  
 2451 of nuisance parameters  $\vec{\beta}$ . Existing knowledge about the value of the nuisance parameter is  
 2452 introduced by extending the likelihood function with a constraint  $\mathcal{G}(\vec{\beta})$  that expresses the  
 2453 knowledge about the nuisance parameters. The templates  $\mu_{ji}$  can depend on the nuisance  
 2454 parameters in two ways. Normalization uncertainties introduce multiplicative nuisance fac-  
 2455 tors on the yield of some (sub)set of the sources defined by Equation 8.5. As an example,  
 2456 consider a simplified situation where there are only two sources:  $Z \rightarrow \tau^+ \tau^-$  and  $W + \text{jets}$   
 2457 events. Both sources are sensitive to the efficiency of the muon trigger. Only the  $Z \rightarrow \tau^+ \tau^-$   
 2458 sample is sensitive to the efficiency of the hadronic tau identification algorithm, as the  
 2459 fake-rate in  $W + \text{jets}$  is measured in data. We would then introduce two multiplicative nui-

2460 sance parameters,  $\beta_\mu$  and  $\beta_\tau$ , which respectively correspond to the two uncertainties. The  
 2461 expected number of events in the  $i$ th bin is then given by a modified form of Equation 8.5

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

eq:ShapeParamet

2462 From Equation 8.6, we can see that  $\beta_\mu$  affects both sources but  $\beta_\tau$  only affects the  $Z \rightarrow \tau^+\tau^-$   
 2463 source.

2464 The shape uncertainties discussed in Section 8.3 are incorporated using a technique  
 2465 called “vertical template morphing.” For each source, different templates are created for  
 2466 three different values of the morphing parameter, corresponding to -1, 0, and +1 standard  
 2467 deviation shifts in the nuisance parameter. To determine the number of expected events in  
 2468 the  $i$ th bin as a function of the morphing parameter, we interpret quadratically between  
 2469 the  $i$ th bin values of the three templates, and extrapolate linearly beyond them. The overall  
 2470 likelihood then, including nuisance parameters, can be written as

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

eq:LikelihoodNui

2471 In Equation 8.7, we have introduced the assumption that the nuisance parameters are  
 2472 uncorrelated.

2473 We wish to determine if the data is compatible with a new signal. To test for the  
 2474 presence of a bump, we examine the likelihood as a function of the signal cross-section. If the  
 2475 presence of a signal is unambiguous, one can simply determine the likelihood as a function  
 2476 of the cross-section. If the signal is known to be non-zero, confidence intervals about the  
 2477 maximum likelihood can be constructed by examining the change in the logarithm  $2\Delta \ln \mathcal{L}$   
 2478 of the likelihood. The Frequentist interpretation of the confidence level  $(1 - \alpha)$  is that if  
 2479 the experiment were repeated  $N$  times, the interval corresponding to  $(1 - \alpha)$  would fail  
 2480 “cover” the *true* value of the observable in  $\alpha \times N$  of the experiments. The correspondence  
 2481 between confidence levels and intervals in  $2\Delta \ln \mathcal{L}$  is given in Table 8.2. Since the mass of  
 2482 a potential Higgs boson is unknown, we repeat this process for different signal masses. If  
 2483 there is not a significant signal, we can set upper bounds on the signal cross-section using

$(1 - \alpha)\%$	$m = 1$	$m = 2$	$m = 3$
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.

(tab:ConfidenceIntervals)

2484 one of several methods, which will be discussed below.

2485 At this point, the likelihood still depends on the nuisance parameters. There are two  
 2486 methods of removing the dependence on the nuisance parameters, “marginalization” and  
 2487 “profiling.” Marginalization is the process of integrating the likelihood of Equation 8.7  
 2488 over the entire range of all nuisance parameters. This effectively averages the effect of  
 2489 the different uncertainties into the marginalized likelihood function. The profiling method  
 2490 maximizes the likelihood in terms of the nuisance parameters. In the profiling method,  
 2491 the “profile likelihood” is created by maximizing the likelihood with respect to all of the  
 2492 nuisance parameters. One way to interpret the profiling method is that the values of nuisance  
 2493 parameters are being measured *in situ*, constrained by the observed data. While not a  
 2494 strictly Bayesian treatment, the profile likelihood method has been shown to give almost  
 2495 identical results to marginalization.

2496 An interesting situation relating to the profiling of the nuisance parameters arises in  
 2497 this analysis in the context of the hadronic tau identification uncertainty. As discussed in  
 2498 Chapter 7, the Higgs–insensitive measurement of the tau identification efficiency has a high  
 2499 uncertainty of 30%. If the bump at the  $Z \rightarrow \tau^+\tau^-$  resonance can be considered to free  
 2500 of contributions from a Higgs boson, the tau identification efficiency can be measured to  
 2501 a much greater precision, approximately 7% [53, 48]. When profiling the likelihood (as a  
 2502 function of cross section) for a Higgs boson with a mass  $m_{A^0} > m_Z$ , the likelihood contains  
 2503 the information that there is no Higgs boson contribution to the  $Z$  resonance. This fact  
 2504 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 boson signal with  $m_{A^0} = m_Z$ , the likelihood cannot distinguish between a potential presence of Higgs boson 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

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

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

In the results presented below we use nuisance parameters corresponding to the systematic errors summarized in Table 8.3.

Source	Method	Magnitude
Muon ID/trigger	Multiplicative	20%
$Z$ cross section	Multiplicative	5%
Jet to $\tau$ 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%
Jet energy scale	Morphing	3%
Unclustered energy scale	Morphing	1%
$Z$ -recoil correction	Morphing	-
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. The  $Z$ -recoil correction factor is varied within the uncertainty on the measurement described in Chapter 7.

(tab-sys)

## 2521 Chapter 9

# 2522 Results

(ch:results)

### 2523 9.1 Selected Events

2524 In the  $36 \text{ pb}^{-1}$  of 7 TeV data collected by CMS in 2010, the analysis selection criteria  
2525 described in Chapter 5 selects a total of 573 events. The expected yields from each of the  
2526 background sources, computed by the Template method (Section 6.3) and verified by the  
2527 Fake-rate method (Section 6.2) are shown in Table 9.1. The total expectation from the  
2528 standard model background is 577. The data agree extremely well with the SM background  
2529 expectation considering the expected statistical variance of the observed data (24 events), as  
2530 well as the uncertainty on the integrated luminosity<sup>1</sup> and the tau identification uncertainty.

2531 The distributions of the visible mass and SVfit mass in the final selected events are  
2532 shown in Figure 9.3. Excellent agreement is seen between the shapes of the distributions. The  
2533  $p_T$  spectrum of the transverse momentum are sensitive to mis-modeling of the kinematics  
2534 and composition of the various background sources. Control plots showing the transverse  
2535 momenta of the muon and visible hadronic tau in the final analysis selection are shown  
2536 in Figure 9.1. The  $\eta$  and  $\phi$  distributions of the muon and tau objects are sensitive to  
2537 detector effects, and the presence of cosmic muons. For example, muons from cosmic events  
2538 will preferentially be produced in the  $\phi = 0$  direction. Spurious candidates resulting from  
2539 poorly model noise in one of the CMS subdetectors will in general be localized in  $\eta - \phi$ . The  
2540  $\eta$  and  $\phi$  distributions of the muon and tau candidates are shown in Figure 9.2 and show  
2541 excellent agreement.

<sup>1</sup>The 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 <sup>-1</sup>
$t\bar{t} + \text{jets}$	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  $H \rightarrow \tau^+ \tau^- \rightarrow \mu + \tau_{had}$  candidate events passing the selection criteria described in Chapter 5.

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

## 9.2 Limits on Higgs Boson Production

We compute upper limits on the cross section times the branching ratio using the Bayesian method described in Section 8.5. We compute an expected limit in the same manner as an observed limit, but with simulated data generated in “toy” experiments. A large number of pseudo-data sets are generated using the null hypothesis templates using Monte Carlo techniques. The nuisance parameters are sampled within their constraints in the generation of the pseudo-data. The pseudo-data sets are expected to have the same statistical sensitivity as the observed dataset. Upper limits are then computed using the pseudo-data. The process is repeated many times, and the spread of the obtained upper limits determines the expected upper limit band. The expected nominal upper limit, and the  $\pm 1$ , and  $\pm 2$  confidence limits are shown in Table 9.3. The observed limit on the MSSM computed from the

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

2558 573 events selected in this analysis is given in the right column of Table 9.3. The observed  
 2559 limit is compatible with the expected limit, within 1.5 standard deviations. The trend of  
 2560 the expected and observed limits versus the Higgs boson mass using both observables are  
 2561 shown in Figure 9.4. The use of the SVfit reconstructed mass as the observable increases  
 2562 the power of the limit significantly. The limit trend has some interesting features. When  
 2563 the Higgs boson mass is close to the mass of the  $Z$  resonance, the analysis have little power  
 2564 to set a limit on the presence of the Higgs boson. This is due to the large uncertainty on  
 2565 the tau identification efficiency. Essentially, when  $m_{A^0} = m_Z$ , the Higgs boson yield in the  
 2566  $Z$  bump would have been larger than 30% of the  $Z \rightarrow \tau^+ \tau^-$  yield for the profile likelihood to  
 2567 be able to recognize an excess of events. Below this value, the profile likelihood can simple  
 2568 shift the tau identification efficiency scale factor up by 30% and “eat” any potential excess  
 2569 of signal.

### 2570 9.3 Interpretation in the MSSM

(sec:MSSMInterp) 2571 The limits on the cross section times branching ratio are roughly model independent,<sup>3</sup> and  
 2572 could be applied to set limits on the parameter space of a number of models. In this thesis,  
 2573 we interpret the results in the context of the MSSM. Specifically, we exclude a region in the  
 2574  $\tan \beta - m_{A^0}$  parameters space of the MSSM. To find the upper limit band on  $\tan \beta$ , we find  
 2575 the minimum value of  $\tan \beta$  which provides the cross section and branching ratio product  
 2576 found in the corresponding row in Table 9.3.

2577 The mapping between  $m_{A^0}$  and  $\tan \beta$  and the Higgs boson cross section is provided by  
 2578 the LHC Higgs Cross Section working group [63]. The cross sections and branching ratios  
 2579 have been computed for the  $h^0$ ,  $H^0$ , and  $A^0$  MSSM Higgs boson states in both the  $ggA$   
 2580 and  $qqA$  production modes, for a grid of points in  $\tan \beta - m_{A^0}$  space. In order to combine  
 2581 the  $ggA$  and  $qqA$  production modes, what we call our signal cross-section is the sum of the  
 2582 cross-section times branching ratio for both modes, assuming  $\tan \beta = 30$ . Additionally, as  
 2583 discussed in Section 1.2.3, the MSSM Higgs sector consists of two Higgs doublets, yielding

---

<sup>3</sup>This 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.

2584 five physical Higgs bosons. This search is sensitive to the three neutral Higgs particles the  
 2585  $h^0$ ,  $H^0$ , and  $A^0$ . The relative contributions of the three Higgs boson types depends on the  
 2586 mass  $m_{A^0}$  of the CP-odd Higgs boson. An observed signal will have contributions from at  
 2587 least two Higgs boson states. For  $m_{A^0} \leq 130$  GeV/ $c^2$ , the  $A^0$  and  $h^0$  are approximately  
 2588 degenerate in mass and width. In this region the  $H^0$  has a very small relative cross section  
 2589 and a constant mass of  $m_{H^0} \approx 130$  GeV/ $c^2$ . For  $m_{A^0} \geq 130$  GeV/ $c^2$ , the  $h_0$  reaches a limiting  
 2590 mass of  $\approx 130$  GeV/ $c^2$ , and the  $H^0$  and  $A^0$  become mass degenerate.

2591 The region in  $\tan\beta - m_{A^0}$  MSSM parameter space excluded by this analysis at 95%  
 2592 CL is shown in Figure 9.3. The limit is compared to the combined result from Run II of  
 2593 the Tevatron (this result is discussed in detail in Section 1.3.3). The result of this analysis  
 2594 sets a stronger limit than the Tevatron for large values of  $m_{A^0}$ . In the low  $m_{A^0}$  region, the  
 2595 analysis suffers due to the large tau identification efficiency uncertainty. This effect can be  
 2596 mitigated by using the  $e - \mu$  channel. The combined CMS result uses this approach, and  
 2597 will be discussed briefly in the conclusion.

**Fixme:**

check this!

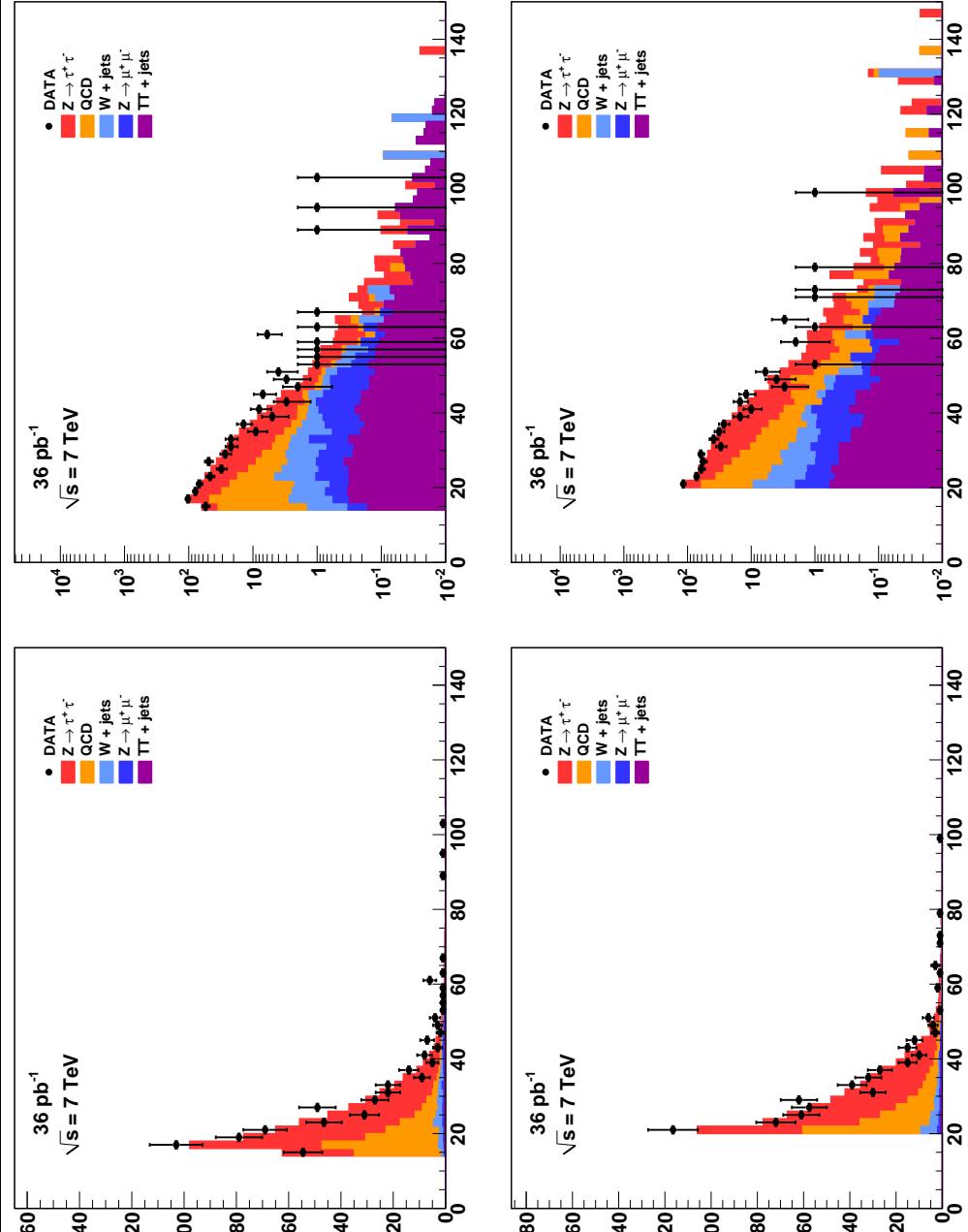


Figure 9.1: Distribution of the transverse momentum of the muon (top) and hadronic tau in  $H \rightarrow \tau^+\tau^- \rightarrow \mu + \tau_{had}$  candidate events passing the selection criteria described in Chapter 5. The distributions are shown in linear (logarithmic) scale on the left (right).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

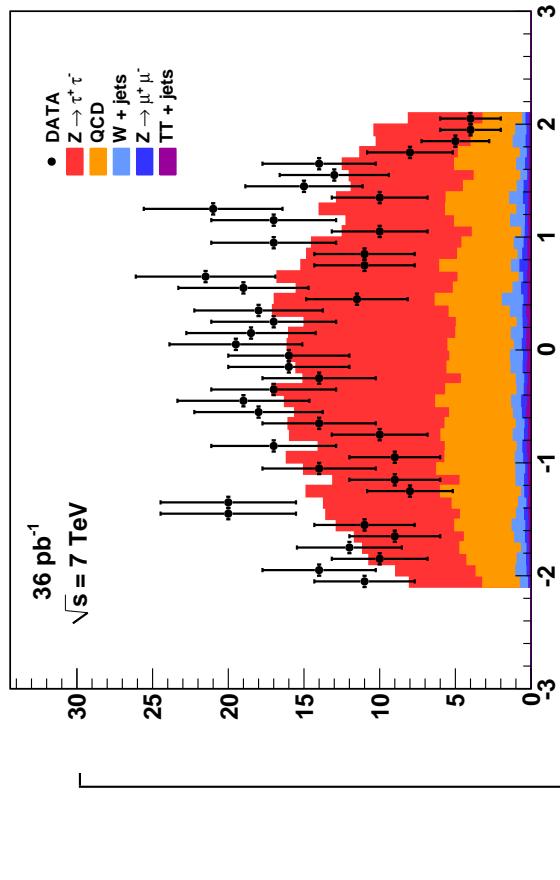
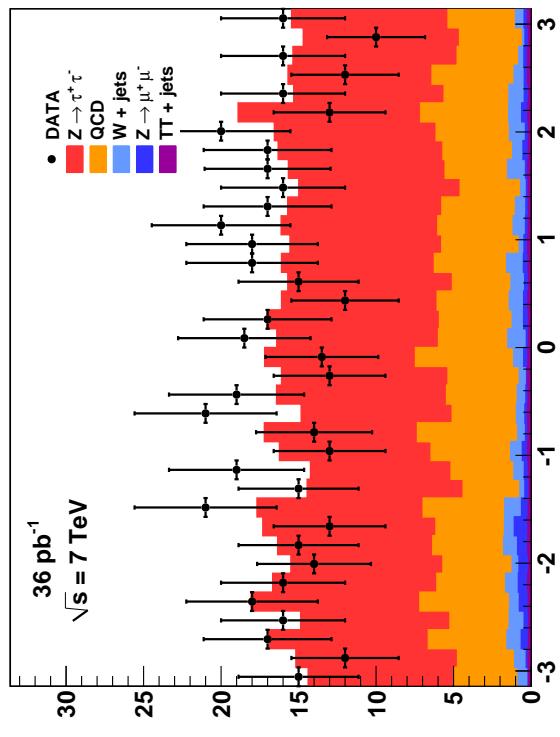
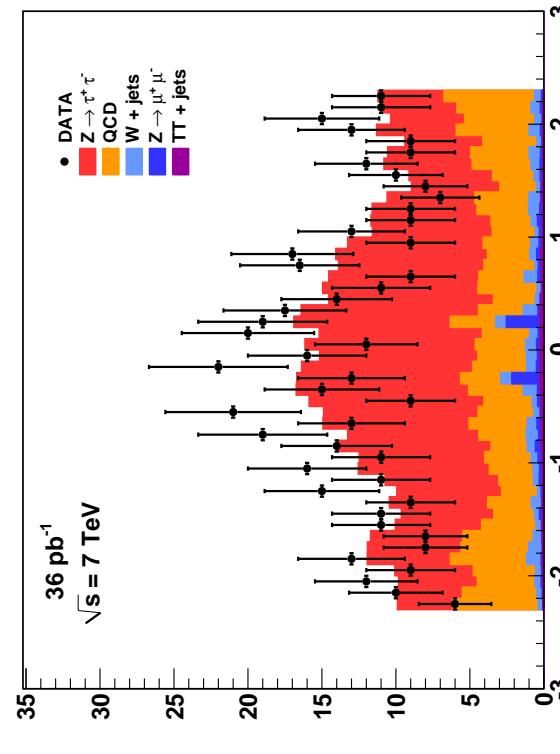
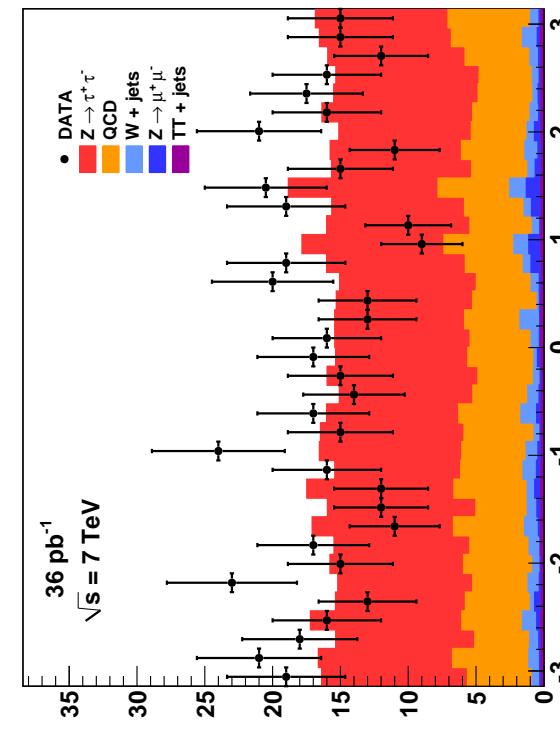
**Muon (final event sample)****Muon (final event sample)****Tau (final event sample)****Tau (final event sample)**

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

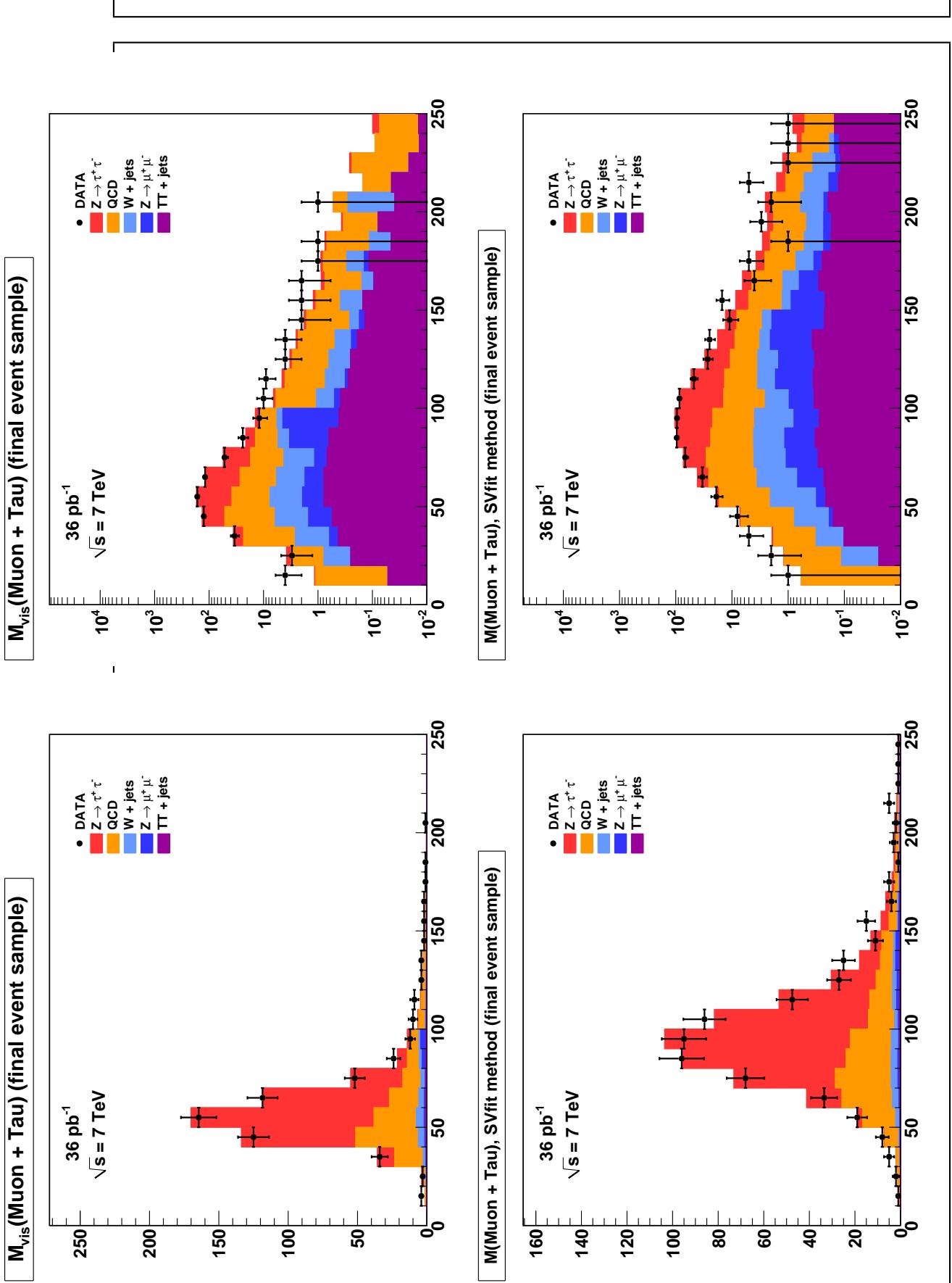


Figure 9.3: Distribution of visible (top) and “full”  $\tau^+\tau^-$  invariant mass reconstructed by the SVfit algorithm (bottom) in  $H \rightarrow \tau^+\tau^- \rightarrow \mu + \tau_{had}$  candidate events passing the selection criteria described in Chapter 5. The distributions are shown in linear (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 boson signal events 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 [63].

resultsLooseAHToMuTau)?

Secondary Vertex Fit 95% CL Upper Limit (pb)						
Mass (GeV/c <sup>2</sup> )	Expected $\sigma_H \times B_\tau$ (pb)					Observed $\sigma_H \times B_\tau$ (pb)
	-2 $\sigma$	-1 $\sigma$	Median	+1 $\sigma$	+2 $\sigma$	
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)						
	-2 $\sigma$	-1 $\sigma$	Median	+1 $\sigma$	+2 $\sigma$	
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.

(tab-exp-limit-TaNC)

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

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

(gsXSectionCombination)?

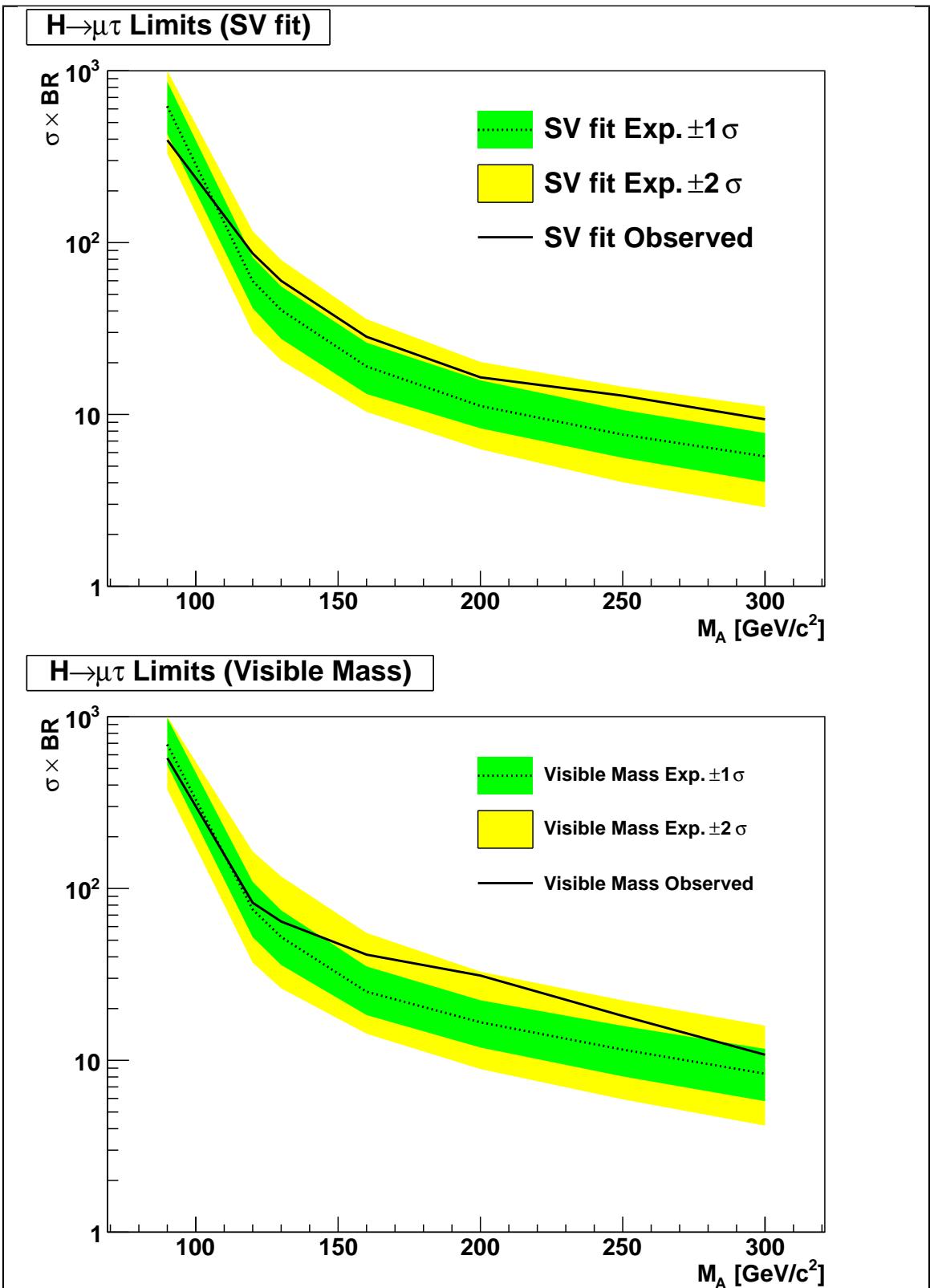


Figure 9.4: Observed and expected limits on the cross section times branching ratio of a Higgs boson versus mass. The top plot gives the limit computed using the SVfit mass as the observable, the bottom plot gives the limit computed using the visible mass. The dashed line gives the nominal expected limit.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

{fig:TanBetaLimit}

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

## 2598 Conclusions

?<ch:conclusions>?  
2599 This analysis has presented a search for MSSM Higgs bosons in the 2010 7 TeV CMS data  
2600 set. Two new experimental methods, the TaNC tau identification algorithm, and the SVfit  
2601 mass reconstruction method have been introduced in this thesis. Both methods increased  
2602 the sensitive of the search. The search was performed using  $36 \text{ pb}^{-1}$  of data. The expected  
2603 event yield from standard model sources is 577 events. In total, 573 events were selected; the  
2604 observed is compatible with the standard model. No signal-like excess of events is observed.  
2605 We set an upper limit on the production of Higgs bosons, and interpret this limit in the  
2606 context of the MSSM.

2607 The analysis presented in this thesis was part of a larger study [1] performed by the  
2608 CMS collaboration searching for the MSSM Higgs boson decaying to tau leptons. The CMS  
2609 analysis used three channels, the  $H \rightarrow \tau\tau \rightarrow e - \tau_h$ ,  $H \rightarrow \tau\tau \rightarrow e - \mu$ , and the  $\mu - \tau_h$   
2610 channel. The  $\mu - \tau_h$  channel search presented in this thesis is very similar to the CMS result.  
2611 While not as pure as the  $\mu - \tau$  channel, the inclusion of the high-statistics  $e - \tau$  channel  
2612 increases the sensitivity of the CMS analysis. The  $e - \mu$  channel has low statistics, but is  
2613 not sensitive to the systematic uncertainty on the hadronic tau identification. The region of  
2614 the MSSM parameter space excluded by combined CMS result is illustrated in Figure 9.3.  
2615 At the time of this writing, the CMS result described in [1] sets the most stringent limits  
2616 on the MSSM using a direct search.

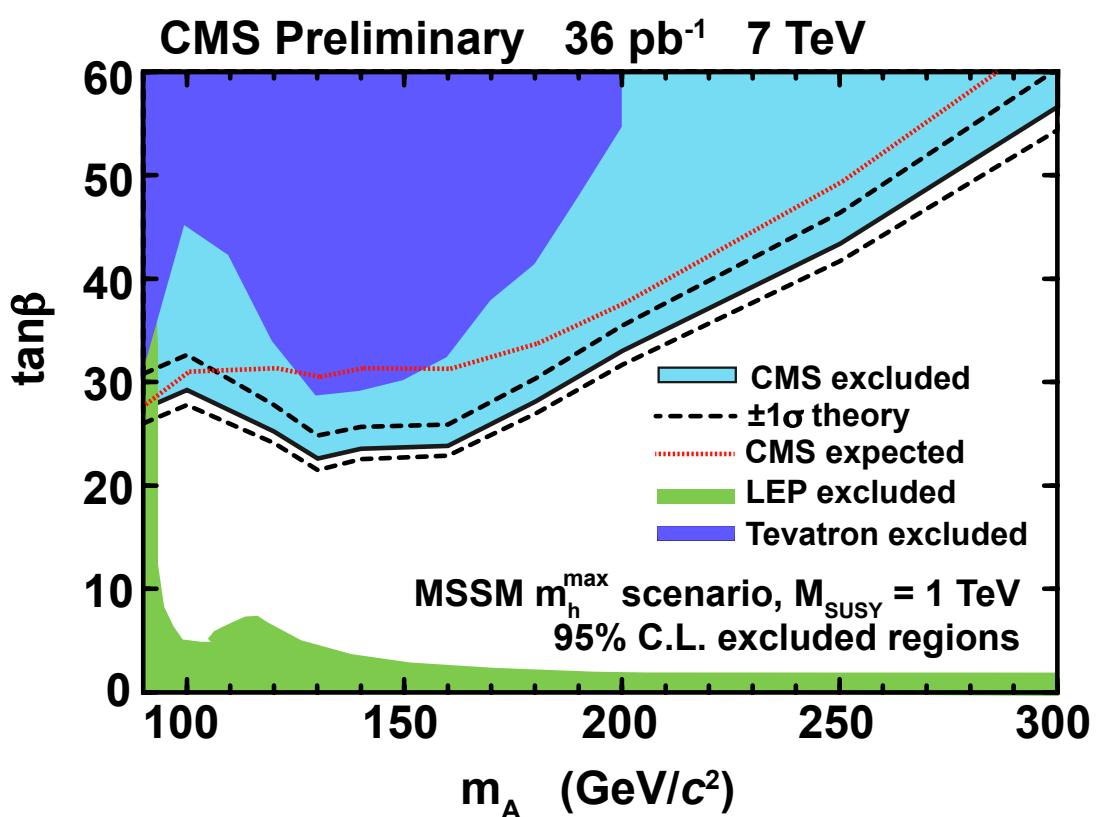


fig:CMSTanBetaExclusion)

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

## Bibliography

- HIG-2617002** [1] CMS Collaboration, “Search For Neutral MSSM Higgs Boson Production via Decays to Tau Pairs in  $pp$  Collisions at  $\sqrt{s} = 7$  TeV”, *to be published* (2011).
- Goldstone:2617eq** [2] J. Goldstone, “Field Theories with Superconductor Solutions”, *Nuovo Cim.* **19** (1961) 154–164. doi:10.1007/BF02812722.
- PhysRev.127.965** [3] J. Goldstone, A. Salam, and S. Weinberg, “Broken Symmetries”, *Phys. Rev.* **127** (Aug, 1962) 965–970. doi:10.1103/PhysRev.127.965.
- PhysRevLett.13.321** [4] F. Englert and R. Brout, “Broken Symmetry and the Mass of Gauge Vector Mesons”, *Phys. Rev. Lett.* **13** (Aug, 1964) 321–323. doi:10.1103/PhysRevLett.13.321.
- PhysRevLett.13.508** [5] P. W. Higgs, “Broken Symmetries and the Masses of Gauge Bosons”, *Phys. Rev. Lett.* **13** (Oct, 1964) 508–509. doi:10.1103/PhysRevLett.13.508.
- PhysRevLett.13.585** [6] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, “Global Conservation Laws and Massless Particles”, *Phys. Rev. Lett.* **13** (Nov, 1964) 585–587. doi:10.1103/PhysRevLett.13.585.
- Glashow:2631tr** [7] S. Glashow, “Partial Symmetries of Weak Interactions”, *Nucl.Phys.* **22** (1961) 579–588. doi:10.1016/0029-5582(61)90469-2.
- Weinberg:2634tr** [8] S. Weinberg, “A Model of Leptons”, *Phys.Rev.Lett.* **19** (1967) 1264–1266. doi:10.1103/PhysRevLett.19.1264.
- Salam:2638rm** [9] A. Salam, “Weak and Electromagnetic Interactions”, Originally printed in \*Svartholm: Elementary Particle Theory, Proceedings Of The Nobel Symposium Held 1968 At Lerum, Sweden\*, Stockholm 1968, 367–377.
- Griffiths:IntroPartikel** [10] D. Griffiths, “Introduction to Elementary Particles”. Wiley-VCH, 2004.
- Fermi:2641tr** [11] E. Fermi, “An attempt of a theory of beta radiation. 1”, *Z. Phys.* **88** (1934) 161–177. doi:10.1007/BF01351864.
- Morii:SMandBSM** [12] S. M. T. Morii, C.S. Lim, “The Physics of the Standard Model and Beyond”. World Scientific, 2004.
- UA1WDDiscovery** [13] UA1 Collaboration, “Experimental observation of isolated large transverse energy electrons with associated missing energy at  $\sqrt{s} = 540$  GeV”, *Phys. Lett.* **B122** (1983) 103–116.
- UA2WDDiscovery** [14] UA2 Collaboration, “Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN  $\bar{p}p$  collider”, *Phys. Lett.* **B122** (1983) 476–485. doi:10.1016/0370-2693(83)91605-2.

- UA1ZDisc<sub>2660y</sub>**  
2651  
2652
- UA2ZDisc<sub>2663y</sub>**  
2654
- Martin:1995asn**  
2656  
2657  
2658  
2659
- FeynmanDiag<sub>2660s</sub>**
- Rainwater:1998kj**  
2662  
2663
- PDG**  
2665
- MSTWXSection<sub>2661s</sub>**
- LHCIG<sub>HiggsXSecG<sub>2664p</sub></sub>**  
2668
- MHMaxBenchmark<sub>2669k</sub>**  
2670  
2671  
2672
- CMSExperiment**  
2674
- CMS-PAS-TRK-10<sub>26694</sub>**  
2676  
2677
- CMS-PAS-TRK-10<sub>26705</sub>**  
2679
- CMS-PT<sub>26771</sub>**  
2681
- CMS-PAS-PFT-08<sub>26801</sub>**  
2683
- CMS-PAS-PFT-09<sub>26801</sub>**  
2685
- TMVA<sub>2686A</sub>**  
2687
- [15] UA1 Collaboration, “Experimental observation of lepton pairs of invariant mass around  $95 \text{ GeV}/c^2$  at the CERN SPS collider”, *Phys. Lett.* **B126** (1983) 398–410. doi:10.1016/0370-2693(83)90188-0.
- [16] UA2 Collaboration, “Evidence for  $Z^0 \rightarrow e^+e^-$  at the CERN  $\bar{p}p$  collider”, *Phys. Lett.* **B129** (1983) 130–140. doi:10.1016/0370-2693(83)90744-X.
- [17] S. P. Martin, “A Supersymmetry Primer”, *arXiv hep-ph* (sep, 1997). 128 pages. Version 5 (December 2008) contains a change in convention that flips the signs of sigma and sigmabar matrices. It also contains a total of about 2 pages of updates, mostly on supersymmetry breaking issues. Errata and a version with larger type (12 pt, 142 pages) can be found at <http://zippy.physics.niu.edu/primer.html>.
- [18] CERN Computing Newsletter.
- [19] D. L. Rainwater, D. Zeppenfeld, and K. Hagiwara, “Searching for  $H \rightarrow \tau^+\tau^-$  in weak boson fusion at the LHC”, *Phys. Rev.* **D59** (1999) 014037, arXiv:hep-ph/9808468. doi:10.1103/PhysRevD.59.014037.
- [20] Particle Data Group Collaboration, “Review of particle physics”, *J. Phys.* **G37** (2010) 075021. doi:10.1088/0954-3899/37/7A/075021.
- [21]
- [22] LHC Higgs Cross Section Working Group Collaboration, “Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables”, arXiv:1101.0593.
- [23] M. Carena, S. Heinemeyer, C. Wagner et al., “MSSM Higgs boson searches at the Tevatron and the LHC: Impact of different benchmark scenarios”, *The European Physical Journal C - Particles and Fields* **45** (2006) 797–814. 10.1140/epjc/s2005-02470-y.
- [24] CMS Collaboration, “The CMS experiment at the CERN LHC”, *JINST* **3** (2008) S08004.
- [25] CMS Collaboration, “Measurement of Momentum Scale and Resolution using Low-mass Resonances and Cosmic-Ray Muons”, *CMS PAS CMS-PAS-TRK-10-004* (2010).
- [26] CMS Collaboration, “Tracking and Primary Vertex Results in First 7 TeV Collisions”, *CMS PAS CMS-PAS-TRK-10-005* (2010).
- [27] G. L. Bayatian et al., “CMS Physics Technical Design Report Volume I: Detector Performance and Software”. Technical Design Report CMS. CERN, Geneva, 2006.
- [28] CMS Collaboration, “CMS Strategies for tau reconstruction and identification using particle-flow techniques”, *CMS PAS CMS-PAS-PFT-08-001* (2008).
- [29] CMS Collaboration, “Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET”, *CMS PAS CMS-PAS-PFT-09-001* (2009).
- [30] A. Hoecker et al., “TMVA - Toolkit for Multivariate Data Analysis”, *ArXiv Physics e-prints* (March, 2007) arXiv:arXiv:physics/0703039.

- [Kolmogorov] [31] A. Kolmogorov, “On the representation of continuous functions of several variables by superposition of continuous functions of one variable and addition”, *Doklady Akademii Nauk SSSR* **114** (1957).
- [DavisBau] [32] J. C. et al., “Size of signal cones and isolation rings in the CMS tau identification algorithm”, *CMS Note* **2008/026** (2008).
- [CMS AN'2010] [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* **2010/082** (2010).
- [CMS-PAS-PFT-10-004] [34] CMS Collaboration, “Study of tau reconstruction algorithms using pp collisions data collected at  $\sqrt{s} = 7$  TeV”, *CMS PAS CMS-PAS-PFT-10-004* (2010).
- [Antikt] [35] G. P. S. M. Cacciari and G. Soyez, “The anti-kt jet clustering algorithm”, *JHEP* **04** (2008) 063, arXiv:0802.1189.
- [CMS-PAS-PFT-10-002] [36] CMS Collaboration, “Commissioning of the Particle–Flow reconstruction in Minimum–Bias and Jet Events from pp Collisions at 7 TeV”, *CMS PAS PFT-10-002* (2010).
- [pythia6.4] [37] S. M. T. Sjöstrand and P. Skands, “PYTHIA 6.4 Physics and Manual”, 2000.
- [tauola] [38] S. Jadach, Z. Was, R. Decker et al., “The Tau Decay Library Tauola: Version 2.4”, *Comput. Phys. Commun.* **76** (1993) 361.
- [g4ant] [39] S. Agostinelli, J. Allison, K. Amako et al., “G4—a simulation toolkit”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** (2003), no. 3, 250 – 303. doi:10.1016/S0168-9002(03)01368-8.
- [CDFMSSMHiggs] [40] CDF Collaboration, “Search for MSSM Higgs decaying to  $\tau$  pairs in  $p\bar{p}$  collision at  $\sqrt{s} = 1.96$  TeV at CDF”, *Phys. Rev. Lett.* **96** (2006).
- [CMS-PTDRII] [41] CMS Collaboration, “CMS technical design report, volume II: Physics performance”, *J. Phys.* **G34** (2007) 995–1579. doi:10.1088/0954-3899/34/6/S01.
- [improvedCollinearApprox] [42] L. Bianchini, “Improved Collinear Approximation for VBF  $H \rightarrow \tau\tau \rightarrow 3\nu + \ell + \tau_{had}$ ”, *CMS Note* **2010/226** (2010).
- [tauDecayPolarization] [43] B. K. Bullock, K. Hagiwara, and A. D. Martin, “Tau Polarization And Its Correlations As A Probe Of New Physics”, *Nucl. Phys.* **B 395** (1993) 499.
- [CMS-PAS-EWK-10-002] [44] CMS Collaboration, “Measurements of Inclusive W and Z Cross Sections in pp Collisions at  $\sqrt{s} = 7$  TeV”, *CMS PAS EWK-10-002* (2010).
- [CMS-PAS-JME-10-005] [45] CMS Collaboration, “MET Performance in Events Containing Electroweak Bosons from pp Collisions at  $\sqrt{s} = 7$  TeV”, *CMS PAS JME-10-005* (2010).
- [CDFrefPzeta] [46] C. C. Almenar, “Search for the neutral MSSM Higgs bosons in the  $\tau\tau$  decay channels at CDF Run II”. PhD thesis, Departament de Física Atomica, Molecular i Nuclear (Universitat de València and IFIC (CSIC - Universitat de València), 2008.

- CMS AN 2010/088** [47] L. Lusito and C. Veelken, “Estimation of Background contributions to Tau analyses via Template Fitting”, *CMS Note* **2010/088** (2010).
- CMS-PAS-TAU-11-001** [48] CMS Collaboration, “Performance of tau reconstruction algorithms in 2010 data collected with CMS”, *CMS PAS TAU-11-001* (2011).
- CDFFakerateD2tag** [49] D. Jang, “Search for MSSM Higgs decaying to  $\tau$  pairs in  $p\bar{p}$  collision at  $\sqrt{s} = 1.96$  TeV at CDF”, *Ph.D. Thesis, Rutgers University* (2006).
- CMS AN 2011/021** [50] J. Conway, E. Friis, and C. Veelken, “Measurement of the  $Z/\gamma^* \rightarrow \tau^+\tau^-$  Production Cross-section in the  $\mu + \tau_{had}$  final state using the HPS+TaNC Tau id. algorithm”, *CMS Note* **2011/021** (2011).
- CMS AN 2010/099** [51] J. Conway, E. Friis, M. Squires et al., “The Tau Neural Classifier algorithm: tau identification and decay mode reconstruction using neural networks”, *CMS Note* **2010/099** (2010).
- MCEmbedding** [52] T. Früboes and M. Zeise, “The TauAnalysis/MCEmbeddingTools Package”. <https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideTauAnalysisMCEmbeddingTools>.
- CMS-PAS-EWK-10-013** [53] CMS Collaboration, “Measurement of the Inclusive  $Z \rightarrow \tau\tau$  Cross Section in  $pp$  Collisions at  $\sqrt{s} = 7$  TeV”, *to be published* (2011).
- CMS AN 2010/059** [54] S. Bolognesi, M. A. Borgia, R. Castello et al., “Calibration of track momentum using dimuon resonances in CMS”, *CMS Note* **2010/059** (2010).
- CMS-PAS-JME-10-010** [55] CMS Collaboration, “Jet Energy Corrections determination at  $\sqrt{s} = 7$  TeV”, *CMS PAS JME-10-010* (2010).
- CMS AN 2010/382** [56] G. Bauer et al., “Modeling of  $W \rightarrow \ell\nu$  MET with Boson Recoil”, *CMS Note* **2010/382** (2010).
- CMS AN 2010/460** [57] G. Cerati et al., “Search for MSSM neutral Higgs  $\rightarrow \tau^+\tau^-$  Production using the TaNC Tau id. algorithm”, *CMS Note* **2010/460** (2010).
- LHCfII** [58] CMS Collaboration, “Absolute luminosity normalization”, *CMS Detector Performance Summary CMS-DPS-2011-002* (2008).
- EWK/pdfUncertaintyTools** [59] J. Alcaraz. <https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideEWKUtilities>.
- CTEQpdfSet** [60] P. M. Nadolsky et al., “Implications of CTEQ global analysis for collider observables”, *Phys. Rev. D* **78** (2008) 013004, arXiv:0802.0007.
- pdfAccSys01** [61] PDF4LHC Working Group. <http://www.hep.ucl.ac.uk/pdf4lhc/PDF4LHCrecos.pdf>.
- pdfAccSys02** [62] R. C. G. D. Bourilkov and M. R. Whalley, “LHAPDF: PDF use from the Tevatron to the LHC”, arXiv:0605.0240.
- LHC-HGWSWG** [63] LHC Higgs Cross Section Working Group Collaboration *arXiv:hep-ex/1101.0593* (2011).