

# **Search for Neutral 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



---

## **3 Abstract**

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

---

<sup>17</sup> **Acknowledgments**

<sup>18</sup>

Hooray for everybody.

---

# <sup>19</sup> Table of Contents

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

56	3.6 Electron and Muon Rejection . . . . .	79
57	<b>4 Mass Reconstruction: The Secondary Vertex Fit</b>	<b>81</b>
58	4.1 Existing mass reconstruction algorithms . . . . .	81
59	4.2 The Secondary Vertex fit . . . . .	83
60	4.3 Parametrization of tau decays . . . . .	84
61	4.4 Likelihood for tau decay . . . . .	85
62	4.4.1 Likelihood for reconstructed missing transverse momentum . . . . .	85
63	4.4.2 Likelihood for tau lepton transverse momentum balance . . . . .	87
64	4.4.3 Secondary vertex information . . . . .	88
65	4.5 Performance . . . . .	90
66	<b>5 Analysis Selections</b>	<b>94</b>
67	5.1 High Level Trigger . . . . .	94
68	5.2 Particle Identification . . . . .	95
69	5.2.1 Muons . . . . .	95
70	5.2.2 Hadronic Taus . . . . .	96
71	5.2.3 Missing Transverse Energy . . . . .	96
72	5.3 Event Selections . . . . .	97
73	<b>6 Data–Driven Background Estimation</b>	<b>102</b>
74	6.1 Background Enriched Control Regions . . . . .	103
75	6.2 The Fake–rate Method . . . . .	105
76	6.2.1 Parameterization of Fake–rates . . . . .	108
77	6.2.2 Measurement of Fake–rates . . . . .	109
78	6.2.3 Application of Fake–rates . . . . .	110
79	6.2.4 “Simple” weight method . . . . .	111
80	6.2.5 “CDF–type” weights . . . . .	113
81	6.2.6 k–Nearest Neighbor Fake–rate Calculation . . . . .	117
82	6.2.7 Results of Background Estimation . . . . .	118
83	6.3 Template method . . . . .	120
84	<b>7 Monte Carlo Corrections</b>	<b>128</b>
85	7.1 Muon Identification Efficiency . . . . .	128
86	7.2 Hadronic Tau Identification Efficiency . . . . .	131
87	7.3 Muon and Tau Momentum Scale . . . . .	133
88	7.4 Missing Transverse Energy Correction . . . . .	134
89	7.5 Pile-up Event Weighting . . . . .	135
90	<b>8 Systematics and Limit Extraction</b>	<b>137</b>
91	8.1 Signal normalization uncertainties . . . . .	139
92	8.2 Background normalization uncertainties . . . . .	139
93	8.3 Shape uncertainties . . . . .	140
94	8.4 Theory uncertainties . . . . .	141
95	8.5 Limit Extraction Method . . . . .	141
96	<b>9 Results</b>	<b>148</b>
97	9.1 Selected Events . . . . .	148
98	9.2 Limits on Higgs Production . . . . .	149

99	9.3 Interpretation in the MSSM . . . . .	150
100	<b>Conclusions</b>	<b>159</b>
101	<b>Bibliography</b>	<b>159</b>

---

## 102 List of Figures

103	1.1	Fermi contact interaction diagram . . . . .	7
104	1.2	Muon decaying through intermediate gauge boson . . . . .	7
105	1.3	QCD Feynman Diagrams . . . . .	17
106	1.4	Loop corrections to Higgs mass . . . . .	19
107	1.5	Higgstrahlung production diagram at $e^+e^-$ colliders . . . . .	24
108	1.6	Gluon fusion Higgs production diagram . . . . .	24
109	1.7	Vector boson fusion Higgs production diagram . . . . .	24
110	1.8	Parton luminosity comparison of the LHC and Tevatron . . . . .	25
111	1.9	SM Higgs cross sections at the LHC . . . . .	26
112	1.10	SM Higgs branching fractions . . . . .	27
113	1.11	Cross sections of interest at hadron colliders . . . . .	28
114	1.12	MSSM Higgs production with association $b$ -quarks . . . . .	29
115	1.13	MSSM Higgs cross sections at the LHC . . . . .	29
116	1.14	LEP Standard Model Higgs limit plot . . . . .	31
117	1.15	Tevatron low mass Standard Model Higgs limit plot . . . . .	32
118	1.16	Tevatron high mass Standard Model Higgs limit plot . . . . .	33
119	1.17	LEP MSSM exclusion limits . . . . .	34
120	1.18	Tevatron MSSM exclusion limits . . . . .	34
121	2.1	Schematic drawings of the CMS detector . . . . .	39
122	2.2	Material budget of the CMS tracker . . . . .	42
123	2.3	Momentum and impact parameter resolutions of CMS tracker . . . . .	43
124	2.4	Energy resolution of the CMS ECAL . . . . .	45
125	2.5	Muon system material budget and identification efficiency . . . . .	47
126	3.1	Visible invariant mass of $\tau$ lepton decay products . . . . .	54
127	3.2	Invariant mass photon pairs in reconstructed $\pi^0$ mesons . . . . .	56
128	3.3	Neutral energy fraction in visible $\tau$ decays . . . . .	57
129	3.4	Tau decay mode reconstruction performance . . . . .	59
130	3.5	Kinematic dependence of decay mode reconstruction . . . . .	60
131	3.6	Neural network over-training validation plots . . . . .	63
132	3.7	Kinematic weighting of training sample . . . . .	65
133	3.8	Neural network output in each decay mode . . . . .	68
134	3.9	Performance curves for the neural networks used in the TaNC . . . . .	69
135	3.10	Tau Neural Classifier performance curves for different $p_T$ ranges . . . . .	70
136	3.11	Tau Neural Classifier transformation performance . . . . .	72
137	3.12	Transformed neural network output . . . . .	73
138	3.13	Tau Neural Classifier performance comparison . . . . .	74
139	3.14	Tau Neural Classifier kinematic performance . . . . .	75

140	3.15 Invariant mass distribution of PF photon pairs . . . . .	77
141	4.1 Coordinate system of the SVfit . . . . .	86
142	4.2 Effect of $p_T$ -balance term on SVfit performance . . . . .	89
143	4.3 Effect of the visible $p_T$ requirements on muon and hadronic $\tau$ decays . . . . .	90
144	4.4 Effect of the visible $p_T$ requirements for $Z$ and Higgs events . . . . .	91
145	4.5 Comparison of SVfit with the Collinear Approximation algorithm . . . . .	92
146	4.6 Comparison of SVfit with the visible mass observable . . . . .	93
147	5.1 Distributions of $M_T$ and muon isolation discriminants . . . . .	99
148	5.2 Reconstruction and distribution of $P_\zeta$ discriminant . . . . .	100
149	6.1 Visible mass distribution of the backgrounds in the signal and control regions	106
150	6.2 SVfit mass distribution of the backgrounds in the signal and control regions	107
151	6.3 $p_T$ and $\eta$ dependency of tau ID performance . . . . .	108
152	6.4 Comparison of fake-rate contribution from genuine taus in the simple and CDF methods . . . . .	116
154	6.5 Muon transverse momentum in the Fake-rate method . . . . .	122
155	6.6 Tau-jet transverse momentum in the Fake-rate method . . . . .	123
156	6.7 Visible mass in the Fake-rate method . . . . .	124
157	6.8 $k$ -Nearest Neighbor classifier example . . . . .	125
158	6.9 Comparison of visible mass and SVfit mass . . . . .	125
159	6.10 Comparison of background shapes in the signal and control regions . . . . .	126
160	6.11 Visible mass distribution in the final fit of the Template Method . . . . .	127
161	7.1 Tag-probe muon isolation method . . . . .	130
162	7.2 Muon isolation correction factors . . . . .	132
163	7.3 Measurement of hadronic tau identification efficiency . . . . .	133
164	7.4 $Z$ -recoil $E_T^{\text{miss}}$ resolution correction . . . . .	135
165	7.5 Distribution of number of reconstructed primary vertices per event . . . . .	136
166	9.1 Transverse momentum distributions of muon and tau in the final selected events . . . . .	152
167	9.2 Distributions of the $\eta$ and $\phi$ of the muon and tau candidates in the final selected events . . . . .	153
168	9.3 Distributions of final selected events . . . . .	154
170	9.4 Observed and expected limits on Higgs $\sigma \times \text{BR}$ . . . . .	157
171	9.5 Excluded regions of MSSM $\tan\beta - m_{A^0}$ parameter space . . . . .	158
172	9.6 CMS combined exclusion of MSSM $\tan\beta - m_{A^0}$ parameter space . . . . .	160

---

## 174 List of Tables

175	1.1	Chiral supermultiplets in the MSSM . . . . .	21
176	1.2	Gauge supermultiplets in the MSSM . . . . .	21
177	1.3	Higgs search channels at LEP . . . . .	30
178	1.4	Decay modes of the $\tau$ lepton . . . . .	35
179	3.1	Decay mode performance – naive reconstruction . . . . .	58
180	3.2	Decay mode performance – TaNC reconstruction . . . . .	58
181	3.3	Neural network training event statistics . . . . .	62
182	3.4	Variables used in the different TaNC neural networks . . . . .	80
183	5.1	High Level Trigger paths used to select $\mu + \tau_h$ events . . . . .	95
184	5.2	Analysis backgrounds that include fake taus . . . . .	97
185	5.3	Event selection summary . . . . .	101
186	6.1	Criteria used to select background enriched control regions . . . . .	104
187	6.2	Comparison of background control region yields in data and the prediction from simulation . . . . .	105
188	6.3	Fake-rate “simple” method closure test results . . . . .	112
189	6.4	Fake-rate “CDF” method closure test results . . . . .	115
190	6.5	Fake-rate method results . . . . .	119
191	6.6	Fake-rate Method predicted yields in like-sign control region . . . . .	120
192	6.7	Background yields measured using the Template Method . . . . .	121
193	7.1	Muon trigger, identification, and isolation correction factors . . . . .	131
194	8.1	Effect of normalization uncertainties on signal efficiency times acceptance .	142
195	8.2	Correspondence of confidence levels and $2\Delta \ln \mathcal{L}$ intervals . . . . .	144
196	8.3	Summary of systematic uncertainties represented by nuisance parameters in the likelihood, their representation method and magnitudes. . . . .	147
197	9.1	Final analysis yield and background expectations . . . . .	149
198	9.2	Expected signal yields at $\tan \beta = 30$ . . . . .	155
199	9.3	Expected and observed 95% CL $\sigma \times \text{BR}$ upper limits . . . . .	156
200	9.4	Contributions of different MSSM Higgs boson types at different $m_{A^0}$ . . . . .	156

---

## 203 Introduction

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

211 Chapter 1 begins with an introduction to the Standard Model (SM) of particle physics.  
 212 Emphasis is given to Electroweak Symmetry Breaking and the Higgs Mechanism, the the-  
 213 oretical phenomena that motivate the presence of a Higgs boson. The theoretical issues  
 214 which motivated the development of the MSSM are discussed, and a brief introduction is  
 215 given. Finally, the phenomenology of Higgs bosons in the SM and MSSM is discussed, with  
 216 an overview of Higgs searches performed at LEP and the Tevatron.

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

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

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

233

## Chapter 1

---

### 234 The Standard Model and Beyond

(ch:theory)

#### 235 §1.1 The Standard Model

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

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

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

### 261 §1.1.1 Quantum Electrodynamics and Gauge Invariance

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

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

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

$$274 p^\mu p_\mu - m^2 c^2 = 0. \quad (1.1) \quad \text{[eq:EnergyPRelat]}$$

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

$$275 p_\mu \rightarrow i\hbar\partial_\mu \quad (1.2) \quad \text{[eq:QuantizeMom]}$$

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

$$276 (\gamma^\kappa p_\kappa + mc)(\gamma^\mu p_\mu - mc) = 0, \quad (1.3) \quad \text{[eq:DiracEquation]}$$

---

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

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

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

- <sup>274</sup> The solutions  $\psi$  of the Dirac equation are called “Dirac spinors,” and represent the quantum  
<sup>275</sup> mechanical state of spin 1/2 particles.

The Lagrangian corresponding to the Dirac equation (1.4) is

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

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

$$\psi' \rightarrow e^{i\theta}\psi. \quad (1.6) \quad \text{[eq:U1GaugeTran]}$$

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

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

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

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

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

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

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

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

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

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

### §1.1.2 The Weak Interactions

*(sec:WeakInteractions)* The theory of Weak Interactions was created to describe the physics of radioactive decay.

The first formulation of the theory was done by Fermi [11] to explain the phenomenon of the  $\beta$  decay of neutrons. The initial theory was a four-fermion “contact” theory. In a contact theory, all four fermions come involved in the  $\beta$ -decay are connected at a single vertex. The Fermi theory Hamiltonian for the  $\beta$ -decay of a proton is then [12]

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

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

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

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

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

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

The Fermi interaction can describe both nuclear  $\beta$  decay ( $p \rightarrow n + e^+ + \bar{\nu}_e$ ) as well as the decay of a muon into an electron ( $\mu \rightarrow \nu_\mu + e + \bar{\nu}_e$ , Figure 1.1). Furthermore, the

**Fixme:** check handedness is correct

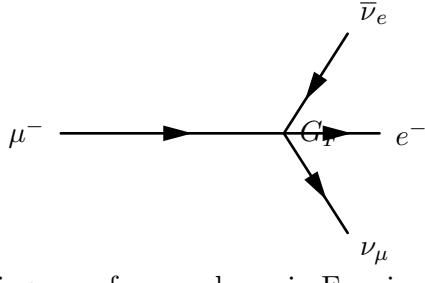


Figure 1.1: Feynmann diagram of muon decay in Fermi contact interaction theory.

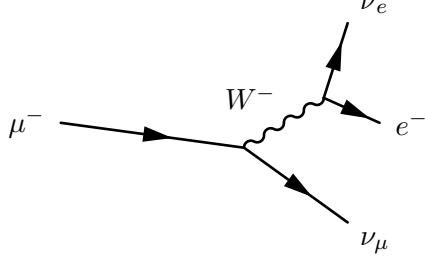


Figure 1.2: Feynmann diagram of muon decay proceeding through an intermediate gauge boson  $W^-$ .

onDecayFeynmanDiagram)

coupling constant  $G$  is found to be a *universal* constant in weak interactions, in that it is the same for interactions regardless of the particle species participating in the interaction. That is,  $G_\mu = G_e = G_F$ . Using an Hamiltonian analogous to Equation 1.8 for muon decay, the decay amplitude  $M$  is found to be

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

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

The first attempt to solve the problems with the Fermi theory was made by introducing an intermediate weak boson [7]. The contact interaction is replaced by a massive propagator, the  $W^\pm$  bosons. The decay of a muon to an electron and two neutrinos then proceeds as pictured in Figure 1.2 with an amplitude given [12] by

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

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

---

<sup>2</sup>Renormalization of quantum field theories is a broad topic beyond the scope of this thesis. Briefly, the process involves "absorbing" infinite divergences that occur in higher-order interactions into physical observables [10].

term of Equation 1.10 is the reason why the contact interaction original formulated by Fermi effectively described low-energy weak phenomenon. When the momentum transfer  $q$  in the interaction is small compared to  $M_W$ , the effect of the propagator is an effective constant. In the low energy limit, the full propagator in Equation 1.10 is equivalent to the Fermi contact interaction in 1.9 as

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

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

### 303 §1.1.3 Spontaneous Symmetry Breaking

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

A symmetry of a Lagrangian is spontaneously broken when the ground state, or vacuum, is at a value about which the Lagrangian is not symmetric. In quantum field theories, a particle is interpreted as quantized fluctuations of its corresponding field about some constant (vacuum) ground state. The “effective” Lagrangian that we observe in the (low energy) laboratory would be the expansion of the Lagrangian about this stable point. The effective Lagrangian no longer obeys the original symmetry, which has been “broken.” We give a brief example of the phenomenological effects of a spontaneously broken symmetry

in a toy model, following the treatment in [12].

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

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

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

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

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

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

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

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

320 If the parameters  $m^2$  and  $\lambda$  are both positive, then the minimum of  $V$  is at the origin ( $\phi_1 =$   
 321  $\phi_2 = 0$ ). In this case, the mass matrix term in Equation 1.13 takes the form  $\left( \frac{\partial^2 V}{\partial \phi_a \partial \phi_b} \right)_0 =$   
 322  $\frac{m^2}{2} \delta_{ab}$ , where  $\delta_{ab}$  is the Kronecker delta function. Therefore the mass matrix is already  
 323 diagonalized, and the  $\phi_1$  and  $\phi_2$  both correspond to particles with mass  $m$ . If the  $m^2$

---

<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  $U(1)$ .

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

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

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

If the vacuum state is defined at the point in Equation 1.16, the symmetry is unbroken and the mass spectra is unchanged. However, the system is unstable at this point, as it is a local maximum. The true global minimum is defined as the set of points which satisfy Equation 1.17, which form a continuous circle in  $\phi_1 - \phi_2$  space (and is therefore infinitely degenerate). We can choose any point on the circle as the vacuum expectation value (VEV).

If the point  $(\phi_1^{min} = \nu, \phi_2^{min} = 0)$ <sup>4</sup> is chosen, evaluating Equation 1.14 yields the mass matrix

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

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

### 331 §1.1.4 The Higgs Mechanism

(sec:HiggsMech)  
 332 As in Section 1.1.1, extending the gauge symmetry requirement to be *locally* invariant  
 333 creates interesting consequences for models that have spontaneously broken symmetry. This  
 334 gives rise to the “Higgs Mechanism,” which we overview here. For simplicity we will again  
 335 consider a model with  $U(1)$  symmetry. The model is identical to the one presented in  
 336 Section 1.1.3, with two exceptions. First, we express the two real fields  $\phi_1$  and  $\phi_2$  as a single  
 337 complex-valued field  $\phi$ . Second, the model is required to be locally  $U(1)$  invariant, and so  
 338 uses the gauge-covariant derivatives, minimal coupling to the gauge field, and contains the

---

<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.

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

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

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

340 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

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

$$\phi \rightarrow \phi' = e^{-i\alpha(x)}\phi$$

$$A_\mu \rightarrow A'_\mu = A_\mu - \frac{1}{2}\partial_\mu\alpha(x).$$

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

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

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

343 “unitary gauge”<sup>5</sup> and have the special forms

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

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

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

346 Lagrangian at the minimum of  $V$  is

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

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

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

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

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

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

352 quiring a mass, the gauge boson  $B_\mu$  has acquired the degree of freedom (as it can now

353 be longitudinally polarized) previously associated to the second degree of freedom in the

---

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

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

### 357 §1.1.5 Electroweak Unification

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

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

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

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

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

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

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

where the  $\tau^i$  are the Pauli matrices. The weak currents  $J_\mu^\pm$  can be written by combining

Equations 1.23–1.25

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

376 Since  $\tau^1$ ,  $\tau^2$ , and  $\tau^3$  are the generators of the  $SU(2)$  group, we can complete the group

377 by adding a neutral current to the charged currents of Equation 1.26. The  $\tau^3$  generator is

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

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

379 Naively one might hope that the neutral current of Equation 1.27 would correspond to the

380 electromagnetic (photon) current of QED. However, this is impossible for two reasons. First,

381 the right-handed component  $e_R$  does not appear in the current, so this interaction violates

382 parity, a known symmetry of the electromagnetic interactions. Second, the current couples to

383 neutrinos, which have no electric charge. Therefore, the “charge” corresponding to the  $SU(2)$

384 gauge symmetry generators  $T^i = \int J_0^i(x)d^3x$  cannot be that of the QED, and the gauge

385 group must be enlarged to include an additional  $U(1)$  symmetry. The generator of the new

386 symmetry must commute with the generators of the  $SU(2)_L$  group. The symmetry cannot

387 be directly extended with  $U(1)_{em}$  as the electromagnetic charge  $Q = \int(e_L^\dagger e_L + e_R^\dagger e_R)d^3x$

388 does not commute with  $T^i$ . The solution is to introduce the “weak hypercharge”  $\frac{Y}{2} = Q - T^3$ ,

389 which commutes the generators of  $SU(2)_L$ . Thus the symmetry group of the electroweak

390 model is  $SU(2)_L \times U(1)_Y$ .

<sup>391</sup> The  $SU(2)_L \times U(1)_Y$  gauge invariant Lagrangian is written

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

<sup>392</sup> As  $R$  is a singlet in  $SU(2)$ , it does not couple to the  $SU(2)$  gauge bosons  $A_\mu^i$ . For this  
<sup>393</sup> Lagrangian to correspond to empirical observations at low energy, the  $SU(2)_L \times U(1)_Y$   
<sup>394</sup> must be broken. As  $U(1)_{em}$  symmetry is observed to be good symmetry at all scales the  
<sup>395</sup> broken Lagrangian must be invariant under  $U(1)_{em}$ .

<sup>396</sup> To accomplish the symmetry breaking, we introduce a new  $SU(2)$  doublet of complex  
<sup>397</sup> Higgs fields  $\phi$  that have hypercharge  $Y = 1$ , and contribute  $\mathcal{L}_S$  to the Lagrangian:

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

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

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

(1.28) [eq:YukawaTerms]

<sup>398</sup> to the Lagrangian which couple the fermions ( $L$  and  $R$ ) to the Higgs field. After symmetry  
<sup>399</sup> breaking these terms will allow the fermions to acquire masses. By choosing the  $m^2$  and  $\lambda$   
<sup>400</sup> parameters of  $V$  appropriately, the new  $\phi$  field acquires a non-zero VEV and the symmetry  
<sup>401</sup> is spontaneously broken.

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

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

The new symmetry of the model can be confirmed by looking at the action of the different  
symmetry generators on the VEV. If the generator acting on the vacuum state has a non-  
zero value, then the corresponding symmetry is broken. It can then be seen that the original  
symmetry generators  $T^+$ ,  $T^-$ ,  $T^3$ , and  $Y$  are all broken. The vacuum *is* invariant under  $Q$ ,

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

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

402 so the broken Lagrangian contains the correct symmetry properties.

The gauge boson content of the electroweak interaction is obtained by parameterizing the Higgs field in the magnitude–phase notation of Equation 1.20 and using the unitary gauge (see Section 1.1.4), where the gauge transformation is chosen so Higgs field is real.

The Higgs scalar doublet is then

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

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

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

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

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

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

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

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

404 term in Equation 1.29 can be written in matrix form and then diagonalized into mass

405 eigenstates

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

giving a massive  $Z$  boson with

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

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

by

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

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

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

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

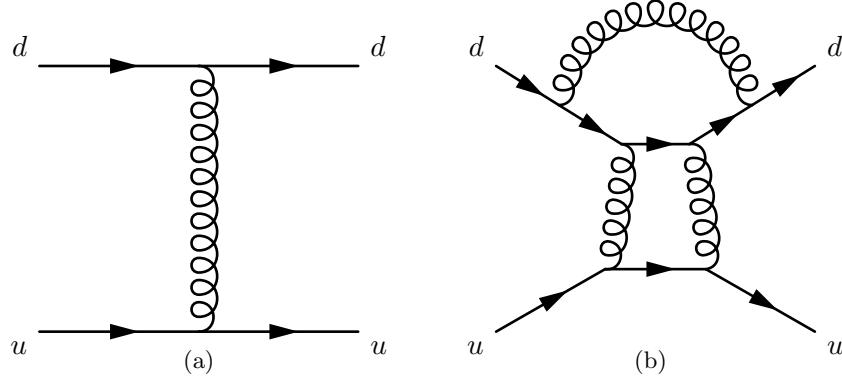
410 The discovery of the  $W$  [13, 14] and  $Z$  [15, 16] at the CERN SPS was a huge triumph for  
 411 the electroweak model.

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

#### 418 §1.1.6 Quantum Chromodynamics

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

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



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.

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

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

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

## 458 §1.2 Beyond the Standard Model

?⟨sec:BSM⟩?<sup>459</sup> The Standard model is one of the most successful theories of the natural world ever created.  
 460 The predictions of the SM have been tested to many orders of magnitude and no experiment  
 461 to date<sup>6</sup> has found a result statistically incompatible with the Standard Model. However,  
 462 there is a general consensus in the physics community that the Standard Model is not  
 463 complete. It is believed that it is only an effective theory that is valid below some energy  
 464 scale  $\Lambda$ . Above this energy, there must exist some other “new physics,” which unifies the  
 465 forces of the Standard Model and correctly describes the natural world at all scales, while  
 466 maintaining equivalence to the Standard Model at low energy. This concept is analogous to  
 467 the relationship between the effective Fermi contact theory of Section 1.1.2 and the unified  
 468 electroweak theory of Section 1.1.5. The size of the cutoff scale  $\Lambda$  is estimated [12] to be  
 469  $\mathcal{O}(10^{15})$  GeV for a unified theory with  $SU(5)$  symmetry and even larger,  $\mathcal{O}(10^{19})$  GeV =  
 470  $M_{\text{planck}}$  if the theory is unified with gravity.

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

---

<sup>6</sup>The Standard Model predicts that lepton number is a good quantum number and that the neutrinos are massless. It has recently been found that the neutrinos do have non-zero mass, and that they undergo oscillations between different neutrino species, violating lepton number.

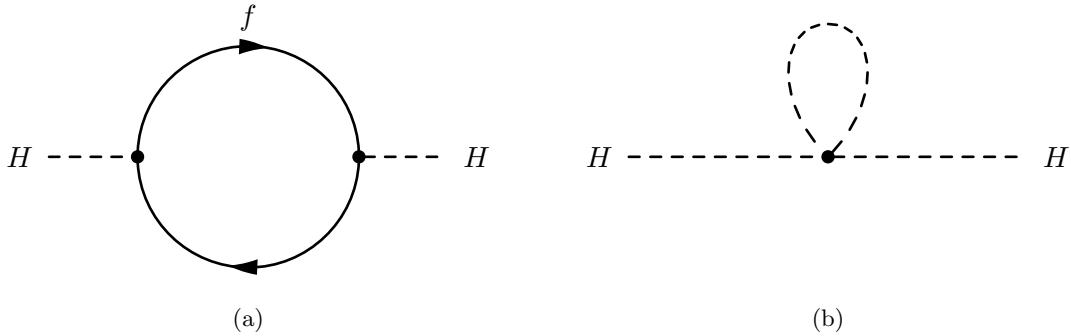


Figure 1.4: Feynman diagram of fermion (a) and scalar (b) loop corrections to Higgs mass.

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

#### 481 §1.2.1 The Hierarchy Problem

The enormous size of the cutoff scale  $\Lambda$  in the Standard Model causes a major theoretical problem in the Standard Model. During renormalization of the Standard Model, amplitudes with divergent integrals are cut off at  $\Lambda$ . These large constant terms are “absorbed” into the physical observables. The cutoff term appears directly in quantum corrections to the Higgs mass [17]. The Yukawa term  $-\lambda_f H \bar{f} f$  coupling the fermion  $f$  to the Higgs  $H$  produces loop corrections to Higgs mass. The two types of corrections due to fermion loops and scalar loops are illustrated in Figure 1.4. The contribution [17] of the loop correction in Figure 1.4(a) to the Higgs mass is

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

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

490  $g$  of the intermediate weak boson theory was compatible with QED. The most promising  
 491 solution to the hierarchy problem is the introduction of a new, “super” symmetry.

### 492 §1.2.2 Supersymmetry

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

### 507 §1.2.3 The Minimal Supersymmetric Model

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

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

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

$\langle \text{tab:chiral} \rangle$

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

Table 1.2: Gauge supermultiplets in the Minimal Supersymmetric Standard Model.

Source: [17]

$\langle \text{tab:gauge} \rangle$

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

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

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

523 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  
 524 Yukawa  $3 \times 3$  matrices which act on the different families. It is important to note that the  
 525 up-type quarks couple to the up-type Higgs  $H_u$ , while the down-type quarks and leptons

<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.

526 couple to the down-type Higgs. This feature has large phenomenological consequences,  
 527 which are discussed in Section 1.3.2. The scalar portion of the  $W_{\text{MSSM}}$  potential defines the  
 528 spontaneous symmetry breaking. Similar to the scalar potential  $V$  symmetry breaking of  
 529 Section 1.1.3, the potential of  $V$  at the minimum is found<sup>8</sup> to be

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

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

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

The ratio of the VEVs is expressed as

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

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

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

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

---

<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.

of each family are

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

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

### 541 §1.3 Searches for the Higgs boson

?<sec:PreviousSearches>? The discovery of the Higgs boson is one of the biggest prizes in science today. Dozens of  
 542 experiments, thousands of scientists and billions of dollars (a human hierarchy problem)  
 543 have been spent in efforts to discovery the Higgs. In this section we discuss how the Higgs  
 544 and the MSSM could appear in modern colliders (with an emphasis on the LHC) and  
 545 the current limits placed on the Higgs by the Large Electron–Positron Collider (LEP) and  
 546 Tevatron experiments.  
 547

#### 548 §1.3.1 Standard Model Higgs boson phenomenology

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

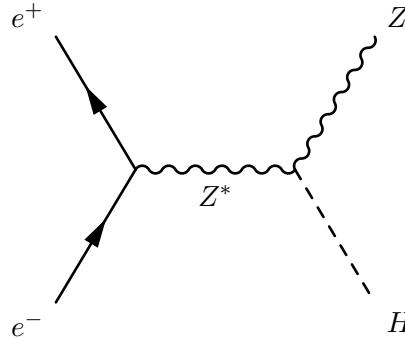
$$\mathcal{L}_e = -\frac{G_e}{\sqrt{2}}(\nu + H(x))\bar{e}e = -\frac{G_e\nu}{\sqrt{2}}\bar{e}e - \frac{G_e}{\sqrt{2}}H(x)\bar{e}e. \quad (1.34)$$

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

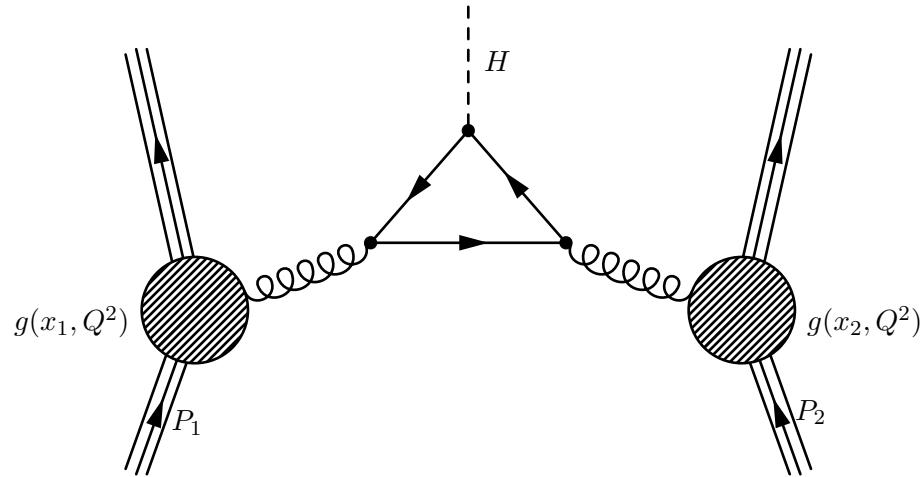
$$\frac{G_e\nu}{\sqrt{2}} = \frac{m_e}{\nu}. \quad (1.35)$$

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

553 The dominant modes of Higgs boson production depend on the type of experiment.  
 554 In general, Higgs production is favored through high-mass intermediate states, due to the

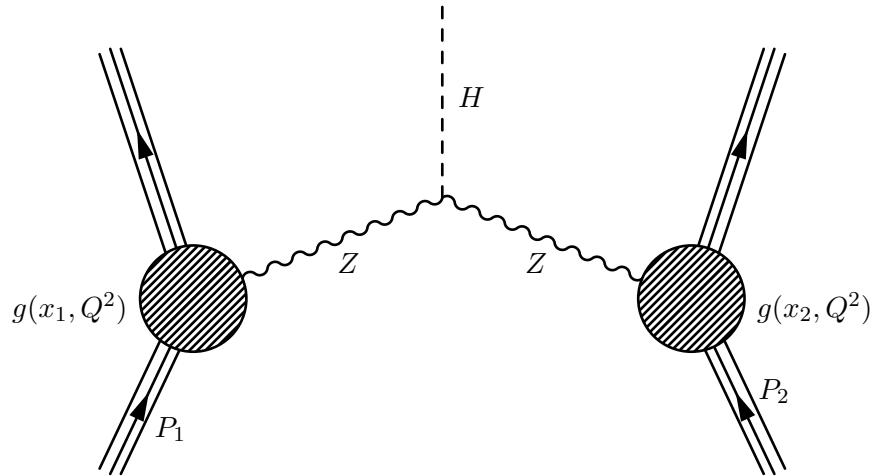


{fig:HiggsStrahlung}

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

{fig:GluonFusion}

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



{fig:VBFProdDiagram}

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

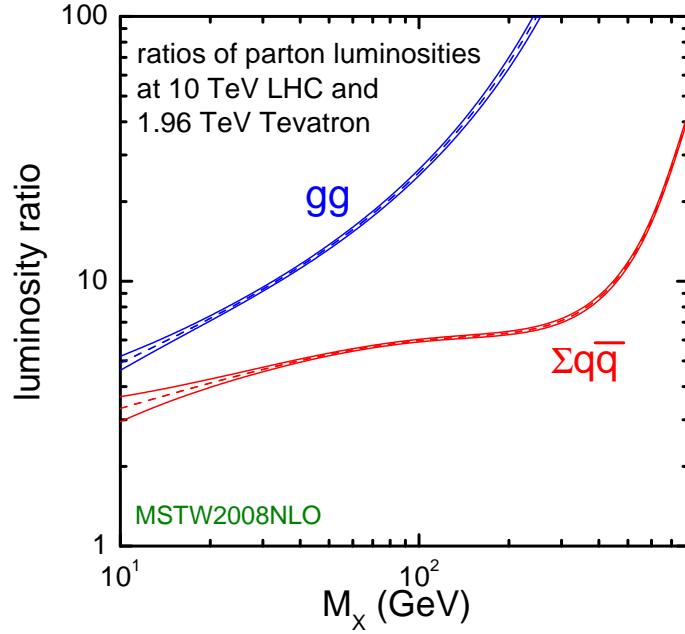


Figure 1.8: Ratio of the parton luminosity (the amount of luminosity contributed by the different species that compose the proton) of the LHC (at  $\sqrt{s} = 10$  TeV) and the Tevatron. The large increase in gluon–gluon luminosity affects the favored production mechanisms of the Higgs boson.

(fig:GluonLumiRatio)

mass<sup>2</sup> proportional coupling. At the Tevatron and LEP experiments, which will be introduced in the next section, the dominant SM Higgs production mode is “Higgstrahlung,” where a virtual  $W^\pm$  or  $Z$  gauge boson is produced and then radiates a Higgs boson. Higgstrahlung is illustrated in Figure 1.3.1. At the Large Hadron Collider, higher gluon luminosities (see Figure 1.8) result in the favored cross section being “gluon fusion,” (illustrated in Figure 1.3.1) where two gluons from the incident protons combine in a quark (dominated by the massive top quark) loop which then radiates a Higgs boson. Another important channel [19] is “vector boson fusion,” (Figure 1.3.1) where weak gauge bosons ( $W^\pm$  or  $Z$ ) are radiated from the incoming quarks and fuse to produce a Higgs. This is a notable channel due to the lack of “color-flow” (gluons) between the two protons, producing an event with low central jet activity and two “tag-jets” in the forward and backward regions. The theoretical cross sections for the SM Higgs at the LHC are shown for the various production mechanisms in Figure 1.9.

The branching fractions of the different decay modes of the SM Higgs boson depend strongly on the mass of the Higgs boson. In general, the Higgs prefers (due to the Yukawa

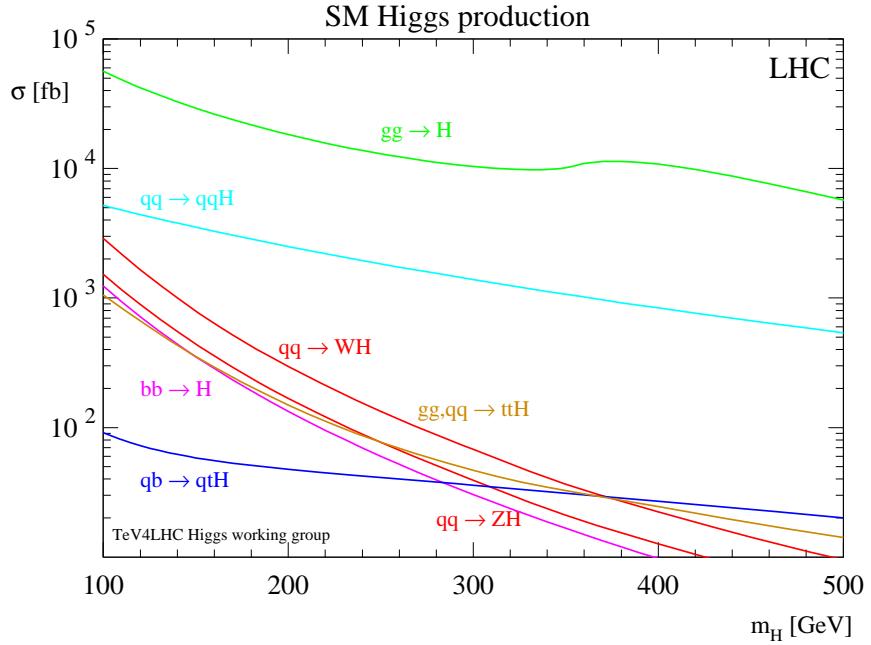


Figure 1.9: Cross section of the Standard Model Higgs boson versus the Higgs boson mass. The different curves give the contribution to the cross section from different production mechanisms. Source: [20].

(fig:LHCSMHiggsXsec)

570    couplings) to decay pairs of the particles with the highest mass possible. Below the threshold  
 571    to decay to pairs of weak bosons ( $M_H < 160 \text{ GeV}/c^2$ ), the Higgs decays predominantly  
 572    to either  $b$ -quarks ( $b\bar{b}$ , 90%) or a pair of  $\tau$  leptons ( $\tau^+\tau^-$ ,  $\approx 10\%$ ). Above the  $W^\pm W^\mp$   
 573    threshold, decays to vector bosons ( $H \rightarrow W^\pm W^\mp$  and  $H \rightarrow ZZ$ ) dominate. The dependence  
 574    of branching fraction on  $M_H$  and the other rare decay modes are illustrated in Figure 1.10.  
 575    For low mass Higgs, the  $\tau^+\tau^-$  decay mode plays a particularly important role. The dominant  
 576    decay mode  $H \rightarrow b\bar{b}$  suffers from enormous backgrounds from QCD jet production. It  
 577    is important to understand the magnitude of difference between expected Higgs boson  
 578    production and the rates of various backgrounds. Figure 1.11 illustrates the cross sections  
 579    for different SM processes at hadron colliders. The rate of Higgs production is many orders  
 580    of magnitude ( $\mathcal{O}(10^{-7})$ ) smaller than that of QCD production. It is important to therefore  
 581    design searches to use handles that can reject the vast majority of the uninteresting events  
 582    at hadron colliders.

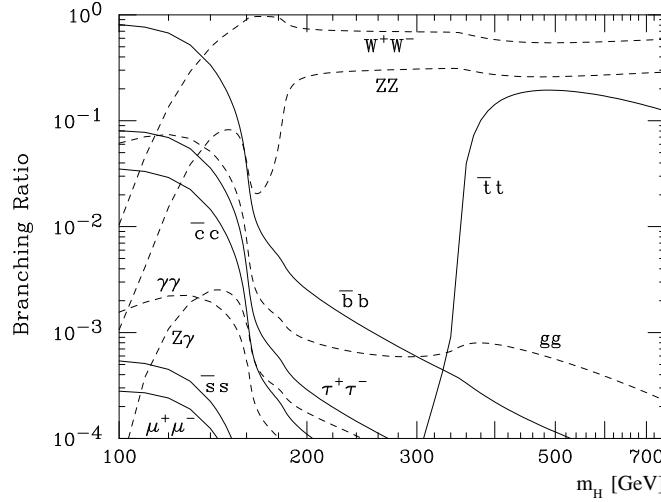


Figure 1.10: Branching fraction of the Standard Model Higgs bosons for different values of  $M_H$ . Source: [20].

`(fig:SMHiggsBR)`

### 583 §1.3.2 MSSM Higgs Phenomenology

`(sec:MSSMHiggsPhenom)` The phenomenology of the Higgs sector of the MSSM is similar to the Standard Model in some respects, but differs in some key aspects which have important implications for final states involving  $\tau$  leptons and  $b$  quarks. When the parameter  $\tan\beta$  is large, the coupling factor between the Higgs and the down-type quarks and leptons (effectively the  $\tau$  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  $m_{A^0}$  and  $\tan\beta$ . For a reasonably high  $\tan\beta$ , there is always one CP-even Higgs ( $h^0$  or  $H^0$ ) which is mass-degenerate with the  $A^0$ . When  $\tan\beta$  and  $m_{A^0}$  are both large, associated  $b$  production dominates the total cross section [22]. The cross sections of the different MSSM neutral Higgs bosons are shown in Figure 1.13. The  $\tan\beta$  enhancement of the MSSM Higgs coupling to the  $b$  quarks and  $\tau$  leptons causes the branching fraction of all neutral MSSM Higgs states to be  $H \rightarrow b\bar{b}$  (90%) and  $H \rightarrow \tau^+\tau^-$  (10%) across the entire range of  $m_{A^0}$ . The enhanced production rate and the high branching fraction to  $\tau$  leptons make the MSSM Higgs decaying to  $\tau$  leptons an exciting and promising channel to search for Higgses and Supersymmetric physics at colliders.

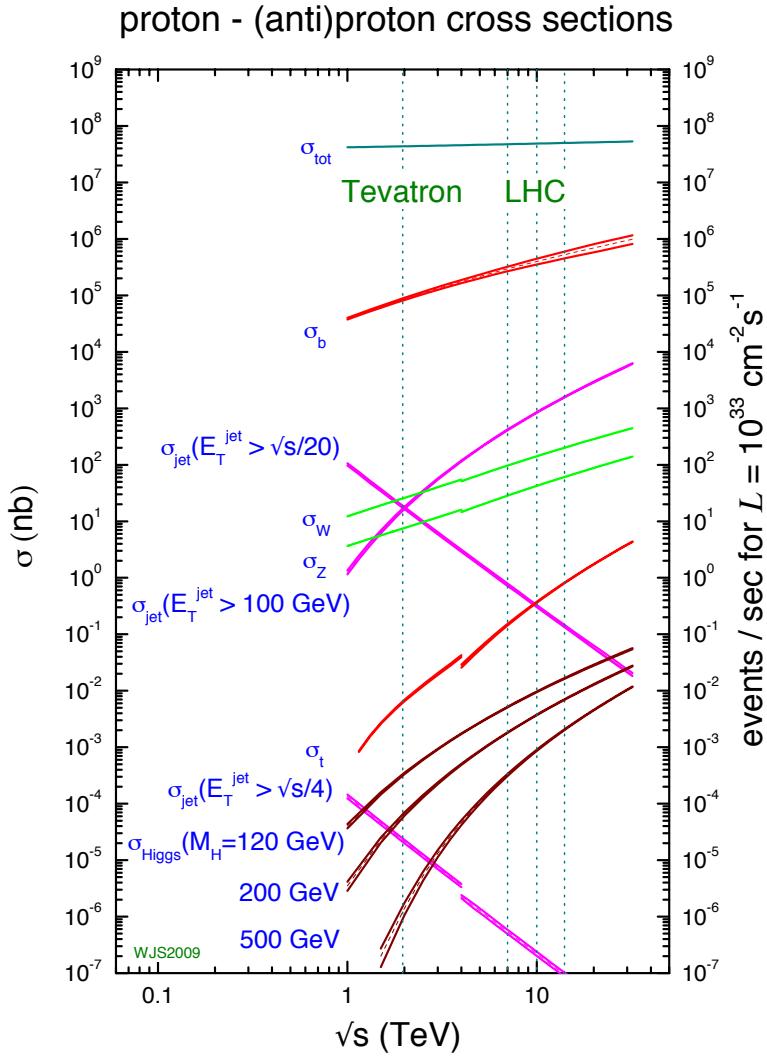
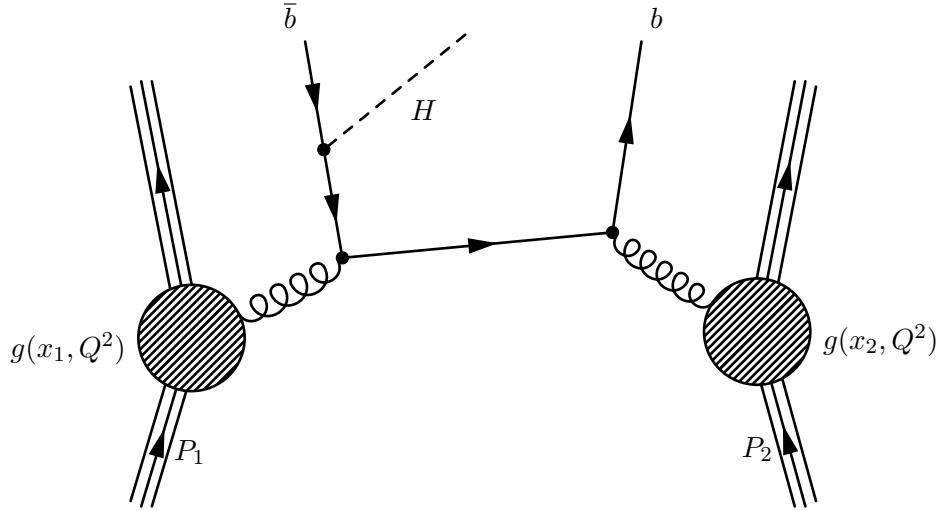


Figure 1.11: Cross sections of various processes at hadron colliders. The horizontal axis represents the center of mass energy of the collision. Of note is the vast difference in scales between Higgs production (maroon lines,  $\mathcal{O}(10^{-2} \text{ nb})$ ) and the QCD cross section to produce 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

hadronColliderCrossSections



g:AssociatedBProduction)

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

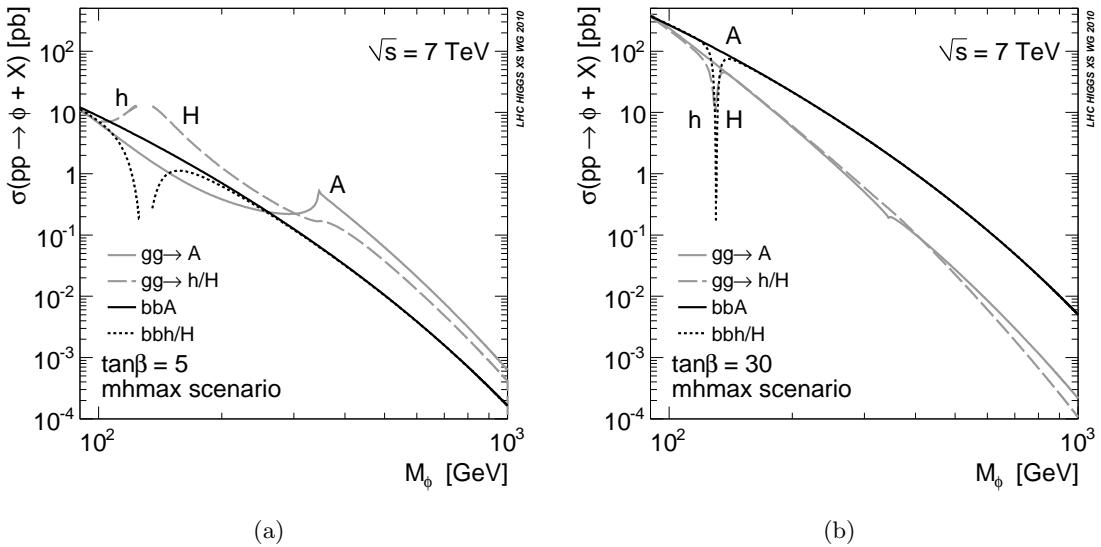


Figure 1.13: Cross sections for the different MSSM Higgs bosons versus  $m_{A^0}$  in the  $m_{h^{max}}$  benchmark scenario [23] scenario for  $\tan\beta = 5$  (a) and  $\tan\beta = 30$  (b). Source: [22]

MSSMXSectionsTanBeta)

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

(tab:LEPModes)

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

### 601    §1.3.3 Results from LEP and Tevatron

(sec:lepAndTevatron)  
 602    The LEP and Tevatron experiments have both set limits on the existence of the Standard  
 603    Model and MSSM Higgs boson. Additionally, precision electroweak measurements give ad-  
 604    ditional hints on the prospects for both models.

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

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

615    The Tevatron is a proton–antiproton collider with a center-of-mass energy of  $\sqrt{s} = 1.96$  TeV ■  
 616    There are two general purpose detectors at the Tevatron, CDF and DØ. The dominant Higgs  
 617    production modes at the Tevatron are Higgstrahlung and gluon fusion (see Figure 1.3.1).  
 618    For low mass ( $m_H < 135$  GeV/c<sup>2</sup>) Higgs bosons the dominant channel at the Tevatron is the  
 619    Higgstrahlung production mode and  $H \rightarrow b\bar{b}$  decays. Large multi-jet backgrounds prevent

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

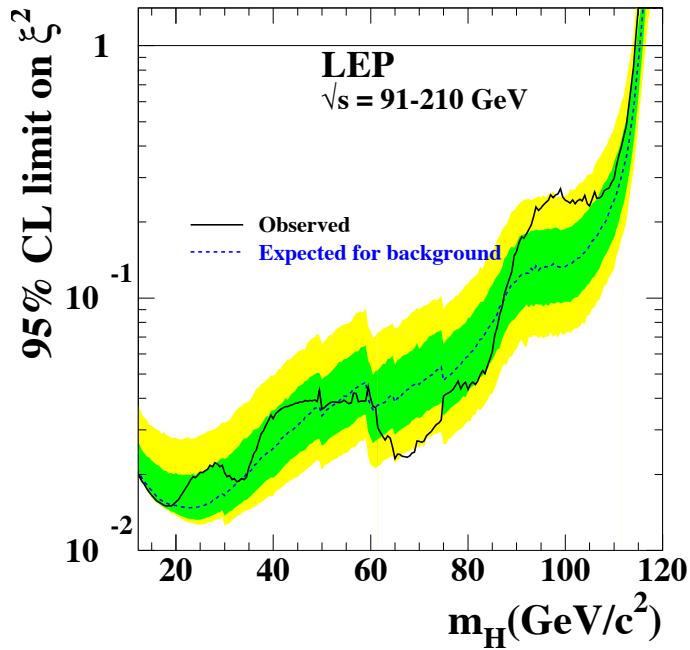


Figure 1.14: Combined LEP upper limit set on the quantity  $\xi^2 = (g_{HZZ}/g_{HZZ})^2$  at 95% confidence level. Regions where the observed ratio is less than one exclude the Standard Model. The dashed line gives the expected limit for the null (background only) hypothesis, with the green and yellow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

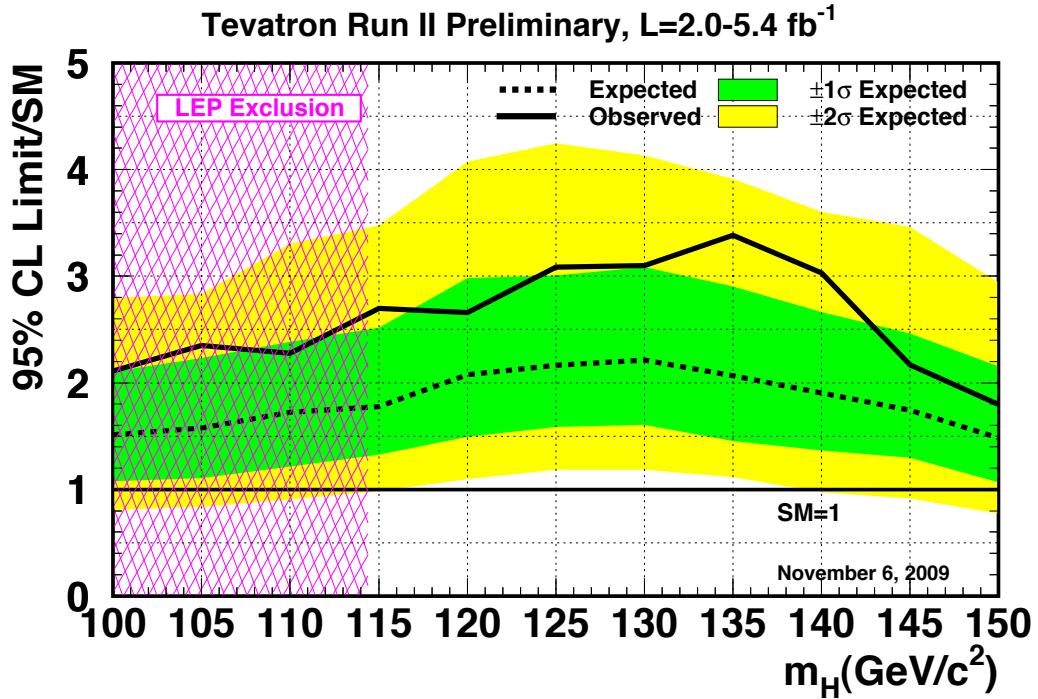


Figure 1.15: Combined CDF and DØ RunII upper limit on the cross section of a Standard Model-like Higgs boson. The LEP limit is shown in pink. Reference: [20]

atronLowMassHiggsLimit)

the  $H \rightarrow b\bar{b}$  decay mode from being useful for searching for Higgs bosons produced by gluon fusion. The  $H \rightarrow \tau^+\tau^-$  and  $H \rightarrow \gamma\gamma$  decays are additionally used in an inclusive search at low mass, but do not dominate the search sensitivity. The combined low-mass limit on the Standard Model Higgs from both Tevatron experiments is shown in Figure 1.15. The Tevatron currently sets an upper limit on the SM Higgs cross section of about 2.5 times the Standard Model expectation.

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

Analyses at LEP and Tevatron have also addressed excluded regions of the MSSM. At LEP, the dominant production modes of the MSSM Higgs bosons are Higgstrahlung and

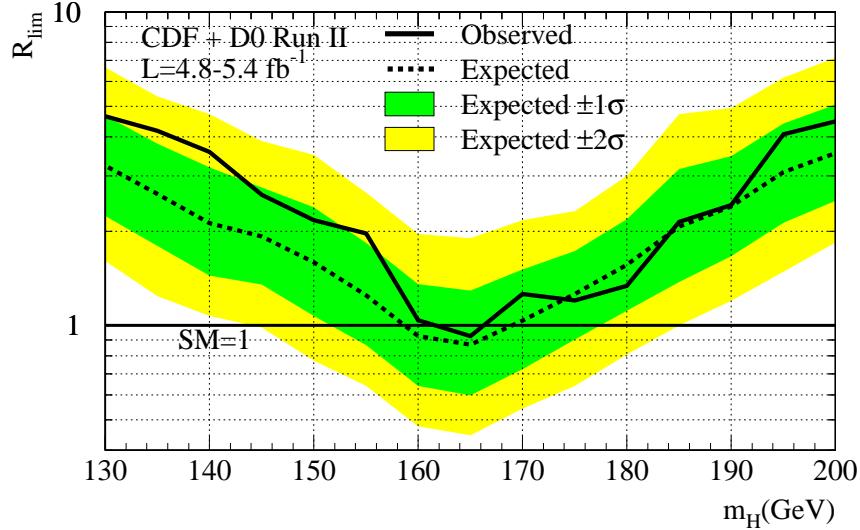


Figure 1.16: Combined CDF and DØ RunII upper limit on the cross section of a Standard Model-like Higgs boson using the  $H \rightarrow W^+W^-$  decay mode. The Standard Model is excluded for Higgs boson masses between 162 and 166  $\text{GeV}/c^2$ . Reference: [20]

tronHighMassHiggsLimit)

634 pair production, where  $e^+e^- \rightarrow h^0A^0$  or  $H^0A^0$ . For the Higgstrahlung production mode,  
 635 the Standard Model search can be reinterpreted in terms of the MSSM. To address the pair  
 636 production mode, searches were performed in the  $e^+e^- \rightarrow h^0A^0 \rightarrow b\bar{b}b\bar{b}$  and  $\tau^+\tau^-q\bar{q}$  decay  
 637 modes. Finally, LEP is also sensitive to associated MSSM Higgs production at low  $m_{A^0}$  and  
 638 high  $\tan\beta$  to  $e^+e^- \rightarrow \{\bar{\phi}\}$ , where the associated fermions  $\{\cdot\}$  are  $b$ -quarks or tau leptons.  
 639 The combined limits from LEP in the  $m_{A^0} - \tan\beta$  plane are shown in Figure 1.17.

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

## 646 §1.4 The Physics of the Tau Lepton

647 As discussed in Sections 1.3.1 and 1.2.3, the  $\tau$  lepton is an important probe of Higgs physics.  
 648 The  $\tau$  lepton has some unusual properties which make it particularly challenging at hadron  
 649 colliders. With a mass of 1.78  $\text{GeV}/c^2$ , the  $\tau$  lepton is heaviest of the leptons. The nominal

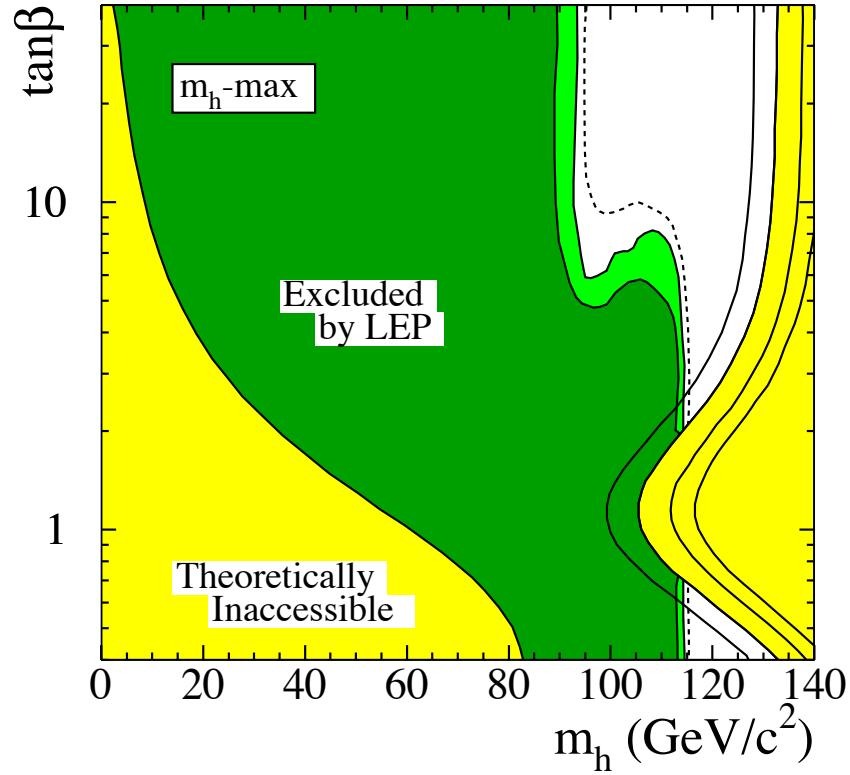


Figure 1.17: Combined LEP limits on the MSSM. The results are interpreted in the context of the  $m_h$ -max benchmark [23] scenario of the MSSM. Reference: [20]

(fig:LEPMSSMLimits)

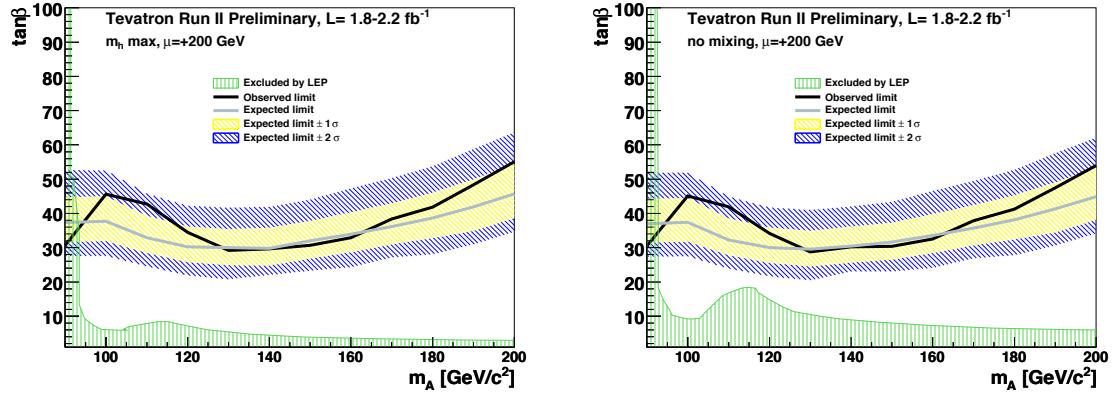


Figure 1.18: Combined Tevatron limits on the MSSM. The grey line and blue and yellow bands give the expected limit and its one and two sigma contours. The black line is the observed limit. The results are interpreted in the context of the  $m_h$ -max benchmark (left) and “no mixing” (right) MSSM scenarios.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

(fig:TevMSSMLimits)

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.

650 decay distance  $c\tau$  of the  $\tau$  lepton is 87  $\mu\text{m}$ , which in practice means that the  $\tau$  will always  
 651 decay before reaching the first layer of the detector. Tau decays can be effectively classified  
 652 into two types. “Leptonic” decays consist of a  $\tau$  decaying to a light lepton ( $\ell = e, \mu$ ) and two  
 653 neutrinos  $\tau^+ \rightarrow \ell^+ \nu_\tau \bar{\nu}_\ell$ . “Hadronic” decays consist of a low-multiplicity collimated group of  
 654 hadrons, typically  $\pi^\pm$  and  $\pi^0$  mesons. The hadronic decays of the  $\tau$  lepton compose approx-  
 655 imately 65% of the  $\tau$  lepton branching fraction, with the remainder shared approximately  
 656 equally by the leptonic decays. The branching fractions for the leptonic and most common  
 657 hadronic decays are shown in Table 1.4.

658 The tau is also a challenging object in that the decay of the tau always includes neu-  
 659 trinos. The associated neutrinos are weakly interacting and do not create a signal in any  
 660 detector at CMS. The only sign that the neutrinos are there is an imbalance in the total  
 661 transverse<sup>11</sup> energy in the event. This thesis will describe a novel way to reconstruct the  
 662 neutrinos associated to tau decays in Chapter 4.

---

<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.

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

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

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

668

## Chapter 2

669

# The Compact Muon Solenoid Experiment

&lt;ch:detector&gt;

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

674

- Good muon identification and momentum resolution over a wide range of momenta and angles, good dimuon mass resolution ( $\approx 1\%$  at 100 GeV/ $c^2$ ), and the ability to determine unambiguously the charge of muons with  $p < 1$  TeV/ $c$ ;

677

- Good charged-particle momentum resolution and reconstruction efficiency in the inner tracker. Efficient triggering and offline tagging of  $\tau$ ’s and  $b$ -jets, requiring pixel detectors close to the interaction region;

680

- Good electromagnetic energy resolution, good diphoton and dielectron mass resolution ( $\approx 1\%$  at 100 GeV/ $c^2$ ), wide geometric coverage,  $\pi^0$  rejection, and efficient photon and lepton isolation at high luminosities;

683

- Good missing-transverse-energy and dijet-mass resolution, requiring hadron calorimeters with a large hermetic geometric coverage and with fine lateral segmentation.

685

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

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

## 703 §2.1 The Large Hadron Collider

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

---

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

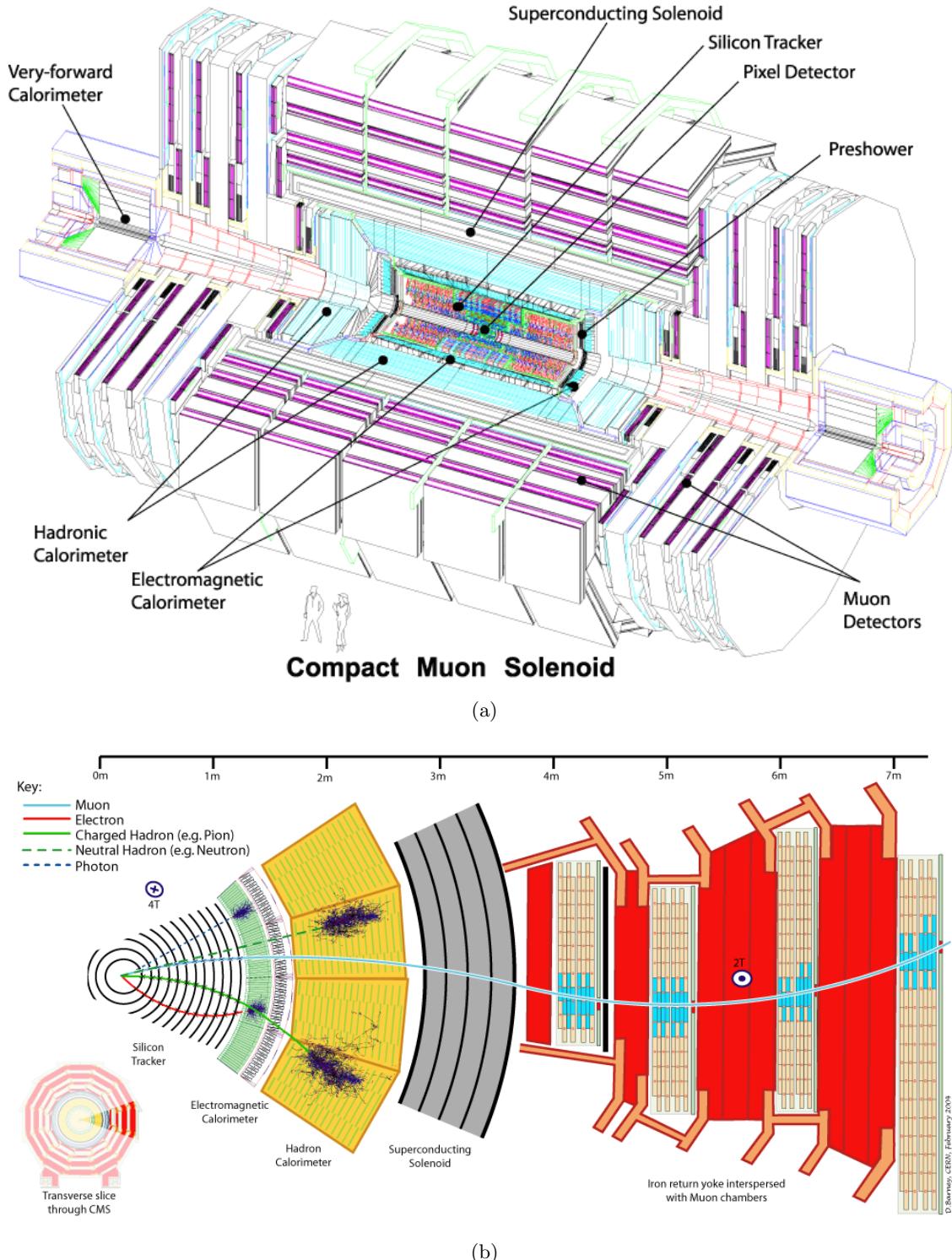


Figure 2.1: Figure (a), top, shows a schematic drawing of the CMS detector. The individual sub-detectors are labeled. Two humans are shown in the foreground for scale. Figure (b) shows a radial cross section of the detector and demonstrates how the (non-)destructiveness of different sub-detectors facilitates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

`(fig:AllCMSCutaways)`

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

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

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

726 The CMS solenoid is extremely large. The radial bore of the magnet is 6.3 meters; the  
 727 magnet is 12.5 meters in length and weighs 220 tons. The large bore of the magnet allows  
 728 the tracker and calorimeter systems to be located inside the solenoid. The internal windings  
 729 of solenoid is arranged in four layers to increase the total field strength and are cooled by  
 730 liquid helium to a temperature of 4.5 Kelvin. The windings are magnetically coupled to  
 731 the support superstructure. This coupling allows the magnetic to heat uniformly during a  
 732 “quench” event<sup>2</sup>, reducing localized stresses. The nominal current at full field of the solenoid  
 733 is 19.14 kA. The solenoid itself is surrounded by an iron return yoke with a total mass of  
 734 10,000 tons. The return yoke surrounding the solenoid minimizes the fringing field. The  
 735 muon detector system is interspersed inside the yoke, and takes advantage of the return  
 736 field in the yoke to measure the momentum and charge of muons.

---

<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.

## 737 §2.3 Charged Particle Tracking Systems

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

751 Both the pixel and strip trackers are silicon based. The principle of operation is similar  
 752 to that of a charged-coupled discharge (CCD) in a modern digital camera. The sensitive  
 753 portion of the detector is a silicon chip that is arranged with diode junctions formed by  
 754 a  $p$ –doped layer and an  $n$ –doped layer<sup>3</sup>. Each  $p - n$  junction is electrically isolated from  
 755 adjacent layers. The size of each junction region determines<sup>4</sup> the spatial resolution of the  
 756 sensor. In the pixel detector, each sensor region “pixel” is  $100 \times 150 \mu\text{m}^2$ . In the strip  
 757 tracker, The rear side of the chip is mounted to read-out electronics. During operation, a  
 758 high-voltage reverse bias is applied to each  $p - n$  junction to achieve full depletion. When  
 759 a charged particle passes through the detector, the diode-junction breaks down and the  
 760 readout system registers the hit.

**Fixme:** right

acronym

761 The tracking system has been specifically designed for the high radiation environment  
 762 around the interaction point. The detector is cooled to  $-27^\circ\text{C}$  during operation to minimize

---

<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  $10^{-4}$ .

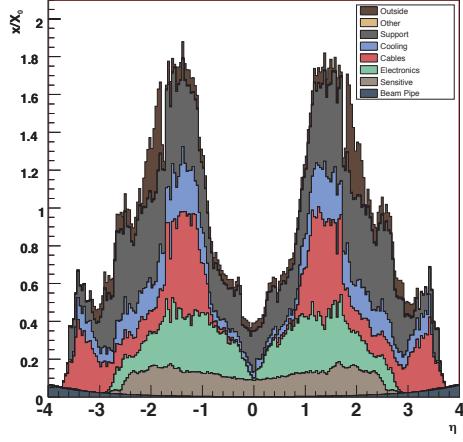


Figure 2.2: Material budget of the CMS tracker in units of radiation lengths  $X_0$  versus pseudorapidity. The material budget is broken down into the contributions from the different components of the tracker. The amount of material is largest in the “transition region” between the barrel and endcap.

Fig:TrackerMaterialBudget

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

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

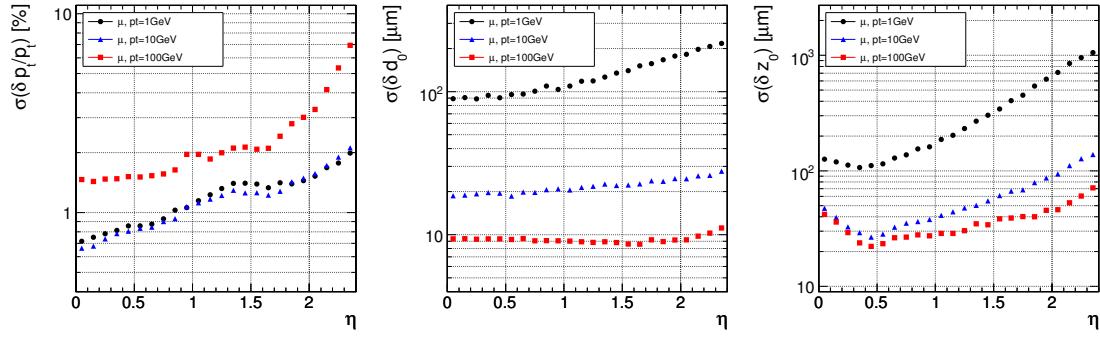


Figure 2.3: Expected resolutions of reconstructed transverse momentum (left), transverse impact parameter (center), and longitudinal impact parameter (right) versus absolute pseudorapidity  $|\eta|$ . The resolution is shown for three different cases of particle  $p_T$ , 1  $\text{GeV}/c$  (black), 10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

(fig:ExpectedTrackerRes)

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* detector, and functions by counting the number of photons produced in an electromagnetic shower inside a crystal. Upon entering the crystal, a charged particle or photon will interact electromagnetically with the crystal, producing a shower of electrons and photons. The

<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.

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

803 The ECAL uses lead tungstate ( $\text{PbWO}_4$ ) crystals as the scintillation medium. The  
 804 crystals have a very large density, which allows the calorimeter to be relatively compact.  
 805 To be able to correctly measure the energy of electrons and photons, an incident photon or  
 806 electron must be completely stopped by interactions with the calorimeter. The quantities  
 807 that determine if an electron or photon will be completely contained is the total depth of  
 808 the crystal, the crystal density, and the radiation length property  $X_0$  of the crystal. The  
 809 radiation length  $X_0$  is defined as the mean distance (normalized to material density) after  
 810 which an electron will have lost  $(1 - \frac{1}{e})$  of its energy. The  $\text{PbWO}_4$  crystals of the CMS  
 811 ECAL have a density of 8.28 g/cm<sup>2</sup> and a depth of 230 mm. A single crystal thus has a  
 812 total radiation length of 25.8  $X_0$ , and will capture on average 99.9993% of the energy of an  
 813 incident electron. The front face of the crystal is 22 mm × 22 mm, which corresponds to an  
 814  $\eta - \phi$  area of  $0.00174 \times 0.00174$ . The Molière radius of a material is the average radial profile  
 815 size of an electromagnetic shower, and for  $\text{PbWO}_4$  is 2.2 cm. The fact that the Molière  
 816 radius is larger than the size of the individual crystals improves the spatial resolution of  
 817 the measurement. As the shower is shared between multiple crystals, the relative amounts  
 818 deposited in each crystal allows the true impact point to be determined with a resolution  
 819 smaller than the individual crystal size.

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

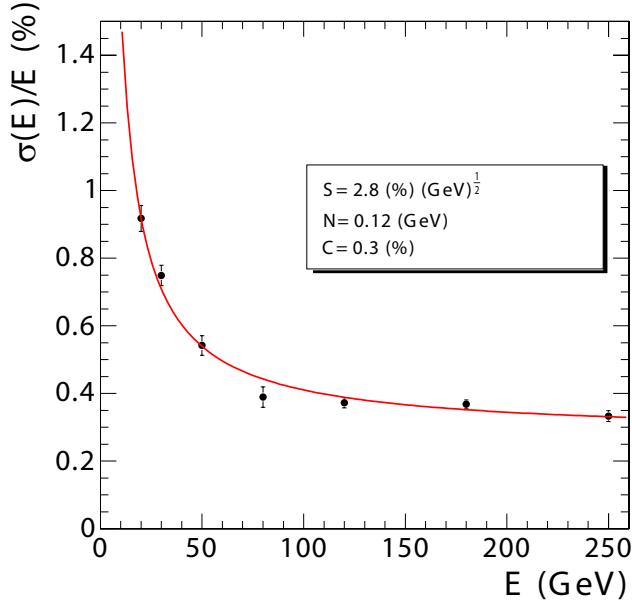


Figure 2.4: Energy resolution (in %) of the CMS ECAL measured at an electron test beam. The resolution depends on the incident energy of the electron. The points are fitted to function with the form given in Equation 2.2. The fitted parameters are given in the legend.

(fig:ECALResolution)

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

The energy resolution of the ECAL is given by

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

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

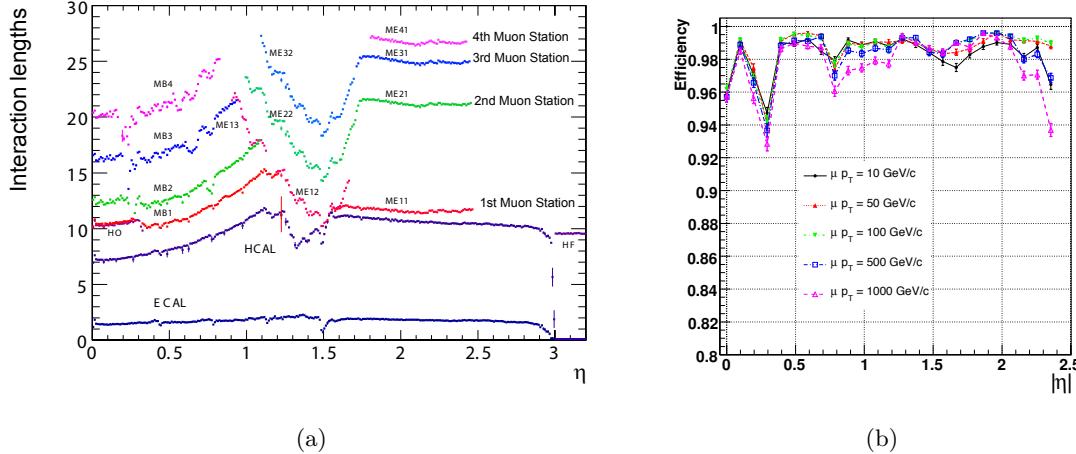
## §2.5 Hadronic Calorimeter

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

839 layer of calorimetry outside of the magnet. The hadronic calorimeter measures the energy  
 840 of charged and neutral hadronic particles. The HCAL is a *sampling* calorimeter. Layers of  
 841 plastic scintillating tiles are interspersed between brass absorber plates. An incident hadron  
 842 produces a hadronic shower as it passes through the absorber. The particles in the shower  
 843 produce light as they pass through the scintillating tiles. Measuring the light produced in  
 844 each layer of tile allows the reconstruction of the radial profile of the shower which can be  
 845 related to the deposited energy. The response of the scintillator tiles are calibrated using a  
 846 radioactive source, either Cs<sup>137</sup> or Co<sup>60</sup>. Small stainless tubes permit the radioactive sources  
 847 to be moved into the center of the tile during calibration. The granularity of the HCAL is  
 848 0.087 × 0.087 and 0.17 × 0.17 in  $\eta - \phi$  in the barrel ( $|\eta| < 1.6$ ) and endcap ( $|\eta| > 1.6$ ),  
 849 respectively.

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

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



detectorChapterMuonShit)?

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.

## 867 §2.6 Muon System

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

871 • Muons are stable particles, for the typical energies and distances considered at a  
 872 collider.

873 • Muons have non-zero charge, so their trajectories can be measured.

874 • Muons are heavy enough that they are “minimum ionizing particles,” in that they  
 875 lose very little energy as they pass through material.

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

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

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

905 Cathode strip chambers (CSCs) are used in the endcap muon system, providing cov-  
 906 erage in the pseudorapidity range  $0.9 < |\eta| < 2.4$ . A cathode strip chamber consists of a  
 907 chamber filled with inert gas that with a number of internal wires held at a high voltage. A  
 908 number of cathode strips are installed perpendicular<sup>6</sup> induced to the wires on the walls of  
 909 the chamber. When a muon passes through the CSC, it ionizes some of the gas. The high  
 910 voltage on a nearby wire causes this ionized gas to break down, forming a conductive pas-

---

<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.

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

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

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

## 926 §2.7 Trigger System

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

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

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

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

The muon component of all the triggers used in this analysis is based on the “L1 seed trigger” L1\_SingleMu7, which nominally selects event which contain a muon with  $p_T > 7$  GeV/c. The L1 muon trigger decision is determined by the Global Muon Trigger (GMT), which combines information from the DT, CSC, and RPC sub-detectors, and is able to trigger muons up to a pseudorapidity of  $|\eta| < 2.1$ . Each sub-detector has a “local trigger,”

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

---

## Chapter 3

---

976

977

978

# Tau Identification: The Tau Neural Classifier

&lt;ch:tanc&gt;

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

989       Tau leptons are unique in that they are the only type of leptons which are heavy enough  
 990 to decay to hadrons. The hadronic decays compose approximately 65% of all tau decays, the  
 991 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  
 992 decays are typically composed of one or three charged pions and zero to two neutral pions.  
 993 The neutral pions decay almost instantaneously to pairs of photons.

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

### 1000 §3.1 Geometric Tau Identification Algorithms

1001 ⟨sec:GeometricTauId⟩ The tau identification strategies used in previously published CMS analyses are fully de-  
 1002 scribed in [28]. A summary of the basic methods and strategies is given here. There are  
 1003 two primary methods for selecting objects used to reconstruct tau leptons. The CaloTau  
 1004 algorithm uses tracks reconstructed by the tracker and clusters of hits in the electromag-  
 1005 netic and hadronic calorimeter. The other method (PFTau) uses objects reconstructed by  
 1006 the CMS particle flow algorithm, which is described in [29]. The particle flow algorithm  
 1007 provides a global and unique description of every particle (charged hadron, photon, elec-  
 1008 tron, etc.) in the event; measurements from sub-detectors are combined according to their  
 1009 measured resolutions to improve energy and angular resolution and reduce double counting.

1010 All of the tau identification strategies described in this thesis use the particle flow objects.

1011 Both methods typically use an “leading object” and an isolation requirement to reject  
 1012 quark and gluon jet background. Quark and gluon jets are less collimated and have a higher  
 1013 constituent multiplicity and softer constituent  $p_T$  spectrum than a hadronic tau decay of  
 1014 the same transverse momentum. The “leading track” requirement is applied by requiring a  
 1015 relatively high momentum object near the center of the jet; typically a charged track with  
 1016 transverse momentum greater than 5 GeV/c within  $\Delta R < 0.1$  about the center of the jet  
 1017 axis. The isolation requirement exploits the collimation of true taus by defining an isolation  
 1018 annulus about the kinematic center of the jet and requiring no detector activity about a  
 1019 threshold in that annulus. This approach yields a misidentification rate of approximately 1%  
 1020 for QCD backgrounds and a hadronic tau identification efficiency of approximately 50% [28].

### 1021 §3.2 Decay Mode Tau Identification: Motivation

1022 The tau identification strategy described previously can be extended by looking at the dif-  
 1023 ferent hadronic decay modes of the tau individually. The dominant hadronic decays of taus  
 1024 consist of a one or three charged  $\pi^\pm$  mesons and up to two  $\pi^0$  mesons and are enumerated  
 1025 in Table 1.4. The majority of these decays proceed through intermediate resonances and  
 1026 each of these decay modes maps directly to a tau final state multiplicity. Each intermediate  
 1027 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.

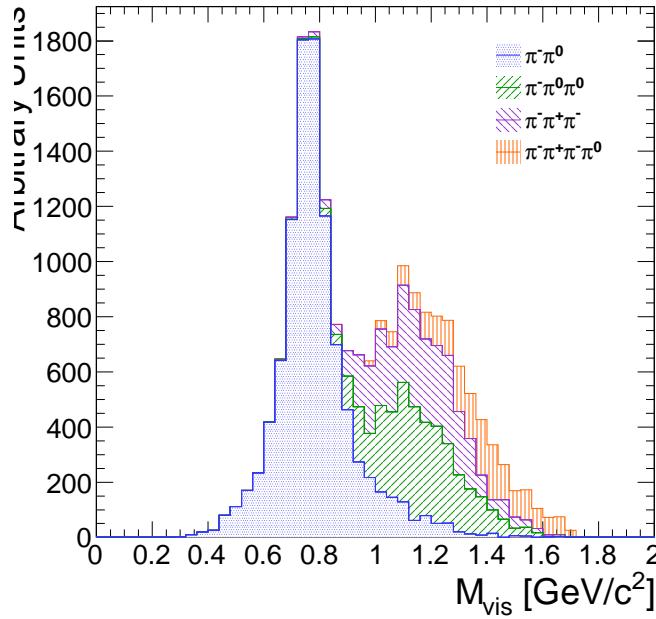


Figure 3.1: The invariant mass of the visible decay products in hadronic tau decays. The decay mode  $\tau^- \rightarrow \pi^- \nu_\tau$  is omitted. The different decay modes have different invariant masses corresponding to the intermediate resonance in the decay.

`(fig:trueInvMass)`

### §3.3 The Tau Neural Classifier

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

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

### §3.3.1 Decay Mode Reconstruction

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

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

#### Photon Merging

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

1070 best pairs are then tagged as  $\pi^0$  mesons, removing lower-ranking candidate  $\pi^0$ s as necessary  
 1071 to ensure that no photon is included in more than one  $\pi^0$  meson.

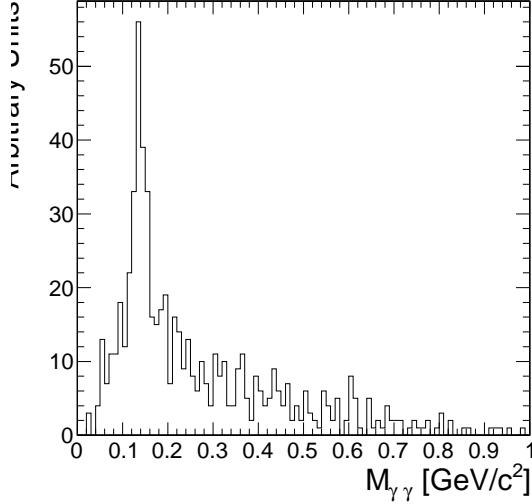


Figure 3.2: Invariant mass of the photon pair for reconstructed tau-candidates with two reconstructed photons in the signal region that are matched to generator level  $\tau^- \rightarrow \pi^- \pi^0 \nu_\tau$  decays.

mDiPhotonsForTrueDM1)

## 1072 Quality Requirements

1073 Photons from the underlying event and other reconstruction effects cause the number of  
 1074 reconstructed photons to be greater than the true number of photons expected from a given  
 1075 hadronic tau decay. Photons that have not been merged into a  $\pi^0$  meson candidate are  
 1076 recursively filtered by requiring that the fraction of the transverse momentum carried by  
 1077 the lowest  $p_T$  photon be greater than 10% with respect to the entire (tracks,  $\pi^0$  candidates,  
 1078 and photons) tau-candidate. In the case that a photon is not merged but meets the minimum  
 1079 momentum fraction requirement, it is considered a  $\pi^0$  candidate. This requirement removes  
 1080 extraneous photons, while minimizing the removal of single photons that correspond to a  
 1081 true  $\pi^0$  meson (see Figure 3.3). A mass hypothesis with the nominal [20] value of the  $\pi^0$   
 1082 is applied to all  $\pi^0$  candidates. All objects that fail the filtering requirements are moved to  
 1083 the isolation collection.

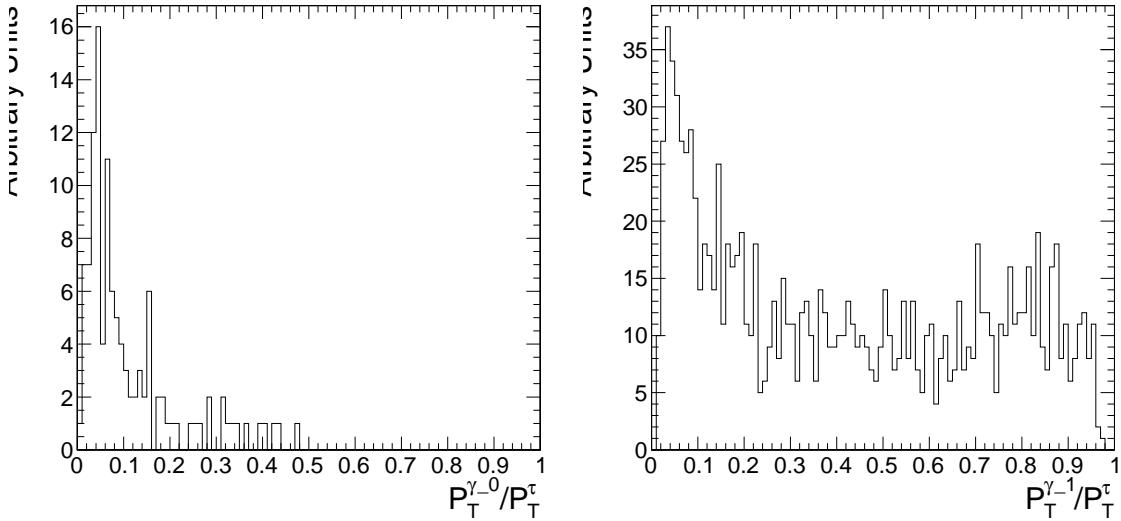


Figure 3.3: Fraction of total  $\tau$ -candidate transverse momenta carried by the photon for reconstructed taus containing a single photons for two benchmark cases. On the left, the reconstructed tau-candidate is matched to generator level  $\tau^- \rightarrow \pi^- \nu_\tau$  decays, for which no photon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

#### 1084 Performance

1085 The performance of the decay mode reconstruction can be measured for tau-candidates that  
 1086 are matched to generator level hadronically decaying tau leptons by examining the correla-  
 1087 tion of the reconstructed decay mode to the true decay mode determined from the Monte  
 1088 Carlo generator level information. Figure 3.4 compares the decay mode reconstruction per-  
 1089 formance of a naive approach where the decay mode is determined by simply counting  
 1090 the number of photons to the performance of the photon merging and filtering approach  
 1091 described in Section 3.3.1. The correlation for the merging and filtering algorithm is much  
 1092 more diagonal, indicating higher performance. The performance is additonally presented for  
 1093 comparison in tabular form in Table 3.3.1 (merging and filtering approach) and Table 3.3.1  
 1094 (naive approach).

1095 The performance of the decay mode reconstruction is dependent on the transverse  
 1096 momentum and  $\eta$  of the tau-candidate and is shown in Figure 3.5. The  $p_T$  dependence  
 1097 is largely due to threshold effects; high multiplicity decay modes are suppressed at low

1098 transverse momentum as the constituents are below the minimum  $p_T$  quality requirements.  
 1099 In the forward region, nuclear interactions and conversions from the increased material  
 1100 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.

### 1101 §3.3.2 Neural Network Classification

{sec:tanc'nn'training}

#### 1102 Neural Network Training

1103 The samples used to train the TaNC neural networks are typical of the signals and back-  
 1104 grounds found in common physics analyses using taus. The signal-type training sample is

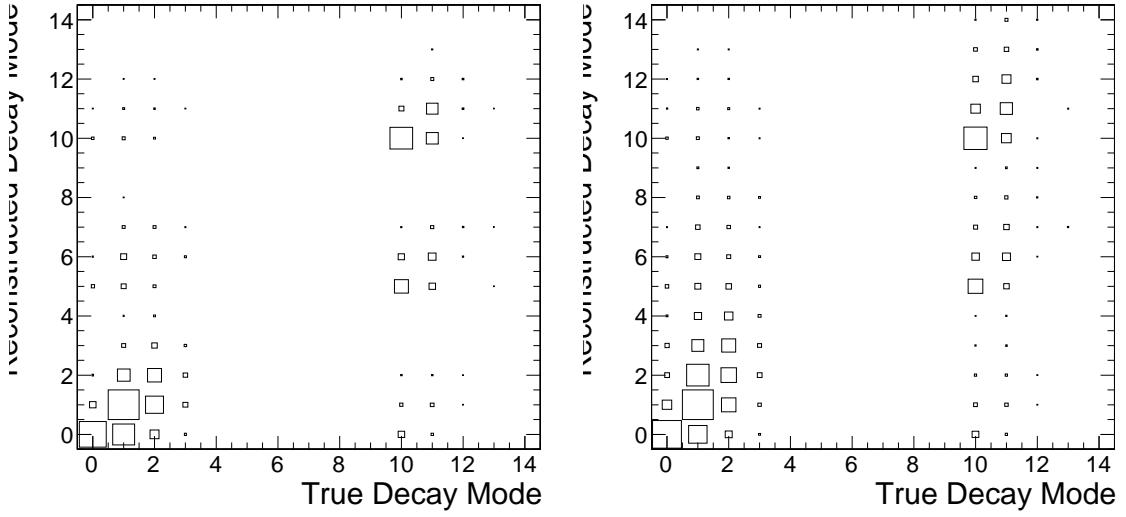


Figure 3.4: Correlations between reconstructed tau decay mode and true tau decay mode for hadronic tau decays in  $Z \rightarrow \tau^+\tau^-$  events. The correlation when no photon merging or filtering is applied is shown on the right, and the correlation for the algorithm described in Section 3.3.1 is on the right. The horizontal and vertical axis are the decay mode indices of the true and reconstructed decay mode, respectively. The decay mode index  $N_{DM}$  is defined as  $N_{DM} = (N_{\pi^\pm} - 1) \cdot 5 + N_{\pi^0}$ . The area of the box in each cell is proportional to the fraction of tau-candidates that were reconstructed with the decay mode indicated on the vertical axis for the true tau decay on the horizontal axis. The performance of a decay mode reconstruction algorithm can be determined by the spread of the reconstructed number of  $\pi^0$  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)

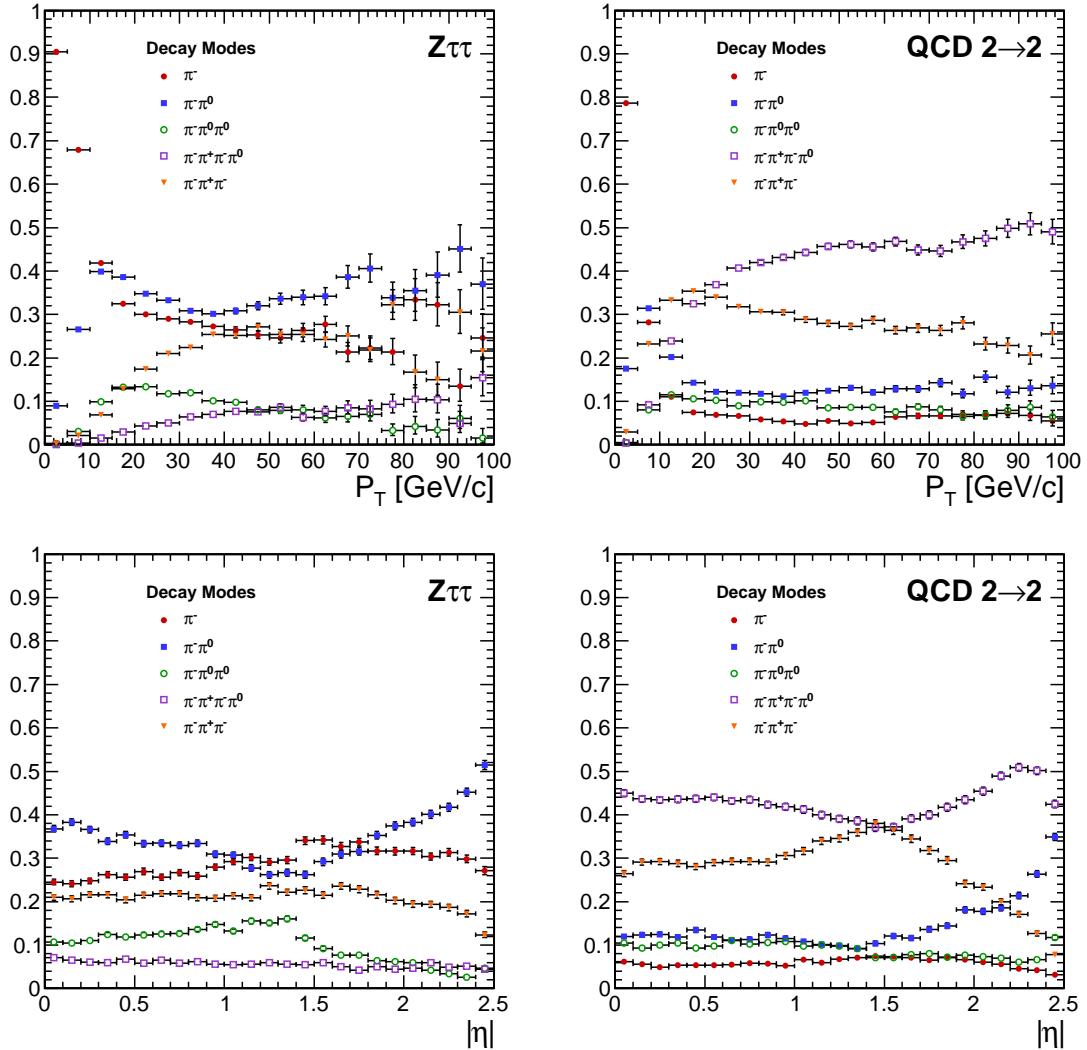


Figure 3.5: Kinematic dependence of reconstructed decay mode for tau-candidates in  $Z \rightarrow \tau^+\tau^-$  (left) and QCD di-jets (right) events versus transverse momentum (top) and pseudo-rapidity (bottom). Each curve is the probability for a tau-candidate to be reconstructed with the associated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

(fig:dmKinematics)

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

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

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

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

	Signal	Background
Total number of tau-candidates	874266	9526176
Tau-candidates passing preselection	584895	644315
Tau-candidates with $W(p_T, \eta) > 0$	538792	488917
Decay Mode	Training Events	
$\pi^-$	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.

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

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

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

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

The neural network error  $E$  is defined [30] as

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

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

No evidence of over-training is observed.

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

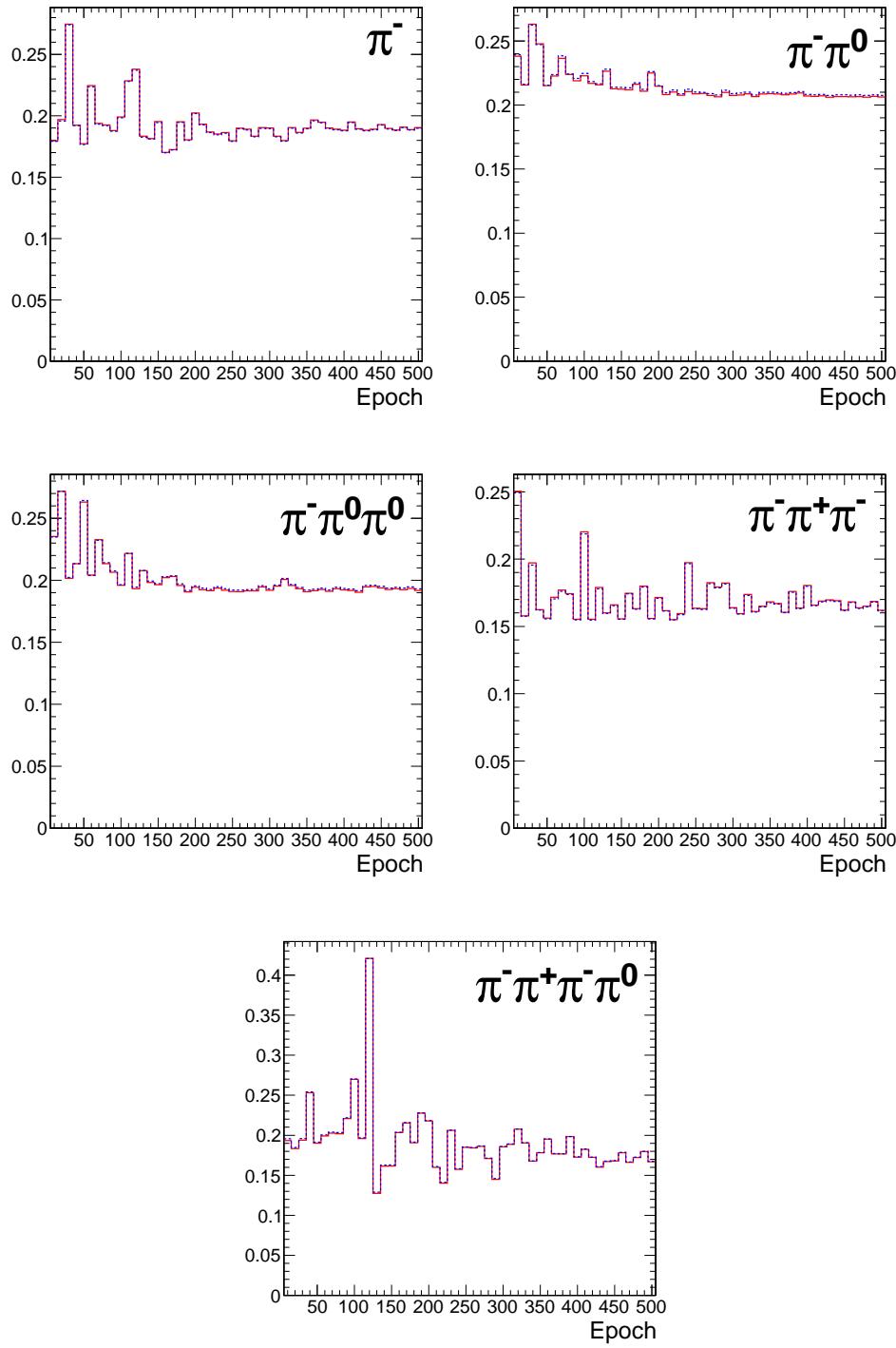


Figure 3.6: Neural network classification error for training (solid red) and testing (dashed blue) samples at ten epoch intervals over the 500 training epochs for each decay mode neural network. The vertical axis represents the classification error, defined by equation 3.2. N.B. that the choice of hyperbolic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

{fig:overTrainCheck}

1141 while for background events the correlation is very small [32]. In the training signal and  
 1142 background samples, there is significant discrimination power in the  $p_T$  spectrum. However,  
 1143 it is desirable to eliminate any systematic dependence of the neural network output on  $p_T$   
 1144 and  $\eta$ , as in practice the TaNC will be presented with tau-candidates whose  $p_T - \eta$  spectrum  
 1145 will be analysis dependent. The dependence on  $p_T$  and  $\eta$  is removed by applying a  $p_T$  and  
 1146  $\eta$  dependent weight to the tau-candidates when training the neural nets.

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

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

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

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

1147 where  $p_{sig}(p_T, \eta)$  and  $p_{bkg}(p_T, \eta)$  are the probability densities of the signal and background  
 1148 samples after the “leading pion” and dominant decay mode selections. Figure 3.7 shows the  
 1149 signal and background training  $p_T$  distributions before and after the weighting is applied.

## 1150 Discriminants

(sec:tanc'nn'discriminants)  
 1151 Each neural network corresponds to a different decay mode topology and as such each  
 1152 network uses different observables as inputs. However, many of the input observables are  
 1153 used in multiple neural nets. The superset of all observables is listed and defined below.  
 1154 Table 3.4 maps the input observables to their associated neural networks. In three prong  
 1155 decays, the definition of the “main track” is important. The main track corresponds to the  
 1156 track with charge opposite to that of the total charge of the three tracks. This distinction is  
 1157 made to facilitate the use of the “Dalitz” observables, allowing identification of intermediate  
 1158 resonances in three-body decays. This is motivated by the fact that the three prong decays  
 1159 of the tau generally proceed through  $\tau^- \rightarrow a1^- \nu_\tau \rightarrow \pi^- \rho^0 \nu_\tau \rightarrow \pi^- \pi^+ \pi^- \nu_\tau$ ; the oppositely  
 1160 charged track can always be identified with the  $\rho^0$  decay.

## 1161 ChargedOutlierAngleN

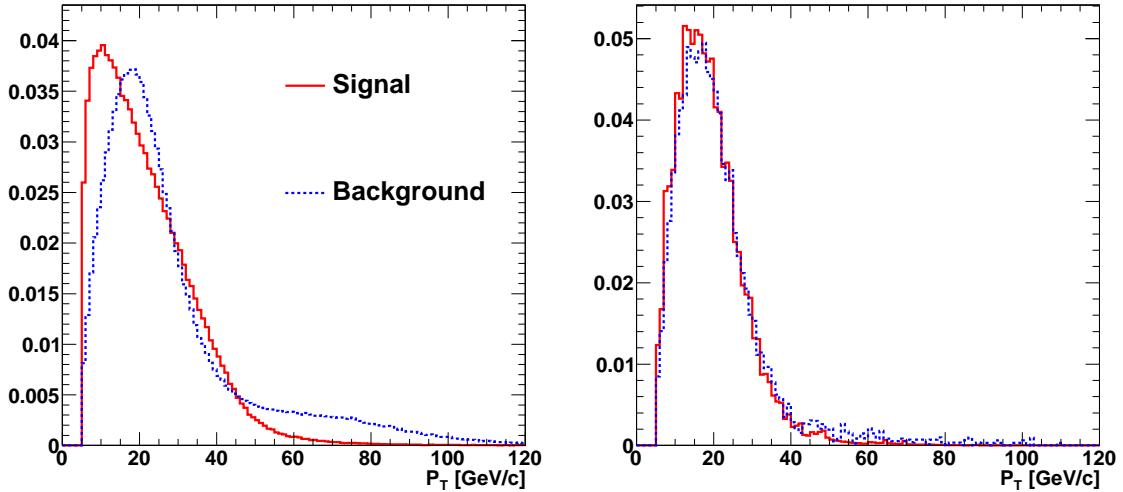


Figure 3.7: Transverse momentum spectrum of signal and background tau-candidates used in neural net training before (left) and after (right) the application of  $p_T - \eta$  dependent weight function. Application of the weights lowers the training significance of tau-candidates in regions of  $p_T -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

(fig:nnTrainingWeights)

1162        $\Delta R$  between the Nth charged object (ordered by  $p_T$ ) in the isolation region and the  
 1163           tau-candidate momentum axis. If the number of isolation region objects is less than  
 1164           N, the input is set at one.

#### 1165   **ChargedOutlierPtN**

1166       Transverse momentum of the Nth charged object in the isolation region. If the number  
 1167           of isolation region objects is less than N, the input is set at zero.

#### 1168   **DalitzN**

1169       Invariant mass of four vector sum of the “main track” and the Nth signal region  
 1170           object.

#### 1171   **Eta**

1172       Pseudo-rapidity of the signal region objects.

#### 1173   **InvariantMassOfSignal**

1174       Invariant mass of the composite object formed by the signal region constituents.

1175 **MainTrackAngle**

1176      $\Delta R$  between the “main track” and the composite four–vector formed by the signal  
 1177     region constituents.

1178 **MainTrackPt**

1179     Transverse momentum of the “main track.”

1180 **OutlierNCharged**

1181     Number of charged objects in the isolation region.

1182 **OutlierSumPt**

1183     Sum of the transverse momentum of objects in the isolation region.

1184 **PiZeroAngleN**

1185      $\Delta R$  between the Nth  $\pi^0$  object in the signal region (ordered by  $p_T$ ) and the tau–  
 1186     candidate momentum axis.

1187 **PiZeroPtN**

1188     Transverse momentum of the Nth  $\pi^0$  object in the signal region.

1189 **TrackAngleN**

1190      $\Delta R$  between the Nth charged object in the signal region (ordered by  $p_T$ ) and the  
 1191     tau–candidate momentum axis, exclusive of the main track.

1192 **TrackPtN**

1193     Transverse momentum of the Nth charged object in the signal region, exclusive of the  
 1194     main track.

1195 Neural Network Performance

neuralNetworkPerformance) 1196     The classification power of the neural networks is unique for each of the decay modes.  
 1197     The performance is determined by the relative separation of the signal and background  
 1198     distributions in the parameter space of the observables used as neural network inputs. A  
 1199     pathological example is the case of tau–candidates with the reconstructed decay mode of  
 1200      $\tau^- \rightarrow \pi^- \nu_\tau$ . If there is no isolation activity, the neural net has no handle with which it

1201 can separate the signal from the background. The neural net output for tau-candidates in  
 1202 the testing sample (independent of the training and validation samples) for each of the five  
 1203 decay mode classifications is shown in Figure 3.8.

1204 When a single neural network is used for classification, choosing an operating point is  
 1205 relatively straightforward: the requirement on neural network output is tuned such that the  
 1206 desired purity is attained. However, in the case of the TaNC, multiple neural networks are  
 1207 used. Each network has a unique separation power (see Figure 3.9) and each neural network  
 1208 is associated to a reconstructed decay mode that composes different relative fractions of the  
 1209 signal and background tau-candidates. Therefore, a set of five numbers is required to define  
 1210 an “operating point” (the signal efficiency and background misidentification rate) in the  
 1211 TaNC output. All points in this five dimensional cut-space map to an absolute background  
 1212 fake-rate and signal efficiency rate. Therefore there must exist a 5D “performance curve”  
 1213 which for any attainable signal efficiency gives the lowest fake-rate. A direct method to  
 1214 approximate the performance curve is possible using a Monte Carlo technique.

1215 The maximal performance curve can be approximated by iteratively sampling points in  
 1216 the five-dimensional cut space and selecting the highest performance points. The collection  
 1217 of points in the performance curve are ordered by expected fake rate. During each iteration,  
 1218 the sample point is compared to the point before the potential insertion position of the  
 1219 sample in the ordered collection. The sample point is inserted into the collection if it has  
 1220 a higher signal identification efficiency than the point before it. The sample point is then  
 1221 compared to all points in the collection after it (i.e. those with a larger fake rate); any point  
 1222 with a lower signal efficiency than the sample point is removed. After the performance curve  
 1223 has been determined, the set of cuts are evaluated on an independent validation sample  
 1224 to ensure that the measured performance curve is not influenced by favorable statistical  
 1225 fluctuations being selected by the Monte Carlo sampling. The performance curves for two  
 1226 different transverse momentum ranges are shown in Figure 3.10.

The 5D performance curve can also be parameterized by using the probability for a tau-candidate to be identified for a given decay mode. An artificial neural network maps a point in the space of input observables to some value of neural network output  $x$ . The neural network training error is given by Equation 3.2. A given point in the vector space

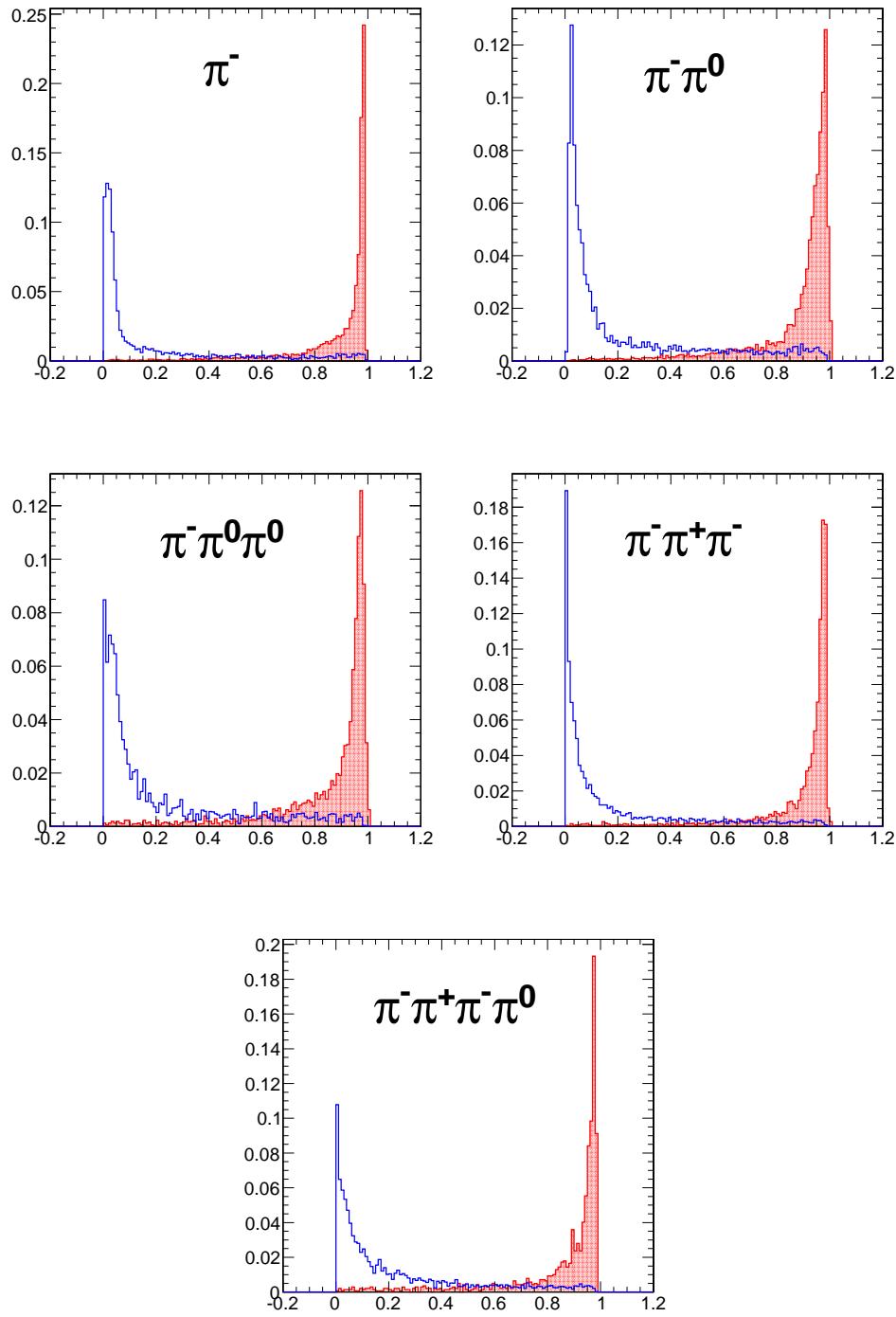


Figure 3.8: Neural network output distributions for the five reconstructed tau–candidate decay modes used in the TaNC for  $Z \rightarrow \tau^+\tau^-$  events (red) and QCD di-jet events (blue).

fig:NNoutputDistributions

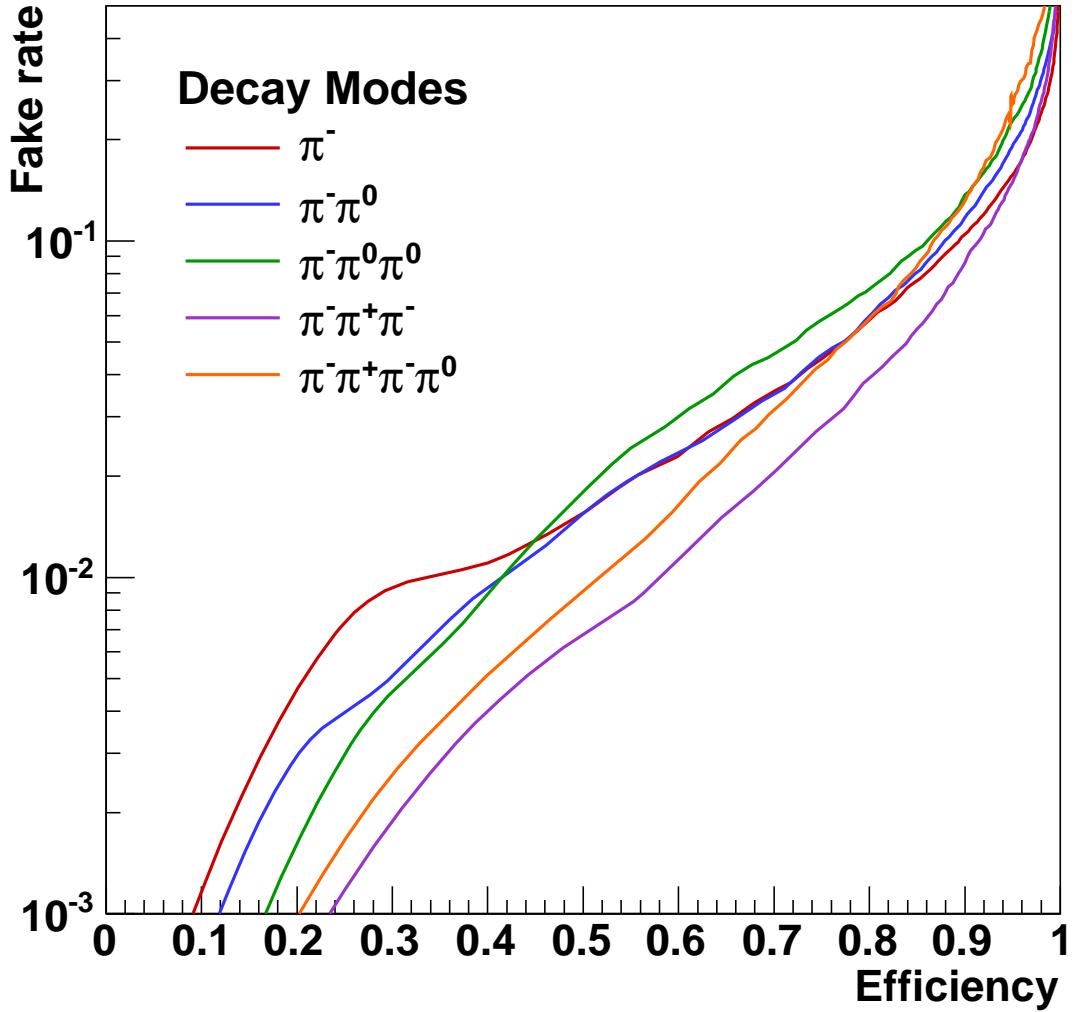


Figure 3.9: Performance curves for the five neural networks used by the TaNC for tau-candidates with transverse momentum greater than 20 GeV/c. Each curve represents the signal efficiency (on the horizontal axis) and background misidentification rate (vertical axis) for a scan of the neural network selection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

{fig:nnPerfCurves}

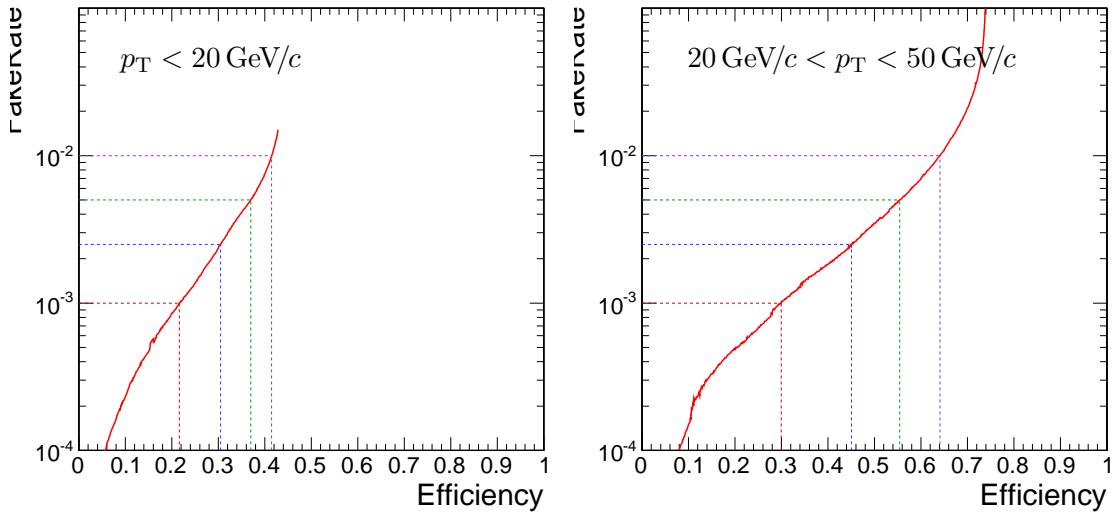


Figure 3.10: Tau Neural Classifier performance curves for tau-candidates with  $p_T < 20 \text{ GeV}/c$  (left) and  $20 < p_T < 50 \text{ GeV}/c$  (right). The vertical axis represents the expected fake-rate of QCD jets and the horizontal axis the expected signal efficiency for hadronic tau decays.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

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

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

1227 where  $\rho^\tau(\rho^{QCD})$  denotes the training sample density of the  $\tau$  signal and QCD–jet back-  
1228 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:

$$\frac{\partial E'}{\partial x} = 0$$

$$0 = -2(1 - x) \cdot \rho^\tau + 2x \cdot \rho^{QCD}$$

$$x = \frac{\rho^\tau}{\rho^\tau + \rho^{QCD}} \quad (3.3) \boxed{\text{eq:probFracToX}}$$

$$\rho^{QCD} = \frac{1}{x} - 1 \quad (3.4) \boxed{\text{eq:rawTransform}}$$

1229 The ratio  $\frac{\rho^{QCD}}{\rho^\tau}$  corresponds to the ratio of the normalized probability density functions of  
1230 signal and background input observable distributions, i.e.  $\int \rho^\tau d\vec{x} = 1$ .

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

$$x'_j = \frac{\rho^\tau \cdot p_j^\tau}{\rho^\tau \cdot p_j^\tau + \rho^{QCD} \cdot p_j^{QCD}} = \frac{p_j^\tau}{p_j^\tau + \frac{\rho^{QCD}}{\rho^\tau} \cdot p_j^{QCD}} \quad (3.5) \boxed{\text{eq:probFracToXV}}$$

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

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

1231 In this manner a single number (the “transform cut”) given by Equation 3.6 can be used  
1232 to specify any point on the performance curve. The training sample neural network output  
1233 after the transformation has been applied is shown in Figure 3.12. The performance curve

1234 for the cut on the transformed output is nearly identical to the optimal performance curve  
 1235 determined by the Monte Carlo sampling technique.

1236 The discriminator output of the TaNC algorithm is a continuous quantity, enabling  
 1237 analysis specific optimization of the selection to maximize sensitivity. For the convenience  
 1238 of the user, four operating point benchmark selections are provided in addition to the  
 1239 continuous output. The four operating points are chosen such that for tau-candidates with  
 1240 transverse momentum between 20 and 50 GeV/c, the expected QCD di-jet fake rate will be  
 1241 0.1%, 0.25%, 0.50% and 1.0%, respectively.

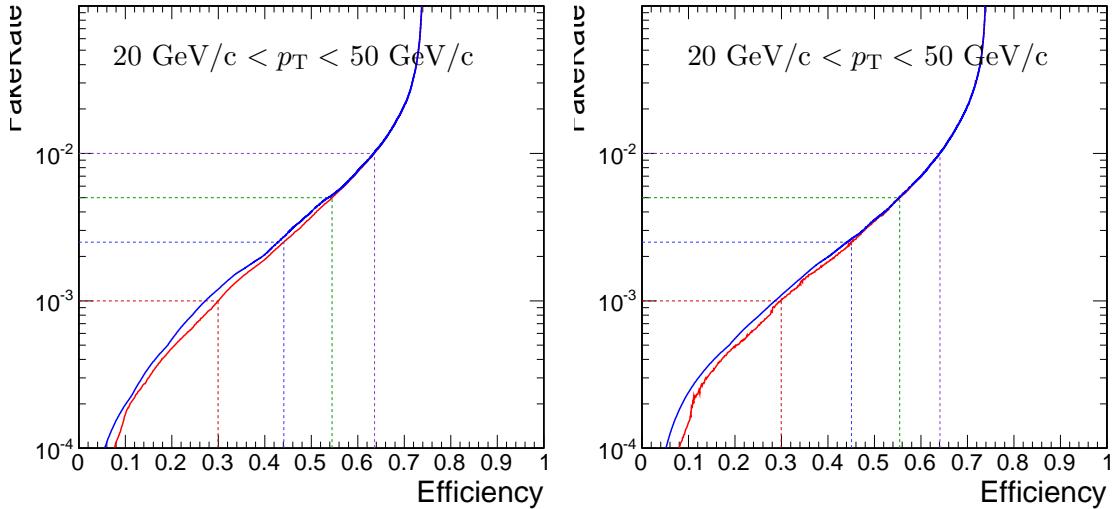


Figure 3.11: Tau Neural Classifier performance curves for tau-candidates with  $20 < p_T < 50 \text{ GeV}/c$ . The figure on the left compares the optimal performance curve determined by the Monte Carlo sampling method (red) to the performance curve obtained by scanning the “transform cut” (blue) defined in Equation 3.6 from zero to one. The figure on the right is the same set of cuts (and cut transformation values) applied on an independent sample to remove any biases introduced by the Monte Carlo sampling. The four dashed lines indicate the performance for the four benchmark points.

## 1242 §3.4 Summary

1243 The Tau Neural classifier introduces two complimentary new techniques for tau lepton  
 1244 physics at CMS: reconstruction of the hadronic tau decay mode and discrimination from  
 1245 quark and gluon jets using neural networks. The decay mode reconstruction strategy pre-  
 1246 sented in Section 3.3.1 significantly improves the determination of the decay mode. This

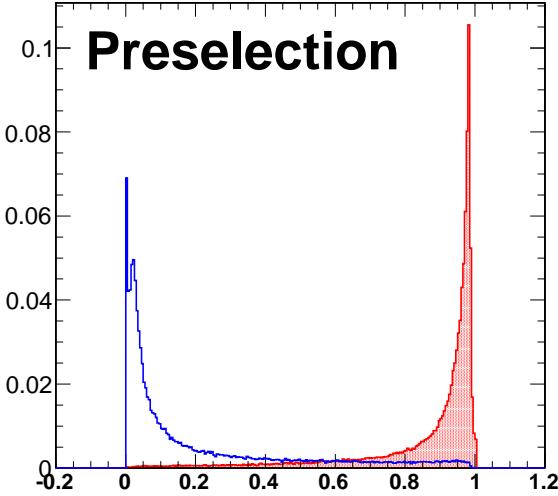


Figure 3.12: Transformed TaNC neural network output for tau-candidates with transverse momentum between 20 and 50 GeV/c that pass the pre-selection criteria. The neural network output for each tau-candidate has been transformation according to Equation 3.6. The decay mode probabilities  $\rho_i^{bkg}$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

`fig:transformedNNOutput`

1247 information has the potential to be useful in studies of tau polarization and background  
 1248 estimation.

1249 The Tau Neural classifier tau identification algorithm significantly improves tau dis-  
 1250 crimination performance compared to isolation-based approaches [28] used in previous CMS  
 1251 analyses. Figure 3.13 compares the performance of the “shrinking cone” isolation tau-  
 1252 identification algorithm [28] to the performance of the TaNC for a scan of requirements  
 1253 on the transformed neural network output. The signal efficiency and QCD di-jet fake rate  
 1254 versus tau-candidate transverse momentum and pseudo-rapidity for the four benchmark  
 1255 points and the isolation based tau identification are show in Figure 3.14. For tau-candidates  
 1256 with transverse momentum between 20 and 50 GeV/c, the TaNC operating cut can be cho-  
 1257 sen such that the two methods have identical signal efficiency; at this point the TaNC  
 1258 algorithm reduces the background fake rate by an additional factor of 3.9. This reduction  
 1259 in background will directly improve the significance of searches for new physics using tau  
 1260 leptons at CMS.

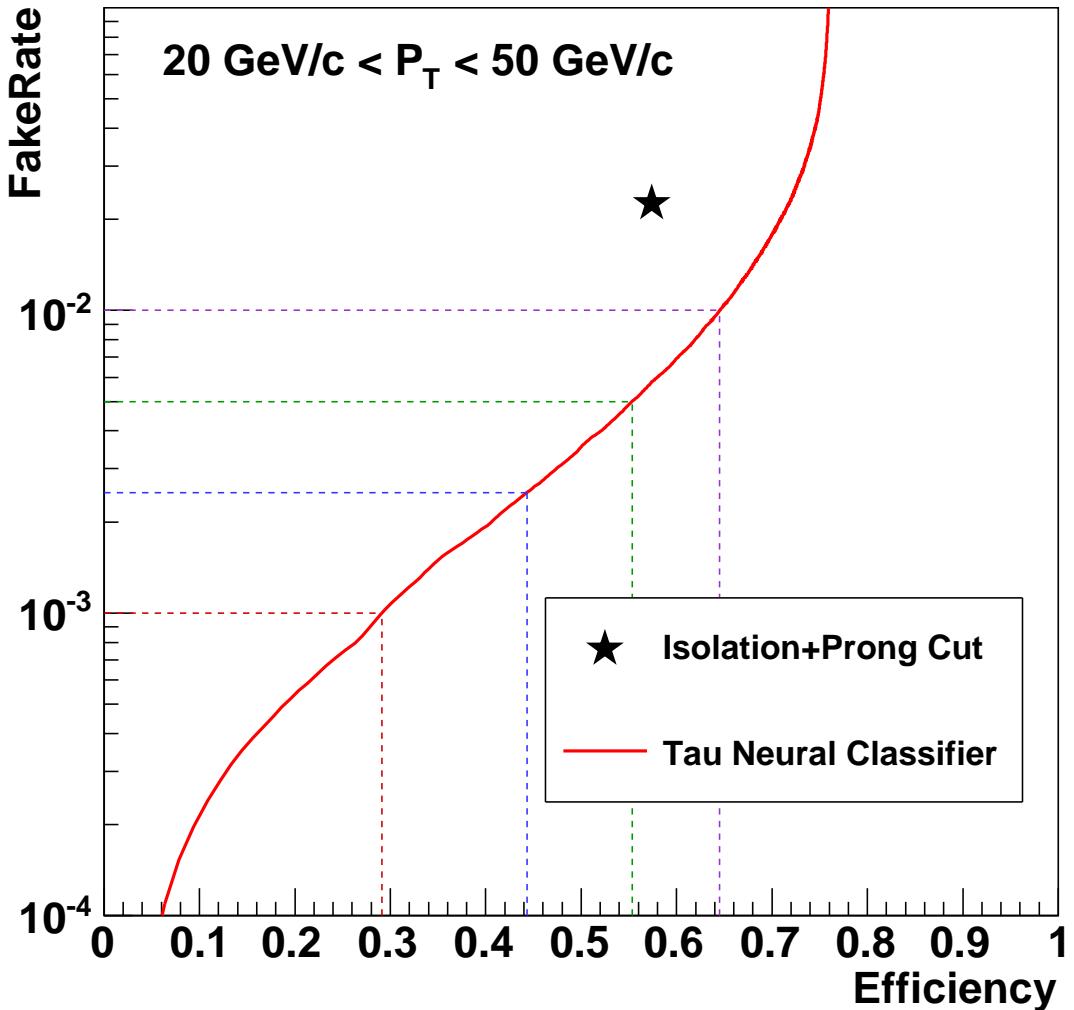


Figure 3.13: Performance curve (red) of the TaNC tau identification for various requirements on the output transformed according to Equation 3.6. The horizontal axis is the efficiency for true taus with transverse momentum between 20 and 50 GeV/c to satisfy the tau identification requirements. The vertical axis gives the rate at which QCD di-jets with generator-level transverse momentum between 20 and 50 GeV/c are incorrectly identified as taus. The performance point for the same tau-candidates using the isolation based tau-identification [28] used in many previous CMS analyses is indicated by the black star in the figure. An additional requirement that the signal cone contain one or three charged hadrons (typical in a final physics analysis) has been applied to the isolation based tau-identification to ensure a conservative comparison.

{fig:finalPerfCurve}

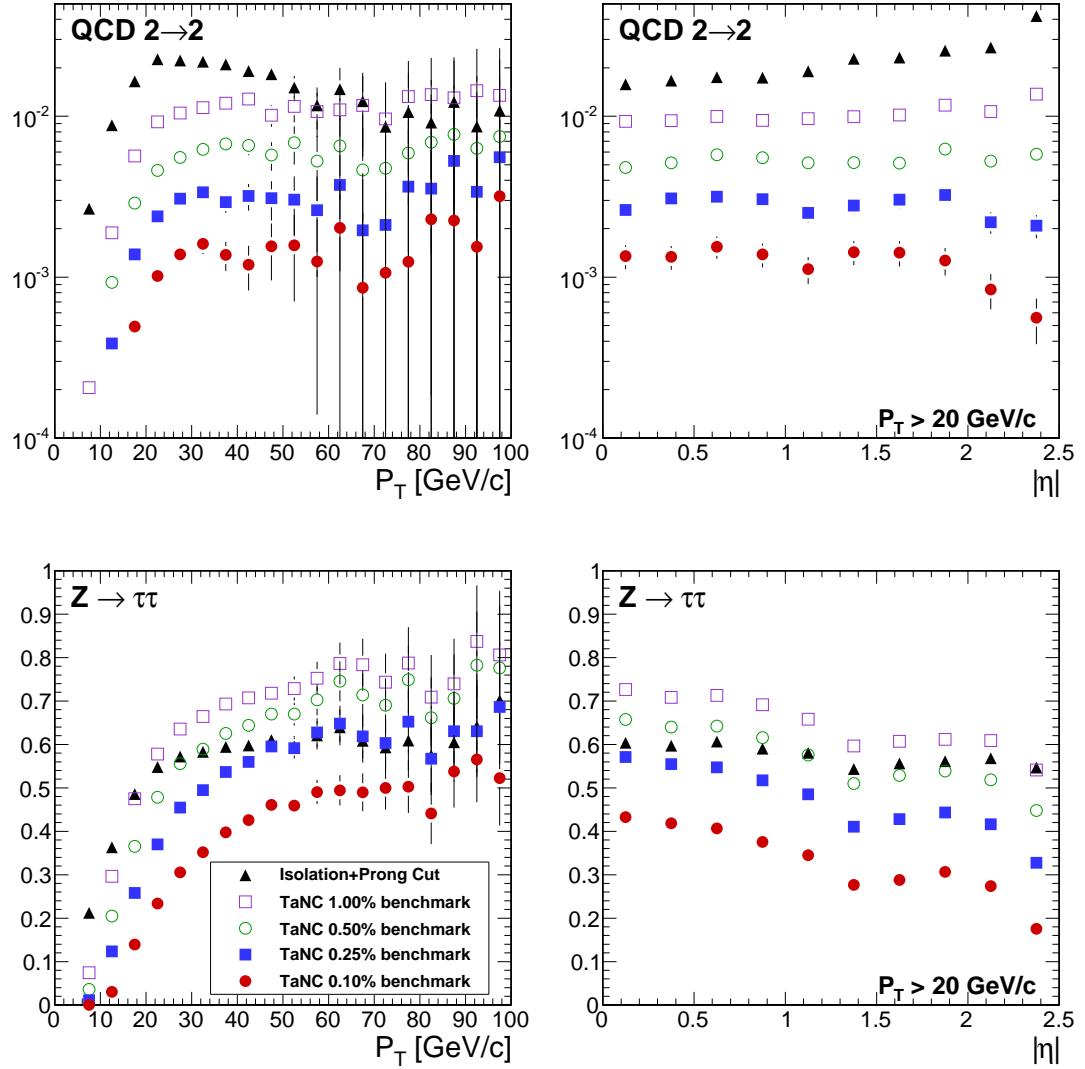


Figure 3.14: Comparison of the identification efficiency for hadronic tau decays from  $Z \rightarrow \tau^+\tau^-$  decays (bottom row) and the misidentification rate for QCD di-jets (top row) versus tau-candidate transverse momentum (left) and pseudo-rapidity (right) for different tau identification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

fig:kinematicPerformance)

## 1261 §3.5 HPS+TaNC: A Hybrid Algorithm

1262 `<sec:TauId>` The techniques used in the TaNC have been hybridized with techniques used by the “Hadrons plus Strips” (HPS) algorithm [33]. The combined algorithm is referred to “Hadrons plus Strips and Tau Neural Classifier” (HPS + TaNC) identification algorithm. The algorithm combines the HPS methods of constructing the signal components of the tau candidate and the discrimination methods of the TaNC algorithm. Both algorithms are based on reconstructing individual tau lepton hadronic decay modes, which has been demonstrated to improve the tau identification performance significantly with respect to previously used cone isolation based algorithms [34]. The HPS + TaNC algorithm first reconstructs the hadronic decay mode of the tau, and applies different discriminants based on the reconstructed decay mode. Identification of hadronic tau decays by the HPS + TaNC algorithm proceeds in two stages: first, the hadronic decay mode of the tau is reconstructed and then different discriminators are applied, based on the reconstructed decay mode. In the decay mode reconstruction particular attention is paid to the reconstruction of neutral pions, which are expected for the majority of hadronic decay modes.

### 1276 §3.5.1 Decay mode reconstruction

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

- 1285 • The  $\pi^0$  candidate is in the ECAL barrel region ( $|\eta| < 1.5$ ) and has invariant mass  $|m_{\gamma\gamma} - m_{\pi^0}| < 0.05 \text{ GeV}/c^2$ .
- 1287 • The  $\pi^0$  candidate is in the ECAL endcap region ( $|\eta| > 1.5$ ) and has invariant mass  $m_{\gamma\gamma} < 0.2 \text{ GeV}/c^2$ .

- 1289     • The  $\pi^0$  candidate contains two or more photons within an  $\eta - \phi$  strip of size  $0.05 \times 0.20$ .  
 1290     • Photons not satisfying any of the other categories are considered as unresolved  $\pi^0$   
 1291        candidates in case they have  $p_T > 1.0$  GeV/c.

1292     The symbol  $m_{\pi^0}$  denotes the nominal neutral pion mass [20]. The size of the invariant mass  
 1293     windows in the ECAL endcap and barrel regions is motivated by the resolution of the  $\pi^0$   
 1294     mass (illustrated in Figure 3.15) measured during the commissioning of the particle-flow  
 1295     algorithm in early CMS data [36]. Multiple  $\pi^0$  candidates in the same category are ranked  
 1296     in quality according to the difference of the reconstructed photon pair mass to the nominal  
 1297      $\pi^0$  mass. After the  $\pi^0$  candidates are ranked, the highest ranked candidate is selected for  
 1298     the final collection. The photon constituents of the highest ranked candidate are removed  
 1299     from remaining  $\pi^0$  candidates not yet selected for the final collection in order to prevent  
 1300     photons from entering more than one  $\pi^0$  candidate. The rank of remaining  $\pi^0$  candidates  
 1301     is reevaluated and the  $\pi^0$  candidate with the next highest rank is selected for the output  
 collection. The process is repeated until no more  $\pi^0$  candidates are remaining.

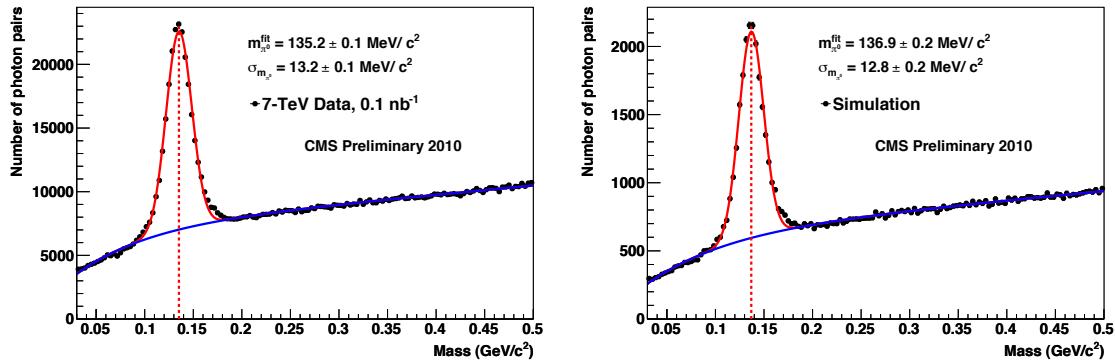


Figure 3.15: Invariant mass distribution of photon pairs reconstructed by the particle-flow in 2010 CMS minimum bias events (left), and predicted by the simulation (right). A clear resonant pick corresponding to the  $\pi_0$  meson is visible above the combinatoric background. Reference: [36]

(fig:PFPiZeroRes)  
 1302

1303     Once the final collection of  $\pi^0$  candidates is determined, tau reconstruction in the  
 1304     HPS + TaNC algorithm proceeds by building tau candidates from reconstructed  $\pi^0$  candi-  
 1305     dates and charged hadrons reconstructed by the particle-flow algorithm. A combinatoric  
 1306     approach is again employed for the tau candidate building. A tau candidate hypothesis is

1307 built for every combination of jet constituents ( $\pi^0$  candidates plus charged hadrons) which  
 1308 has a multiplicity consistent with a hadronic tau decay. The tau candidates are ranked anal-  
 1309 ogous to the ranking utilized for the  $\pi^0$  reconstruction, but with the following categorical  
 1310 rankings:

- 1311     • In each decay mode category, the tau candidate with the highest neural network  
     1312       output is selected.
- 1313     • The tau candidate has unit charge.
- 1314     • The tau candidate passes the “lead pion” criteria, requiring that there is a photon or  
     1315       charged pion candidate with  $p_T > 5$  GeV/c.
- 1316     • The tau candidate passes the HPS invariant mass and collimation<sup>1</sup> requirements.

1317 In case multiple tau candidates satisfy all four categorical requirements, the tau candidate  
 1318 with the highest energy sum of charged and neutral pions is selected as the highest ranking  
 1319 one.

### 1320 §3.5.2 Hadronic tau discrimination

1321 The final level of discrimination is performed by an ensemble of neural networks, with each  
 1322 neural network corresponding to a specific decay mode, analogously to the method used  
 1323 original TaNC algorithm (Section 3.3.2). The inputs of each neural network are different and  
 1324 correspond to the observables (invariant mass, Dalitz masses) available for its associated  
 1325 decay mode. The neural networks are trained on samples simulated  $Z \rightarrow \tau^+\tau^-$  events  
 1326 (“signal”) and QCD di-jet events selected in the 7 TeV data collected by CMS in 2010  
 1327 (“background”). All of the tau hypothesis from a given jet reconstructed in data are used  
 1328 for training. The  $Z \rightarrow \tau^+\tau^-$  signal sample is generated by PYTHIA [37] which has been  
 1329 interfaced to TAUOLA [38] for the purpose of generating the decays of the tau leptons using  
 1330 the full matrix element information. After generation, the events passed through the “full”  
 1331 GEANT [39] based simulation of the CMS detector. Only tau candidates which have been

---

<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].

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 3.3.2. 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.

### §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 in [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 requiring that there is a HCAL energy deposit with a magnitude that is greater than 10% of the momentum of the leading track in the tau.

---

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

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

Table 3.4: Input observables used for each of the neural networks implemented by the Tau Neural Classifier. The columns represents the neural networks associated to various decay modes and the rows represent the superset of input observables (see Section 3.3.2) used in the neural networks. A dot in a given row and column indicates that the observable in that row is used in the neural network corresponding to that column.

(tab:nn`var`table)

## 1355 Chapter 4

# 1356 Mass Reconstruction: The Secondary Vertex Fit

1355  
 1356     The dominant background in the search for a Higgs decaying to a  $\tau^+\tau^-$  pair are Standard  
 1357     Model  $Z \rightarrow \tau^+\tau^-$  events. The most “natural” observable to discriminate between a Higgs  
 1358     signal and the  $Z$  background is the invariant mass of the di-tau system, utilizing the fact  
 1359     that the  $Z$  resonance is well known ( $m_Z = 91.1876 \pm 0.0021$  GeV/ $c^2$ ) and has a narrow width  
 1360     ( $\Gamma_Z = 2.4952 \pm 0.0023$  GeV) [20]. The experimental complication in this approach is due to  
 1361     the neutrinos produced in the tau lepton decays, which escape detection and carry away an  
 1362     unmeasured amount of energy, making it difficult to reconstruct the tau lepton four-vectors.  
 1363     In this chapter we give an overview of techniques used in previous literature [19, 40, 41] to  
 1364     construct an observable related to the tau pair mass. We then introduce a new algorithm,  
 1365     called the Secondary Vertex (SV) fit. The SVfit reconstructs the “full” tau pair mass, and  
 1366     provides increased performance with respect to techniques previously used in the literature.

### 1369 §4.1 Existing mass reconstruction algorithms

1370     The simplest observable related to the  $\tau^+\tau^-$  mass is one can construct that is sensitive to  
 1371     new particle content is the invariant mass of the visible (reconstructible) decay products  
 1372     associated with each tau decays. This quantity, referred in this document as the “Visible  
 1373     Mass,” has the advantages of simplicity and lack of exposure to systematic errors associ-  
 1374     ated with the reconstruction of the  $E_T^{\text{miss}}$ . However, no attempt is made to reconstruct the  
 1375     neutrinos in the event. The reconstructed mass is thus systematically smaller than mass of  
 1376     the resonance which produced the tau leptons. The visible mass is typically on the order of  
 1377     1/2 of the resonance mass, depending on the kinematic requirements applied to the visible  
 1378     products of the tau decays.

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

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

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

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

1379 The Collinear Approximation suffers from two problems. The approximation can fail  
 1380 (yielding unphysical negative energies for the reconstructed neutrinos) when the missing  
 1381 transverse energy is mis-measured. The events with unphysical solutions must be removed  
 1382 from the analysis, leading to a dramatic reduction in acceptance (on the order of 50% in  
 1383 this analysis). Improvements to the collinear approximation algorithm have recently been  
 1384 made which aim to recover part of the events with unphysical solutions [42]. But even with  
 1385 these improvements, no physical solution is still found for a large fraction of signal events.  
 1386 Additionally, the method is numerically sensitive when the two  $\tau$  lepton are nearly back-  
 1387 to-back in azimuth. In these cases the  $\sin(\phi_2 - \phi_1)^{-1}$  term in Equation 4.2 is very large

---

<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.

1388 and small mis-measurements of the missing transverse energy can produce a large tail on  
 1389 the reconstructed mass. This tail is particularly large for low-mass resonances. The large  
 1390 tail for high mass is predominantly due to the fact (discussed in Section 4.4.2) that the  
 1391 kinematic requirements<sup>2</sup> applied on the visible decay products preferentially selects events  
 1392 where the visible decay products carry the majority of the energy of the original  $\tau$  lepton,  
 1393 reducing the amount of true missing energy in the event.

## 1394 §4.2 The Secondary Vertex fit

1395 A novel algorithm is presented in the following, which succeeds in finding a physical solution  
 1396 for every event. As an additional benefit, the new algorithm is found to improve the di-tau  
 1397 invariant mass resolution, making it easier to separate the Higgs signal from the  $Z \rightarrow \tau^+\tau^-$   
 1398 background.

1399 The novel Secondary Vertex fit (SVfit) algorithm for di-tau invariant mass reconstruc-  
 1400 tion that we present in the following utilizes a likelihood maximization to fit a  $\tau^+\tau^-$  in-  
 1401 variant mass hypothesis for each event. The likelihood is composed of separate terms which  
 1402 represent probability densities of:

- 1403     • tau decay kinematics
- 1404     • matching between the momenta of neutrinos produced in the tau decays and the  
       1405       reconstructed missing transverse momentum
- 1406     • a regularization “ $p_T$ -balance” term which accounts for the effects on the di-tau in-  
       1407       variant mass of acceptance cuts on the visible tau decay products
- 1408     • the compatibility of tau decay parameters with the position of reconstructed tracks  
       1409       and the known tau lifetime of  $c\tau = 87 \mu\text{m}$  [20].

1410 The likelihood is maximized as function of a set of parameters which fully describe the tau  
 1411 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.

### 1412 §4.3 Parametrization of tau decays

1413 The decay of a tau with visible four-momentum  $p_{vis}$  measured in the CMS detector (“laboratory”)  
 1414 frame can be parametrized by three variables. The invisible (neutrino) momentum  
 1415 is fully determined by these parameters.

1416 The “opening-angle”  $\theta$  is defined as the angle between the boost direction of the tau  
 1417 lepton and the momentum vector of the visible decay products in the rest frame of the  
 1418 tau. The azimuthal angle of the tau in the lab frame is denoted as  $\bar{\phi}$  (we denote quantities  
 1419 defined in the laboratory frame by a overline). A local coordinate system is defined such  
 1420 that the  $\bar{z}$ -direction lies along the visible momentum and  $\bar{\phi} = 0$  lies in the plane spanned  
 1421 by the momentum vector of the visible decay products and the proton beam direction. The  
 1422 third parameter,  $m_{\nu\nu}$ , denotes the invariant mass of the invisible momentum system.

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

$$E^{vis} = \frac{m_\tau^2 + m_{vis}^2 - m_{\nu\nu}^2}{2m_\tau} \quad (4.3) \quad \text{[eq:restFrameM...]$$

1423 Note that for hadronic decays,  $m_{\nu\nu}$  is a constant of value zero, as only a single neutrino is  
 1424 produced. Consequently, the magnitude of  $P^{vis}$  depends on the reconstructed mass of the  
 1425 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:

$$\begin{aligned} p_{\perp}^{vis} &= \bar{p}_{\perp}^{vis} \\ \Rightarrow \sin \bar{\theta} &= \frac{p^{vis} \sin \theta}{\bar{p}^{vis}} \end{aligned} \quad (4.4) \quad \text{[eq:labFrameOpen...]$$

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:

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

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

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

1426 The energy of the tau lepton in the laboratory frame as function of the measured visible  
 1427 momentum depends on two of the three parameters only - the rest frame opening angle  $\theta$  and  
 1428 the invariant mass  $m_{\nu\nu}$  of the neutrino system. The direction of the tau lepton momentum  
 1429 vector is not fully determined by  $\theta$  and  $m_{\nu\nu}$ , but is constrained to lie on the surface of a  
 1430 cone of opening angle  $\bar{\theta}$  (given by Equation 4.4), the axis of which is given by the visible  
 1431 momentum vector. The tau lepton four-vector is fully determined by the addition of the  
 1432 third parameter  $\bar{\phi}$ , which describes the azimuthal angle of the tau lepton with respect to the  
 1433 visible momentum vector. The spatial coordinate system used is illustrated in Figure 4.1.

## 1434 §4.4 Likelihood for tau decay

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:

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

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

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

1435 In the present implementation of the SVfit algorithm, the matrix element is assumed to be  
 1436 constant, so that the likelihood depends on the phase-space volume of the decay only <sup>3</sup>.

### 1437 §4.4.1 Likelihood for reconstructed missing transverse momentum

1438 Momentum conservation in the plane perpendicular to the beam axis implies that the  
 1439 vectorial sum of the momenta of all neutrinos produced in the decay of the tau lepton pair

---

<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 and  $Z$  decays [43].

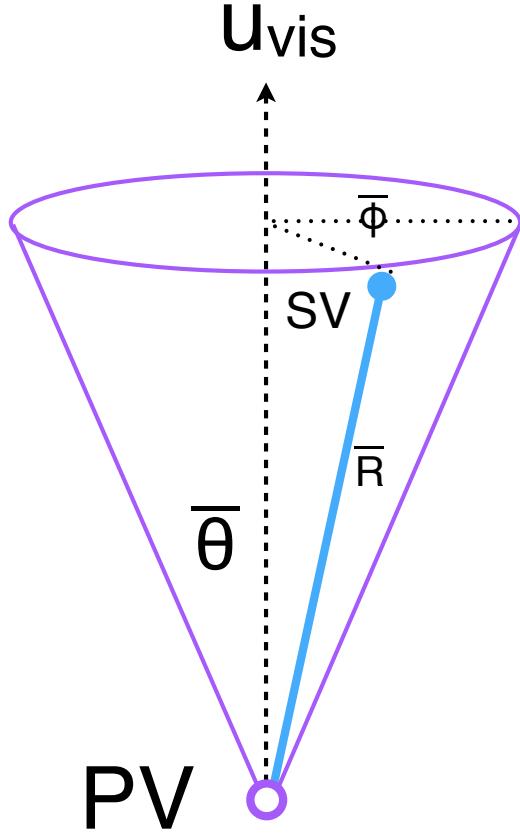


Figure 4.1: Illustration of the coordinate system used by the SVfit to describe the decays of tau leptons.

ig:svFitDecayParDiagram)

- 1440 matches the reconstructed missing transverse momentum. Differences are possible due to
- 1441 the experimental resolution and finite  $p_T$  of particles escaping detection in beam direction
- 1442 at high  $|\eta|$ .

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:

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

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

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

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

1443 where  $q_T$  denotes the transverse momentum of the tau lepton pair.

#### 1444 §4.4.2 Likelihood for tau lepton transverse momentum balance

(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 Higgs  $\rightarrow \tau^+\tau^-$  events as a function of the Higgs mass. The functional form of the parametrization is taken to be the sum of two terms. The first term, denoted by  $p^*(p_T|M)$ , is derived by assuming an isotropic two-body decay, that is

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

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

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

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

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

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 Higgs  $\rightarrow \tau^+\tau^-$  events. The relative

weight  $a$  of the two terms is also determined in the fit. Replacing the integrand in Equation 4.7 by its Taylor expansion so that the integration can be carried out analytically, keeping polynomial terms up to fifth order, and assuming the fit parameters to depend at most linearly on the Higgs mass, we obtain the following numerical values for the parameters:

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

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

#### §4.4.3 Secondary vertex information

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

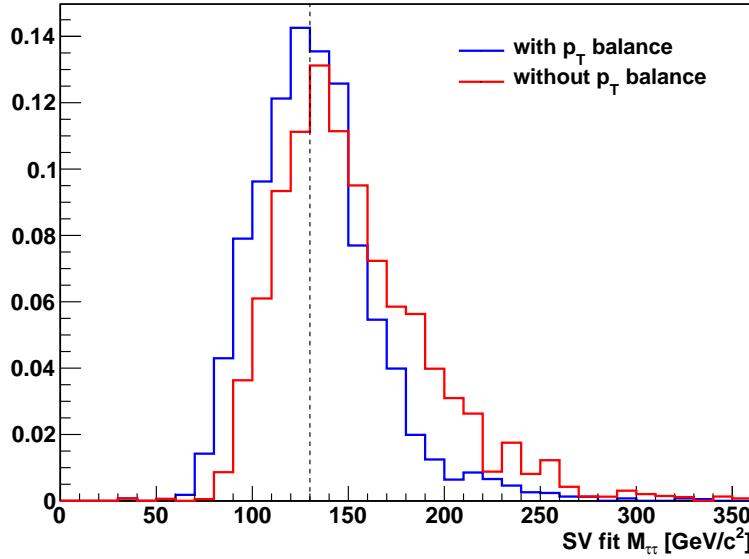


Figure 4.2: Distribution of di-tau invariant mass reconstructed by the SVfit algorithm in simulated Higgs events with  $m_{A^0} = 130 \text{ GeV}/c^2$ . The SVfit algorithm is run in two configurations, with (blue) and without (red) the  $p_T$ -balance likelihood term included in the fit.

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

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

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

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

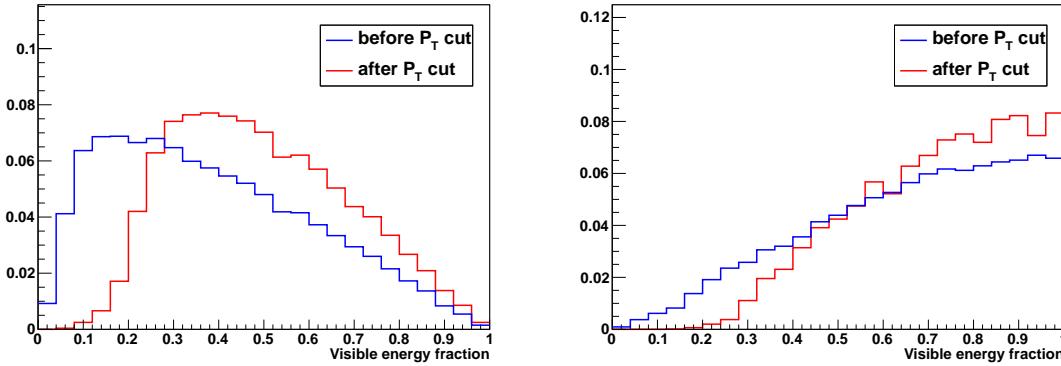


Figure 4.3: Normalized distributions of the fraction of total tau decay energy carried by the muon (left) and hadronic constituents (right) in simulated Higgs events with  $m_{A^0} = 130 \text{ GeV}/c^2$ . The distribution is shown before (blue) and after (red) the requirement on the  $p_T$  of the visible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

(fig:ptBalancePtVisCuts)

## 1474 §4.5 Performance

1475 The tau pair mass reconstructed by the Secondary Vertex fit (“SVfit mass”) provides  
 1476 the observable with the largest separation between signal Higgs events and the dominant  
 1477  $Z \rightarrow \tau^+ \tau^-$  background. The mean of the SVfit mass is located at the true mass of the di-tau  
 1478 pair. The SVfit algorithm has a higher acceptance and better resolution than the Collinear  
 1479 Approximation algorithm. The SVfit always finds a physical solution, improving on the effi-  
 1480 ciency of the collinear approximation by a factor of two. Additionally, it has a much better  
 1481 resolution. The collinear approximation reconstructed mass distribution has a large tail at  
 1482 high mass due to events with poorly measured  $E_T^{\text{miss}}$ . The shape of the SVfit distribution  
 1483 is nearly Gaussian. The comparison is illustrated in Figure 4.5. Previous searches for Higgs  
 1484 bosons decaying to tau leptons [40] have in general used the “visible mass” as the observ-  
 1485 able used to search for new resonances. The SVfit method has the obvious difference that it  
 1486 reconstructs the “full” tau pair mass, which is the most natural observable corresponding  
 1487 to a particle decaying to tau leptons. In addition, the relative resolution<sup>4</sup> of the SV fit is  
 1488 superior to that of the visible mass. This feature is illustrated in Figure 4.6. In Figure 4.6,  
 1489 the visible mass distribution is scaled by an arbitrary number such that the scaled mean  
 1490 of the distribution matches the true mass of the tau pair (and the SVfit mass). The width

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

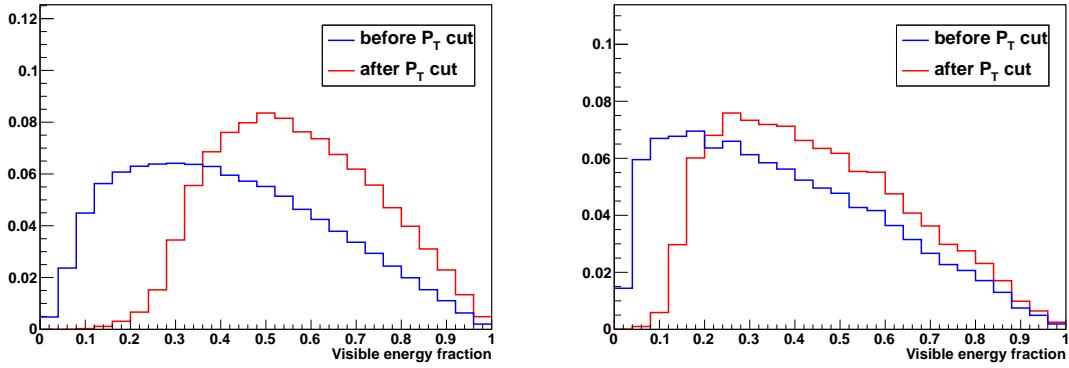


Figure 4.4: Normalized distributions of the fraction of total tau decay energy carried by the muon in simulated  $Z \rightarrow \tau^+\tau^-$  (left) and Higgs events with  $m_{A^0} = 200$   $\text{GeV}/c^2$  (right). The distribution is shown before (blue) and after (red) the requirement that the  $p_T$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

PtVisCutsCompareMasses

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

1491     of the SVfit distribution is smaller than that of the scaled visible mass distribution, indi-  
 1492     cating better performance. The increase in relative resolution allows a “bump,” due to the  
 1493     presence of signal events, to be more easily distinguished from the  $Z \rightarrow \tau^+\tau^-$  background.  
 1494     This increases the power of the search for the new signal.

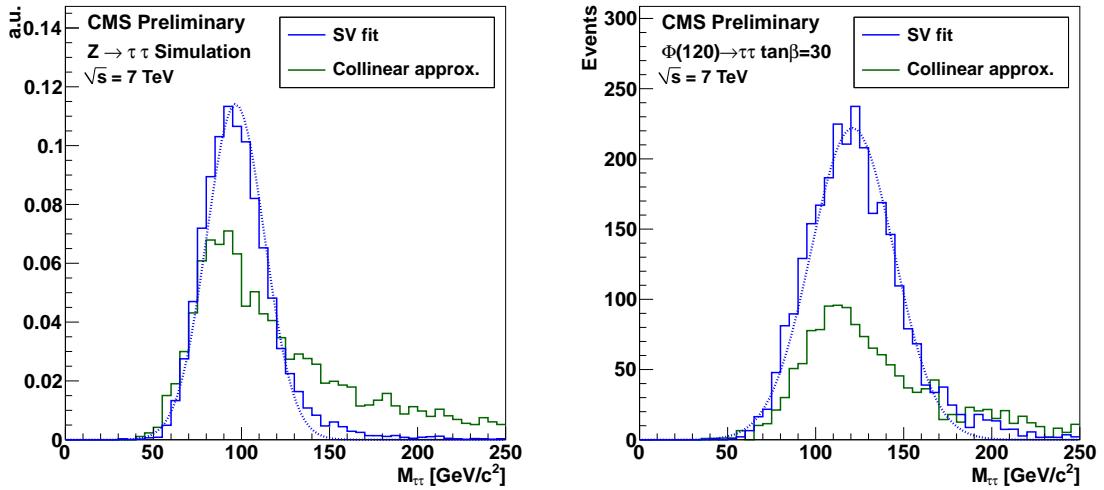


Figure 4.5: Comparison of the reconstructed tau pair mass spectrum in  $Z \rightarrow \tau^+\tau^-$  (left) and MSSM  $H(120) \rightarrow \tau^+\tau^-$  (right) events after the selections described in chapter 5. The mass spectrum reconstructed by the Secondary Vertex fit is shown in blue, the result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

`(fig:SVversusCollinear)`

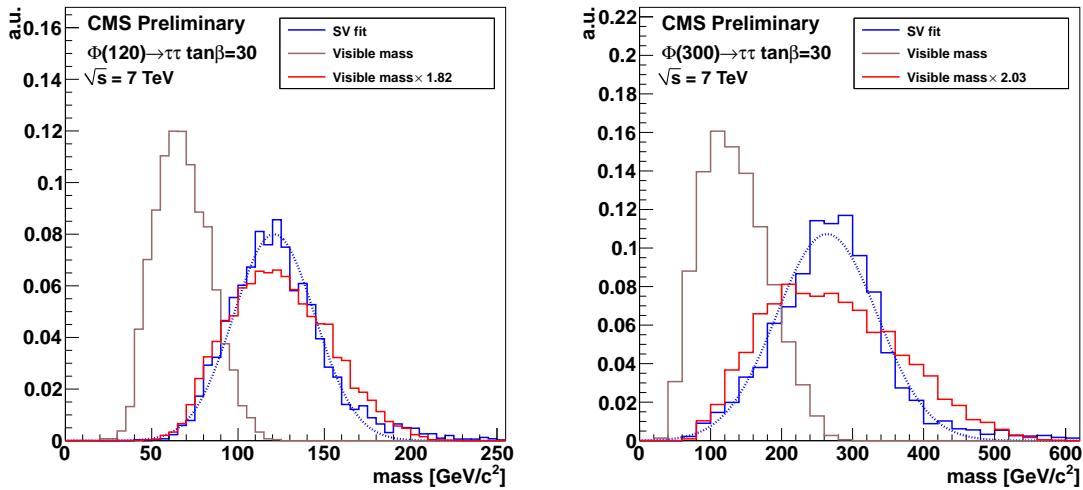


Figure 4.6: Comparison of the invariant mass of the muon and  $\tau_{\text{jet}}$  (the “visible mass”) with the full  $\tau^+\tau^-$  mass reconstructed by the SVfit. The spectrum is shown for two simulated MSSM Higgs samples, with  $m_{A^0} = 120 \text{ GeV}/c^2$  (left), and  $m_{A^0} = 200 \text{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

1495

## Chapter 5

---

1496

# Analysis Selections

&lt;ch:selections&gt;

1497 The selections applied to data in this analysis are designed to maximize the significance of  
 1498 Higgs signal events in the final set of selected events. This analysis presented in this thesis  
 1499 is an inclusive analysis, meaning that no preference is given to any single Higgs production  
 1500 mechanism. The analysis looks specifically at the channel in which one tau decays to a  
 1501 muon and the other decays to hadrons. Therefore the first step in the analysis selection is  
 1502 to find High Level Trigger selection that is highly efficient for the signal and is not highly  
 1503 prescaled<sup>1</sup>. After the trigger selection, events are required to contain at least a good muon  
 1504 and a good tau. Vetoos on extra leptons are applied to reduce backgrounds from di-muon  
 1505 events. Finally, kinematic and charge selections on the are applied to the event to reduce  
 1506  $W + \text{jets}$  and QCD backgrounds.

1507

## §5.1 High Level Trigger

1508 As only data which passes the HLT is recorded, it is thus critical that an appropriate trig-  
 1509 ger path is found. The events in this analysis are triggered by a combination of muon and  
 1510 muon + tau-jet “cross-channel” triggers. For the muon triggers, the HLT paths with low-  
 1511 est  $p_T$  thresholds are used as long as the path remained unprescaled (see Table 5.1). The  
 1512 muon + tau-jet “cross-channel” trigger paths increase the trigger efficiency for events con-  
 1513 taining muons of transverse momenta close to the  $p_T^\mu > 15$  GeV/c cut threshold. The trigger  
 1514 efficiency is measured in data via the tag-and-probe technique. Details of the muon trigger  
 1515 efficiency measurement are given in Section 7.1. Monte Carlo simulated events are required

---

<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–jet “cross–channel” trigger paths utilized to trigger events in different data–taking periods.

{tab:AHToMuTauTriggers}

1516 to pass the HLT\_Mu9 trigger path. Weights are applied to simulated events to account for the  
 1517 difference between the simulated HLT\_Mu9 efficiency and the combined efficiency of the set  
 1518 HLT\_Mu9, HLT\_IsoMu9, HLT\_Mu11, HLT\_IsoMu13, HLT\_Mu15, HLT\_IsoMu9\_PFTau15  
 1519 and HLT\_Mu11\_PFTau15 used to trigger the data.

## 1520 §5.2 Particle Identification

### 1521 §5.2.1 Muons

1522 {sec:MuonId} Muon candidates are required to be reconstructed as global and as tracker muons, meaning  
 1523 that a full track is reconstructed in the muon system and is well matched to a track in the  
 1524 silicon strip and pixel trackers. Additionally, they are required to pass the “Vector Boson  
 1525 Task Force” (VBTF) muon identification criteria developed for the  $Z \rightarrow \mu^+ \mu^-$  cross–section  
 1526 measurement [44]:

- 1527 •  $\geq 1$  Pixel hits
- 1528 •  $\geq 10$  hits in silicon Pixel + Strip detectors
- 1529 •  $\geq 1$  hit(s) in muon system
- 1530 •  $\geq 2$  matched segments

- 1531     •  $\chi^2/DoF < 10$  for global track fit
- 1532     • transverse impact parameter of “inner” track  $d_{\text{IP}} < 2$  mm with respect to beam-spot

1533     In order to reduce background contributions from muons originating from heavy quark  
 1534     decays in QCD multi-jet events, muons are required to be isolated. Isolation is computed  
 1535     as the  $p_{\text{T}}$  sum of charged and neutral hadrons plus photons reconstructed by the CMS  
 1536     particle-flow algorithm [29] within a cone of size  $\Delta R_{iso} = 0.4$  around the muon direction  
 1537     divided by the muon  $p_{\text{T}}$ . The innermost region of size  $\Delta R_{veto} = 0.08$  (0.05) is excluded  
 1538     from the computation of the isolation  $p_{\text{T}}$  sum with respect to neutral hadrons (photons),  
 1539     in order to avoid energy deposits in the electromagnetic and hadronic calorimeters which  
 1540     are due to the muon to enter the sum. In order to reduce pile-up effects, particles entering  
 1541     the isolation  $p_{\text{T}}$  sum are required to have transverse momenta  $p_{\text{T}} > 1.0$  GeV/c. Charged  
 1542     particles are additionally required to originate from the same vertex as the muon. The  
 1543     muons are required to be isolated with respect to charged hadrons of  $p_{\text{T}} > 1.0$  GeV/c and  
 1544     photons of  $p_{\text{T}} > 1.5$  GeV/c as reconstructed by the particle-flow algorithm [29] in a cone  
 1545     of size  $\Delta R = 0.4$  around the direction of the muon. The distribution of the muon isolation  
 1546     discriminant is shown in Figure 5.1(a).

### 1547     §5.2.2 Hadronic Taus

1548     Hadronic decays of taus are identified by the HPS + TaNC hybrid algorithm described in  
 1549     Section 3.5. The “loose” working point is used, corresponding to an expected QCD fake-rate  
 1550     of about 1%.  $Z \rightarrow \mu^+ \mu^-$  background contributions are largely due to muons which failed  
 1551     to get reconstructed as global muons (thus failing the muon identification requirement) and  
 1552     are misidentified as tau-jet candidates. These muons are typically isolated and have a large  
 1553     chance to pass the hadronic tau ID discriminators. To reject these events, hadronic taus are  
 1554     additionally required to pass an anti-muon veto described in Section 3.6.

### 1555     §5.2.3 Missing Transverse Energy

1556     The missing transverse energy  $E_{\text{T}}^{\text{miss}}$ , in the event is reconstructed based on the vectorial  
 1557     momentum sum of particle candidates reconstructed by the particle-flow algorithm [29, 45].  
 1558     In the ideal case, the  $E_{\text{T}}^{\text{miss}}$  corresponds to the vector sum of the transverse components of

Background	Cross Section ( pb)
QCD Heavy Flavor	84679 <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.

(tab:FakeBackgrounds)

1559 all neutrinos in the event. The  $E_T^{\text{miss}}$  resolution in simulated  $Z \rightarrow \mu^+\mu^-$  events is found  
 1560 to be smaller (better) than in the data. The reconstructed  $E_T^{\text{miss}}$  in the simulated events is  
 1561 “smeared” by a correction factor such that the data and simulation are in agreement. The  
 1562 “Z-recoil”  $E_T^{\text{miss}}$  correction procedure is described in Section 7.4.

### 1563 §5.3 Event Selections

1564 The selections applied to the analysis are designed to reject large fractions of the background  
 1565 while maintaining a high efficiency for identifying signal Higgs events. The backgrounds  
 1566 can be divided into two classifications: “fake” backgrounds, in which there is at least one  
 1567 misidentified hadronic tau decay, and the irreducible  $Z \rightarrow \tau^+\tau^-$  background, which cannot<sup>2</sup>  
 1568 be distinguished from the potential presence of a Higgs boson of the same mass. Strategies  
 1569 for dealing with the irreducible  $Z$  background will be discussed in the Chapter 9. The  
 1570 different fake backgrounds, their cross section, and the basic removal strategies are outlined  
 1571 in Table 5.2.

1572 Events are selected by requiring a muon of  $p_T^\mu > 15$  GeV/c within  $|\eta_\mu| < 2.1$  and a tau-  
 1573 jet candidate of  $p_T^{\tau-\text{jet}} > 20$  GeV/c within  $|\eta_{\tau-\text{jet}}| < 2.3$ . The  $\eta$  requirement on the muon  
 1574 ensures that it is within the fiducial region of the muon trigger system. The  $\eta$  requirement  
 1575 on the hadronic tau ensures it is well within the fiducial region of the tracker ( $|\eta| < 2.5$ )  
 1576 and minimizes exposure to large QCD backgrounds in the very forward region.

---

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

1577        The muon and tau–jet candidate are required to be of opposite charge, as the Higgs is  
 1578        neutral and charge is conserved. The muon is required to be pass the identification criteria  
 1579        described in Section 5.2.1. The tau–jet candidate is required to pass the “loose” TaNC tau  
 1580        identification discriminator.

1581        Additional event selection criteria are applied to reduce contributions of specific back-  
 1582        ground processes. In order to reject  $Z \rightarrow \mu^+ \mu^-$  background, a dedicated discriminator  
 1583        against muons is applied [28]. Remaining di-muon background is suppressed by reject-  
 1584        ing events which have a track of  $p_T > 15$  GeV/c and for which the sum of energy deposits  
 1585        in ECAL plus HCAL is below  $0.25 \cdot P$  within a cylinder of radius 15 cm (ECAL) and  
 1586        25 cm (HCAL), respectively. Contamination from  $Z \rightarrow \tau^+ \tau^-$  events in which the recon-  
 1587        structed tau–jet candidate is due to a  $\tau \rightarrow e \nu \nu$  decay is reduced by applying a dedicated  
 1588        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

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

1589        where  $\Delta\phi$  is the angle between the muon and the reconstructed  $E_T^{\text{miss}}$  in the transverse  
 1590        plane. The  $M_T$  quantity is much higher in events  $W \rightarrow \mu\nu$  decays than in signal Higgs  
 1591        events. In  $W \rightarrow \mu\nu$  decays, the neutrino expected to be produced in the opposite to the  
 1592        muon in azimuth. In signal events, there are three neutrinos produced, with the majority  
 1593        (two) of the neutrinos being associated to the  $\tau \rightarrow \mu\nu\nu$  decay. Accordingly, we expect  
 1594        that the  $E_T^{\text{miss}}$  is on average collinear with the muon in signal events. The  $M_T$  distribution  
 1595        immediately before the  $M_T$  cut is applied is illustrated in Figure 5.1(b) for the different  
 1596        background sources and 2010 data.

1597        The  $P_\zeta$  variable is another quantity with discriminant power against  $W + \text{jets}$  and  
 1598         $t\bar{t}$  backgrounds. The observable has been introduced in the search for  $H \rightarrow \tau^+ \tau^-$  events  
 1599        performed by the CDF collaboration [46]. The observable is motivated by the fact that in  
 1600         $\tau^+ \tau^-$  signal events all neutrinos are produced in direction very close of the visible decay  
 1601        products of one of the two tau leptons. It is therefore expected that the direction of the  
 1602        missing transverse energy vector in these events point in a direction somewhere between

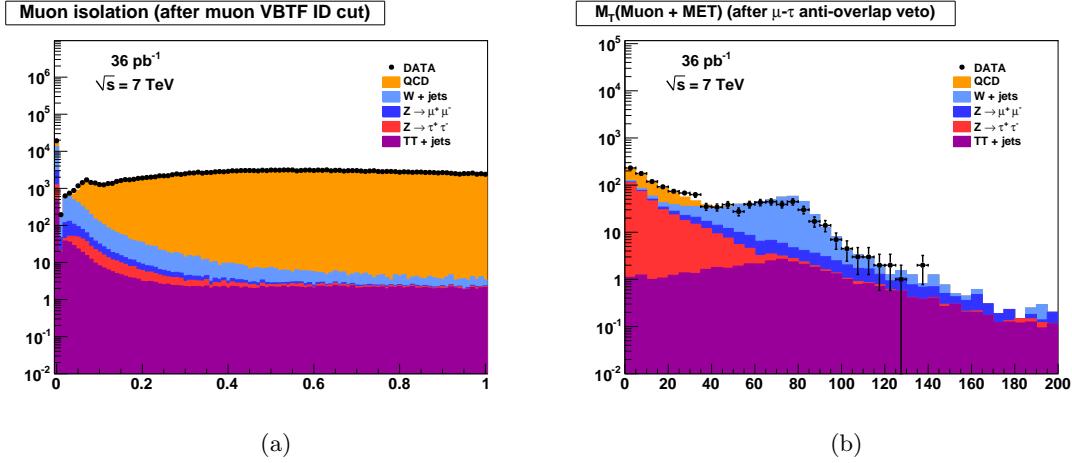


Figure 5.1: Distributions of the muon isolation (a) and  $M_T$  (b) discriminant variables. The muon isolation discriminant rejects the QCD background at a high rate. The  $M_T$  cut is designed to reject  $W + \text{jets}$  and  $t\bar{t} + \text{jets}$  backgrounds. The distributions shown are computed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

1603 the visible  $\tau$ -lepton decay products. This event topology is not preferred in  $W + \text{jets}$ ,  $t\bar{t}$   
 1604 and QCD background events. The observable is computed as difference of the projections:

$$P_\zeta = \vec{P}_T^{vis_1} + \vec{P}_T^{vis_2} + E_T^{\text{miss}}$$

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.

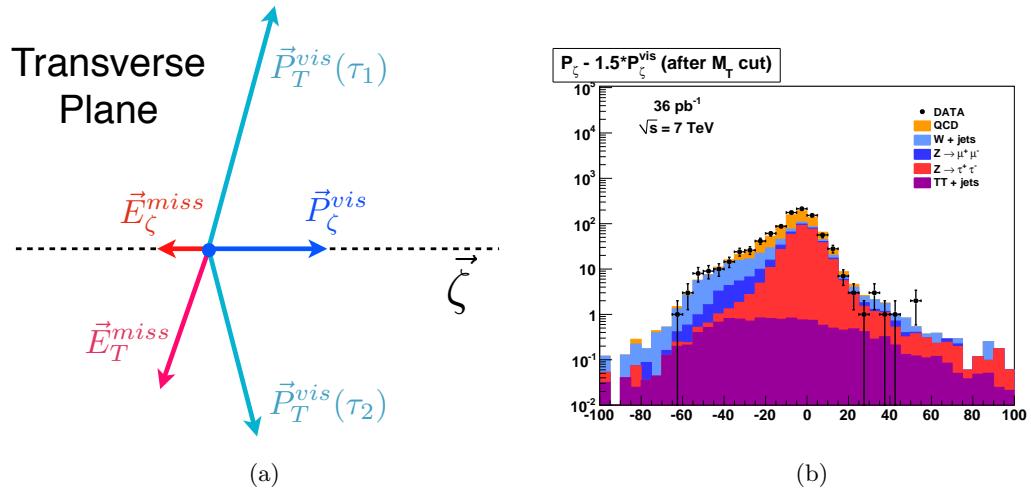


Figure 5.2: The vector quantities used in construction of the quantity  $P_\zeta - 1.5 \cdot P_\zeta^{\text{vis}}$  are illustrated in Figure (a). Image credit: [46] The distribution of the  $P_\zeta$  variable in the different background sources and 2010 data after the  $M_T$  cut has been applied is shown in Figure 5.2(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-jet Candidate	reconstructed by HPS + TaNC combined Tau ID algorithm TaNC “medium” Tau ID discriminator and discriminators against electrons and muons passed, calorimeter muon rejection passed
Muon + Tau-jet	$\text{charge}(\text{Muon}) + \text{charge}(\text{Tau-jet}) = 0$ , $\Delta R(\text{Muon}, \text{Tau-jet}) > 0.5$
Kinematics	$M_T(\text{Muon-MET}) < 40 \text{ GeV}$ $P_\zeta - 1.5 \cdot P_\zeta^{\text{vis}} > -20 \text{ GeV}$

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

1609

## Chapter 6

1610

# Data–Driven Background Estimation

&lt;ch:backgrounds&gt;

1611 For the result of this analysis to be reliable, it is of paramount importance that the back-  
 1612 grounds be well understood. The CMS experiment has adopted a policy that if possible, all  
 1613 background processes should be measured in a “data–driven” way. By requiring that the  
 1614 background comes from data, biases due to incorrectly modeling the background processes  
 1615 in simulation can be minimized or eliminated. In general, the data–driven methods also have  
 1616 the advantage that they are independent of the uncertainty on the integrated luminosity.  
 1617 This analysis measures the backgrounds using two complementary methods, the “Template  
 1618 Method” and the “Fake–rate method.” In both cases, predictions are made about back-  
 1619 grounds in the signal region using measurements obtained in background enriched control  
 1620 regions of the data. The Template Method fits the sum of background shape templates to  
 1621 the  $M_{vis}$  spectrum of events selected in the final analysis and is described in Section 6.3.  
 1622 The Fake–rate Method is based on applying probabilities for quark and gluon jets to be  
 1623 misidentified as hadronic tau decays to events passing all event selection criteria except  
 1624 the tau identification requirements. The probabilities with which jets fake hadronic tau sig-  
 1625 natures are measured in data. Contrary to the Template Method, The Fake–rate Method  
 1626 estimates the sum of the contributions of backgrounds that contain incorrectly identified  
 1627 taus. The Fake–rate method is detailed in Section 6.2. The two methods are complementary  
 1628 as the Template Method uses only information about the different visible mass distribution  
 1629 shapes of the backgrounds, while the Fake–rate method uses only information about the  
 1630 hadronic tau fake–rate.

## 1631 §6.1 Background Enriched Control Regions

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

1642 All control regions are selected from the 2010 CMS muon primary datasets using single  
 1643 muon HLT trigger paths. The set of triggers and run–ranges used to select events in the  
 1644 background enriched control samples is the same as for the analysis (see Table 5.1). The  
 1645 Monte Carlo simulated events used for comparison with the control region selections are  
 1646 required to pass the HLT\_Mu9 trigger path and are weighted according to the description  
 1647 in Chapter 7 to account for the difference in efficiency between HLT\_Mu9 and the trigger  
 1648 paths required to have passed in the data.

1649 QCD di–jet events containing a muon (originating from the leptonic decay of a  $b$  or  $c$   
 1650 quark) are selected by applying an *anti*–isolation requirement on the jet containing a muon.  
 1651  $W + \text{jets}$  and  $t\bar{t} + \text{jets}$  are selected by requiring an isolated muon, and inverting the transverse  
 1652 mass ( $M_T$ ) and  $P_\zeta$  selections. Tau–jet candidates considered in the  $Z \rightarrow \mu^+ \mu^-$  sample where  
 1653 the reconstructed tau–jet candidate is faked by a misidentified muon and in the  $t\bar{t} + \text{jets}$   
 1654 control sample are required to pass the “loose” TaNC discriminator. For the Template  
 1655 Method, the  $Z \rightarrow \mu^+ \mu^-$  sample where the reconstructed tau–jet candidate is faked by a  
 1656 misidentified quark or gluon jet, the  $W + \text{jets}$  and the QCD enriched control samples have a  
 1657 loose hadronic tau “preselection” applied. The tau–jet candidates are required to pass the  
 1658 “very loose”, but fail the “loose” TaNC discriminator. The criteria applied to select events  
 1659 in the different background enriched control samples are summarized in Table 6.1. The goal

Requirement	Enriched background process				
	$Z \rightarrow \mu^+ \mu^-$		$W + \text{jets}$	$t\bar{t} + \text{jets}$	QCD
	Muon fake	Jet fake			
Muon rel. iso.	< 0.15	< 0.1	< 0.1	< 0.1	> 0.10 && < 0.30
Muon Track IP	-	-	-	-	-
Tau TaNC discr.	-	1	1	medium passed	1
Tau 1  3-Prong	-	-	-	-	-
Charge(Tau) = $\pm 1$	-	-	-	-	-
Tau $\mu$ -Veto	inverted	applied	applied	applied	applied
Charge(Muon+Tau)	applied	-	-	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
central Jet Veto	-	-	2	-	-
b-Tagging	-	-	-	3	-

<sup>1</sup> vloose passed && loose failed <sup>2</sup> no Jets of  $E_T > 20$  GeV within  $|\eta| < 2.1$  (other than the  $\tau$ -jet candidate)

<sup>3</sup> min. two Jets of  $E_T > 40$  GeV, at least one of which with  $E_T > 60$  GeV and at least of which with “TrackCountingHighEff” discriminator  $> 2.5$

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

MuTauBgControlRegions

1660 of the background enriched selection process is to select different background processes with  
 1661 high purity. A highly pure background control sample improves the stability of inferences  
 1662 about the signal region made using information in the enriched control region. The purity  
 1663 of the control regions (estimated using simulation) are summarized in Table 6.2.

1664 The number of events observed in the different control samples is compared to the  
 1665 Monte Carlo expectation in table 6.2. Except for the contribution of  $Z \rightarrow \mu^+ \mu^-$  events  
 1666 in which the reconstructed tau-jet candidate is due to a misidentified quark or gluon jet,  
 1667 good agreement between data and Monte Carlo simulation is observed. Differences observed  
 1668 between data and simulation will be accounted for as systematic uncertainties.

1669 The distributions of visible and “full”  $\tau^+ \tau^-$  invariant mass reconstructed by the SVfit  
 1670 algorithm (see Chapter 4) observed in the background enriched control regions is compared

Enriched Selection	Data	Contribution from						Purity
		$\Sigma$ SM	$Z \rightarrow \tau^+\tau^-$	$Z \rightarrow \mu^+\mu^-$	$W + \text{jets}$	$t\bar{t} + \text{jets}$	QCD	
$Z \rightarrow \mu^+\mu^-$								
Muon fake	15156	17109.8	331.6	16586.6	55.1	80.4	35.0	96.9%
Jet fake	85	62.7	2.5	55.5	0.5	1.4	2.4	88.5%
$W + \text{jets}$	514	642.4	17.9	22.9	581.7	0.8	16.7	90.6%
$t\bar{t} + \text{jets}$	26	39.7	0.7	< 0.1	0.6	38.4	< 1.0	96.7%
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  $Z \rightarrow \tau^+\tau^-$ ,  $Z \rightarrow \mu^+\mu^-$ ,  $W + \text{jets}$ ,  $t\bar{t} + \text{jets}$  and QCD processes. The expected purity of each control sample is computed as the ratio of contribution of the enriched process to  $\Sigma$  SM.

MuTauBgControlRegions)

to the Monte Carlo simulation in Figures 6.1 and 6.2. The template for the  $W + \text{jets}$  background has been corrected for the bias on the  $M_{vis}^{\mu\tau_{had}}$  shape caused by the  $M_T^{\mu E_T^{\text{miss}}} < 50 \text{ GeV}/c^2$  and  $P_\zeta - 1.5 \cdot P_\zeta^{vis} > -20 \text{ GeV}$  requirements applied in the final analysis via the reweighting procedure described in [47]. In the  $t\bar{t} + \text{jets}$  enriched control region a peak at the  $Z$  mass is observed in data, which is not modeled by the Monte Carlo samples considered. The peak could be due to  $Z \rightarrow \mu^+\mu^-$  events produced in association with  $b$  quarks. On the other hand, the contribution from  $t\bar{t} + \text{jets}$  events to that sample seems to be overestimated. The origin of the  $Z$  mass peak merits further investigation, but overall the  $t\bar{t} + \text{jets}$  is a negligible background contribution.

## §6.2 The Fake-rate Method

The probabilities with which quark and gluon jets get misidentified as tau-jets may be utilized to obtain an estimate of background contributions in physics analyses. As an illustrative example and in order to demonstrate the precision achievable with the method, we introduce the method in the context of a “closure test,” using a simulated samples, a simple method of computing the fake-rate, and a simpler<sup>1</sup> hadronic tau identification algorithm. The closure test demonstrates that the method is self-consistent, and that the

<sup>1</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].

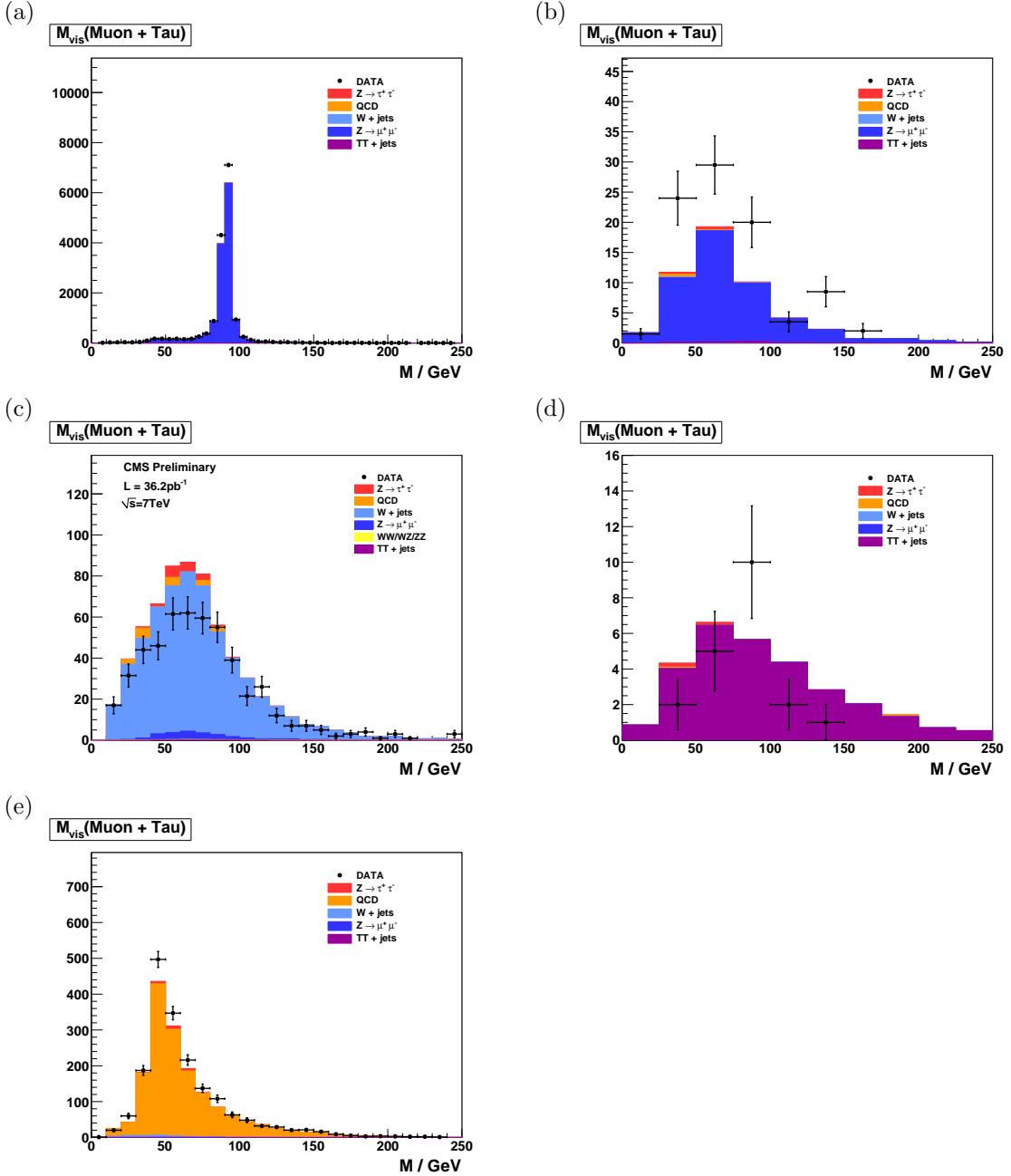


Figure 6.1: Distribution of visible mass of muon plus the tau–jet candidate reconstructed in the background enriched control samples for  $Z \rightarrow \mu^+ \mu^-$  (a) and (b),  $W + \text{jets}$  (c),  $t\bar{t} + \text{jets}$  (d) and QCD multi-jet (e) backgrounds. In (a) reconstructed tau–jet candidates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

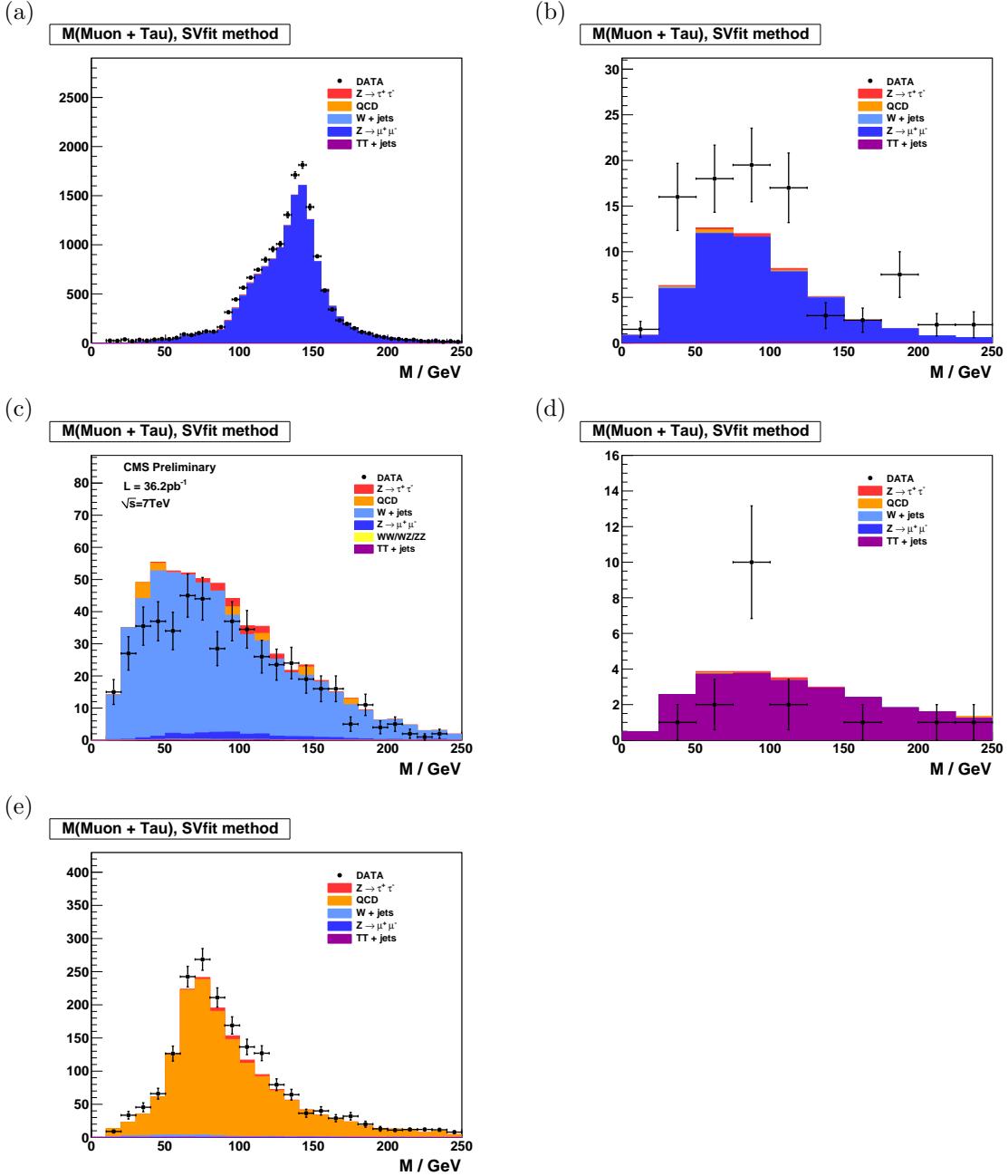


Figure 6.2: Distribution of “full” invariant mass reconstructed by the SVfit algorithm in the background enriched control samples for  $Z \rightarrow \mu^+ \mu^-$  (a) and (b),  $W + \text{jets}$  (c),  $t\bar{t} + \text{jets}$  (d) and QCD multi-jet (e) backgrounds. In (a) reconstructed tau-jet candidates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

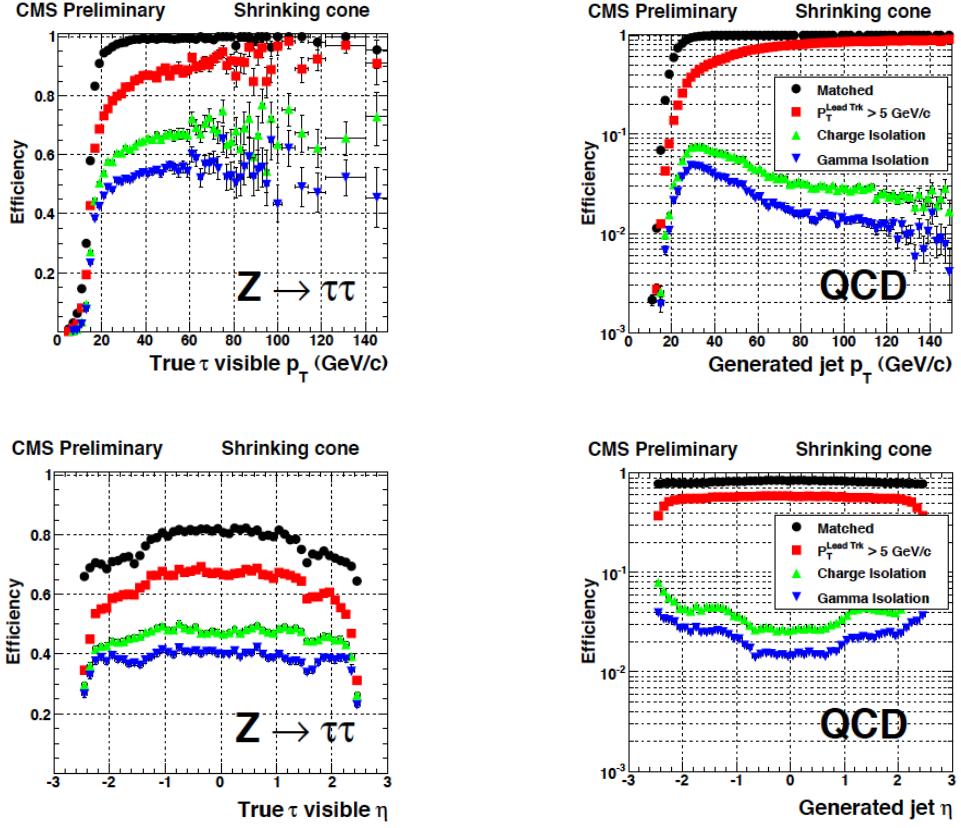


Figure 6.3: Cumulative efficiencies (left) and fake-rates (right) of successively applied tau identification cuts of the “shrinking signal cone” particle-flow based tau identification algorithm described in [28] as function of  $p_T^{jet}$  (top) and  $\eta^{jet}$  (bottom) of tau-jet candidates. The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

EfficienciesAndFakeRates)

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

### 1693 §6.2.1 Parameterization of Fake-rates

FakeRateParametrization)  
 1694 Efficiencies and fake-rates of the tau identification algorithm based on requiring no tracks  
 1695 of  $p_T > 1 \text{ GeV}/c$  and ECAL energy deposits of  $p_T > 1.5 \text{ GeV}/c$  reconstructed within

1696 an “isolation cone” of size  $\Delta R_{iso} = 0.5$  and outside of a “shrinking signal cone” of size  
 1697  $\Delta R_{sig} = 5.0/E_T$  as it is used in the  $Z \rightarrow \tau^+\tau^- \rightarrow \mu + \tau$ -jet analysis [44] are displayed  
 1698 in Figure 6.3. In order to account for the visible  $p_T$  and  $\eta$  dependence, we parametrize  
 1699 the fake-rates in bins of transverse momentum and pseudo-rapidity. As we will show in  
 1700 section 6.2.3, the parametrization of the fake-rates by  $p_T$  and  $\eta$  makes it possible to not  
 1701 only estimate the total number of background events contributing to physics analyses, but  
 1702 to model the distributions of kinematic observables with a precision that is sufficient to  
 1703 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

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

1704 where  $E(\eta^2)$ ,  $E(\phi^2)$  is the second  $\eta$ ,  $\phi$  moment of the jet constituents, weighted by constituent  
 1705 transverse energy. Analyses performed by the CDF collaboration [40, 49, 46] found  
 1706 that systematic uncertainties on background estimates obtained from the fake-rate method  
 1707 are reduced in case differences between quark and gluon jets are accounted for in this way.

### 1708 §6.2.2 Measurement of Fake-rates

Efficiencies and fake-rates are obtained by counting the fraction of tau-jet candidates passing all tau identification cuts and discriminators in a given bin<sup>2</sup> of  $p_T^{jet}$ ,  $\eta_{jet}$  and  $R_{jet}$ :

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

1709 The pre-selection in the denominator of equation 6.1 in general refers to  $p_T$  and  $\eta$  cuts,  
 1710 which are applied with thresholds matching those applied on the final analysis level, but  
 1711 may include loose tau identification criteria (which may be applied e.g. already during event  
 1712 skimming). It is critical that the selection used in the denominator be identical to that of  
 1713 the final analysis to ensure the fake-rates are not biased by different selections.

---

<sup>2</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.

1714        Different sets of fake–rates are determined for the highest  $p_T$  and for the second highest  
 1715         $p_T$  jet in QCD di–jet events, for jets in a QCD event sample enriched by the contribution  
 1716        of heavy quarks and gluons by requiring the presence of a muon reconstructed in the final  
 1717        state, and for jets in “electroweak” events selected by requiring a  $W$  boson in the final state.

### 1718        §6.2.3 Application of Fake–rates

⟨sec:FakeRateApplication⟩  
 1719        Knowledge of the tau identification efficiencies and fake–rates as function of the parameters  
 1720         $p_T^{jet}$ ,  $\eta_{jet}$  and  $R_{jet}$  as defined by equation 6.1 is utilized to obtain an estimate for the contri-  
 1721        butions of background processes to physics analyses involving tau lepton hadronic decays  
 1722        in the final state. The basic idea is to replace tau identification cuts and discriminators by  
 1723        appropriately chosen weights.

1724        Application of the fake–rate technique consists of two stages. The first stage consists of  
 1725        loosening the tau identification cuts and discriminators and applying only the preselection  
 1726        requirements defined by the denominator of Equation 6.1, in order to obtain an event  
 1727        sample dominated by contributions of background processes. After disabling the selections  
 1728        on hadronic tau identification, the relative contributions of the backgrounds are expected  
 1729        to increase by the inverse of the (average) fake–rate, typically by a factor  $\mathcal{O}(100)$ . In the  
 1730        second stage, weights are applied to all events in the background dominated control sample,  
 1731        according to the probabilities  $P_{fr}(p_T^{jet}, \eta_{jet}, R_{jet})$  for jets to fake the signature of a hadronic  
 1732        tau decay. After application of the weights, an estimate for the total number of background  
 1733        events passing the tau identification cuts and discriminators and thus contributing to the  
 1734        final analysis sample is obtained.

1735        The fake–rate technique works best if all background contributions to the analysis arise  
 1736        from misidentification of quark and gluon jets as hadronic tau decays. Corrections to the  
 1737        estimate obtained from the fake–rate technique are needed in case of background processes  
 1738        contributing to the final analysis sample which either produce genuine tau leptons in the  
 1739        final state (e.g.  $t\bar{t} + \text{jets}$ ) or in which tau–jet candidates are due to misidentified electrons  
 1740        or muons (e.g.  $Z \rightarrow \mu^+\mu^-$ ,  $Z \rightarrow e^+e^-$ ), as the latter may fake signatures of hadronic tau  
 1741        decays with very different probabilities than quark and gluon jets.

1742 In the “simple” fake–rate method described in detail in the next section, the correc-  
 1743 tions are taken from Monte Carlo simulations. Corrections based on Monte Carlo are needed  
 1744 also to compensate for signal contributions to the background dominated control sample.  
 1745 An alternative to Monte Carlo based corrections is to utilize additional information con-  
 1746 tained in the background dominated control sample. The modified version is described in  
 1747 section 6.2.5. It has been used to estimate background contributions in searches for Higgs  
 1748 boson production with subsequent decays into tau lepton pairs performed by the CDF col-  
 1749 laboration in TeVatron Run II data [40, 49, 46]. We will refer to the modified version as  
 1750 “CDF–type” method in the following.

#### 1751 §6.2.4 “Simple” weight method

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

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

1752 These weights are applied to all jets in the background dominated control sample which  
 1753 pass the preselection defined by the denominator of Equation 6.1. Note that the weights  
 1754 defined by Equation 6.2 can be used to estimate the contributions of background processes  
 1755 to distributions of tau–jet related observables. They cannot be used as event weights.

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

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

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

Background Process	Expectation	Estimate obtained by applying weights of type:				Average fake-rate estimate
		QCD lead jet	QCD second jet	QCD $\mu$ -enriched	$W + \text{jets}$	
$W + \text{jets}$	$163.0 \pm 7.1$	$157.2 \pm 2.8$	$140.9 \pm 2.7$	$129.9 \pm 2.5$	$177.9 \pm 3.2$	$151.5^{+26.6}_{-21.8}$
QCD	$246.4 \pm 31.8$	$269.2 \pm 14.0$	$246.5 \pm 14.3$	$219.7 \pm 11.8$	$300.8 \pm 15.2$	$259.1^{+44.9}_{-41.7}$
$t\bar{t} + \text{jets}$	$12.2 \pm 0.6$	$14.3 \pm 0.3$	$12.6 \pm 0.3$	$11.6 \pm 0.3$	$16.5 \pm 0.3$	$13.8^{+2.7}_{-2.2}$
$Z \rightarrow \mu^+ \mu^-$	$68.6 \pm 2.9$	$58.2 \pm 1.3$	$51.2 \pm 1.2$	$48.5 \pm 1.1$	$65.8 \pm 1.4$	$55.9^{+10.0}_{-7.5}$
$\Sigma$ Background	$490.4 \pm 32.7$	$499.9 \pm 14.4$	$451.2 \pm 14.6$	$409.7 \pm 12.1$	$561.1 \pm 15.6$	$480.2^{+82.7}_{-71.9}$
$Z \rightarrow \tau^+ \tau^-$	—	$284.3 \pm 3.7$	$269.0 \pm 3.9$	$256.5 \pm 3.3$	$325.3 \pm 4.2$	$283.3^{+42.2}_{-27.1}$

Table 6.3: 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  $Z \rightarrow \tau^+ \tau^- \rightarrow \mu + \tau$ -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.

FakeRate'frSimpleResults`

“composite particles”, depending on  $p_T^{jet}$ ,  $\eta_{jet}$ ,  $R_{jet}$  of its tau-jet candidate constituent, according to:

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

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

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

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

1767

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

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

### §6.2.5 “CDF–type” weights

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

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

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

as:

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

1794 For the derivation of equation 6.5 for the “CDF–type” weights assigned to tau–jet candi-  
 1795 dates, we will use the following notation: Let  $n_\tau$  ( $n_{QCD}$ ) denote the total number of tau–jets  
 1796 (quark and gluon jets) in a certain bin of transverse momentum  $p_T^{jet}$ , pseudo–rapidity  $\eta_{jet}$   
 1797 and jet–width  $R_{jet}$  and  $n_\tau^{sel}$  ( $n_{QCD}^{sel}$ ) denote the number of tau–jets (quark and gluon jets)  
 1798 in that bin which pass all tau identification cuts and discriminators. By definition of the  
 1799 tau identification efficiency  $\varepsilon := \varepsilon(p_T^{jet}, \eta_{jet}, R_{jet})$  and fake–rate  $f := f(p_T^{jet}, \eta_{jet}, R_{jet})$ :

$$n_\tau^{sel} = \varepsilon \cdot n_\tau$$

$$n_{QCD}^{sel} = f \cdot n_{QCD}. \quad (6.6) \boxed{\text{eqBgEstFakeRate}}$$

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

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

while averaging to zero for genuine hadronic tau decays:

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

The latter equation yields the relation:

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

1800 associating the two types of weights. By inserting relation 6.8 into equation 6.7 we obtain:

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

and

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

1801 which matches exactly equation 6.5 for the “CDF–type” weights applied to tau–jet candi-  
 1802 dates given in section 6.2.5.

1803 Event weights and the weights assigned to “composite particles” are computed in the

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.

keRate'frCDFtypeResults)

1804 same way as for the “simple” weights, based on the weights assigned to the tau-jet candi-

1805 dates:

$$W_{\text{event}}^{\text{CDF}} := \sum w_{\text{jet}}^{\text{CDF}}$$

$$w_{\text{comp-part}}^{\text{CDF}}(p_{\text{T}}^{\text{jet}}, \eta_{\text{jet}}, R_{\text{jet}}) := w_{\text{jet}}^{\text{CDF}}(p_{\text{T}}^{\text{jet}}, \eta_{\text{jet}}, R_{\text{jet}}), \quad (6.10) \boxed{\text{eqBgEstFakeRate}}$$

1806 where the sums extend over all jets in the background dominated control sample which pass

1807 the preselection defined by the denominator of equation 6.1.

1808 The effect of the negative weights to compensate the positive weights in case the “CDF-  
1809 type” fake-rate method is applied to signal events containing genuine hadronic tau decays is  
1810 shown in Table 6.4 and illustrated in Figure 6.4. As expected, positive and negative weights  
1811 do indeed cancel in the statistical average.

1812 Figures 6.5, 6.6 and 6.7 demonstrate that an unbiased estimate of the background  
1813 contribution by the “CDF-type” weights is maintained. Overall, the estimates obtained  
1814 are in good agreement with the contributions expected for different background processes,  
1815 indicating that the adjustment of negative and positive weights works as expected for the  
1816 background as well.

1817 Results obtained by the “CDF-type” fake-rate method are summarized in table 6.4,

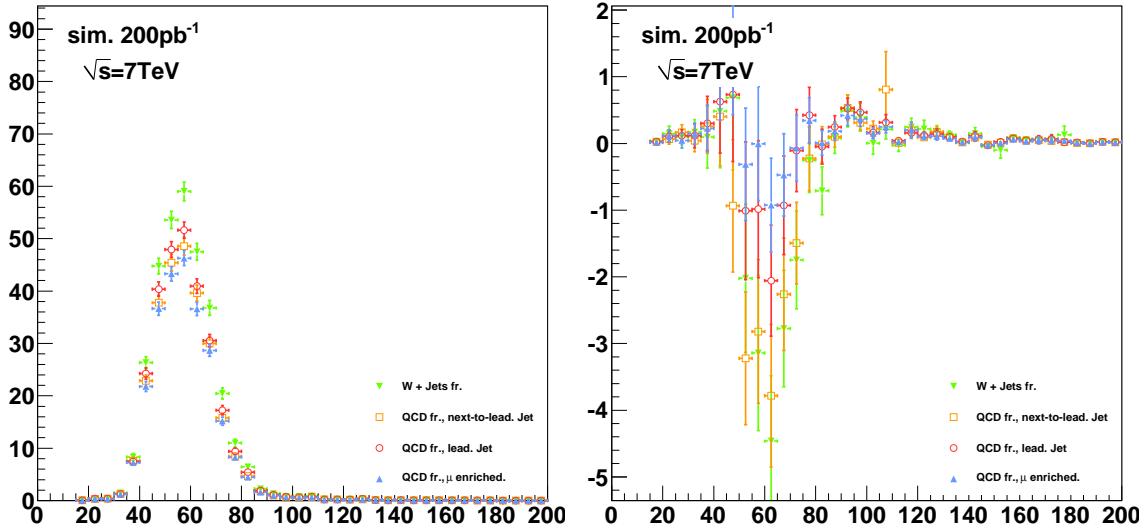


Figure 6.4: Distributions of visible invariant mass of muon plus tau-jet in  $Z \rightarrow \tau^+\tau^-$  signal events weighted by “simple” weights computed according to Equation 6.4 (left) and “CDF-type” weights computed according to Equation 6.10 (right). The signal contribution to the background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

peResults.mVisibleSignal)

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

1829    §6.2.6 k-Nearest Neighbor Fake-rate Calculation

1830     $\langle \text{sec:KNN} \rangle$  For the fake-rate method to give correct results, care must be taken that the measured  
 1831    fake-rate is well defined in all of the regions of phase space where it will be used. In the  
 1832    closure test described above, the computation of the fake-rate was accomplished by binning  
 1833    the numerator (tau ID passed) and denominator (tau ID passed and failed) distributions  
 1834    in the three dimensions of the parameterizations. This method has the disadvantage that  
 1835    the determination of the optimal binning is extremely difficult to determine, and that any  
 1836    bins with no entries in the denominator distribution caused the fake-rate to be undefined  
 1837    in those regions.

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

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

1838    where  $n_{\text{sig}}, n_{\text{bkg}}$  are the observed number of signal and background events, respectively. By  
 1839    construction,  $k = n_{\text{sig}} + n_{\text{bkg}}$ . The principle of operation is illustrated in Figure 6.8

1840    The classification feature of a  $k$ NN can be trivially adapted to parameterize a fake-rate  
 1841    such that it is defined everywhere. Examining the form of Equation 6.11, it is clear that  
 1842    by replacing  $n_{\text{sig}}$  with  $n_{\text{passed}}$  and  $n_{\text{bkg}}$  with  $n_{\text{failed}}$ , the equation is equivalent to the tau-  
 1843    fake rate. We thus “train” the  $k$ NN with tau-candidates which pass the tau identification as  
 1844    signal events and those which fail as background events. The resulting classifier is a function  
 1845    which returns the expected fake-rate for any point in the space of the parameterization.  
 1846    The choice of  $k$  must be optimized. When  $k$  is low, the small number of neighbors causes  
 1847    large counting fluctuations in the fake rate. If  $k$  is too large, the  $k$ NN effectively averages  
 1848    over a large area of the space of the variables<sup>3</sup>. For the training statistics available in the  
 1849    2010 data,  $k = 20$  is found to be the optimal choice.

---

<sup>3</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.

1850    §6.2.7 Results of Background Estimation

1851    An independent estimate of the background contributions to the analysis presented in this  
 1852    thesis is obtained by applying the fake-rate method in a manner analogous to the closure  
 1853    test. Fake-rates in QCD multi-jet events (light quark enriched sample), QCD events con-  
 1854    taining muons (heavy quark and gluon enriched sample) and  $W + \text{jets}$  events are measured  
 1855    in data [34, 48] and applied to events which pass all the event selection criteria listed in  
 1856    table 5.3, with the exceptions of

- 1857         • the “medium” HPS + TaNC discriminator, and  
 1858         • the requirement that the tau have unit charge.

1859         No assumption is made on the composition of  $Z \rightarrow \mu^+ \mu^-$ ,  $W + \text{jets}$ ,  $t\bar{t} + \text{jets}$  and QCD  
 1860    backgrounds contributing to the event sample selected by the analysis. Differences between  
 1861    fake-rates obtained for QCD multi-jet, QCD muon enriched and  $W + \text{jets}$  background events  
 1862    are attributed as systematic uncertainties of the fake-rate method. Per jet and per event  
 1863    weights have been computed by the “simple” and “CDF-type” weights as described in the  
 1864    closure test and the results are found to be compatible within statistical and systematic  
 1865    uncertainties. In the following, we present results for “CDF-type” weights. The “CDF-type”  
 1866    weights have the advantage that the background estimate obtained does not change, whether  
 1867    there is MSSM Higgs  $\rightarrow \tau^+ \tau^-$  signal present in the data or not.

1868         Tau identification efficiencies need to be known when using “CDF-type” weights. Ded-  
 1869    icated studies have checked the tau identification efficiencies in data [48]. Statistical and  
 1870    systematic uncertainties of these studies are still sizeable at present, in the order to 20–30%.  
 1871    No indication has been found, however, that the Monte Carlo simulation does not correctly  
 1872    model hadronic tau decays in data. For the purpose of computing fake-rate weights via  
 1873    the “CDF-type” method, tau identification efficiencies are taken from the Monte Carlo  
 1874    simulation of hadronic tau decays in  $Z \rightarrow \tau^+ \tau^-$  events. Systematic uncertainties on the  
 1875    background estimate obtained by the fake-rate method are determined by varying the tau  
 1876    identification efficiencies by  $\pm 30\%$  relative to the value obtained from the Monte Carlo  
 1877    simulation.

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-jet candidates to pass the “medium” tight TaNC discriminator and have unit charge by fake-rates measured in QCD multi-jet, QCD muon enriched and  $W + \text{jets}$  data samples.

IuTauFakeRateResultsOS)

1878     The results of applying the fake-rate method to the mu + tau channel are summarized  
 1879    in Table 6.5. The background prediction has been corrected for the expected<sup>4</sup> contribution  
 1880    of  $13.1^{+2.8}_{-0.6}$  events from  $Z \rightarrow \mu^+ \mu^-$  background events in which the reconstructed tau-jet is  
 1881    due to a misidentified muon. The obtained estimate is in good agreement with the Monte  
 1882    Carlo expectation.

1883     As an additional cross-check of the method, a sample of events containing a muon  
 1884    plus a tau-jet of like-sign charge is selected in data and compared to the background  
 1885    prediction obtained by applying the fake-rate method to the like-sign sample. The like-sign  
 1886    sample is expected to be dominated by the contributions of  $W + \text{jets}$  and QCD background  
 1887    processes and allows to verify the fake-rate method in a practically signal free event sample.  
 1888    The background estimate obtained by the fake-rate method is compared to the number of  
 1889    events observed in the like-sign data sample in Table 6.6. The number of events expected  
 1890    in the like-sign control sample from Monte Carlo simulation is indicated in the caption. All  
 1891    numbers are in good agreement.

1892     The fake-rate method does not only allow to estimate the total number of background  
 1893    events, but allows to model the distributions of background processes as well. The capability  
 1894    to model distributions is illustrated in Figure 6.9, which shows good agreement between the  
 1895    distributions observed in the like-sign data sample and the predictions obtained by the

---

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

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	216

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

MuTauFakeRateResultsSS)

1896 fake-rate method for the distributions of muon plus tau-jet visible mass and of the “full”  
 1897 invariant mass reconstructed by the SVfit algorithm.

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

### 1898 §6.3 Template method

⟨sec:template⟩ 1899 Shape templates for the  $\mu + \tau_{had}$  visible mass  $M_{vis}$  are obtained from data, using a set of  
 1900 dedicated control regions which are chosen to select a high purity sample of one particular  
 1901 background process each. The number of events selected in each control region and com-  
 1902 parisons to the predictions from Monte Carlo simulations are summarized in Table 6.2. The  
 1903 template  $M_{vis}$  shapes obtained from data in the background enriched control regions are  
 1904 compared to the signal region shapes obtained by Monte Carlo simulation in figure 6.10.  
 1905 The  $M_{vis}$  spectrum observed in the final analysis is fitted to the sum of these templates. Es-  
 1906 timates for background yields are obtained from the normalization factor of each template,  
 1907 determined by the fit. Further details of the method can be found in [47] and [50].

1908 The TaNC (Section 3.3, [51]) discriminators used in [50] are replaced by the correspond-  
 1909 ing discriminators of the HPS + TaNC algorithm (Section 3.5, [33]). The  $Z/\gamma^* \rightarrow \tau^+\tau^-$  sig-  
 1910 nal shape is obtained via the  $Z/\gamma^* \rightarrow \mu^+\mu^-$  embedding technique [52]. The  $\mu + \tau_{had}$  visible  
 1911 mass spectrum observed in the final analysis is compared to the sum of template shapes  
 1912 scaled by the normalization factors determined by the fit in Figure 6.11. The corresponding  
 1913 estimates for background contributions are summarized in Table 6.7.

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

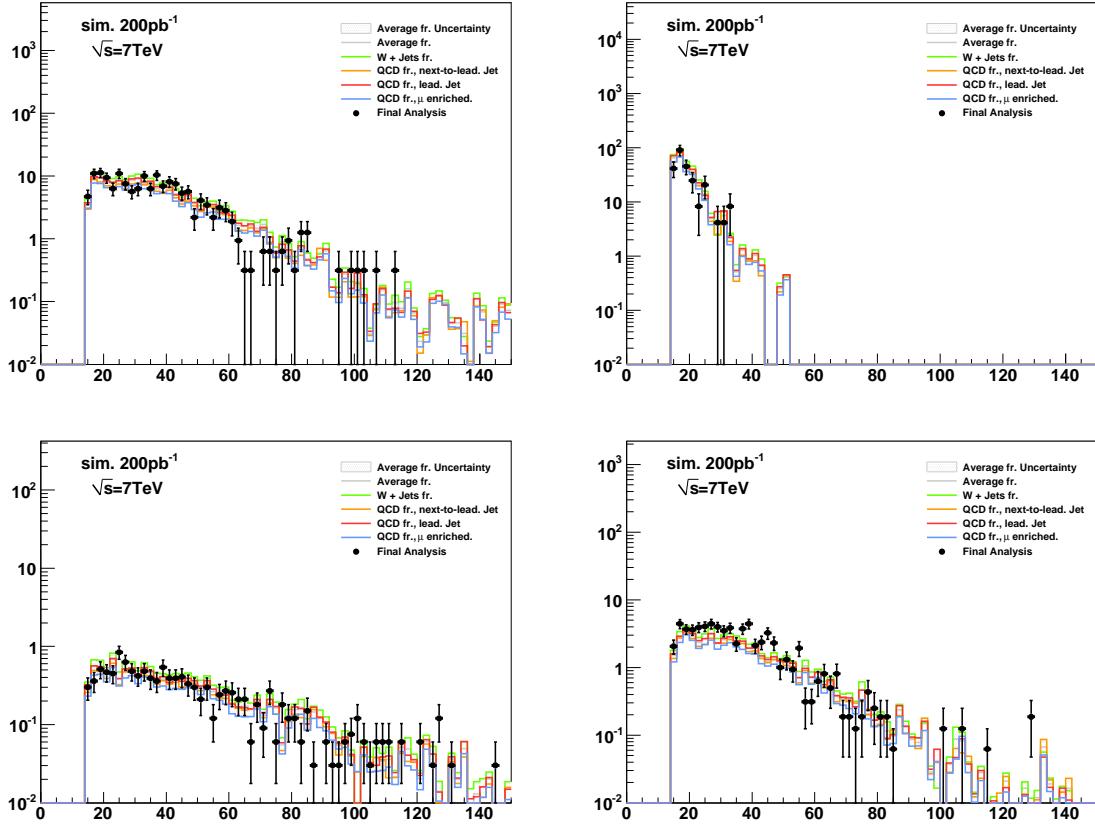


Figure 6.5: Distributions of muon transverse momentum in  $W + \text{jets}$  (top left), QCD (top right),  $t\bar{t} + \text{jets}$  (bottom left) and  $Z \rightarrow \mu^+ \mu^- \rightarrow \mu + \tau\text{-jet}$  cross-section analysis [44] compared to the estimate obtained from the “CDF method” fake-rate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

CDFtypeResults['muonPt']

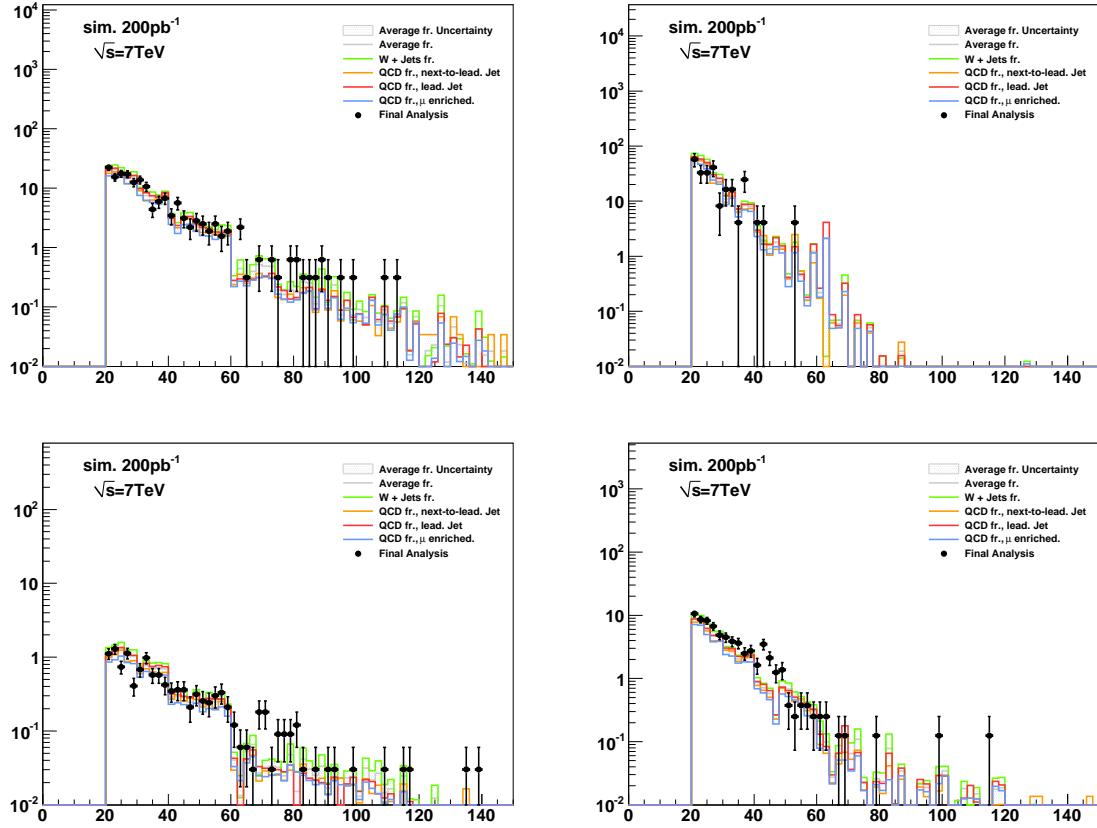


Figure 6.6: Distributions of transverse momenta of the tau-jet candidates in  $W + \text{jets}$  (top left), QCD (top right),  $t\bar{t} + \text{jets}$  (bottom left) and  $Z \rightarrow \tau^+\tau^- \rightarrow \mu + \tau\text{-jet}$  cross-section analysis compared to the estimate obtained from the 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)

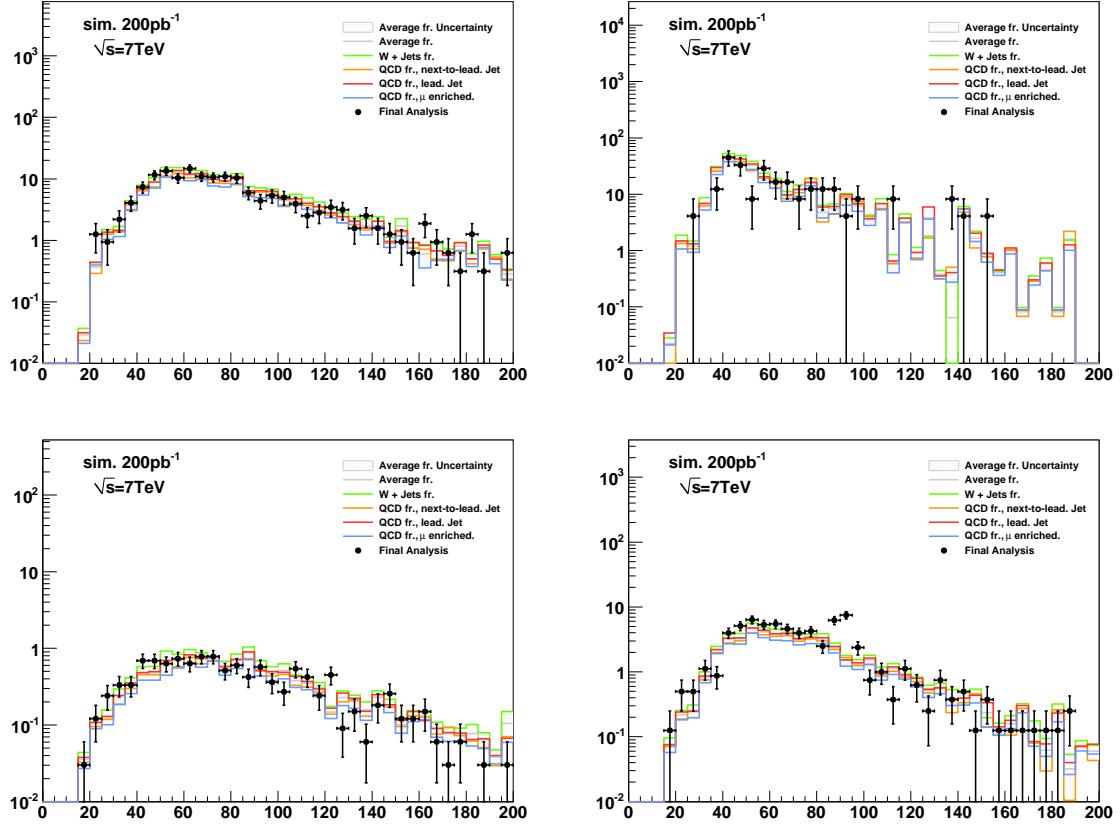


Figure 6.7: Distributions of the visible invariant mass of muon plus tau-jet in  $W + \text{jets}$  (top left), QCD (top right),  $t\bar{t} + \text{jets}$  (bottom left) and  $Z \rightarrow \mu^+\mu^-$  (bottom right) background events which pass all selection criteria of the closure test analysis compared to the estimate obtained from the fake-rate technique, computed according to Equation 6.10. The expected contribution of background processes is indicated by points. Lines of different colors represent the estimates obtained by applying fake-rate weights determined for different compositions of light quark, heavy quark and gluon jets, as described in Section 6.2.1. The maximum (minimum) estimate is interpreted as upper (lower) bound. The difference between the bounds is taken as systematic uncertainty on the estimate obtained from the ‘‘CDF-type’’ fake-rate method and is represented by the gray shaded area.

CDFtypeResults'mVisible)

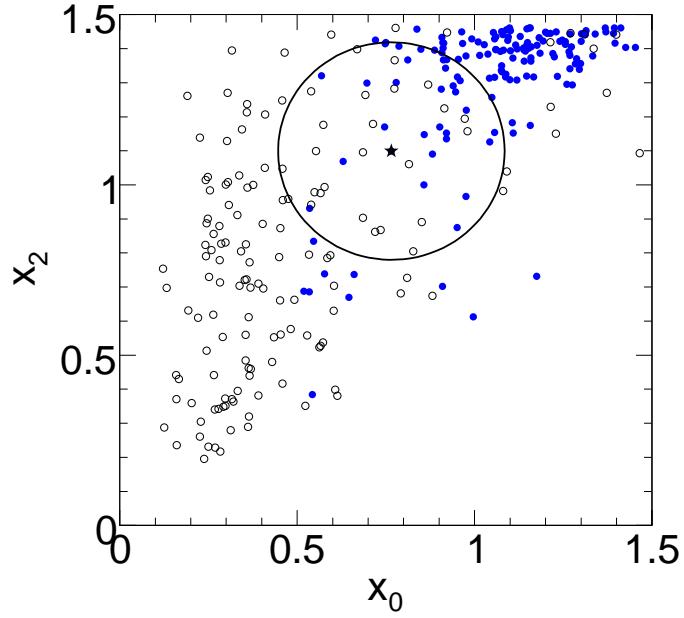


Figure 6.8: Example of the operation of a  $k$ NN classifier. The closest  $k = 50$  neighbors (those inside the circle) to a test point (indicated by the star marker) are selected. The probability that the star marker is a signal event is given the number of signal neighbors (blue markers) in the circle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

(fig:KNN)

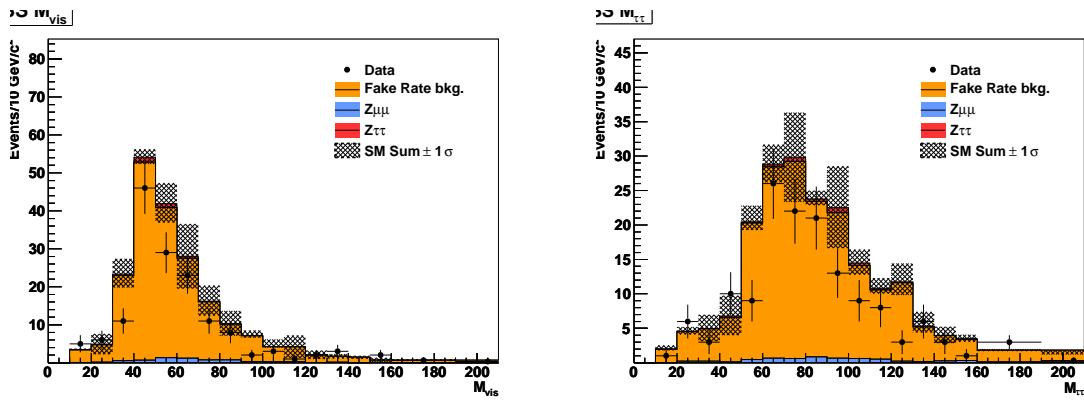


Figure 6.9: Distribution of visible mass (left) and “full” invariant mass reconstructed by the SVfit algorithm (right) observed in the like-sign charge control region compared to the background estimate obtained by the fake-rate method.

MuTauFakeRateResultsSS

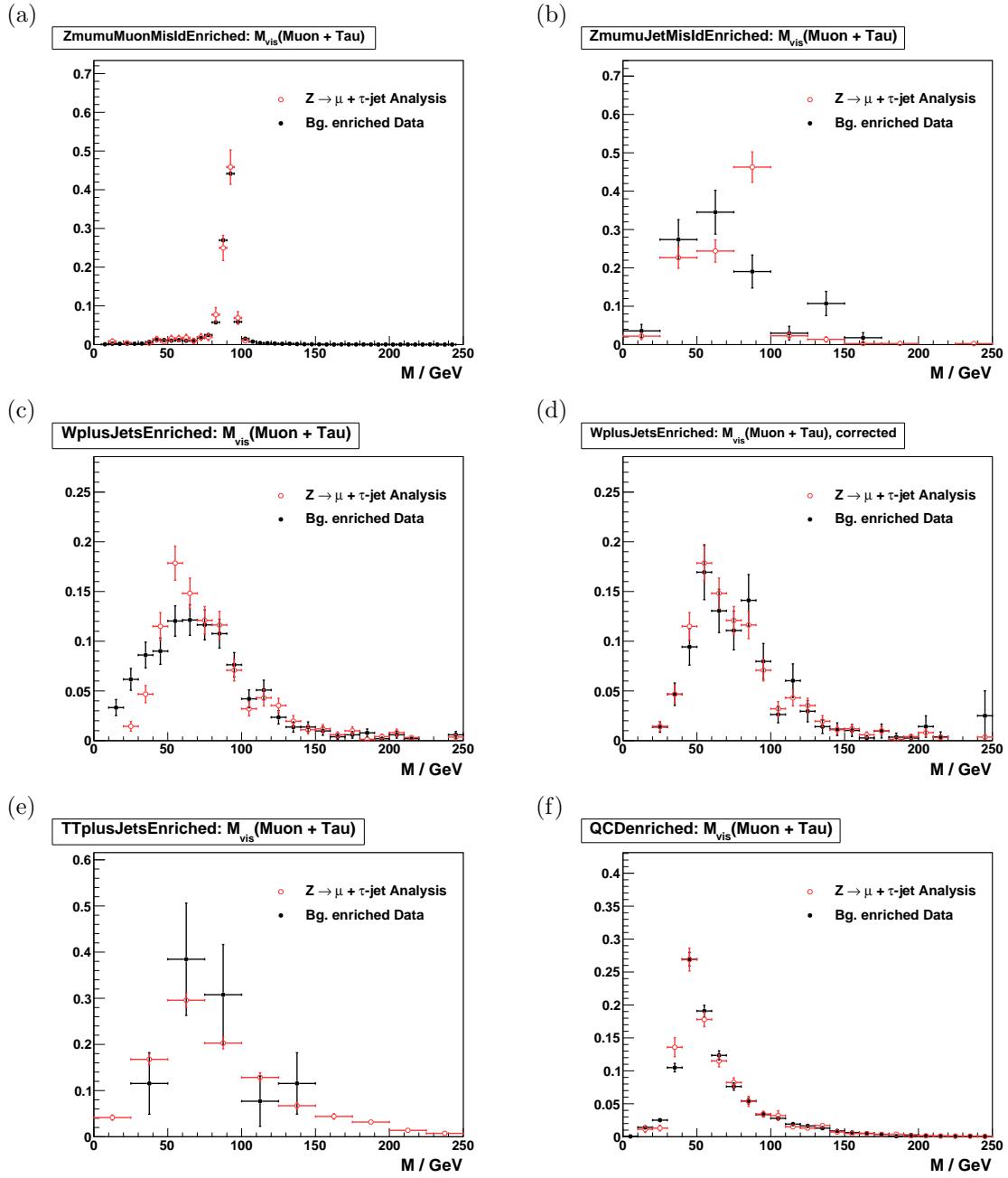


Figure 6.10:  $\mu + \tau_{had}$  shape templates obtained from  $Z \rightarrow \mu^+ \mu^-$  (a) and (b),  $W +$  jets before (c) and after (d) the bias correction explained in Section 6.3,  $t\bar{t} +$  jets (e) and QCD multi-jet (f) backgrounds enriched control regions compared to the expected distribution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

{fig:VisMassTemplates}

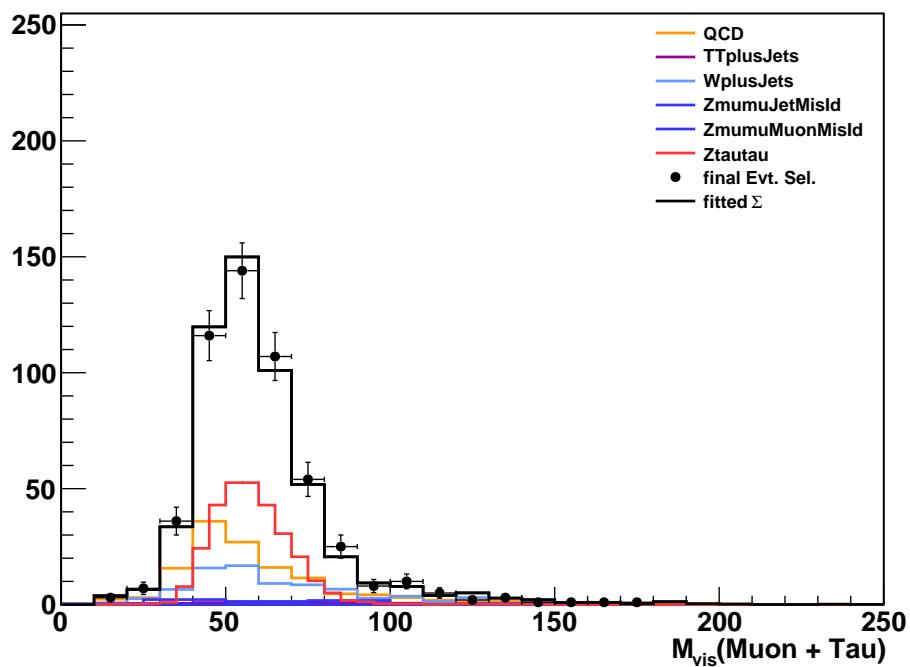


Figure 6.11:  $M_{vis}$  distribution of events selected by the  $Z/\gamma^* \rightarrow \tau^+\tau^- \rightarrow \mu + \tau_{had}$  cross-section analysis compared to the sum of shape templates for signal and background processes scaled by the normalization factors determined by the fit.

g:TemplateFitControlPlot)

1914

## Chapter 7

1915

# Monte Carlo Corrections

&lt;ch:corrections&gt;

1916 One of the most important goals of the analysis is to minimize the effect of potentially  
 1917 incorrect simulation effects on the final result. While the simulated CMS events have been  
 1918 observed to match the 2010 data with surprising results, it is nonetheless critical to measure  
 1919 in real data phenomenon which can have significant effects on the analysis whenever possible.  
 1920 In practice, these measurements are used to apply a correction factor to the corresponding  
 1921 measurement obtained from Monte Carlo. This measured correction factor has an associ-  
 1922 ated uncertainty, and is taken into account as a systematic uncertainty. The application of  
 1923 systematic uncertainties is described in the next chapter.

1924 The corrections measured and used in this analysis can be divided into two categories,  
 1925 efficiency corrections and scale corrections. Identification efficiency corrections scale the  
 1926 expected yield (due to a given identification selection) up or down. Scale corrections sys-  
 1927 tematically scale the energy of a particle (or  $E_T^{\text{miss}}$ ) up or down. In this analysis we apply  
 1928 efficiency corrections for the High Level Trigger muon requirement, all stages of muon iden-  
 1929 tification, and the hadronic tau identification. We apply a momentum scale correction to  
 1930 the muon and tau legs, and to the resolution of the  $E_T^{\text{miss}}$ . Finally, events are simulated with  
 1931 overlapping “pile-up”<sup>1</sup> events. The simulated events are weighted such that the number of  
 1932 pile-up events in the simulation matches that observed in the data.

1933

## §7.1 Muon Identification Efficiency

&lt;sec&gt;ZmumuTagAndProbe&gt;

1934 The identification efficiencies associated with the muon are measured in  $Z \rightarrow \mu^+ \mu^-$  events  
 1935 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.

1936 7 TeV CMS 2010 datasets<sup>2</sup> by requiring that the events pass the “loose” Vector Boson Task  
 1937 Force (VBTF) event selections [44]. In the selected events, we define the “tag” muons as  
 1938 those that have transverse momentum greater than 15 GeV/c and pass the VBTF muon  
 1939 selection. The tag muons are further required to pass the “combined relative isolation” de-  
 1940 scribed in the VBTF paper. We finally require that the tag muon be matched to an HLT  
 1941 object corresponding to the run-dependent requirements listed in table 5.1. The trigger  
 1942 match requirement ensures that the event would be recorded independently of the probe  
 1943 muon. After the tag and probe muon pairs have been collected, we compare the muon identi-  
 1944 fication performance in the probe collection in events selected in data to the performance in  
 1945 simulated  $Z \rightarrow \mu^+ \mu^-$  events. The selection of events and tag muon in the simulated sample  
 1946 is the same as the data sample, with the notable exception that the only HLT requirement  
 1947 applied in MC is that the tag muon is matched to an HLT\_Mu9 object. Any difference in ef-  
 1948 ficiency between the HLT\_Mu9 path and the paths used to select the data (in the tag–probe  
 1949 measurement and in the analysis) will be considered implicitly in the correction factor.

1950 The efficiencies for the muon selections applied in this analysis are measured using  
 1951 the “probe” objects. We measure the following marginal efficiencies, each relative to the  
 1952 previous requirement:

- 1953 • Efficiency of global probe muons to satisfy VBTF muon identification selections.
- 1954 • Efficiency of global probe muons passing the VBTF muon identification selection to  
1955 satisfy the isolation criteria described in Section 5.2.1.
- 1956 • Efficiency of probe muons passing the offline analysis selection defined in Chapter 5  
1957 to pass the HLT selection.

1958 In each case, the invariant mass spectrum of the tag–probe pair is fitted with a Crystal  
 1959 Ball function for the signal ( $Z \rightarrow \mu^+ \mu^-$ ) events and an exponential for the background. The  
 1960 fit is done for two cases; where the probe fails the selection and the where it passes. The  
 1961 method is illustrated in Figure 7.1. The signal yield  $N$  is extracted from each fit and the  
 1962 efficiency is computed as  $N_{\text{pass}} / (N_{\text{pass}} + N_{\text{fail}})$ . Each efficiency is measured in both the data

---

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

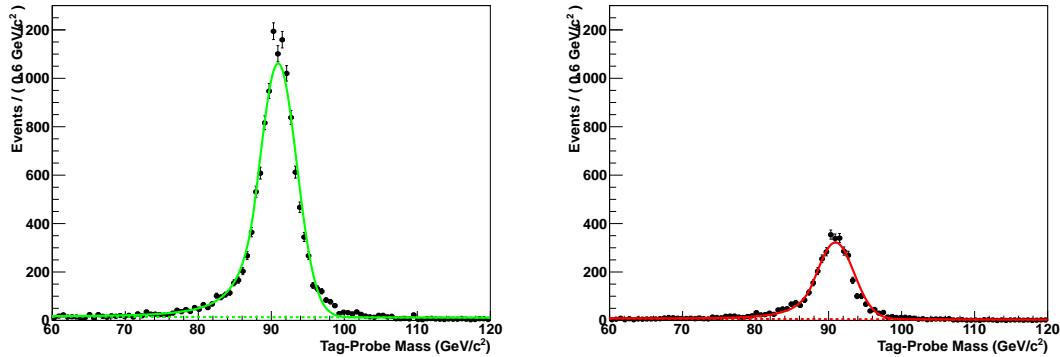


Figure 7.1: The tag–probe dimuon invariant mass spectrum in events in which the probe muon passed (left) and failed (right) the muon isolation requirement. The solid gives the result of the simultaneous fit of the signal (real  $Z \rightarrow \mu^+\mu^-$  events) and background. The fitted background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

and the simulation. The results of the measurements are shown in table 7.1. In the final analysis, the simulated events are weighted by the fractional difference to the measured values; the statistical uncertainty on the weight is taken as the sum in quadrature of the statistical uncertainties for the data and simulation efficiency measurements. The uncertainty on this measurement is taken as systematic uncertainty in the final measurement.

1968

1969        The correction for the trigger efficiency needs to take into account the differences in  
 1970      the HLT selections applied during different operating periods (see table 5.1). To determine  
 1971      the overall correction factor, we measure the trigger efficiency in data for each of the op-  
 1972      erating periods and compare it to the simulated efficiency of the HLT\_Mu9 selection. The  
 1973      overall efficiency in data is taken as the average of the three periods, weighted by integrated  
 1974      luminosity.

1975        The efficiency of the “cross–triggers” used in the run–range period 148822 – 149182  
 1976      (period C) cannot be measured in  $Z \rightarrow \mu^+\mu^-$  events as they require a reconstructed PFTau  
 1977      object at the trigger level. A single muon trigger (HLT\_Mu15) is also used in period C.  
 1978      The contribution of the cross–triggers is taken as a correction to the single muon trigger  
 1979      period C efficiency. The “muon leg” of the cross–triggers have the same requirements as  
 1980      the single muon triggers used in the run–range 147196 – 148058 (period B). The “cross–

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

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

muonTagAndProbeResults)

trigger” contribution is estimated as the difference between the efficiency in period B and the single-muon period C efficiency multiplied by a correction factor of  $0.9 \pm 10\%$  to account for the  $\tau$  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,

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

The only suitable resonance for this method is  $Z \rightarrow \tau^+\tau^-$ . This method has been applied<sup>3</sup> in CMS  $Z \rightarrow \tau^+\tau^-$  cross section analysis ??, and measured a tau identification simulation to data correction factor of  $0.960 \pm 0.067$ .

---

<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. 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.

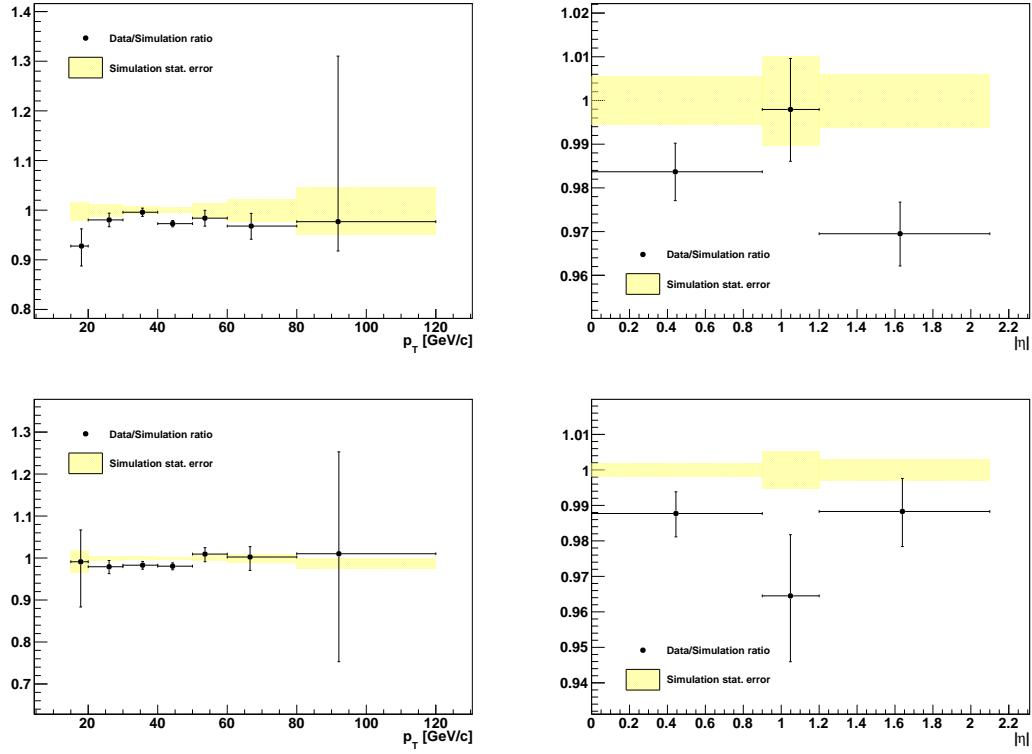


Figure 7.2: Ratio of muon isolation efficiency measured in data compared to simulated  $Z \rightarrow \mu^+\mu^-$  events.

`fig:MuonIsoCorrVersusPt?`

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

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

The solution to this problem is to use a “tag and probe” approach analogous to the muon efficiency measurement of Section 7.1. The tag and probe method is only sensitive to the shapes of the distributions, and will be independent of a Higgs 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].

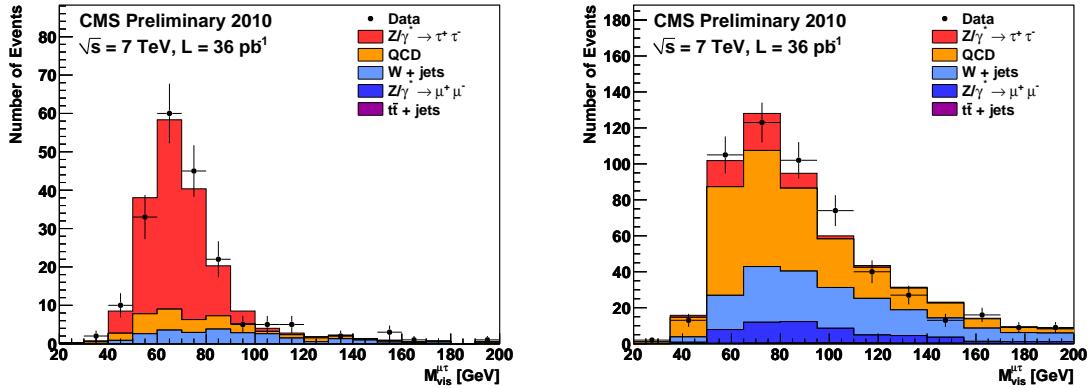


Figure 7.3: Visible mass spectrum of preselected events used to measure the hadronic tau identification efficiency in 2010 CMS 7 TeV data. The figure on the left (right) shows the preselected events that pass (fail) the hadronic tau identification. The different colors indicate the fitted yields of the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

(fig:TauIdEffFits)

1996 The preselected sample is then split into two categories, those that pass the hadronic tau  
 1997 identification and those that fail. The signal and background yields in each category  
 1998 are fitted using the Template Method described in Section 6.3. An illustrative example of  
 1999 the fits for the yields is shown in Figure 7.3. The hadronic tau identification efficiency can  
 2000 then be computed using the relative size of the true tau yields in the passing and failing  
 2001 categories. The efficiency is measured [48] for the loose HPS + TaNC tau identification in  
 2002 the 2010 CMS dataset and is found to be  $1.06 \pm 0.30$ .

### 2003 §7.3 Muon and Tau Momentum Scale

MuonTauMomentumScale<sup>2004</sup>) Muons are one of the best measured objects at CMS. The momentum scale of CMS muons  
 2005 has been measured [54] using the  $J/\psi$ ,  $\psi(2S)$  and  $\Upsilon$  di-muon resonant decays. The muon  
 2006 momentum resolution is found to be 3% or better for muons with  $p_T < 100$  GeV/ $c$ . We  
 2007 apply the muon momentum correction using the ‘‘MuSICleFit’’ algorithm described in [54].  
 2008 The muon momentum correction and uncertainty varies as a function of  
 2009 muon  $p_T$  and  $\eta$ . The effect of the muon momentum correction uncertainty is a small effect  
 2010 in this analysis compared to the  $\tau$  and  $E_T^{\text{miss}}$  scale uncertainties.

2011 The uncertainty on the jet energy scale is determined from an analysis of the  $p_T$  bal-  
 2012 ance between photons and jets in  $\gamma + \text{jets}$  events [55]. The jet energy scale uncertainties

2013 determined by the JetMET group are applied to tau–jets as well as other jets in the event.  
 2014 The tau energy scale correction factor is currently taken to be 1.0 with an uncertainty of  
 2015 3%. The QCD jet energy scale has been measured to within 3% uncertainty. In the future,  
 2016 the energy scale of the tau is expected to be determined to a much better precision, as the  
 2017 neutral hadronic activity of a hadronic tau decay is expected to be zero. The jet energy  
 2018 scale of 3% can be confidently considered [48] an upper limit<sup>4</sup>, and is used in this analysis  
 2019 as the tau energy scale uncertainty.

## 2020 §7.4 Missing Transverse Energy Correction

(sec:ZRecoilCorr) 2021 In practice, the resolution of the reconstructed missing transverse energy is poor as it is  
 2022 sensitive to the mis–measurement of any object in the event. Furthermore, a fraction of the  
 2023 particles produced in the hard collision can be produced in the very forward region, outside  
 2024 of the fiducial region of the calorimeters. The resolution of the  $E_T^{\text{miss}}$  reconstruction can be  
 2025 measured in  $Z \rightarrow \mu^+ \mu^-$  events. The true  $E_T^{\text{miss}}$  in such events is expected to be zero. The  
 2026  $E_T^{\text{miss}}$  resolution in simulated  $Z \rightarrow \mu^+ \mu^-$  events is found to be smaller (better) than in the  
 2027 data.

2028 The  $E_T^{\text{miss}}$  resolution depends on the “recoil” of the  $Z$  boson. The reason for this effect  
 2029 is that for events where the  $Z$  is produced nearly at rest, the associated recoil products have  
 2030 very small transverse momentum and are produced at very high pseudorapidity. The  $E_T^{\text{miss}}$   
 2031 is corrected using a procedure called a “ $Z$ –recoil” correction, as described in [56]. The reso-  
 2032 lution of the  $E_T^{\text{miss}}$  is measured in  $Z \rightarrow \mu^+ \mu^-$  events in simulation and data. The difference  
 2033 in the reconstructed  $E_T^{\text{miss}}$  resolution in both samples is parameterized by the magnitude of  
 2034 the transverse momenta of the particles recoiling against the  $Z$ .<sup>5</sup> The reconstructed  $E_T^{\text{miss}}$   
 2035 in the simulated  $Z \rightarrow \tau^+ \tau^-$ ,  $Z \rightarrow \mu^+ \mu^-$ , and  $W + \text{jets}$  samples is “smeared” by a random  
 2036 amount in each event such that the final resolution matches the observed resolution in the  
 2037 data.

---

<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  $E_T^{\text{miss}}$ .

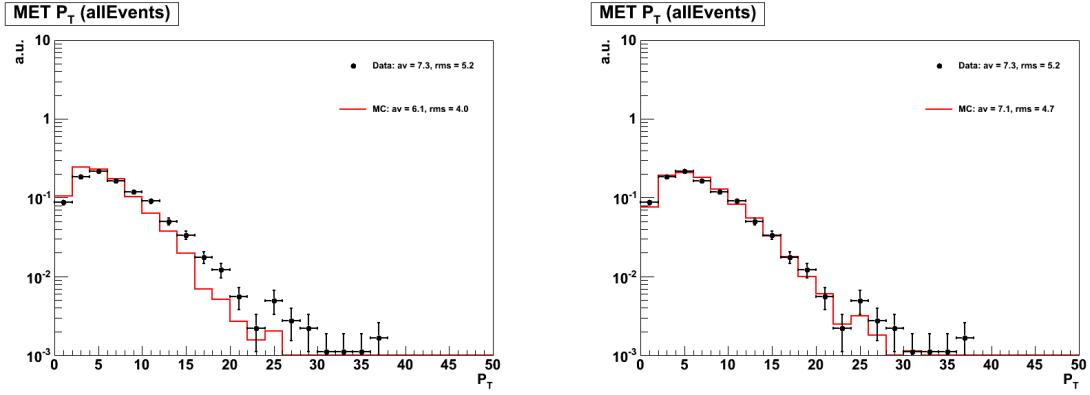


Figure 7.4: Missing transverse energy reconstructed in  $Z \rightarrow \mu^+\mu^-$  events selected in data compared to  $Z \rightarrow \mu^+\mu^-$  events in Monte Carlo simulation before (left) and after (right) the  $Z$ -recoil corrections to the  $E_T^{\text{miss}}$  resolution are applied.

(fig:ZrecoilCorrection)

2038      $Z$ -recoil corrections are determined as described in [56] and applied to simulated  
 2039      $Z \rightarrow \tau^+\tau^-$ ,  $Z \rightarrow \mu^+\mu^-$  and  $W + \text{jets}$  events, in order to correct for residual differences in  
 2040      $E_T^{\text{miss}}$  response and resolution between data and Monte Carlo simulation [57]. The correc-  
 2041     tions are obtained by an unbinned maximum likelihood fit (in data and simulation) of the  
 2042     transverse recoil vector  $\vec{u}_T = -(\vec{q}_T + E_T^{\text{miss}})$  as function of the transverse momentum  $\vec{q}_T$  of  
 2043     the  $Z$ -boson in directions parallel and perpendicular to the  $Z$ -boson transverse momentum  
 2044     vector. The effect of the  $Z$ -recoil correction is illustrated in Figure 7.4. The uncertainty on  
 2045     the  $Z$ -recoil correction factor from the maximum likelihood fit is treated as a systematic  
 2046     uncertainty in the final result.

## 2047     §7.5 Pile-up Event Weighting

?<sec:PUweighting>? The average number of pile-up interactions in the event can effect almost all aspects of  
 2048     the analysis. In general, increasing pile-up lowers particle identification efficiencies and  
 2049     lowers  $E_T^{\text{miss}}$  resolution. It is therefore important that the distribution of pile-up events  
 2050     in the simulation matches the distribution found in the data. Differences in the number  
 2051     of pile-up interactions between the data (averaged over the analyzed run-range) and pile-  
 2052     up Monte Carlo samples produced for “BX156” pile-up conditions are corrected for by  
 2053     reweighting Monte Carlo simulated events according to the number of reconstructed event

---

<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.

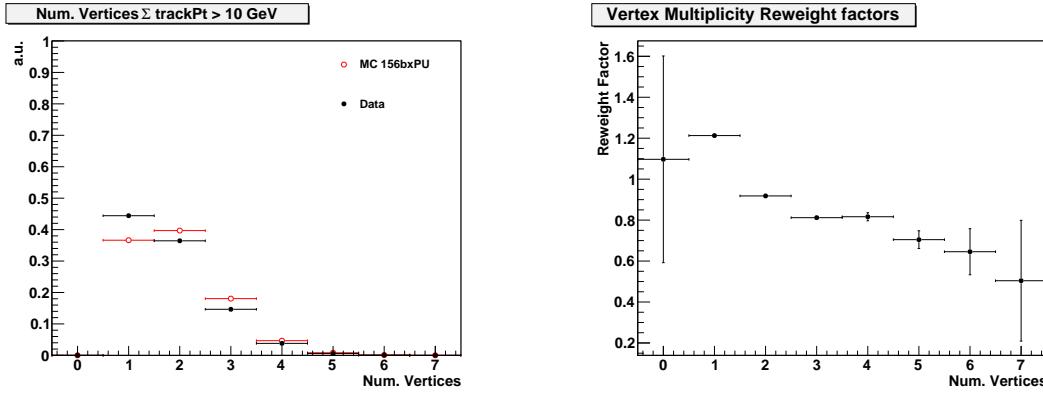


Figure 7.5: Vertex multiplicity distribution measured in the analyzed data-taking period compared to Monte Carlo simulation with “BX156” pile-up conditions (left) and resulting Monte Carlo reweighting factors (right).

{fig:pileUpReweighting}

2055 vertices, in order to match the distribution measured in a  $W \rightarrow \mu\nu$  dataset triggered by  
 2056 the HLT\_Mu15 High Level Trigger path. Vertices considered for this purpose are required  
 2057 to pass  $-24 < z_{vtx} < +24$  cm,  $|\rho| < 2$  cm, nDoF  $> 4$ . In addition, the total transverse  
 2058 momenta of all tracks fitted to the vertex is required to exceed 10 GeV/ $c$ , assuming that  
 2059 “softer” vertices have little or no effect on the “hard” event to pass event selection criteria.  
 2060 The average vertex multiplicity distribution measured in data is compared to Monte Carlo  
 2061 simulation with “BX156” pile-up conditions in Figure 7.5. Both distributions are similar,  
 2062 resulting in Monte Carlo reweighting factors close to unity.

2063

## Chapter 8

2064

# Systematics and Limit Extraction

(ch:systematics)

2065 In this chapter we discuss the systematic uncertainties affecting the search for the Higgs  
 2066 boson and the statistical techniques used to establish an upper limit on the Higgs  $\rightarrow \tau^+\tau^-$   
 2067 branching ratio times cross section ( $\sigma \times \text{BR}_\tau$ ). The limit can be interpreted as the largest<sup>1</sup>  
 2068 signal presence that could exist in the data and still be consistent with the null hypothesis.  
 2069 The limit on  $\sigma \times \text{BR}_\tau$  is roughly independent of the theoretical model<sup>2</sup>. In the conclusion,  
 2070 we will interpret the  $\sigma \times \text{BR}_\tau$  limit result in the context of the MSSM theory.

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

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

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.

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

$$\frac{\delta(\sigma \times BR)}{\sigma \times BR} = \frac{\partial(\sigma \times BR)}{\partial N_{bkg}} \frac{1}{\sigma \times BR} \delta N_{bkg} = \frac{-\delta N_{bkg}}{N_{obs} - N_{bkg}}. \quad (8.2) \quad \text{[eq:CrossSectionError]}$$

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

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

2098 §8.1 Signal normalization uncertainties

2099 The signal normalization uncertainties are due to imperfect knowledge of how improperly  
 2100 modeled effects in the detector affect our “acceptance” model, or the probability that a  
 2101 given signal event will pass one of the selections (detailed in Chapter 5). The general pro-  
 2102 cedure to quantify these uncertainties is to measure the effect in some control region in  
 2103 both the data and Monte Carlo. The ratio of data to Monte Carlo then gives a correc-  
 2104 tion factor which is applied to the simulation. An uncertainty on the measurement of the  
 2105 effect in control region (in data, simulation, or both) is then taken as the systematic un-  
 2106 certainties. The signal normalization uncertainties affecting this analysis on muon trigger,  
 2107 reconstruction, identification and isolation efficiencies are taken from the tag and probe  
 2108 analysis of  $Z \rightarrow \mu^+ \mu^-$  events presented in Section 7.1. A very conservative estimate of 30%  
 2109 is taken for the uncertainty on the tau reconstruction and identification efficiency. The tau  
 2110 identification uncertainty measurement is discussed briefly in 7.2. The dependency of the  
 2111 Higgs signal extraction on the tau identification efficiency has been studied, the result being  
 2112 that uncertainties on the tau identification efficiency affect the limit on cross-section times  
 2113 branching ratio for MSSM Higgs  $\rightarrow \tau^+ \tau^-$  production by a few percent only. An uncertainty  
 2114 of 11% is attributed to the luminosity measurement.

2115 §8.2 Background normalization uncertainties

2116 Uncertainties on the normalization of background processes are obtained from the study of  
 2117 background enriched control regions presented in Chapter 6. The main fake tau backgrounds  
 2118 in this analysis are QCD multi-jet and  $W +$  jets events. For a detailed discussion of the  
 2119 measurements, see Chapter 6. These backgrounds are produced copiously enough for the  
 2120 backgrounds to be studied in control regions dominated by a single background process  
 2121 with a purity exceeding 90% and an event statistics exceeding the expected contribution of  
 2122 that background to the analysis by more than one order of magnitude. Both backgrounds  
 2123 are found to be well modeled by the Monte Carlo simulation. An uncertainty of 10% is  
 2124 attributed to the contribution of QCD and  $W +$  jet backgrounds to the analysis. The  
 2125 cross-section for  $t\bar{t} + jets$  production makes it difficult to select a high purity sample of

2126  $t\bar{t}+jet$  events of high event statistics. From the study of the 19 events selected in the  $t\bar{t}+jets$   
 2127 background enriched control sample we assume an uncertainty on the  $t\bar{t}+jets$  background  
 2128 contribution in the analysis of 30%. The  $Z \rightarrow \mu^+\mu^-$  background has been studied with  
 2129 large statistical precision in two separate control regions, dominated by events in which  
 2130 the reconstructed tau–jet candidate is either due to a misidentified quark or gluon jet or  
 2131 due to a misidentified muon. Good agreement between data and Monte Carlo simulation is  
 2132 found in both cases. Sizeable uncertainties on the  $Z \rightarrow \mu^+\mu^-$  background contribution arise  
 2133 due to the extrapolation from the background enriched control regions to the data sample  
 2134 considered in the analysis, however: the contribution of  $Z \rightarrow \mu^+\mu^-$  background events to  
 2135 the analysis is due to events in which one of the two muons produced in the  $Z$  decay either  
 2136 escapes detection or fakes the signature of a hadronic tau decay. Both cases may be difficult  
 2137 to model precisely in the Monte Carlo simulation. The non–observation of a  $Z$  mass peak in  
 2138 the mu + tau visible mass distribution studied with the fake–rate method on the other hand  
 2139 sets a limit on possible contributions from  $Z \rightarrow \mu^+\mu^-$  background events. Conservatively,  
 2140 we assume an uncertainty of 100% on both types of  $Z \rightarrow \mu^+\mu^-$  background contributions.

### 2141 §8.3 Shape uncertainties

<sec:ShapeUncertainties>  
 2142 Shape uncertainties on the distributions of visible and “full” invariant mass reconstructed by  
 2143 the SVfit algorithm are estimated by varying the electron energy and muon momentum scale,  
 2144 the energy scale of tau–jets and other jets in the event and varying the missing transverse  
 2145 energy in Monte Carlo simulated events. After each variation the complete event is re–  
 2146 reconstructed and passed through the event selection. Shifted visible and “full” invariant  
 2147 mass shapes are obtained for each variation from the events passing all event selection  
 2148 criteria. The difference between shifted shapes and the “nominal” shapes obtained from  
 2149 Monte Carlo simulated events with no variation of energy or momentum scale or of the  
 2150 missing transverse energy applied is then taken as shape uncertainty.

2151 The systematic uncertainties on the muon and tau energy scales have been provided by  
 2152 the muon and tau Physics Object Groups and are described in Section 7.3. The modelling  
 2153 of missing transverse energy in different types of background events has been studied in  
 2154 the background enriched control regions described in Chapter 6. No significant deviations

2155 between data and Monte Carlo simulation have been found. Uncertainties due to missing  
 2156 transverse energy are estimated by varying parameters of the  $Z$ -recoil corrections within  
 2157 the uncertainties obtained when fitting (see Section 7.4) the  $Z$ -recoil correction factor pa-  
 2158 rameters in simulated  $Z \rightarrow \mu^+ \mu^-$  events versus  $Z \rightarrow \mu^+ \mu^-$  events selected in data.

## 2159 §8.4 Theory uncertainties

2160 The signal and background normalization as well as the shape uncertainties are all exper-  
 2161 imental uncertainties in nature. Additional theoretical uncertainties arise from imprecise  
 2162 knowledge of parton-distribution functions (PDFs) and of the exact dependency of signal  
 2163 cross-sections and branching ratios on  $\tan\beta$  and  $m_A$ . The PDFs describe how the energy  
 2164 of the protons is shared between the quarks and gluons. Since the longitudinal boost of  
 2165 the hard collisions depends on the PDFs, the signal acceptance is sensitive to errors in the  
 2166 PDFs. The uncertainties on the signal acceptance due to PDF uncertainties are estimated  
 2167 using tools developed by the CMS Electroweak working group [58]. The acceptance is com-  
 2168 puted with respect to MSSM Higgs  $\rightarrow \tau^+ \tau^-$  decays that have muons of  $p_T^\mu > 15$  GeV/ $c$   
 2169 and  $|\eta_\mu| < 2.1$ , jets produced in hadronic tau decays with visible  $p_T^{vis} > 20$  GeV/ $c$  and  
 2170  $|\eta_{vis}| < 2.3$  on generator level. Acceptance values are computed for the central value and  
 2171 44 eigenvectors of the CTEQ66 PDF set [59]. The systematic uncertainty on the signal  
 2172 acceptance is computed following the PDF4LHC recommendations [60, 61].

2173 The effect of Monte Carlo normalization, shape and theory uncertainties on the signal  
 2174 efficiency times acceptance is summarized in Table 8.1.

## 2175 §8.5 Limit Extraction Method

?<sec:statmethod>? The search for a new signal is performed by examining the observed distribution of the  
 2176 reconstructed di-tau mass  $m_{\tau\tau}$  as reconstructed by the SVfit algorithm. An “bump” in  
 2177 this spectrum would indicate the presence of a new particle. To make a statement about  
 2178 the presence of a bump with confidence, the shape under a potential bump must be well  
 2179 described. The background shape is decomposed into the combination of shapes from the  
 2180 difference background sources, which we refer to here as “templates.” The data and 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–jet identification	$1.00 \pm 0.30$
Shape uncertainties	
Muon momentum scale	$\ll 1\%$
Tau–jet 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)

2182 templates for each background distribution is binned in the observable  $m_{\tau\tau}$  variable. The  
 2183 normalization of each template represents the total yield expected for that source.

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

$$P(k|\mu) = \frac{\mu^k e^{-\mu}}{k!}. \quad (8.3) \quad [\text{eq:PoissonProb}]$$

2184 The total likelihood for observed data given some configuration of templates is then simply

2185 a product of the Poisson probabilities (Equation 8.3) in each of the  $N_{bin}$  bins:

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

2186 where the expected number of events  $\mu_i$  in the bin  $i$  is the sum of the number of events

2187 from all sources

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

2188 The number of expected events in a source, in turn, can be written

$$\mu_{ji} = L \sigma_j \epsilon_{ji} \quad (8.5) \text{[eq:ShapeParameter]}$$

2189 where  $L$  is the integrated luminosity,  $\sigma_j$  is the cross section for source  $j$ , and  $\epsilon_{ji}$  is the  
2190 efficiency for source  $j$  in bin  $i$ .

2191 We incorporate the systematic uncertainties of the analysis by introducing a set of  
2192 “nuisance parameters”  $\vec{\beta}$  into the likelihood function. As the name suggests, we are not  
2193 interested in the actual value of the nuisance parameters. Each nuisance parameter param-  
2194 eterizes some phenomenon in the analysis. The shape templates (which can be defined purely  
2195 in terms of  $\sigma_j$  and  $\epsilon_{ji}$  in Equation 8.5) are now interpreted as functions of the set of nuisance  
2196 parameters  $\vec{\beta}$ . Existing knowledge about the value of the nuisance parameter is introduced  
2197 by extending the likelihood function with a constraint  $\mathcal{G}(\vec{\beta})$  that expresses the uncertainty  
2198 on the nuisance parameters. The templates  $\mu_{ji}$  can depend on the nuisance parameters in  
2199 two ways. Normalization uncertainties introduce multiplicative nuisance factors on the yield  
2200 of some (sub)set of the sources defined by Equation 8.5. As an example, consider a simplified  
2201 situation where there are only two sources:  $Z \rightarrow \tau^+ \tau^-$  and  $W + \text{jets}$  events. Both are sources  
2202 are sensitive to the efficiency of the muon trigger. Only the  $Z \rightarrow \tau^+ \tau^-$  sample is sensitive  
2203 to the efficiency of the hadronic tau identification algorithm, as the fake-rate in  $W + \text{jets}$   
2204 is measured in data. We would then introduce two multiplicative nuisance parameters,  $\beta_\mu$   
2205 and  $\beta_\tau$ , which respectively correspond to the two uncertainties. The expected number of  
2206 events in the  $i$ th bin is then given by a modified form of Equation 8.5

$$\mu_i = L \beta_\mu \sigma^W \epsilon_i^W + L \beta_\mu \beta_\tau \sigma^Z \epsilon_i^Z. \quad (8.6) \text{[eq:ShapeParameter]}$$

2207 From Equation 8.6, we can see that  $\beta_\mu$  affects both sources but  $\beta_\tau$  only affects the  $Z \rightarrow \tau^+ \tau^-$   
2208 source.

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

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

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

(tab:ConfidenceIntervals)

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

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

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

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

**Fixme:** Did I mix this up?

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

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

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

Bayesian posterior PDF is computed as

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

2248 The left hand side of Equation 8.8 represents the probability density for a given signal cross  
 2249 section, given the observed data  $\bar{x}$  and the assumed value  $m_H$  for the Higgs mass. We refer  
 2250 to this method as the “Bayesian” method of setting limits.

2251 In the other method, which is referred to as the “Delta Log Likelihood” (DLL) method,  
 2252 we compute two likelihoods. The first likelihood is computed for the “null hypothesis” case.  
 2253 The likelihood is profiled (maximized) assuming that no signal is present. We then construct  
 2254 the profile likelihood for increasing values of the signal yield. The upper limit is achieved  
 2255 when the logarithm of the profile likelihood is 1.92 units less than the value of the null  
 2256 hypothesis profile likelihood.

2257 In general the limits computed by the Bayesian and DLL methods are similar. However,  
 2258 the effect of upward or downward statistical fluctuations in the observed data affect the two  
 2259 methods in different ways. When the data fluctuates low, the DLL method will produce  
 2260 a more stringent limit than the DLL method. When the data fluctuates high, creating an  
 2261 apparent signal, the Bayesian method will (correctly) set a more stringent limit. In this  
 2262 thesis we report the results of both methods.

2263 In the results presented below we use nuisance parameters corresponding to the sys-  
 2264 tematic errors summarized in Table 8.3.

**Fixme:** Get

MET scale

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	10%
Tau energy scale	Morphing	2%
Missing $E_T$ scale	Morphing	XX%
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.  
 (tab-sys)

2265

## Chapter 9

---

2266

# Results

(ch:results)

## 2267 §9.1 Selected Events

2268 In the  $36 \text{ pb}^{-1}$  of 7 TeV data collected by CMS in 2010, the analysis selection criteria  
 2269 described in Chapter 5 selects a total of 573 events. The expected yields from each of  
 2270 the background sources, computed by the Template method (Section 6.3) and verified by  
 2271 the Fake-rate method (Section 6.2) are shown in Table 9.1. The total expectation from the  
 2272 Standard Model background is 577. The data agrees extremely well with the SM background  
 2273 expectation considering the expected statistical variance of the observed data (24 events), as  
 2274 well as the uncertainty on the integrated luminosity<sup>1</sup> and the tau identification uncertainty.

2275

2276 The distributions of the visible mass and SVfit mass in the final selected events are  
 2277 shown in Figure 9.3. Excellent agreement is seen between the shapes of the distributions. The  
 2278  $p_T$  spectrum of the transverse momentum are sensitive to mis-modeling of the kinematics  
 2279 and composition of the various background sources. Control plots showing the transverse  
 2280 momenta of the muon and visible hadronic tau in the final analysis selection are shown  
 2281 in Figure 9.1. The  $\eta$  and  $\phi$  distributions of the muon and tau objects are sensitive to  
 2282 detector effects, and the presence of cosmic muons. For example, muons from cosmic events  
 2283 will preferentially be produced in the  $\phi = 0$  direction. Spurious candidates resulting from  
 2284 poorly model noise in one of the CMS subdetectors will in general be localized in  $\eta - \phi$ . The  
 2285  $\eta$  and  $\phi$  distributions of the muon and tau candidates are shown in Figure 9.2 and show  
 2286 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% [62]. 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 Higgs  $\rightarrow \tau^+ \tau^- \rightarrow \mu + \tau_{had}$  candidate events passing the selection criteria described in Chapter 5.

ResultsLooseAHtoMuTau)

The expected yields from an MSSM Higgs boson signal for  $\tan \beta = 30$ <sup>2</sup> 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 mass of  $m_{A^0} = 160$  GeV/c<sup>2</sup>, a total of 17 events are expected at  $\tan \beta = 30$ .

**Fixme:**

Get cite

## §9.2 Limits on Higgs Production

We compute upper limits on the cross section times the branching ratio using the Bayesian method described in Section . 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 pseudo-data sets are expected to have the same statistical sensitivity as the observed dataset. The 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 413 events selected in this analysis is given in the right column of Table 9.3. The observed

**Fixme:** at what value of the nuisance?

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

2303 limit is compatible with the expected limit, within 1.5 standard deviations. The trend of  
 2304 the expected and observed limits versus the Higgs mass using both observables are shown in  
 2305 Figure 9.4. The use of the SVfit reconstructed mass as the observable increases the power  
 2306 of the limit significantly. The limit trend has some interesting features. When the Higgs  
 2307 mass is close to the mass of the  $Z$  resonance, the analysis have little power to set a limit  
 2308 on the presence of the Higgs. This is due to the large uncertainty on the tau identification  
 2309 efficiency. Essentially, when  $m_{A^0} = m_Z$ , the Higgs yield in the  $Z$  bump would have be larger  
 2310 than 30% of the  $Z \rightarrow \tau^+\tau^-$  yield for the profile likelihood to be able to recognize an excess  
 2311 of events. Below this value, the profile likelihood can simple shift the tau identification  
 2312 efficiency scale factor up by 30% and “eat” any potential excess of signal.

### 2313 §9.3 Interpretation in the MSSM

(sec:MSSMInterp) 2314 The limits on the cross section times branching ratio are roughly model independent,<sup>3</sup> and  
 2315 could be applied to set limits on the parameter space of a number of models. In this thesis,  
 2316 we interpret the results in the context of the MSSM. Specifically, we exclude a region in the  
 2317  $\tan\beta - m_{A^0}$  parameters space of the MSSM. To find the upper limit band on  $\tan\beta$ , we find  
 2318 the minimum value of  $\tan\beta$  which provides the cross section and branching ratio product  
 2319 found in the corresponding row in Table 9.3.

2320 The mapping between  $m_{A^0}$  and  $\tan\beta$  and the Higgs cross section is provided by the  
 2321 LHC Higgs Cross Section working group . The cross sections and branching ratios have  
 2322 been computed for the  $h^0$ ,  $H^0$ , and  $A^0$  MSSM Higgs states in both the  $ggA$  and  $qqA$   
 2323 production modes, for a grid of points in  $\tan\beta - m_{A^0}$  space. In order to combine the  
 2324  $ggA$  and  $qqA$  production modes, what we call our signal cross-section is the sum of the  
 2325 cross-section times branching ratio for both modes, assuming  $\tan\beta = 30$ . Additionally, as  
 2326 discussed in Section 1.2.3, the MSSM Higgs sector consists of two Higgs doublets, yielding  
 2327 five physical Higgs bosons. This search is sensitive to the three neutral Higgs particles the  
 2328  $h^0, H^0$ , and  $A^0$ . The relative contributions of the three Higgs types depends on the mass

**Fixme:**  
get ref

---

<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.

2329  $m_{A^0}$  of the CP-odd Higgs. An observed signal will have contributions from at least two  
 2330 Higgs states. For  $m_{A^0} \leq 130$  GeV/ $c^2$ , the  $A^0$  and  $h^0$  are approximately degenerate in mass  
 2331 and width. In this region the  $H^0$  has a very small relative cross section and a constant  
 2332 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 mass of  
 2333  $\approx 130$  GeV/ $c^2$ , and the  $H^0$  and  $A^0$  become mass degenerate.

2334 The region in  $\tan\beta - m_{A^0}$  MSSM parameter space excluded by this analysis at 95%  
 2335 CL is shown in Figure 9.3. The limit is compared to the combined result from Run II of  
 2336 the Tevatron (this result is discussed in detail in Section 1.3.3). The result of this analysis  
 2337 sets a stronger limit than the Tevatron for large values of  $m_{A^0}$ . In the low  $m_{A^0}$  region, the  
 2338 analysis suffers due to the large tau identification efficiency uncertainty. This effect can be  
 2339 mitigated by using the  $e - \mu$  channel. The combined CMS result uses this approach, and  
 2340 will be discussed briefly in the conclusion.

**Fixme:**

check this!

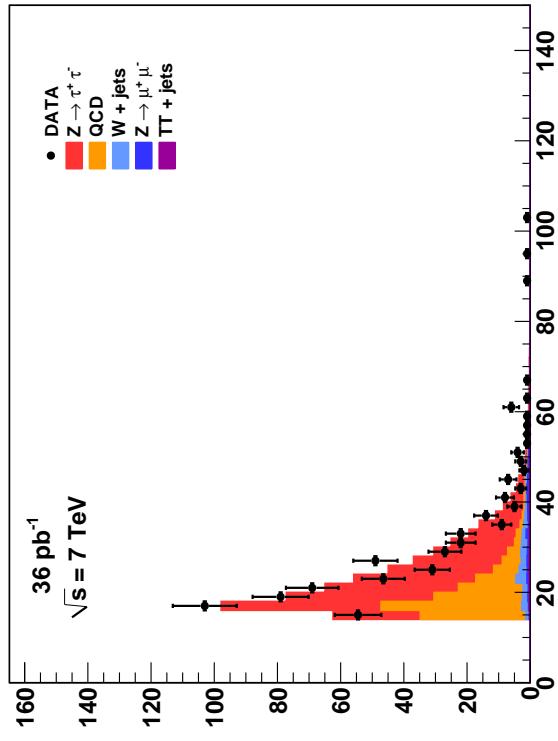
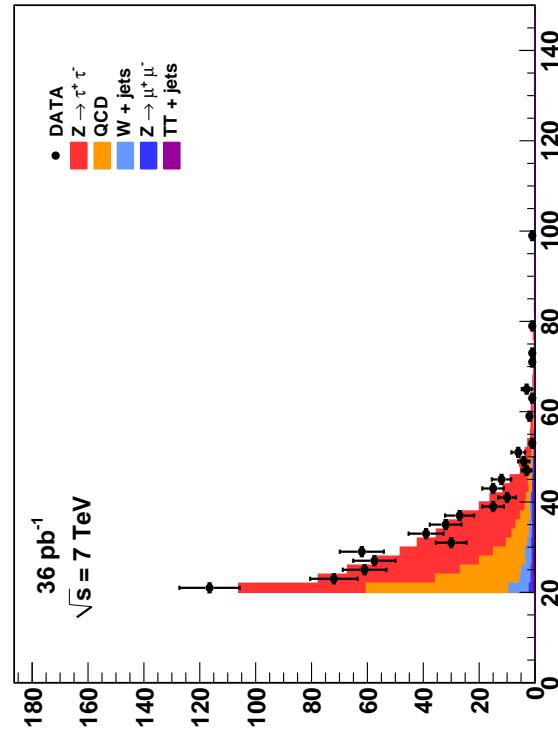
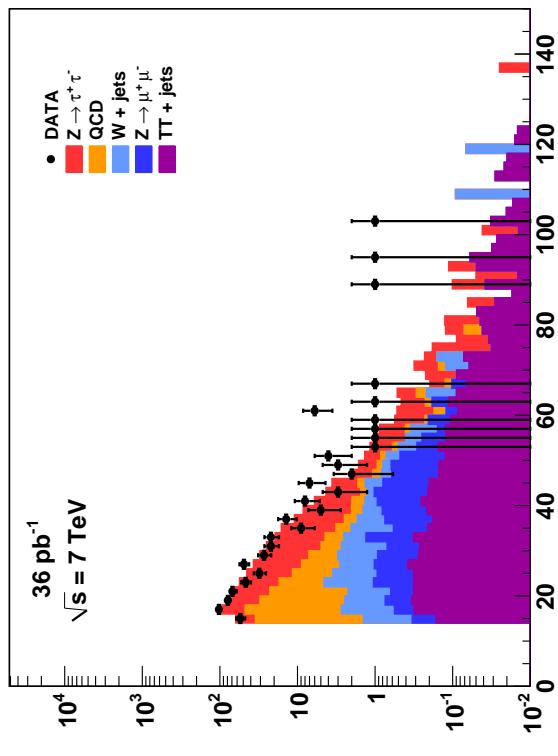
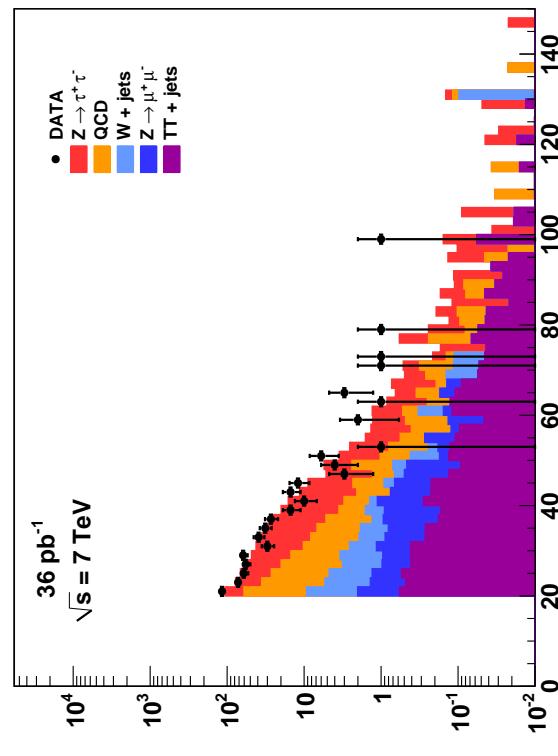
**Muon (final event sample)****Tau (final event sample)****Muon (final event sample)****Tau (final event sample)**

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

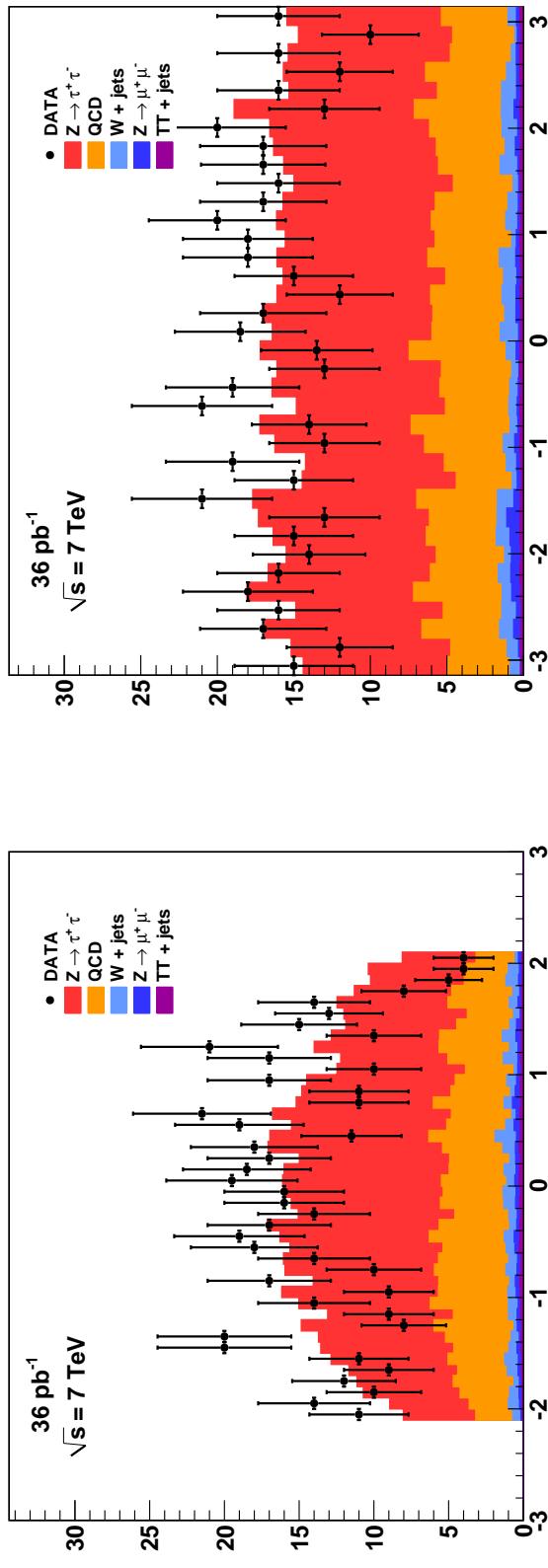
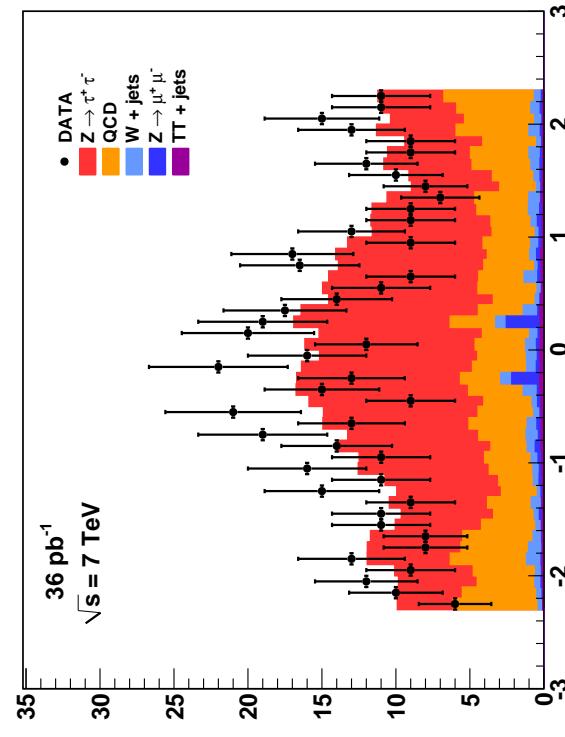
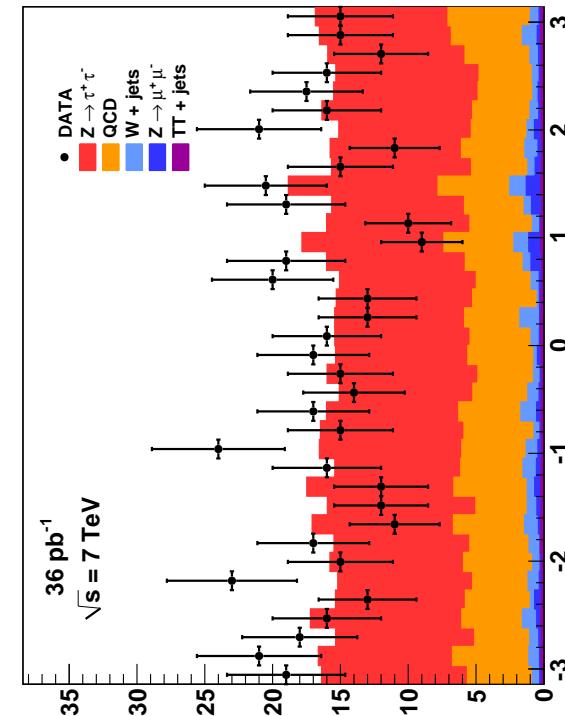
**Muon (final event sample)****Tau (final event sample)****Tau (final event sample)**

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

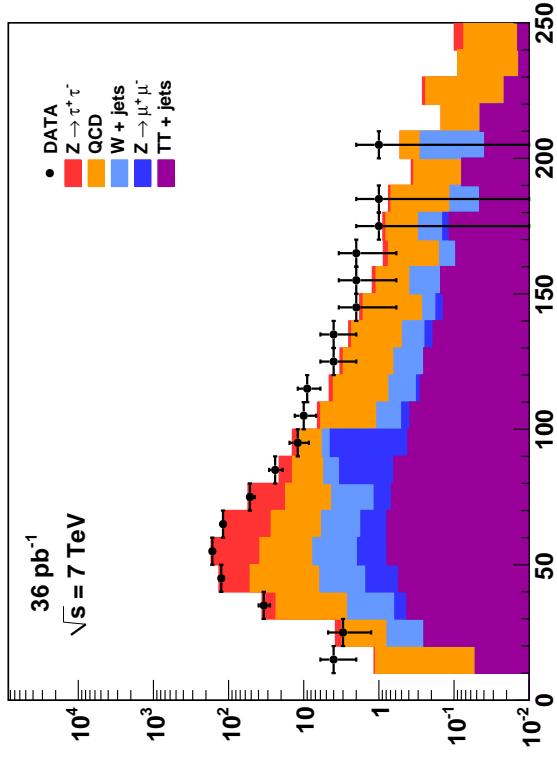
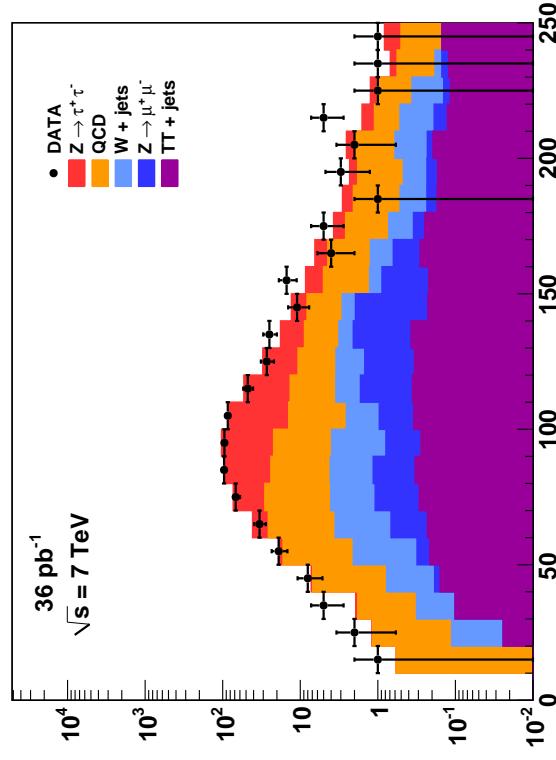
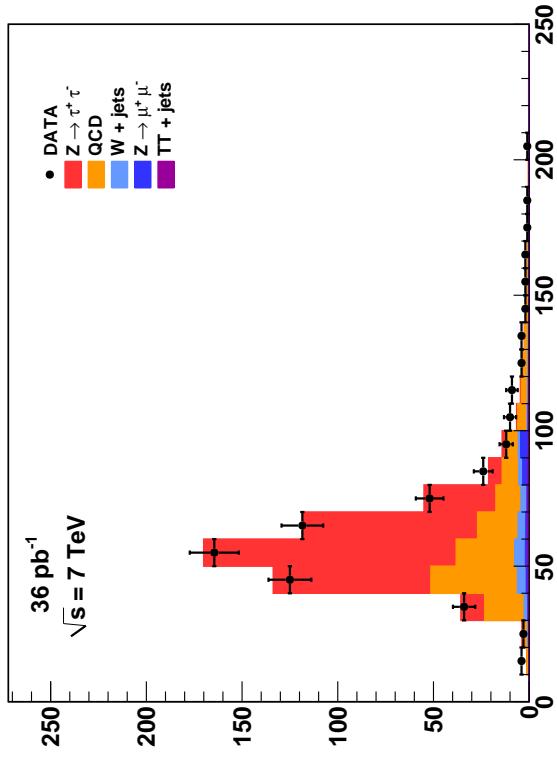
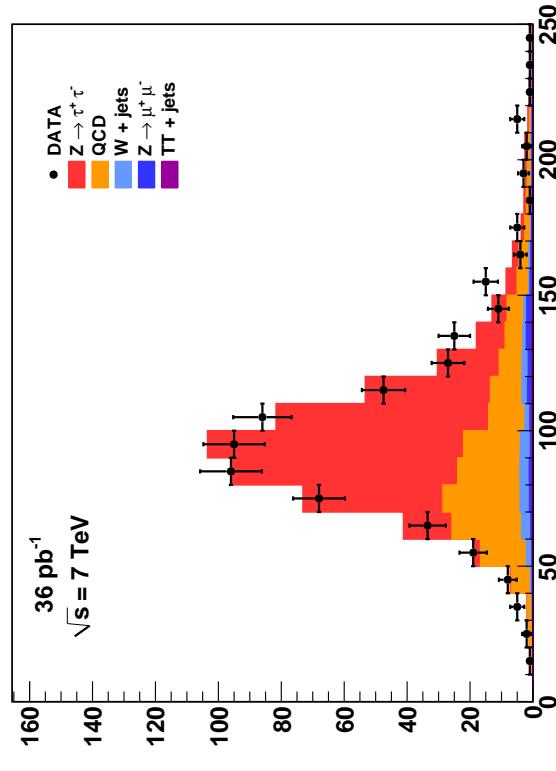
**M<sub>vis</sub>(Muon + Tau) (final event sample)****M(Muon + Tau), SVfit method (final event sample)****M<sub>vis</sub>(Muon + Tau) (final event sample)****M(Muon + Tau), SVfit method (final event sample)**

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

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

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

Secondary Vertex Fit Limit						
Mass	$-2\sigma$	$-1\sigma$	Median	$+1\sigma$	$+2\sigma$	Observed
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 Limit						
Mass	$-2\sigma$	$-1\sigma$	Median	$+1\sigma$	$+2\sigma$	Observed
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		
	$m_{A^0} < 130 \text{ GeV}/c^2$	$m_{A^0} = 130 \text{ GeV}/c^2$	$m_{A^0} > 130 \text{ GeV}/c^2$
$A^0$	yes	yes	yes
$H^0$	yes	yes	no
$h^0$	no	yes	yes

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

(ggsXSectionCombination)?

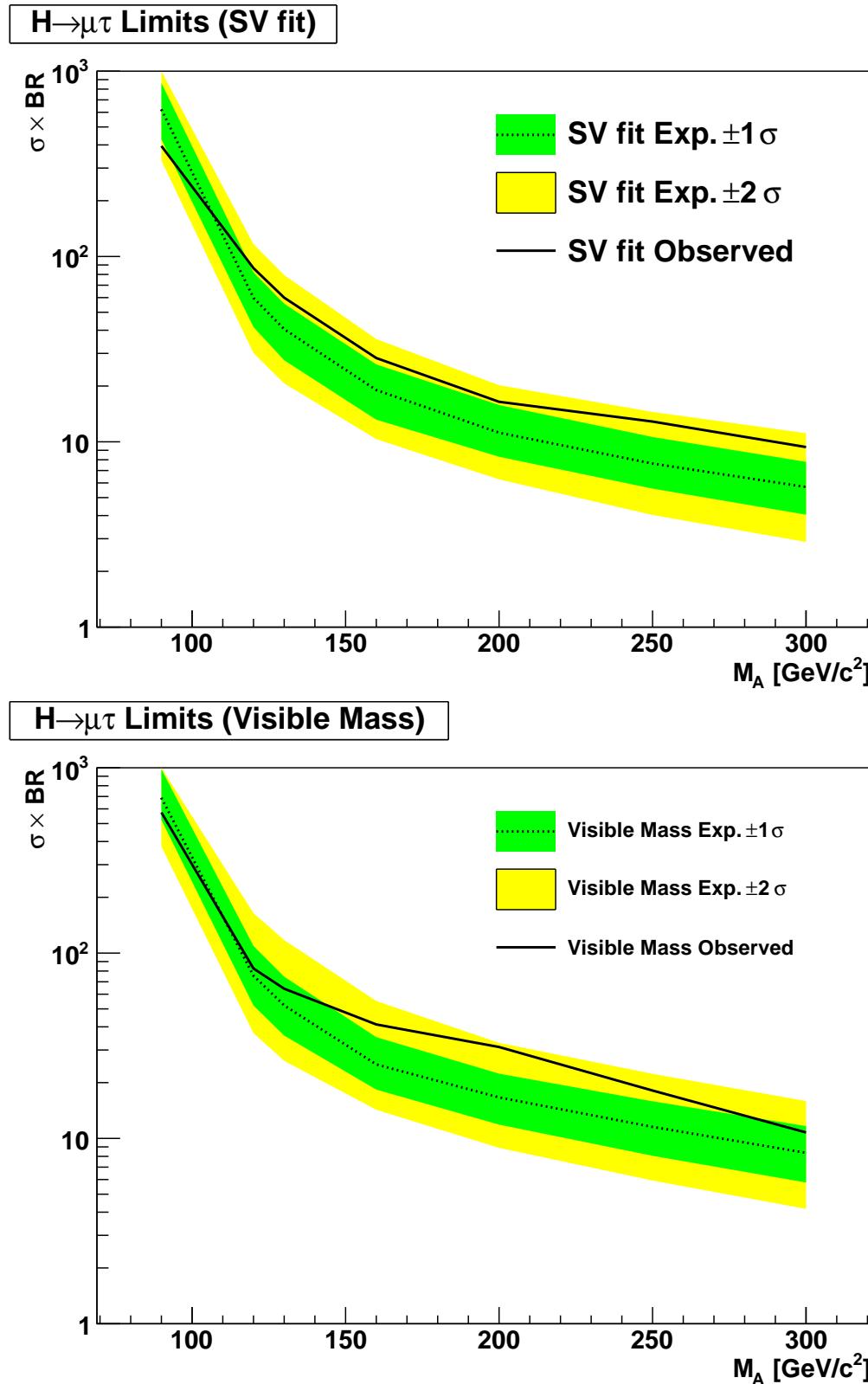


Figure 9.4: Observed and expected limits on the cross section times branching ratio of a Higgs boson versus Higgs mass. The top plot gives the limit computed using the SVfit mass as the observable, the bottom plot gives the limit computed using the visible mass. The dashed line gives the nominal expected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

(fig:SVXSecLimits)

{fig:TanBetaLimit}

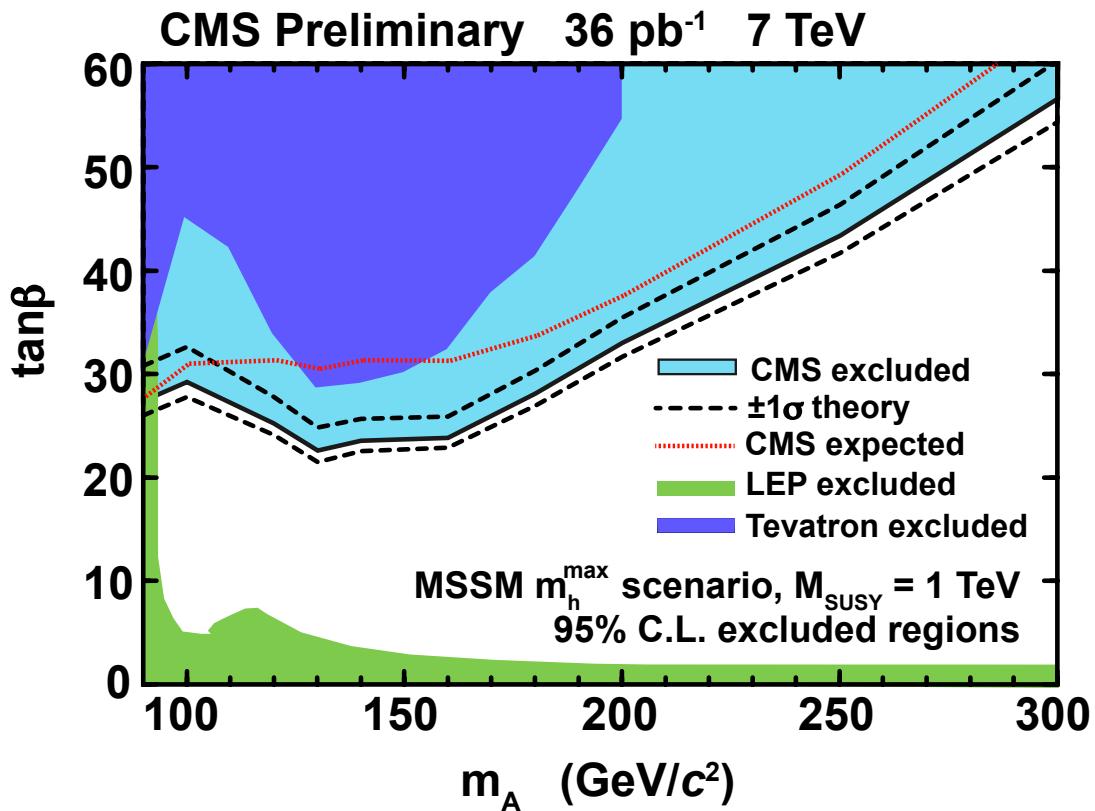
Figure 9.5: Region of MSSM  $\tan \beta - m_{A^0}$  parameter space excluded by this analysis.

---

## 2341 Conclusions

?<ch:conclusions>?  
 2342 This analysis has presented a search for MSSM Higgs bosons in the 2010 7 TeV CMS data  
 2343 set. Two new experimental methods, the TaNC tau identification algorithm, and the SVfit  
 2344 mass reconstruction method have been introduced in this thesis. Both methods increased  
 2345 the sensitive of the Higgs search. The search was performed using  $36 \text{ pb}^{-1}$  of data. The  
 2346 expected event yield from Standard Model sources is 577 events. In total, 573 events were  
 2347 selected; the observed is compatible with the Standard Model. No signal-like excess of  
 2348 events is observed. We set an upper limit on the production of Higgs bosons, and interpret  
 2349 this limit in the context of the MSSM.

2350 The analysis presented in this thesis was part of a larger study [1] performed by the  
 2351 CMS collaboration searching for the MSSM Higgs boson decaying to tau leptons. The CMS  
 2352 analysis used three channels, the  $H \rightarrow \tau\tau \rightarrow e - \tau_h$ ,  $H \rightarrow \tau\tau \rightarrow e - \mu$ , and the  $\mu - \tau_h$   
 2353 channel. The  $\mu - \tau_h$  channel search presented in this thesis is very similar to the CMS result.  
 2354 While not as pure as the  $\mu - \tau$  channel, the inclusion of the high-statistics  $e - \tau$  channel  
 2355 increases the sensitivity of the CMS analysis. The  $e - \mu$  channel has low statistics, but is  
 2356 not sensitive to the systematic uncertainty on the hadronic tau identification. The region of  
 2357 the MSSM parameter space excluded by combined CMS result is illustrated in Figure 9.3.  
 2358 At the time of this writing, the CMS result described in [1] sets the most stringent limits  
 2359 on the MSSM using a direct search.



ig:CMSTanBetaExclusion)

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

---

## 2360 Bibliography

- [HIG-~~2000~~002] [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:~~1961~~eq] [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:~~1961~~tr] [7] S. Glashow, “Partial Symmetries of Weak Interactions”, *Nucl.Phys.* **22** (1961) 579–588. doi:10.1016/0029-5582(61)90469-2.
- [Weinberg:~~1967~~tq] [8] S. Weinberg, “A Model of Leptons”, *Phys.Rev.Lett.* **19** (1967) 1264–1266. doi:10.1103/PhysRevLett.19.1264.
- [Salam:~~1968~~rm] [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:IntroParticule] [10] D. Griffiths, “Introduction to Elementary Particles”. Wiley-VCH, 2004.
- [Fermi:~~1934~~hr] [11] E. Fermi, “An attempt of a theory of beta radiation. 1”, *Z. Phys.* **88** (1934) 161–177. doi:10.1007/BF01351864.
- [Morii:SMand~~BSM~~] [12] S. M. T. Morii, C.S. Lim, “The Physics of the Standard Model and Beyond”. World Scientific, 2004.
- [UA1WD~~Discovery~~] [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.
- [UA2WD~~Discovery~~] [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.

- UA1ZDiscovery** [15] UA1 Collaboration, “Experimental observation of lepton pairs of invariant mass around 95 GeV/ $c^2$  at the CERN SPS collider”, *Phys. Lett.* **B126** (1983) 398–410.  
2394  
2395 doi:10.1016/0370-2693(83)90188-0.
- UA2ZDiscovery** [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.  
2397
- Martin:1997gn** [17] S. P. Martin, “A Supersymmetry Primer”, *arXiv hep-ph* (sep, 1997). 128 pages.  
2399 Version 5 (December 2008) contains a change in convention that flips the signs of  
2400 sigma and sigmabar matrices. It also contains a total of about 2 pages of updates,  
2401 mostly on supersymmetry breaking issues. Errata and a version with larger type (12  
2402 pt, 142 pages) can be found at <http://zippy.physics.niu.edu/primer.html>.
- FeynmanDiag2003s** [18] CERN Computing Newsletter.
- Rainwater:1998kj** [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.  
2405  
2406
- PDG** [20] Particle Data Group Collaboration, “Review of particle physics”, *J. Phys.* **G37** (2010) 075021. doi:10.1088/0954-3899/37/7A/075021.  
2408
- MSTWXSectionPlots** [21]
- LHCHiggsXSecGroup** [22] LHC Higgs Cross Section Working Group Collaboration, “Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables”, arXiv:1101.0593.  
2411
- MHMaxBenchmark** [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.  
2413  
2414  
2415
- CMSExperiment** [24] CMS Collaboration, “The CMS experiment at the CERN LHC”, *JINST* **3** (2008) S08004.  
2417
- CMS-PAS-TRK-10-004** [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).  
2419  
2420
- CMS-PAS-TRK-10-005** [26] CMS Collaboration, “Tracking and Primary Vertex Results in First 7 TeV Collisions”, *CMS PAS CMS-PAS-TRK-10-005* (2010).  
2422
- CMS-PTDR-I** [27] G. L. Bayatian et al., “CMS Physics Technical Design Report Volume I: Detector Performance and Software”. Technical Design Report CMS. CERN, Geneva, 2006.  
2424
- CMS-PAS-PFT-08-001** [28] CMS Collaboration, “CMS Strategies for tau reconstruction and identification using particle-flow techniques”, *CMS PAS CMS-PAS-PFT-08-001* (2008).  
2426
- CMS-PAS-PFT-09-001** [29] CMS Collaboration, “Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET”, *CMS PAS CMS-PAS-PFT-09-001* (2009).  
2428
- TMVA** [30] A. Hoecker et al., “TMVA - Toolkit for Multivariate Data Analysis”, *ArXiv Physics e-prints* (March, 2007) arXiv:arXiv:physics/0703039.  
2430

- [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).
- [DavisTau] [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-2010-082] [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).
- [pythia64] [37] S. M. T. Sjöstrand and P. Skands, “PYTHIA 6.4 Physics and Manual”, 2000.
- [taula] [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-PTDR-II] [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).
- [CDFrefP2463a] [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).
- [CDFFakerateD2473] [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”.  
2481 <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-2010-332] [56] G. Bauer et al., “Modeling of  $W \rightarrow \ell\nu$  MET with Boson Recoil”, *CMS Note* **2010/332** (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).
- [EWK-pdfUncertaintyTools] [58] J. Alcaraz. <https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideEWKUtilities>.
- [CTEQQp2494] [59] P. M. Nadolsky et al., “Implications of CTEQ global analysis for collider observables”, *Phys. Rev.* **D 78** (2008) 013004, arXiv:0802.0007.
- [pdfAccS24961] [60] PDF4LHC Working Group. <http://www.hep.ucl.ac.uk/pdf4lhc/PDF4LHCrecom.pdf>.
- [pdfAccS24962] [61] R. C. G. D. Bourilkov and M. R. Whalley, “LHAPDF: PDF use from the Tevatron to the LHC”, arXiv:0605.0240.
- [Ld98II] [62] CMS Collaboration, “Absolute luminosity normalization”, *CMS Detector Performance Summary* **CMS-DPS-2011-002** (2008).