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

By

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

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Physics

in the

OFFICE OF GRADUATE STUDIES

of the

1 UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Professor John Conway (Chair)

Professor Robin Erbacher

Professor Mani Tripathi

Committee in Charge

2011

for Alvin A Klose

Abstract

blah blah blah

Acknowledgments

Hooray for everybody.

Table of Contents

List of Figures

List of Tables

Chapter 1

The Standard Model and Beyond

?⟨ch:theory⟩?

§1.1 The Standard Model

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

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

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

§1.1.1 Quantum Electrodynamics and Gauge Invariance

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

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

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

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

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

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

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

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

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

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

$$i\hbar\gamma^\mu\partial_\mu\psi - mc\psi = 0. \tag{1.4}$$ eq:DiracEquatior

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

The Lagrangian corresponding to the Dirac equation (1.4) is

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

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

$$\psi' \rightarrow e^{i\theta}\psi. \tag{1.6}$$ eq:U1GaugeTran

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

$$\partial_\mu \rightarrow D_\mu = \partial_\mu - \frac{ie}{\hbar}A^\mu \tag{1.7}$$ {?}

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

$$\mathcal{L} = \overline{\psi}(i\hbar c\gamma^\mu D_\mu - mc^2)\psi. \tag{1.8}$$ eq:LocalQEDLag

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

$$\mathcal{L}_{int} = \frac{e}{\hbar}\overline{\psi}\gamma^\mu\psi A_\mu. \tag{1.9}$$ {?}

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

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

The gauge symmetry group of QED is $U(1)$, the unitary group of degree 1. This symmetry can be visualized as a rotation of a two–dimensional unit vector. (The application of the gauge transformation $e^{i\theta}$ rotates a number in the complex plane.) In 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 [?] to explain the phenomenon of the β decay of neutrons. The initial theory was a four-fermion "contact" theory. In a contact theory, all four fermions come involved in the β–decay are connected at a single vertex. The Fermi theory Hamiltonian for the β–decay of a proton is then [10]

$$H = \frac{G_{\beta}}{\sqrt{2}}\left[\overline{\psi}_p\gamma_{\mu}(1 - g_A\gamma_5)\psi_n\right]\left[[\overline{\psi}_e\gamma^{\mu}(1 - \gamma_5)\psi_{\nu}\right] + h.c., \tag{1.11}$$ eq:FermiTheoryH

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

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

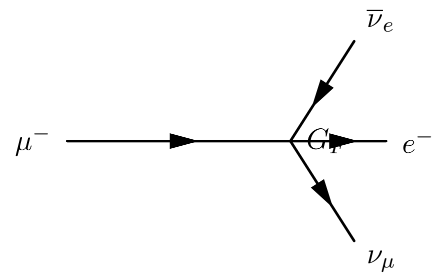


Figure 1.1: Feynamn diagram of muon decay in Fermi contact interaction theory.

⟨fig:MuonDecayContact⟩

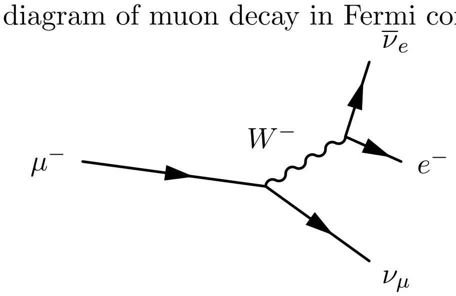


Figure 1.2: Feynamn diagram of muon decay proceeding through an intermediate gauge boson W^-.

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

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

eq:ContactAmpli

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

The first attempt to solve the problems with the Fermi theory was made by introducing an intermediate weak boson [6]. The contact interaction is replaced by a massive propagator, the $W^\pm$ bosons. The decay of a muon to an electron and two neutrinos then proceeds as pictured in Figure 1.2 with an amplitude given [10] by

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

eq:WeakPropaga

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

[2] 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 [9].

term of Equation 1.13 is the reason why the contact interaction original formulated by Fermi effectively described low–energy weak phenomenon. When the momentum transfer q in the interaction is small compared to M_W, the effect of the propagator is an effective constant. In the low energy limit, the full propagator in equation 1.13 is equivalent to the Fermi contact interaction in 1.12 as

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

eq:ContactVersus

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

§1.1.3 Spontaneous Symmetry Breaking

⟨sec:SSB⟩

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

A symmetry of a Lagrangian is spontaneously broken when the ground state, or vacuum, is at a value which about which the Lagrangian is not symmetric. In quantum field theories, a particle is interpreted as quantized fluctuations of its corresponding field about some constant (vacuum) ground state. The "effective" Lagrangian that we observe in the (low energy) laboratory would be the expansion of the Lagrangian about this stable point. The effective Lagrangian no longer obeys the original symmetry, which has been "broken". We give a brief example of the phenomenological effects of a spontaneously broken symmetry in a toy model, following the treatment in [10].

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

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

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

Since at the minimum the partial derivative of V is zero with respect to all fields, the second term in equation 1.16 is zero. The third term determines the masses of the particles in the theory. Since a mass term for a particle corresponding to a field ϕ_n in the Lagrangian appears as $\frac{1}{2}m^2\phi_n\phi_n$, we can identify

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

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

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

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

If the parameters m^2 and λ are both positive, then the minimum of V is at the origin ($\phi_1 = \phi_2 = 0$). In this case, the mass matrix term in Equation 1.16 takes the form $\left(\frac{\partial^2 V}{\partial\phi_a\partial\phi_b}\right)_0 = \frac{m^2}{2}\delta_{ab}$, where δ_{ab} is the Kronecker delta function. Therefore the mass matrix is already diagonalized, and the ϕ_1 and ϕ_2 both correspond to particles with mass m. If the m^2

[3] Technically, the symmetric transformation is

$$\begin{pmatrix}\phi_1\\\phi_2\end{pmatrix} \to \begin{pmatrix}\phi_1'\\\phi_2'\end{pmatrix} = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}\phi_1\\\phi_2\end{pmatrix},$$

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

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

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

eq:WignerPoint

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

eq:NambuGoldst

If the vacuum state is defined at the point in Equation 1.19, the symmetry is unbroken and the mass spectra is unchanged. However, the system is unstable at this point, as it is a local maximum. The true global minimum is defined as the set of points which satisfy Equation 1.20, which form a continuous circle in $\phi_1 - \phi_2$ space (and is therefore infinitely degenerate). We can choose any point on the circle as the vacuum expectation value (VEV). If the point $(\phi_1^{min} = \nu, \phi_2^{min} = 0)$[4] is chosen, evaluating Equation 1.17 yields the mass matrix

FiXme: *check matrix*

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

Breaking the symmetry has changed the mass spectrum of the physical particles in the model. There is now a massive particle with $m = v^2$ and a massless particle. This massless particle is called the "Goldstone boson." Goldstone found [1] that a massless particle appears for each generator in the symmetry group that is broken.

§1.1.4 The Higgs Mechanism

⟨sec:HiggsMech⟩

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

[4] The point chosen for the VEV here is not arbitrary. One can chose any point thats satisfies Equation 1.20 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.

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

$$\begin{aligned}\mathcal{L} &= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + (D_\mu\phi^*)(D^\mu\phi) - V(\phi^*\phi) && (1.21)\\ V(\phi^*\phi) &= -\mu^2\phi^*\phi + \lambda(\phi^*\phi)^2, && (1.22)\end{aligned}$$

eq:LocalInvariant eq:PotentialLoca

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

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

{?}

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

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

eq:HiggsMechani

If the gauge transform is chosen to be $\alpha(x) = \theta(x)/\nu$, the fields of are defined in the so–called "unitary gauge"[5] and have the special forms

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

eq:AfterUnitaryG

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

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

eq:HiggsMechani

The breaking of the original symmetry has dramatically altered the physical consequences of the model. In its unbroken form, the model described by Equation 1.21 would produce two real massive particles and one massless gauge boson mandated by local gauge invariance. After symmetry breaking, the effective Lagrangian in Equation 1.26 contains a massive scalar η with $m = \sqrt{2\mu^2}$ and a *massive* gauge boson B_μ with mass $m = \sqrt{2}e\nu$. By acquiring a mass, the gauge boson B_μ has acquired the degree of freedom (as it can now be longitudinally polarized) previously associated to the second degree of freedom in the

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

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

§1.1.5 Electroweak Unification

ec:ElectroweakUnification⟩

In the 1960s, the ideas of local gauge invariance in field theories, spontaneous symmetry breaking, and the Higgs mechanism were combined by Glashow [6], Weinberg [7] and Salam [8] to form the unified theory of electroweak interactions, the nucleus of the Standard Model. This model successfully unified the electromagnetic and weak interactions into a unified theory with a larger symmetry group. The reason for the empirically observed difference in scales between two interactions is due to the larger, unified symmetry group being broken. This broken symmetry creates heavy gauge bosons via the Higgs mechanism, whose large mass decreases the strength of "weak" interactions at low energy, as discussed in Section 1.1.2. The model successfully predicted the existence and approximate masses of the weak force carriers, the $W^{\pm}$ and Z bosons. These particles were later observed [11, 12, 13, 14] with the predicted masses at the UA1 and UA2 experiments.

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

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

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

eq:EWDoubletFo

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

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

?eq:Su2Generator

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

eq:Su2Generator

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

$$J_\mu^\pm = \overline{L}\gamma_\mu \tau^\pm L. \tag{1.30}$$

eq:WeakCurrentI

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

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

eq:EWNeutralCu

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

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

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

?eq:FermionAndC

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

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

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{1.33}$$

{?}

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

{?}

where D_μ is the gauge covariant derivative containing couplings to both the $SU(2)_L$ and $U(1)_Y$ gauge fields, and V has a form analogous to V in Equation 1.22. At this point we also add $SU(2)_L \times U(1)_Y$ invariant "Yukawa" terms

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

eq:YukawaTerms

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

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

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

{?}

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

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

so the broken Lagrangian contains the correct symmetry properties.

The gauge boson content of the electroweak interaction is obtained by parameterizing the Higgs field in the magnitude–phase notation of Equation 1.24 and using the unitary gauge (see Section 1.1.4), where the gauge transformation is chosen so Higgs field is real. The Higgs scalar doublet is then

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

?eq:HiggsFieldPa

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

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

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

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

eq:GaugeBosonM

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

$$W^{\pm}_\mu = \frac{A'^1_\mu \mp iA'^2_\mu}{2}. \tag{1.39}$$

{?}

This is the familiar $W^{\pm}$ boson of β and muon decay, and has mass $M_W = \frac{1}{2}g\nu$. The third term in Equation 1.38 can be written in matrix form and then diagonalized into mass eigenstates

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

{?}

giving a massive Z boson with

$$M_Z = \frac{\nu}{2}\sqrt{g^2 + g'^2} \tag{1.41}$$

eq:ZBosonMass

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

by

$$M_Z \equiv \frac{M_W}{\cos\theta_W}, \tag{1.42} \{?\}$$

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

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

$$M_Z \approx \frac{76\,\text{GeV}}{\sin 2\theta_W} > 76\,\text{GeV}. \tag{1.44} \{?\}$$

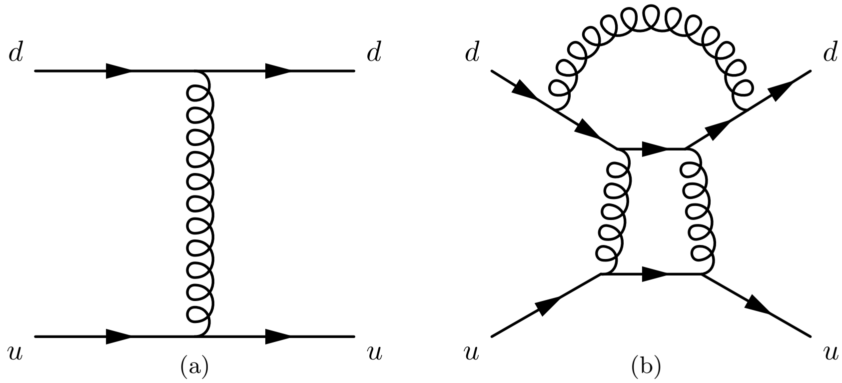
The discovery of the W [11, 12] and Z [13, 14] at the CERN SPS was a huge triumph for the electroweak model.

The model that is presented in this section assumes only one species of leptons, the electron and its associated neutrino. The electroweak model is trivially extended [10] to include the other species (μ, τ) of leptons and the three families of quarks. The masses of the fermions are determined by the Yukawa terms in Equation 1.35. Each particle species has a Yukawa term relating the Higgs VEV to its mass that is not constrained by the theory, and must be determined by experiment.

§1.1.6 Quantum Chromodynamics

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

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



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.

colored object exists in nature. The corollary of this is that it is believed to be impossible for a single quark or gluon to be observed. The mechanism that gives rise to 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 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. In practice what is done is to "factorize" QCD interaction amplitudes into a perturbative (high–energy) part and a non–perturbative part. The perturbative portion is calculable using the Feynman calculus; the non–perturbative must be estimated from parameterization functions that are experimentally measured.

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

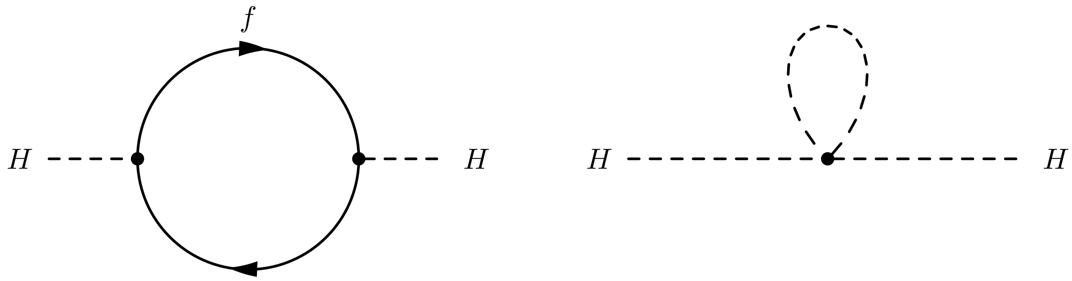
§1.2 Beyond the Standard Model

?⟨sec:BSM⟩?

The Standard model is one of the most successful theories of the natural world ever created. The predictions of the SM have been tested to many orders of magnitude and no experiment to date[6] has found a result statistically incompatible with the Standard Model. However, there is a general consensus in the physics community that the Standard Model is not complete. It is believed that it is only an effective theory that is valid below some energy scale Λ. Above this energy, there must exist some other "new physics," which unifies the forces of the Standard Model and correctly describes the natural world at all scales, while maintaining equivalence to the Standard Model at low energy. This concept is analogous to the relationship between the effective Fermi contact theory of Section 1.1.2 and the unified electroweak theory of Section 1.1.5. The size of the cutoff scale Λ is estimated [10] to be $O(10^{15})$ GeV for a unified theory with $SU(5)$ symmetry and even larger, $O(10^{19})$ GeV $= M_{planck}$ if the theory is unified with gravity.

There are many compelling reasons that indicate that the Standard Model is incomplete. One is the fact that the model does not include gravity, which has still not be successfully reformulated into a quantum mechanical theory. Another is that cosmological observations indicate the presences of massive amounts of "dark matter" in the universe. Dark matter is expected to be composed of a stable massive neutral particle which interacts very weakly with other matter; no Standard Model particle fits this description. Finally, there is the "hierarchy," or fine–tuning problem. This problem strongly affects the Higgs sector,

[6] 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.



(a) (b)

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

iggsMassLoopCorrections⟩

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

§1.2.1 The Hierarchy Problem

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

$$m_H^2 = -\frac{|\lambda_f|^2}{8\pi^2}\Lambda^2 + \cdots , \tag{1.45}$$

eq:HiggsMassCor

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

§1.2.2 Supersymmetry

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

§1.2.3 The Minimal Supersymmetric Model

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

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

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

⟨tab:chiral⟩

Names	spin 1/2	spin 1	$SU(3)_C,\ SU(2)_L,\ U(1)_Y$
gluino, gluon	$\widetilde{g}$	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	$\widetilde{B}^0$	B^0	$(\ \mathbf{1},\ \mathbf{1}\ ,\ 0)$

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

⟨tab:gauge⟩

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

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

$$W_{\text{MSSM}} = \overline{u}\mathbf{y_u}QH_u - \overline{d}\mathbf{y_d}QH_d - \overline{e}\mathbf{y_e}LH_d + \mu H_u H_d\ , \tag{1.46}$$ {?}

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

[7] 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.

to the down–type Higgs. This feature has large phenomenological consequences, which are discussed in 1.3.2. The scalar portion of the $W_{\rm MSSM}$ potential defines the spontaneous symmetry breaking. Similar to the scalar potential V symmetry breaking of Section 1.1.3, the potential of V at the minimum is found[8] to be

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

eq:MSSMScalarP

Under suitable choices[9] of the parameters in Equation 1.47, the up–type and down–type neutral Higgs fields acquire a VEV, ν_u and ν_d, respectively. The VEVs are related to the VEV of electroweak symmetry breaking (Equation 1.41) in the SM,

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

The ratio of the VEVs is expressed as

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

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

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

eq:MSSMLittleH

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

?eq:MSSMHiggs0

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

$$m_t = y_t v \sin\beta, \qquad m_b = y_b v \cos\beta, \qquad m_\tau = y_\tau v \cos\beta. \tag{1.50}$$

?eq:YukawaTanB

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

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

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

§1.3 Searches for the Higgs boson

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

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

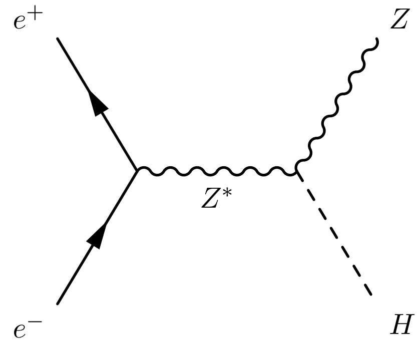
$$\mathcal{L}_e = -\frac{G_e}{\sqrt{2}}(\nu + H(x))\bar{e}e = -\frac{G_e \nu}{\sqrt{2}}\bar{e}e - \frac{G_e}{\sqrt{2}}H(x)\bar{e}e. \tag{1.51}$$ eq:ElectronYukaw

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

$$\frac{G_e \nu}{\sqrt{2}} = \frac{m_e}{\nu}. \tag{1.52}$$ eq:HiggsVEVtoC

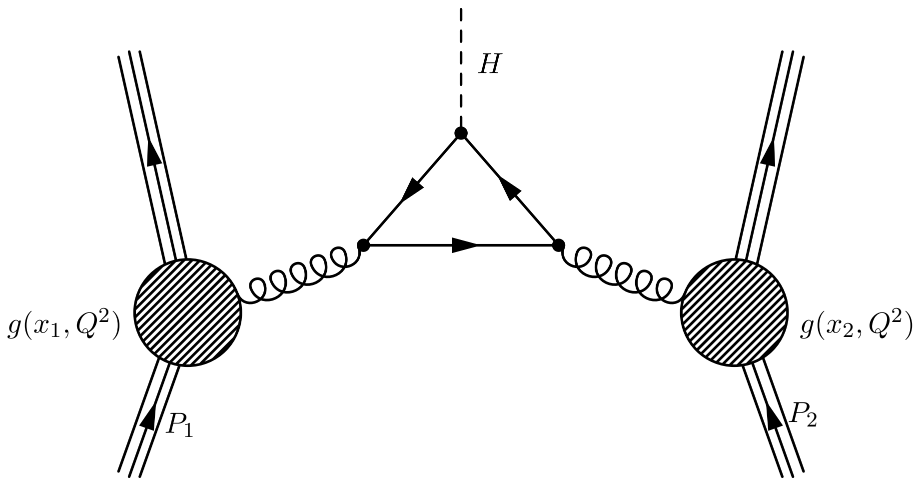
The left–hand side of Equation 1.52 is the same as the constant in the electron–Higgs $H(x)\bar{e}e$ coupling term in Equation 1.51. Therefore the coupling between the fermions and Higgs boson is proportional to their mass! This remarkable fact shapes the possible production modes and the branching fractions of Higgs decays.

The dominant modes of Higgs boson production depend on the type of experiment. In general, Higgs production is favored through high–mass intermediate states, due to the mass–proportional coupling. At the Tevatron and LEP experiments, which will be introduced in the next section, the dominant SM Higgs production mode is "Higgstrahlung,"



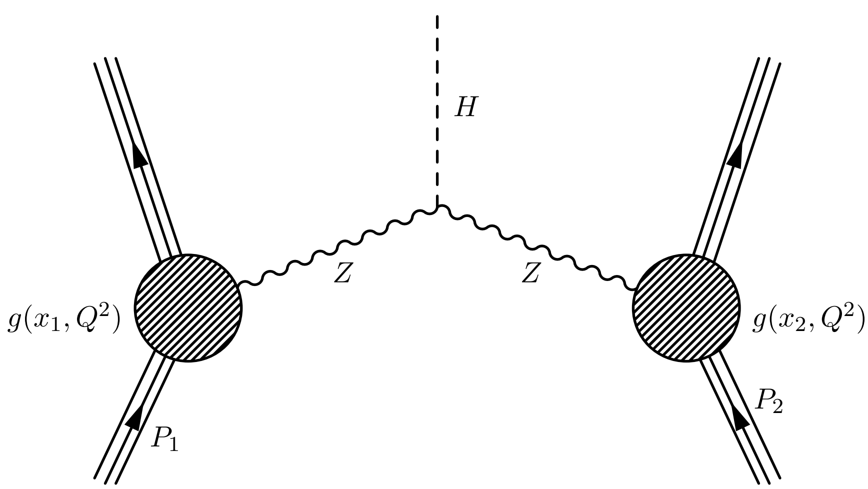
?⟨fig:HiggsStrahlung⟩?

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 heavy quarks in the central loop. Image credit: [**?**]



⟨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.

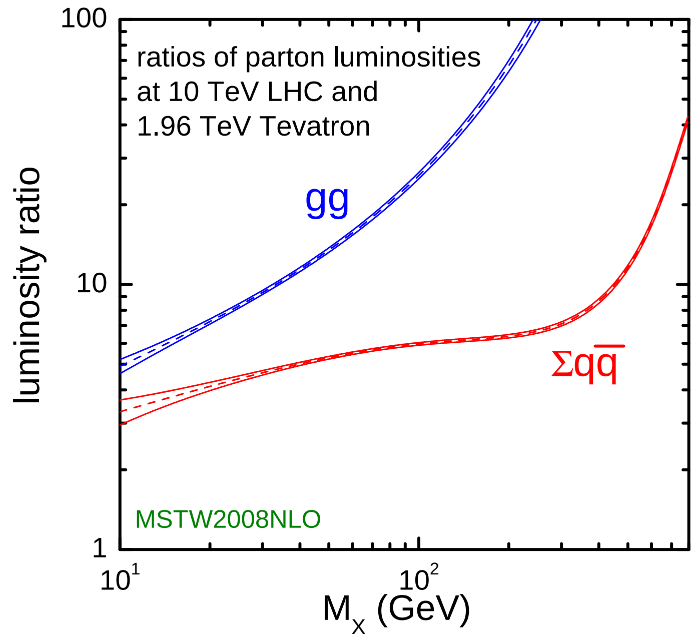


Figure 1.8: Ratio of the parton luminosity (the amount of luminosity contributed by the different species that compose the proton) of the LHC (at $\sqrt{s} = 10$ TeV) and the Tevatron. The large increase in gluon–gluon luminosity affects the favored production mechanisms of the Higgs boson.

⟨fig:GluonLumiRatio⟩

where a virtual $W^{\pm}$ or Z gauge boson is produced and then radiates a Higgs boson. Higgstrahlung is illustrated in Figure **??**. At the Large Hadron Collider, higher gluon luminosities (See Figure 1.8) result in the favored cross section being "gluon fusion," (illustrated in Figure 1.3.1) where two gluons from the incident protons combine in a quark (dominated by the massive top quark) loop which then radiates a Higgs boson. Another important channel **??** is "vector boson fusion," (Figure 1.3.1) where weak gauge bosons ($W^{\pm}$ or Z) are radiated from the incoming quarks and fuse to produce a Higgs. This is a notable channel due to the lack of "color–flow" (gluons) between the two protons, producing an event with low central jet activity and two "tag–jets" in the forward and backward regions. The theoretical cross sections for the SM Higgs at the LHC are shown in Figure 1.9.

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

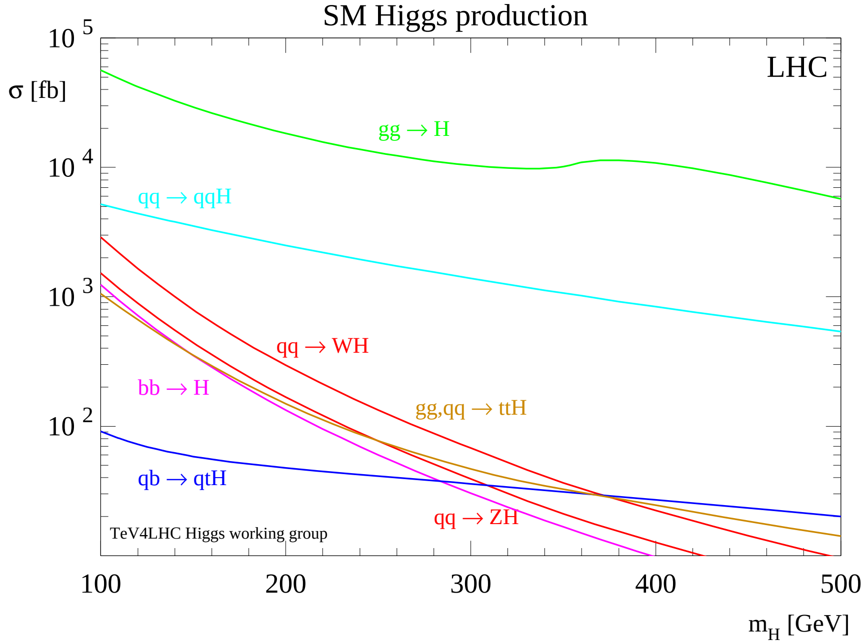


Figure 1.9: Cross section of the Standard Model Higgs boson versus the Higgs boson mass. The different curves give the contribution to the cross section from different production mechanisms. Source: [16].

⟨fig:LHCSMHiggsXsec⟩

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

§1.3.2 MSSM Higgs Phenomenology

⟨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 τ leptons and b quarks. When the parameter $\tan\beta$ is large, the coupling

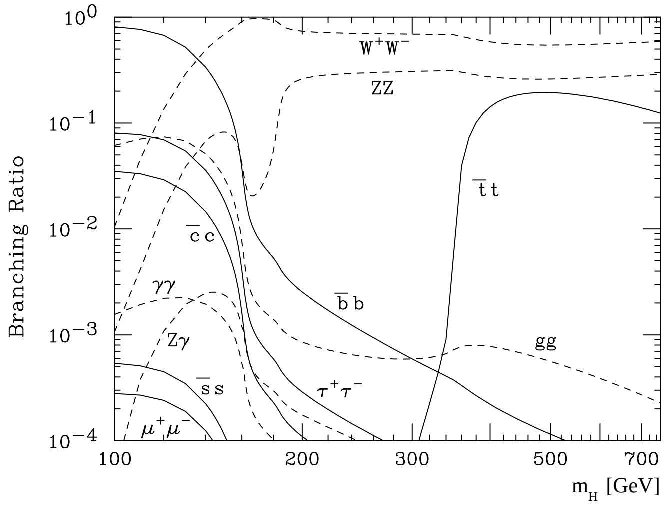


Figure 1.10: Branching fraction of the Standard Model Higgs bosons for different values of M_H. Source: [16].

⟨fig:SMHiggsBR⟩

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

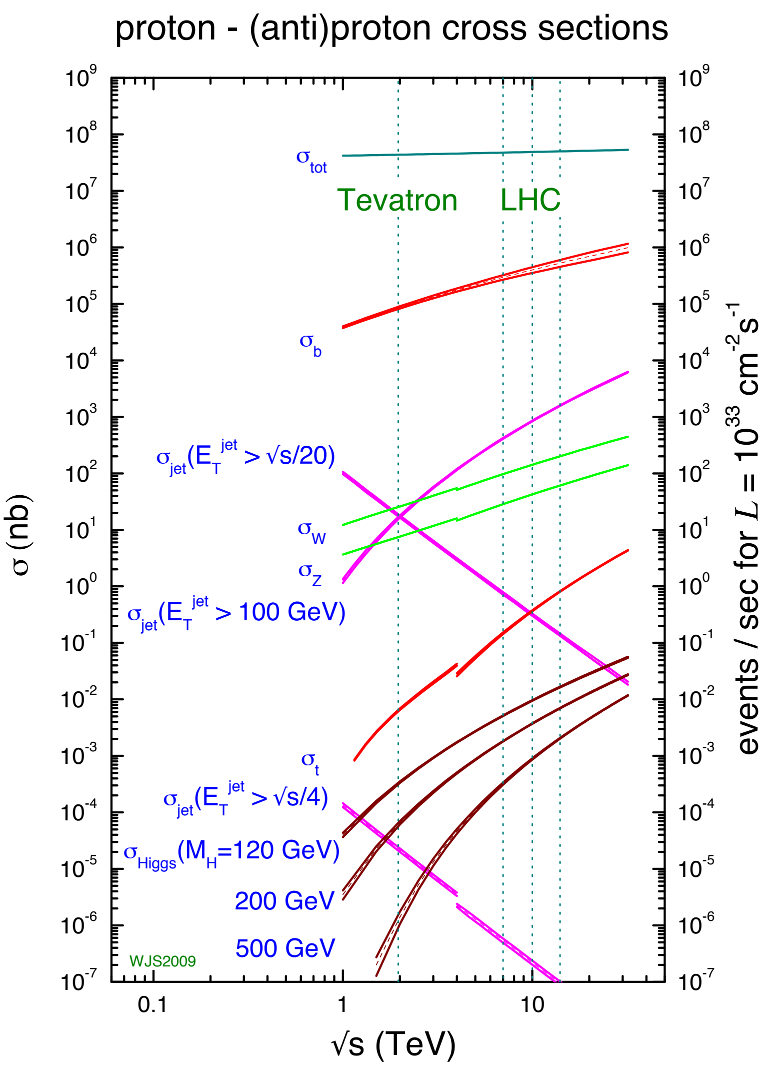
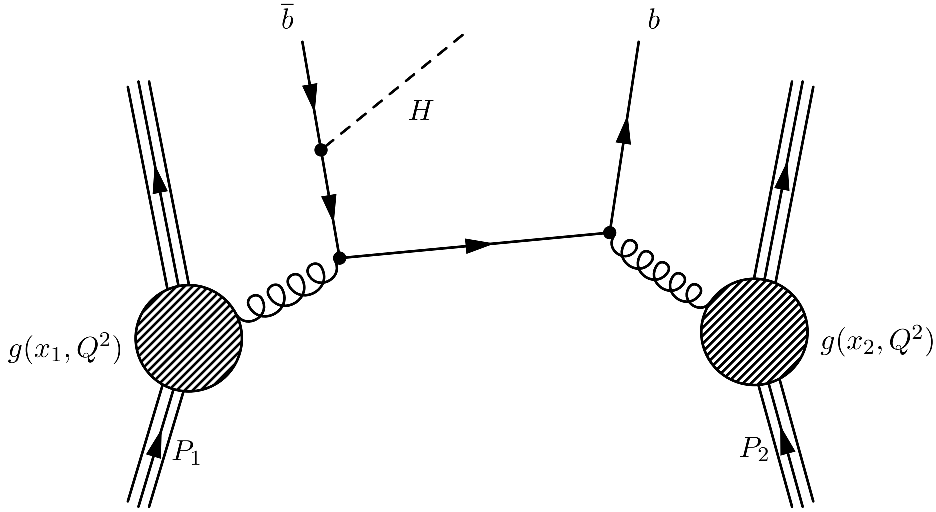


Figure 1.11: Cross sections of various processes at hadron colliders. The horizontal axis represents the center of mass energy of the collision. Of note is the vast difference in scales between Higgs production (maroon lines, $O(10{-}2)$ nb) and the QCD cross section to produce $b\bar{b}$ pairs (red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

dronColliderCrossSections⟩



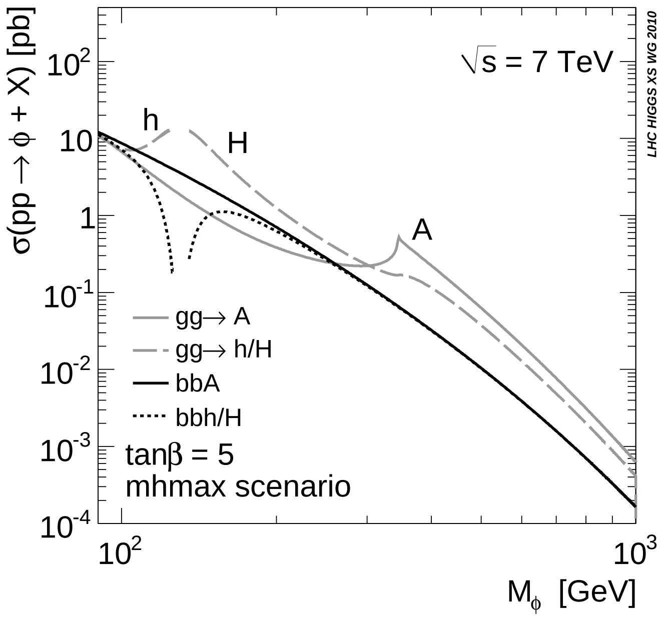
g:AssociatedBProduction⟩

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

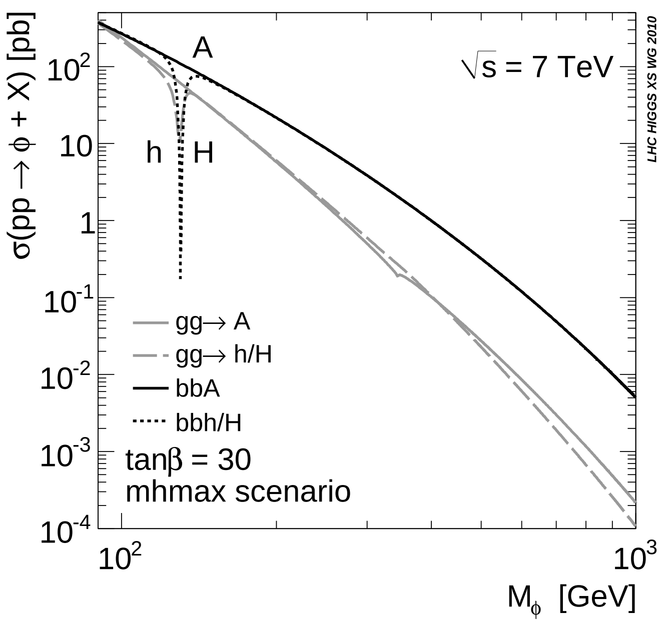
§1.3.3 Results from LEP and Tevatron

§1.4 The physics of the τ lepton

As discussed in sections 1.3.1 and 1.2.3, the τ lepton is an important probe of Higgs physics. The τ lepton has some unusual properties which make it particularly challenging at hadron colliders. With a mass of 1.78 GeV/c^2, the τ lepton is heaviest of the leptons. The nominal decay distance $c\tau$ of the τ lepton is 87 μm, which in practice means that the τ will always decay before reaching the first layer of the detector. Tau decays can be effectively classified into two types. “Leptonic” decays consist of a τ decaying to a light lepton ($\ell = e, \mu$) and two neutrinos $\tau^+ \rightarrow \ell^+ \nu_\tau \overline{\nu_\ell}$. “Hadronic” decays consist of a low–multiplicity collimated group of hadrons, typically $\pi^\pm$and π^0mesons. The hadronic decays of the τ lepton compose approximately 65% of the τ lepton branching fraction, with the remainder shared approximately equally by the leptonic decays. The branching fractions for the leptonic and most common hadronic decays are shown in table **??**.



(a)



(b)

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

MSSMXSectionsTanBeta⟩?

Visible Decay Products	Resonance	Mass (MeV/c^2)	Fraction [16]
Leptonic modes modes			
$e^- \nu_\tau \overline{\nu}_e$	-	0.5	17.8%
$\mu^- \nu_\tau \overline{\nu}_\mu$	-	105	17.4%
Hadronic modes			
$\pi^- \nu_\tau$	-	135	10.9%
$\pi^- \pi^0 \nu_\tau$	ρ	770	25.5%
$\pi^- \pi^0 \pi^0 \nu_\tau$	$a1$	1200	9.3%
$\pi^- \pi^- \pi^+ \nu_\tau$	$a1$	1200	9.0%
$\pi^- \pi^- \pi^+ \pi^0 \nu_\tau$	$a1$	1200	4.5%
Total			94.4%

⟨tab:decay˙modes⟩

Table 1.3: Resonances and branching ratios of the dominant decay modes of the τ lepton. The decay products listed correspond to a negatively charged τ lepton; the table is identical under charge conjugation.

Chapter 2

The Compact Muon Solenoid Experiment

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

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

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

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

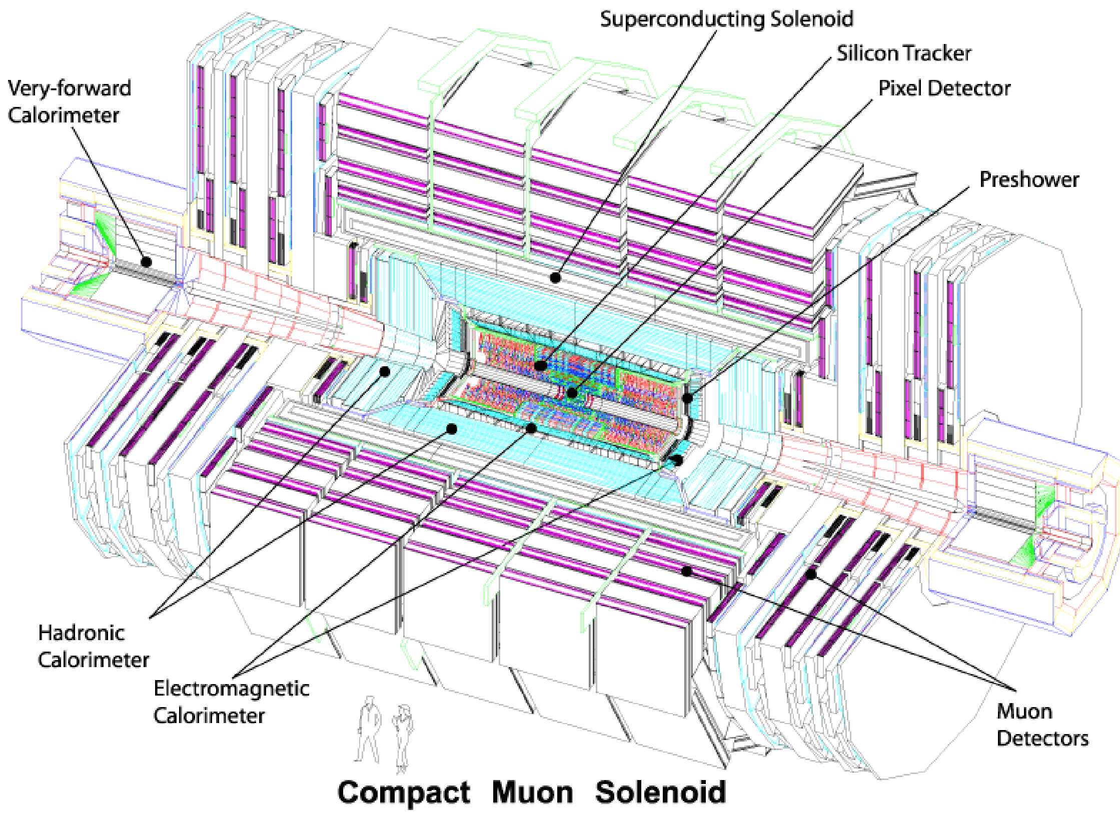
- Good missing–transverse–energy and dijet–mass resolution, requiring hadron calorimeters with a large hermetic geometric coverage and with fine lateral segmentation.

The detector uses a hermetic design that maximizes the solid–angle of the fiducial region to capture as much information about the collisions as possible. The general geometry of the detector is cylindrical. At 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 of different species of particles. The central (closest to interaction point) sub–detector are the charged particle tracking systems (the "tracker"). The tracker is designed to be a *non–destructive* instrument, which means that ideally that the momentum of particles are unchanged after passing through it. Outside of the tracker is the electromagnetic and hadronic calorimeters, which are abbreviated ECAL and HCAL, respectively. The calorimeter is a *destructive* detector, and is designed such that visible incident particles are completely absorbed. The outer layers of CMS are designed to measure muons, the one[1] species of particle that is immune to the effects of the calorimeter. The arrangement of destructive and non–destructive sub–detectors facilitates the identification of different types of particles. This concept is illustrated in Figure 2.1(b). In this chapter we give an brief overview of the LHC machine, and then describe the individual sub–detector systems of CMS.

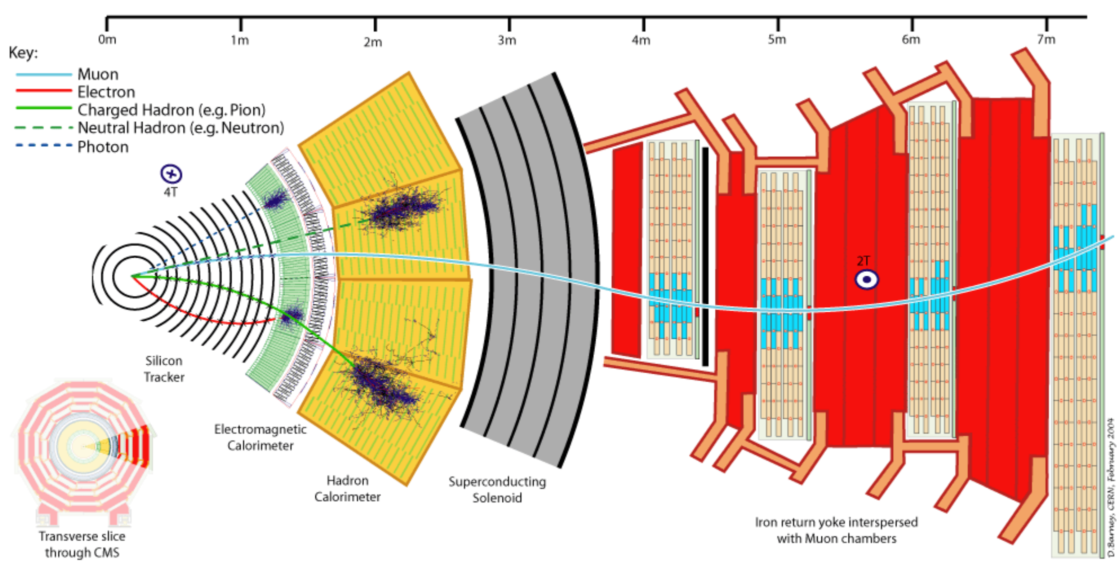
§2.1 The Large Hadron Collider

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

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



(a)



(b)

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

⟨fig:AllCMSCutaways⟩

§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 collisions at the LHC. The momentum of charged particles is measured in the detector by examining the curvature of the particles path as it travels through the magnetic field. The radius of curvature r of a charged particle in a magnetic field is given by

$$r = \frac{p_\perp}{|q|B}, \tag{2.1}$$

eq:LarmorRadius

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

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

[2] 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.

§2.3 Charged Particle Tracking Systems

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

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

FiXme: *right acronym*

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

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

[4] 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}.

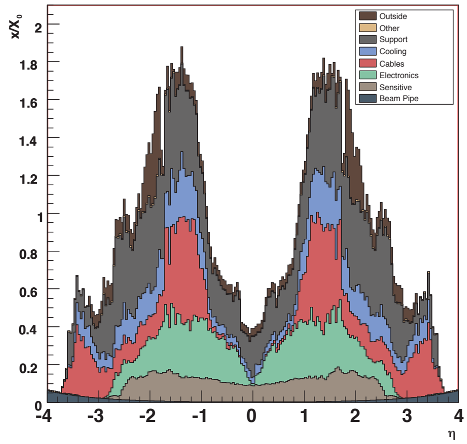


Figure 2.2: Material budget of the CMS tracker in units of radiation lengths X_0. The material budget is broken down into the contributions from the different components of the tracker. The amount of material is largest in the "transition region" between the barrel and endcap.

ig:TrackerMaterialBudget⟩

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

In the ideal case, the tracker would be a non–destructive instrument. However, charged particles can interact with the mass of the tracker (and its support infrastructure). These interactions limit the resolution of the tracker. The amount of matter in the tracker is referred to as the "material budget". The material budget of the CMS tracker depends heavily on the pseudorapidity η and is illustrated in Figure 2.2. The relatively large material budget of the CMS tracker has two effects: charged particles can "multiple scattering," interacting with material in the tracker. This can cause "kinks" in the reconstructed track. Hadronic particles (charged and neutral) can undergo "nuclear interactions," which are

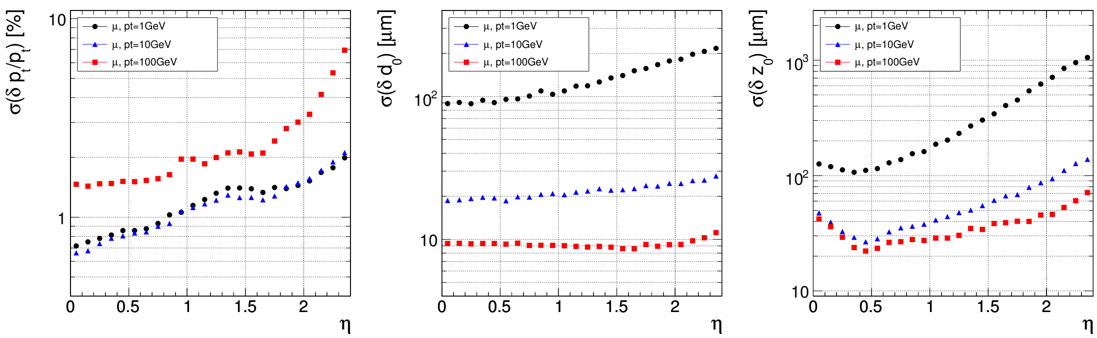


Figure 2.3: Expected resolutions of reconstructed transverse momentum (left), transverse impact parameter (center), and longitudinal impact parameter (right) versus absolute pseudorapidity $|\eta|$. The resolution is shown for three different cases of particle p_{T}, 1 GeV/c (black), 10 GeV/c (blue),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

⟨fig:ExpectedTrackerRes⟩

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

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

§2.4 Electromagnetic Calorimeter

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

[5]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.

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

The ECAL uses lead tungstate ($PbWO_4$) crystals as the scintillation medium. The crystals have a very large density, which allows the calorimeter to be relatively compact. To be able to correctly measure the energy of electrons and photons, an incident photon or electron must be completely stopped by interactions with the calorimeter. The quantities that determine if an electron or photon will be completely contained is the total depth of the crystal, the crystal density, and the radiation length property X_0 of the crystal. The radiation length X_0 is defined as the mean distance (normalized to material density) after which an electron will have lost $(1 - \frac{1}{e})$ of its energy. The $PbWO_4$ crystals of the CMS ECAL have a density of 8.28 g/cm^2 and a depth of 230 mm. A single crystal thus has a total radiation length of 25.8 X_0, and will capture on average 99.9993% of the energy of an incident electron. The front face of the crystal is 22 mm × 22 mm, which corresponds to an $\eta - \phi$ area of 0.00174×0.00174. The Molière radius of a material is the average radial profile size of an electromagnetic shower, and for $PbWO_4$ is 2.2 cm. The fact that the Molière radius is larger than the size of the individual crystals improves the spatial resolution of the measurement. As the shower is shared between multiple crystals, the relative amounts deposited in each crystal allows the true impact point to be determined with a resolution smaller than the individual crystal size.

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

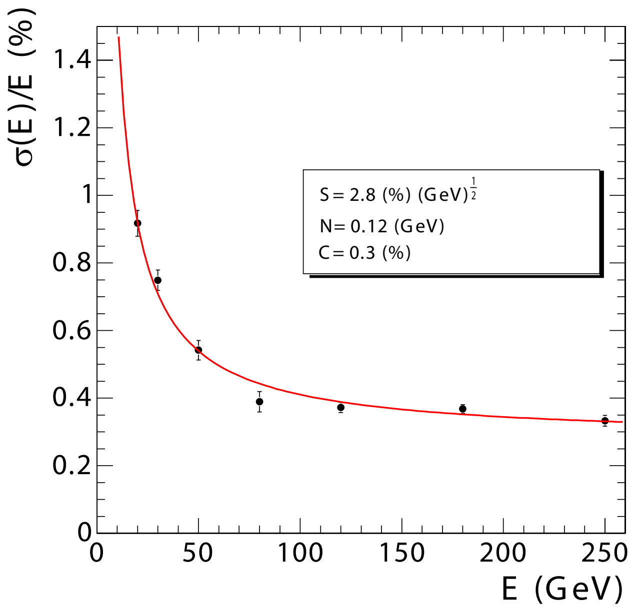


Figure 2.4: Energy resolution (in %) of the CMS ECAL measured at an electron test beam. The resolution depends on the incident energy of the electron. The points are fitted to function with the form given in Equation **??**. The fitted parameters are given in the legend.

⟨fig:ECALResolution⟩

responds to 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 synchronized to pulse between LHC bunch trains so the transparency can be continuously monitored while collisions are occuring.

The energy resolution of the ECAL is given by

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

eq:ECALResolut

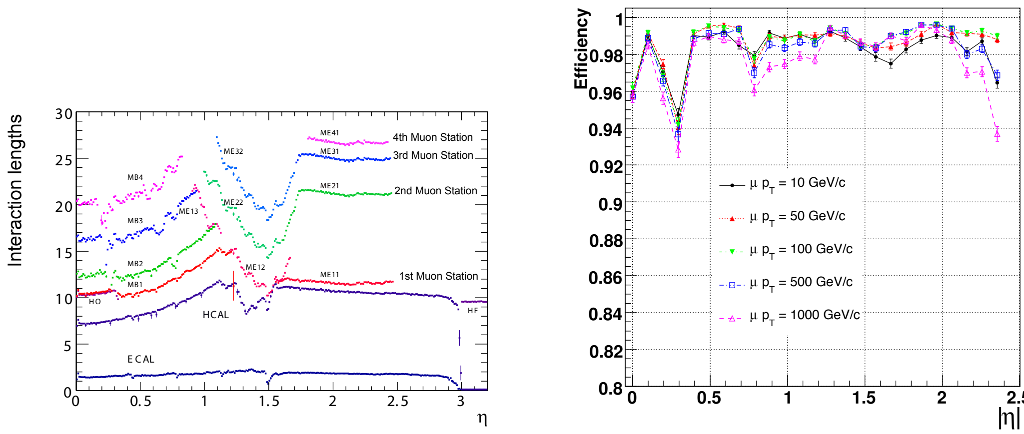
where S is a stochastic noise term (due to photon counting statistics), N is a noise term, and C is a constant term. The parameters of Equation 2.2 have been measured at an electron test–beam (see Figure 2.4). The energy resolution is better than 1% for electron energies greater than 20 GeV.

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

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

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



(a) (b)

etectorChapterMuonShit⟩?

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

§2.6 Muon System

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

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

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

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

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

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

passage in the gas and an "avalanche" current between the wire and a number of the cathode strips. The spatial position of the hit in two dimensions is found taking one coordinate from the wire and the other coordinate from the signal average of the cathode strips.

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

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

§2.7 Trigger System

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

The CMS trigger consists of two stages: a fast Level–1 (L1) trigger and a High–Level Trigger (HLT). The L1 trigger system is built on custom, typically re–programmable electronics 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 μs, requiring that the output from detector electronics be passed through memory pipelines to ensure that no bunch crossings go unanalyzed. The High–Level trigger runs on a farm of about 1000 commercial compute nodes and processes events that pass that are accepted by the L1 trigger. A High–Level trigger 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 Hzas 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 **??**. 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 5.3.

The muon component of all the triggers used in this analysis is based on the "L1 seed trigger" BLAH. 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,

FiXme: *what is it?*

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}, η, ϕ, 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.

§2.8 Particle Flow Reconstruction Algorithm

§2.9 DAQ

Chapter 3

Tau Identification: The Tau Neural Classifier

⟨ch:tanc⟩

§3.1 Introduction

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

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

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

§3.2 Geometric Tau Identification Algorithms

The tau identification strategies used in previously published CMS analyses are fully described in [23]. A summary of the basic methods and strategies is given here. There are two primary methods for selecting objects used to reconstruct tau leptons. The CaloTau algorithm uses tracks reconstructed by the tracker and clusters of hits in the electromagnetic and hadronic calorimeter. The other method (PFTau) uses objects reconstructed by the CMS particle flow algorithm, which is described in [24]. The particle flow algorithm provides a global and unique description of every particle (charged hadron, photon, electron, etc.) in the event; measurements from sub–detectors are combined according to their measured resolutions to improve energy and angular resolution and reduce double counting. The strategies described in this paper use the particle flow objects.

Both methods typically use an "leading object" and an isolation requirement to reject quark and gluon jet background. Quark and gluon jets are less collimated and have a higher constituent multiplicity and softer constituent p_T spectrum than a hadronic tau decay of the same transverse momentum. The "leading track" requirement is applied by requiring a relatively high momentum object near the center of the jet; typically a charged track with transverse momentum greater than 5 GeV/c within $\Delta R < 0.1$ about the center of the jet axis. The isolation requirement exploits the collimation of true taus by defining an isolation annulus about the kinematic center of the jet and requiring no detector activity about a threshold in that annulus. This approach yields a misidentification rate of approximately 1% for QCD backgrounds and a hadronic tau identification efficiency of approximately 50% [23].

§3.3 Decay Mode Tau Identification: Motivation

The tau identification strategy described previously can be extended by looking at the different hadronic decay modes of the tau individually. The dominant hadronic decays of taus consist of a one or three charged $\pi^{\pm}$ mesons and up to two π^0 mesons and are enumerated in table 1.4. The majority of these decays proceed through intermediate resonances and each of these decay modes maps directly to a tau final state multiplicity. Each intermediate resonance has a different invariant mass (see figure 3.1). This implies that the problem of hadronic tau identification can be re-framed from a global search for collimated hadrons satisfying the tau mass constraint into 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.

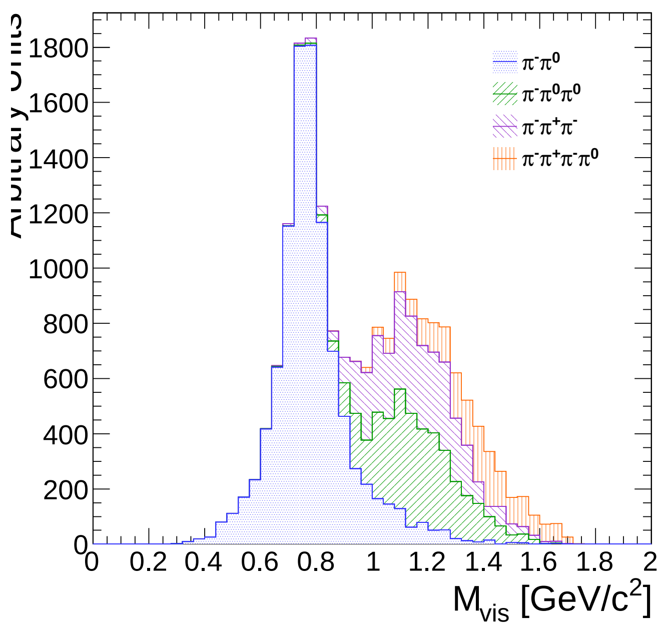


Figure 3.1: The invariant mass of the visible decay products in hadronic tau decays. The decay mode $\tau^- \rightarrow \pi^-\nu_\tau$ is omitted. The different decay modes have different invariant masses corresponding to the intermediate resonance in the decay.
⟨fig:trueInvMass⟩

§3.4 The Tau Neural Classifier

The Tau Neural Classifier algorithm reconstructs the decay mode of the tau–candidate and then feeds the tau–candidate to a discriminator associated to that decay mode to make the classification decision. Each discriminator therefore maps to a reconstructed decay mode in a one-to-one fashion. To optimize the discrimination for each of the different decay modes, the TaNC uses an ensemble of neural nets. Each neural net corresponds to one of the dominant hadronic decay modes of the tau lepton. These selected hadronic decays constitute 95% of all hadronic tau decays. Tau–candidates with reconstructed decay modes not in the set of dominant hadronic modes are immediately tagged as background.

§3.4.1 Decay mode reconstruction

⟨sec:decay˙mode˙reco⟩ The major task in reconstructing the decay mode of the tau is determining the number of π^0 mesons produced in the decay. A π^0 meson decays almost instantaneously to a pair of photons. The photon objects are reconstructed using the particle flow algorithm [24]. The initial collection of photon objects considered to be π^0 candidates are the photons in the signal cone described by using the "shrinking–cone" tau algorithm, described in [23].

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

Photon merging

Photons are merged into composite π^0 candidates by examining the invariant mass of all possible pairs of photons in the signal region. Only π^0 candidates (photon pairs) with a composite invariant mass less than 0.2 GeV/c are considered. The combination of the high granularity of the CMS ECAL and the particle flow algorithm provide excellent energy and angular resolution for photons; the π^0 mass peak is readily visible in the invariant mass spectrum of signal photon pairs (see figure 3.4.1). The π^0 candidates that satisfy the invariant mass requirement are ranked by the difference between the composite invariant mass of the photon pair and the invariant mass of the π^0 meson given by the PDG [16]. The best pairs are then tagged as π^0 mesons, removing lower–ranking candidate π^0s as necessary to ensure that no photon is included in more than one π^0 meson.

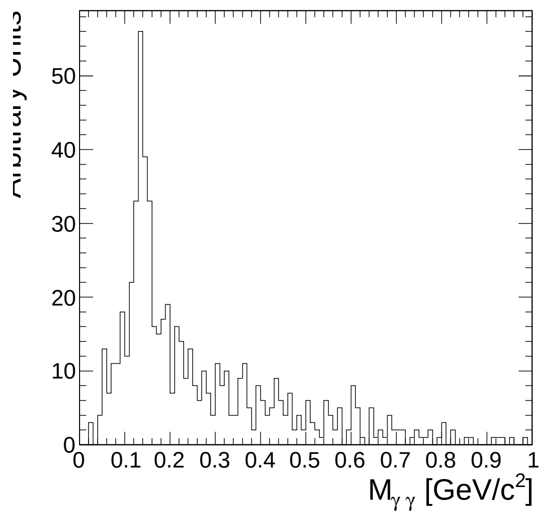


Figure 3.2: Invariant mass of the photon pair for reconstructed tau–candidates with two reconstructed photons in the signal region that are matched to generator level $\tau^- \to \pi^-\pi^0\nu_\tau$ decays.

mDiPhotonsForTrueDM1⟩

Quality requirements

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

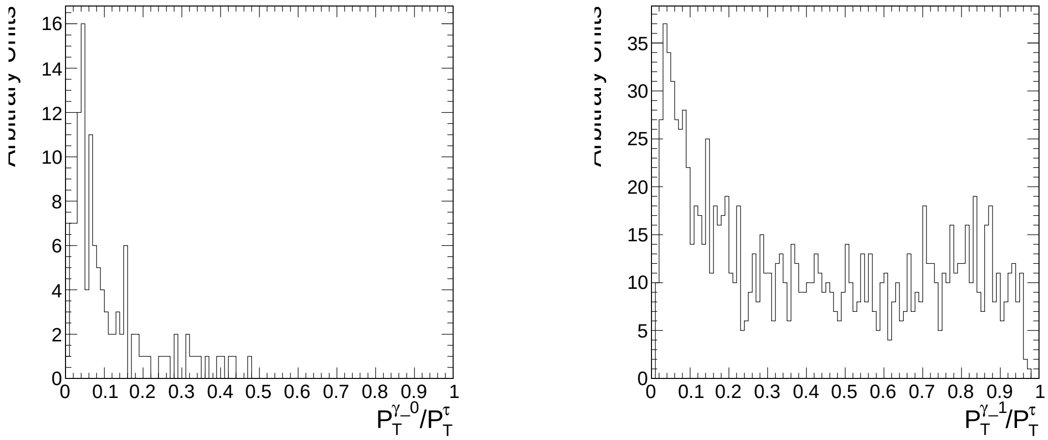


Figure 3.3: Fraction of total τ-candidate transverse momenta carried by the photon for reconstructed taus containing a single photons for two benchmark cases. On the left, the reconstructed tau–candidate is matched to generator level $\tau^- \rightarrow \pi^- \nu_\tau$ decays, for which no photon is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

⟨fig:photonFiltering⟩

Performance

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

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

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

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

: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.

§3.4.2 Neural network classification

Neural Network Training

⟨sec:tanc˙nn˙training⟩

The samples used to train the TaNC neural networks are typical of the signals and backgrounds found in common physics analyses using taus. The signal–type training sample is

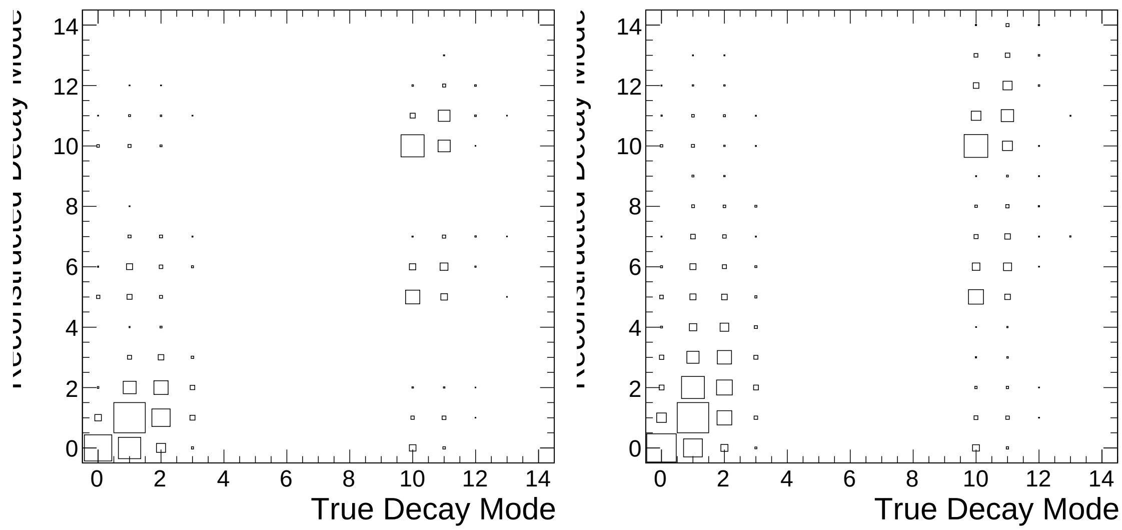


Figure 3.4: Correlations between reconstructed tau decay mode and true tau decay mode for hadronic tau decays in $Z \rightarrow \tau^+\tau^-$ events. The correlation when no photon merging or filtering is applied is shown on the right, and the correlation for the algorithm described in section 3.4.1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

⟨fig:dmResolution⟩

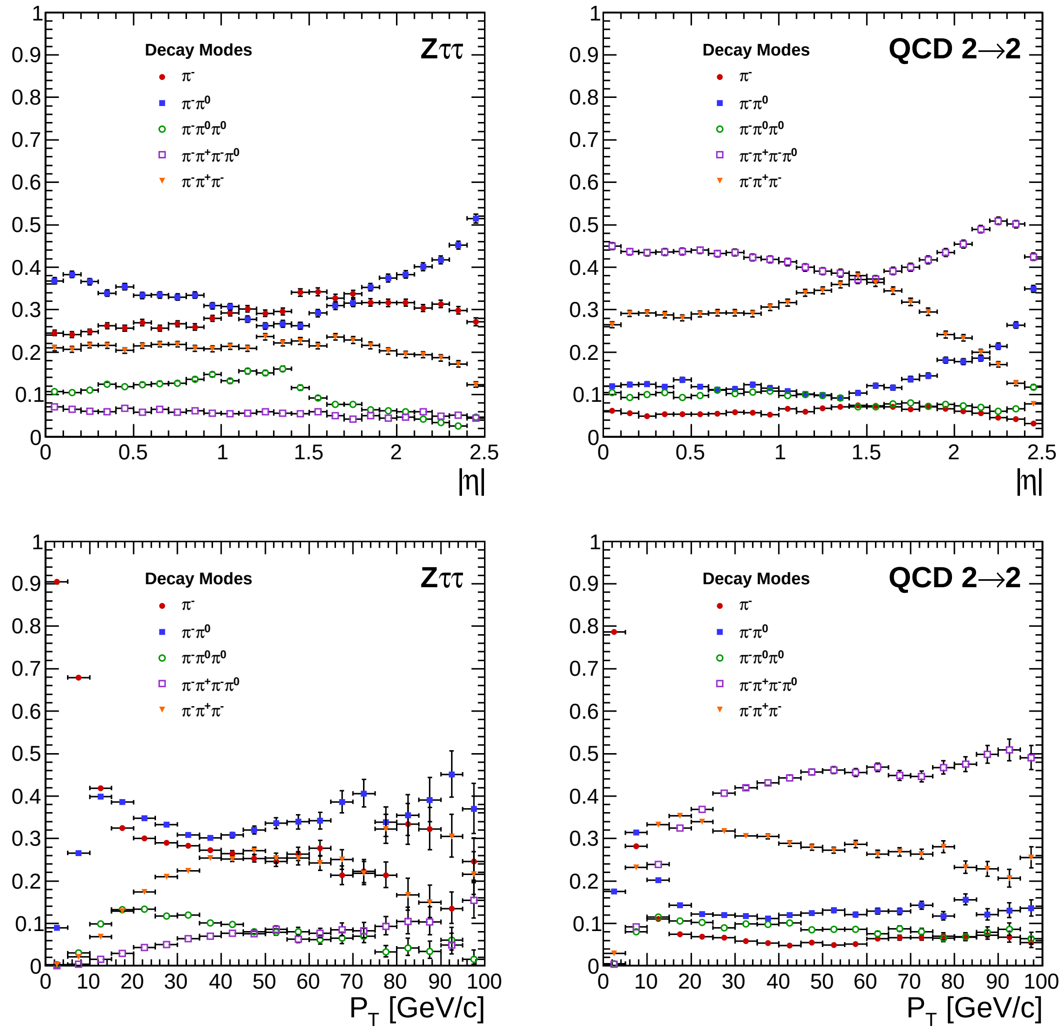


Figure 3.5: Kinematic dependence of reconstructed decay mode for tau–candidates from $Z \to \tau^+\tau^-$ (left) and QCD di–jets (right) versus transverse momentum (top) and pseudo–rapidity (bottom). Each curve is the probability for a tau–candidate to be reconstructed with the associated decay mode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

⟨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 $\hat{p_{\mathrm{T}}}$ 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 by the TMVA software, with half (40%) used as a validation sample used to ensure the neural network is not over–trained. The number of signal and background entries used for training and validation in each decay mode subsample is given in table 3.4.2.

The remaining 20% of the signal and background samples are reserved as a statistically independent sample to evaluate the performance of the neural nets after the training is completed. The TaNC uses the "MLP" neural network implementation provided by the TMVA software package, described in [25]. 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 [26], any continuous function $g(x)$ defined

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

⟨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

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

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

The neural network is trained for 500 epochs. At ten epoch intervals, the neural network error is computed using the validation sample to check for over–training (see figure 3.6). The neural network error E is defined [25] as

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

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

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

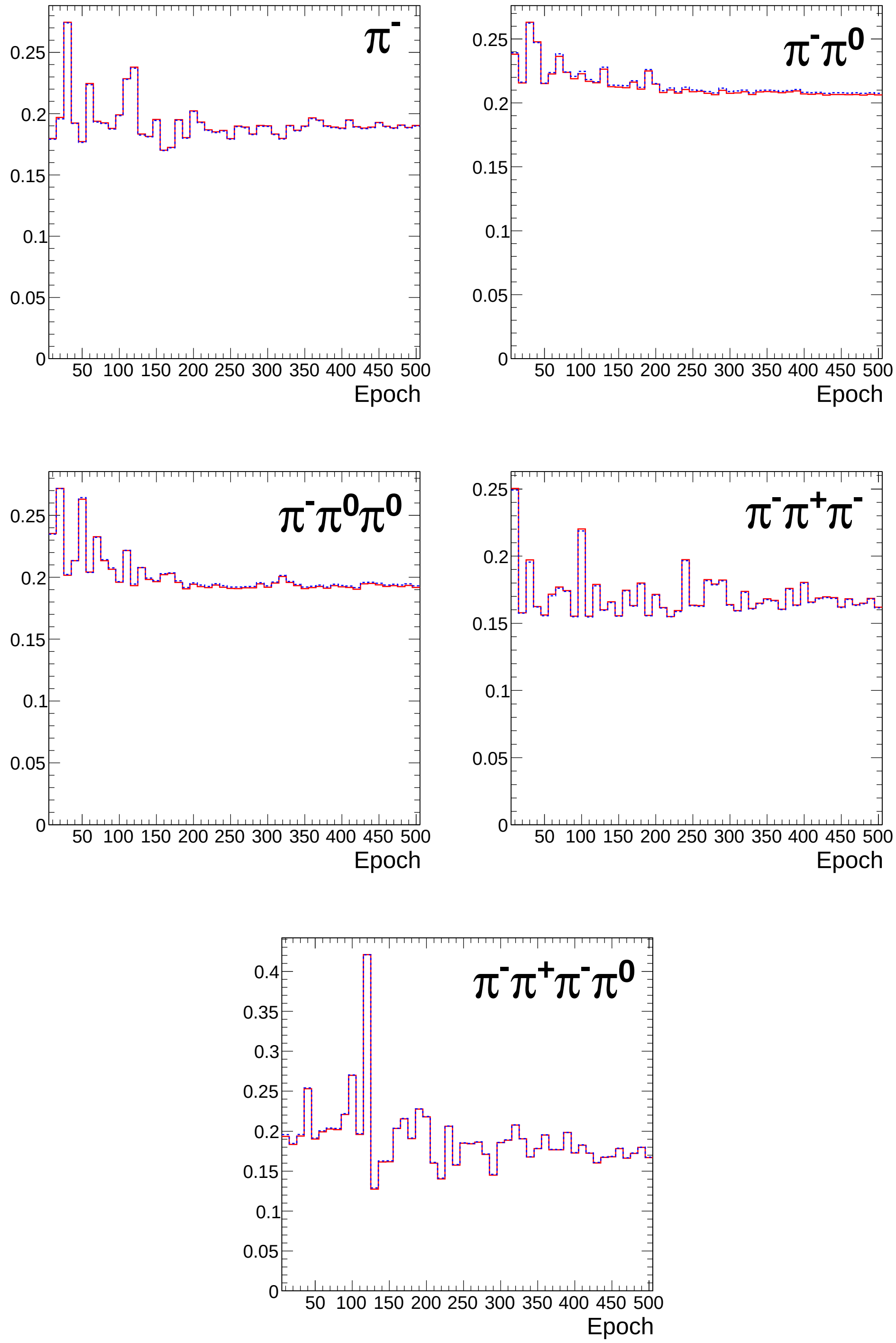


Figure 3.6: Neural network classification error for training (solid red) and testing (dashed blue) samples at ten epoch intervals over the 500 training epochs for each decay mode neural network. The vertical axis represents the classification error, defined by equation 3.2. N.B. that the choice of hyperbolic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

⟨fig:overTrainCheck⟩

while for background events the correlation is very small [?]. In the training signal and background samples, there is significant discrimination power in the p_{T} spectrum. However, it is desirable to eliminate any systematic dependence of the neural network output on p_{T} and η, as in practice the TaNC will be presented with tau–candidates whose $p_{\mathrm{T}} - \eta$ spectrum will be analysis dependent. The dependence on p_{T} and η is removed by applying a p_{T} and η dependent weight to the tau–candidates when training the neural nets.

The weights are defined such that in any region in the vector space spanned by p_{T} and η where the signal sample and background sample probability density functions are different, the sample with higher probability density is weighted such that the samples have identical $p_{\mathrm{T}} - \eta$ probability distributions. This removes regions of $p_{\mathrm{T}} - \eta$ space where the training sample is exclusively signal or background. The weights are computed according to

$$W(p_{\mathrm{T}}, \eta) = \mathrm{less}(p_{sig}(p_{\mathrm{T}}, \eta), p_{bkg}(p_{\mathrm{T}}, \eta))$$

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

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

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

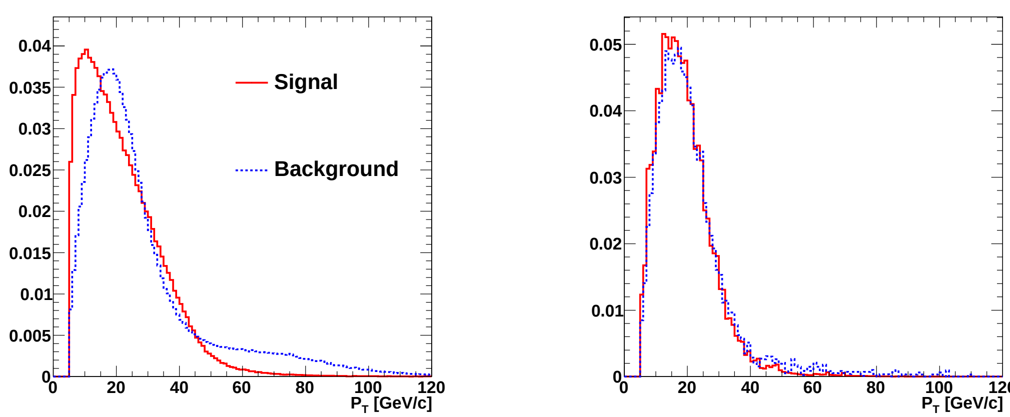


Figure 3.7: Transverse momentum spectrum of signal and background tau–candidates used in neural net training before (left) and after (right) the application of $p_{\mathrm{T}} - \eta$ dependent weight function. Application of the weights lowers the training significance of tau–candidates in regions of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

⟨fig:nnTrainingWeights⟩

Discriminants

⟨sec:tanc˙nn˙discriminants⟩

Each neural network corresponds to a different decay mode topology and as such each network uses different observables as inputs. However, many of the input observables are used in multiple neural nets. The superset of all observables is listed and defined below. Table 3.4 maps the input observables to their associated neural networks. The signal and background distributions of the input observables for tau–candidates in the training sample are shown in appendix **??**. In three prong decays, the definition of the "main track" is important. The main track corresponds to the track with charge opposite to that of the total charge of the three tracks. This distinction is made to facilitate the use of the "Dalitz" observables, allowing identification of intermediate resonances in three–body decays. This is motivated by the fact that the three prong decays of the tau generally proceed through $\tau^- \to a1^-\nu_\tau \to \pi^-\rho^0\nu_\tau \to \pi^-\pi^+\pi^-\nu_\tau$; the oppositely charged track can always be identified with the ρ^0 decay.

ChargedOutlierAngleN

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

ChargedOutlierPtN

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

DalitzN

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

Eta

Pseudo-rapidity of the signal region objects.

InvariantMassOfSignal

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

MainTrackAngle

ΔR between the "main track" and the composite four–vector formed by the signal region constituents.

MainTrackPt

Transverse momentum of the "main track."

OutlierNCharged

Number of charged objects in the isolation region.

OutlierSumPt

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

PiZeroAngleN

ΔR between the Nth π^0 object in the signal region (ordered by p_{T}) and the tau–candidate momentum axis.

PiZeroPtN

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

TrackAngleN

ΔR between the Nth charged object in the signal region (ordered by p_{T}) and the tau–candidate momentum axis, exclusive of the main track.

TrackPtN

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

Neural network performance

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

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

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

The 5D performance curve can also be parameterized by using the probability for a tau–candidate to be identified for a given decay mode. 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

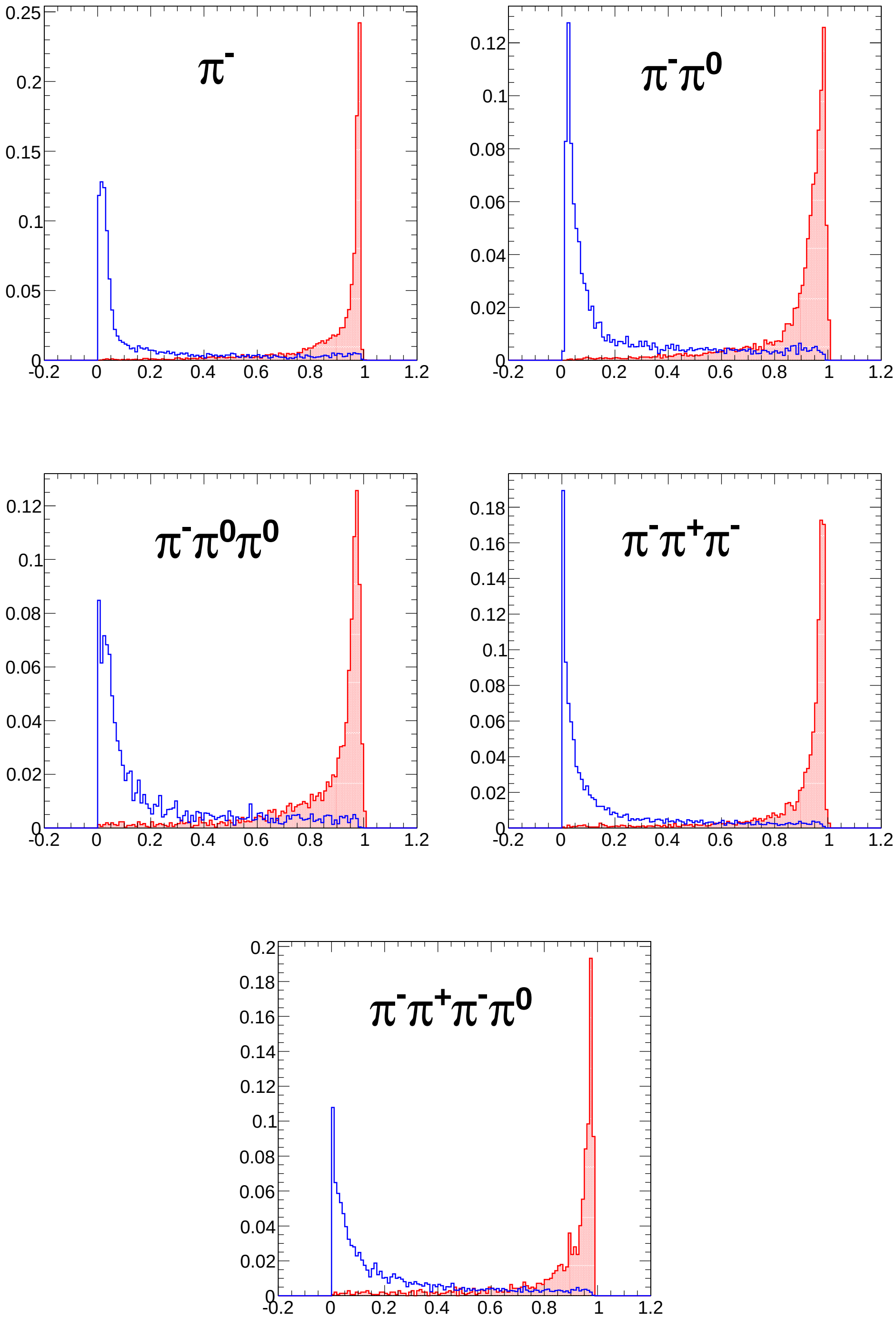


Figure 3.8: Neural network output distributions for the five reconstructed tau–candidate decay modes used in the TaNC for $Z \to \tau^+\tau^-$ events (red) and QCD di–jet events (blue).

fig:NNoutputDisributions⟩

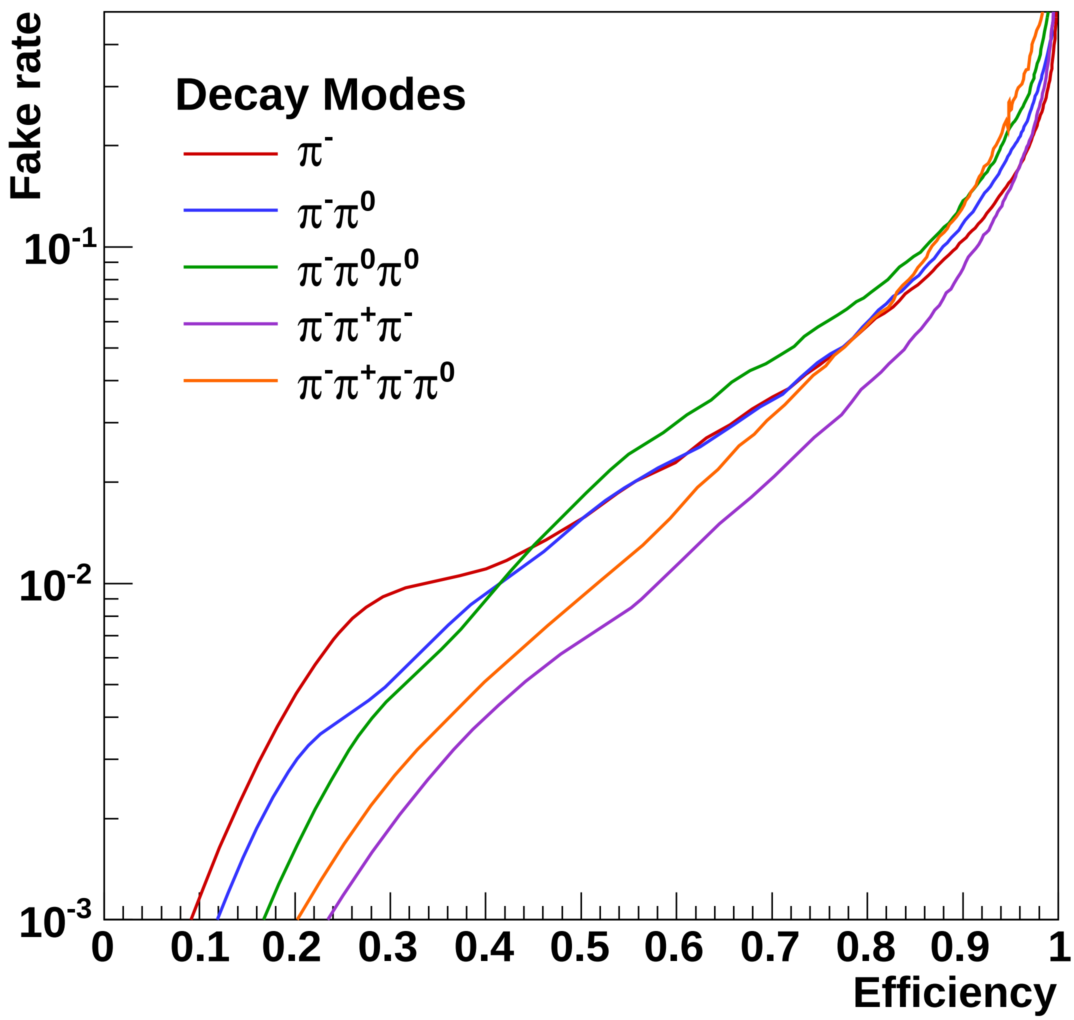


Figure 3.9: Performance curves for the five neural networks used by the TaNC for tau–candidates with transverse momentum greater than 20 GeV/c. Each curve represents the signal efficiency (on the horizontal axis) and background misidentification rate (vertical axis) for a scan of the neural network selection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

⟨fig:nnPerfCurves⟩

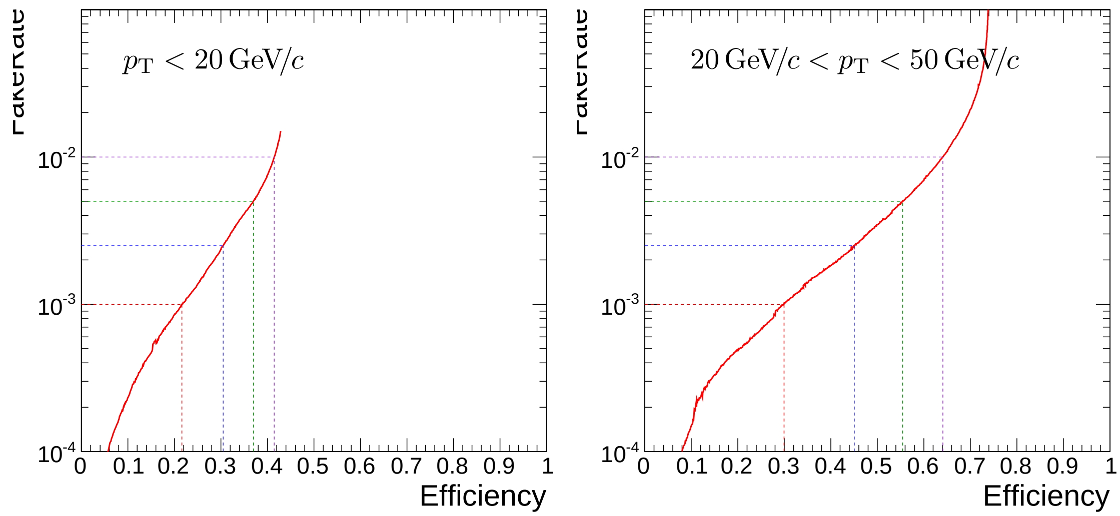


Figure 3.10: Tau Neural Classifier performance curves for tau–candidates with $p_T < 20\,\mathrm{GeV}/c$ (left) and $20\,\mathrm{GeV}/c < p_T < 50\,\mathrm{GeV}/c$ (right). The vertical axis represents the expected fake–rate of QCD jets and the horizontal axis the expected signal efficiency for hadronic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

$$E' = (1-x)^2 \cdot \rho^{\tau} + x^2 \cdot \rho^{QCD} \tag{3.3}$$ {?}

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

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

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

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

$$x = \frac{\rho^{\tau}}{\rho^{\tau} + \rho^{QCD}} \tag{3.4}$$ eq:probFracToX

$$\rho^{\tau} = x(\rho^{\tau} + \rho^{QCD})$$

$$\frac{\rho^{QCD}}{\rho^{\tau}} = \frac{1}{x} - 1 \tag{3.5}$$ eq:rawTransform

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

In the case of multiple neural networks, one can derive a formula that maps the output x_j of the neural network corresponding to decay mode j according to the "prior probabilities" $p_j^{\tau}(p_j^{QCD})$ for true τ lepton hadronic decays (quark and gluon jets) to pass the preselection criteria and be reconstructed with decay mode j. By substituting $\rho^s \rightarrow \rho^s p_j^s$ for $s \in \{\tau, QCD\}$ in equation 3.4, the output x_j can be related to $p_j^{\tau}(p_j^{QCD})$ by

$$x'_j = \frac{\rho^{\tau} \cdot p_j^{\tau}}{\rho^{\tau} \cdot p_j^{\tau} + \rho^{QCD} \cdot p_j^{QCD}} = \frac{p_j^{\tau}}{p_j^{\tau} + \frac{\rho^{QCD}}{\rho^{\tau}} \cdot p_j^{QCD}} \tag{3.6}$$ eq:probFracToXV

Substituting equation 3.5 into equation 3.6 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.

$$x'_j = \frac{p_j^{\tau}}{p_j^{\tau} + \left(\frac{1}{x_j} - 1\right) \cdot p_j^{QCD}} \tag{3.7}$$ eq:TransformCut

In this manner a single number (the "transform cut") given by Equation 3.7 can be used to specify any point on the performance curve. The training sample neural network output after the transformation has been applied is shown in figure 3.12. The performance curve for the cut on the transformed output is nearly identical to the optimal performance curve determined by the Monte Carlo sampling technique.

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

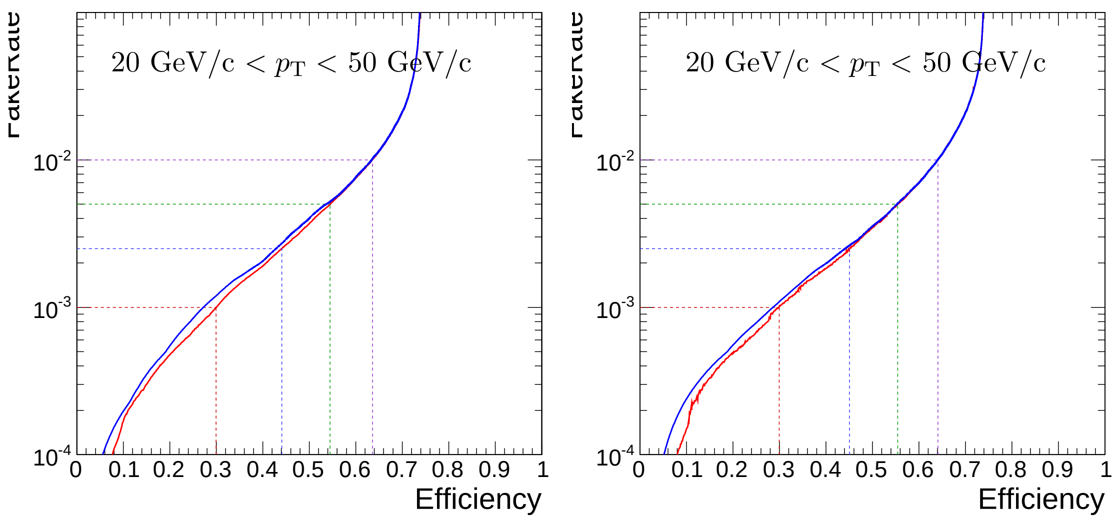


Figure 3.11: Tau Neural Classifier performance curves for tau–candidates with $20\,\mathrm{GeV}/c < p_T < 50\,\mathrm{GeV}/c$. The figure on the left compares the optimal performance curve determined by the Monte Carlo sampling method (red) to the performance curve obtained by scanning the “transform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

ncCurvesWithTransform⟩?

§3.5 Summary

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

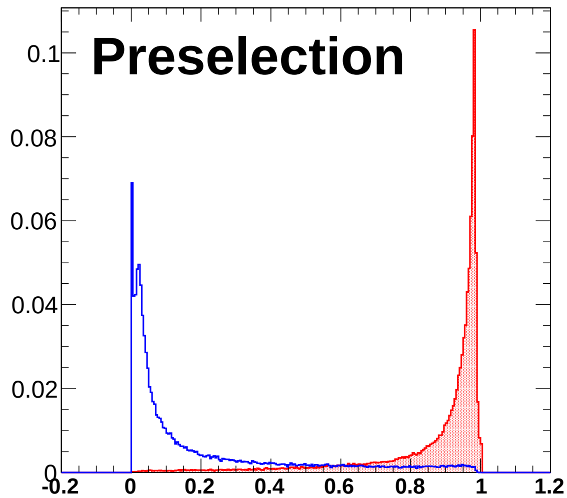


Figure 3.12: Transformed TaNC neural network output for tau–candidates with transverse momentum between 20 and 50 GeV/c that pass the pre–selection criteria. The neural network output for each tau–candidate has been transformation according to equation **??**. The decay mode probabilities ρ_i^{bkg},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

ig:transformedNNOutput⟩

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

The Tau Neural classifier tau identification algorithm significantly improves tau discrimination performance compared to isolation–based approaches [**?**] used in previous CMS analyses. Figure 3.13 compares the performance of the "shrinking cone" isolation tau–identification algorithm [**?**] to the performance of the TaNC for a scan of requirements on the transformed neural network output. The signal efficiency and QCD di–jet fake rate versus tau–candidate transverse momentum and pseudo–rapidity for the four benchmark points and the isolation based tau identification are show in figure 3.14. For tau–candidates with transverse momentum between 20 and 50 GeV/c, the TaNC operating cut can be chosen such that the two methods have identical signal efficiency; at this point the TaNC algorithm reduces the background fake rate by an additional factor of 3.9. This reduction in background will directly improve the significance of searches for new physics using tau leptons at CMS.

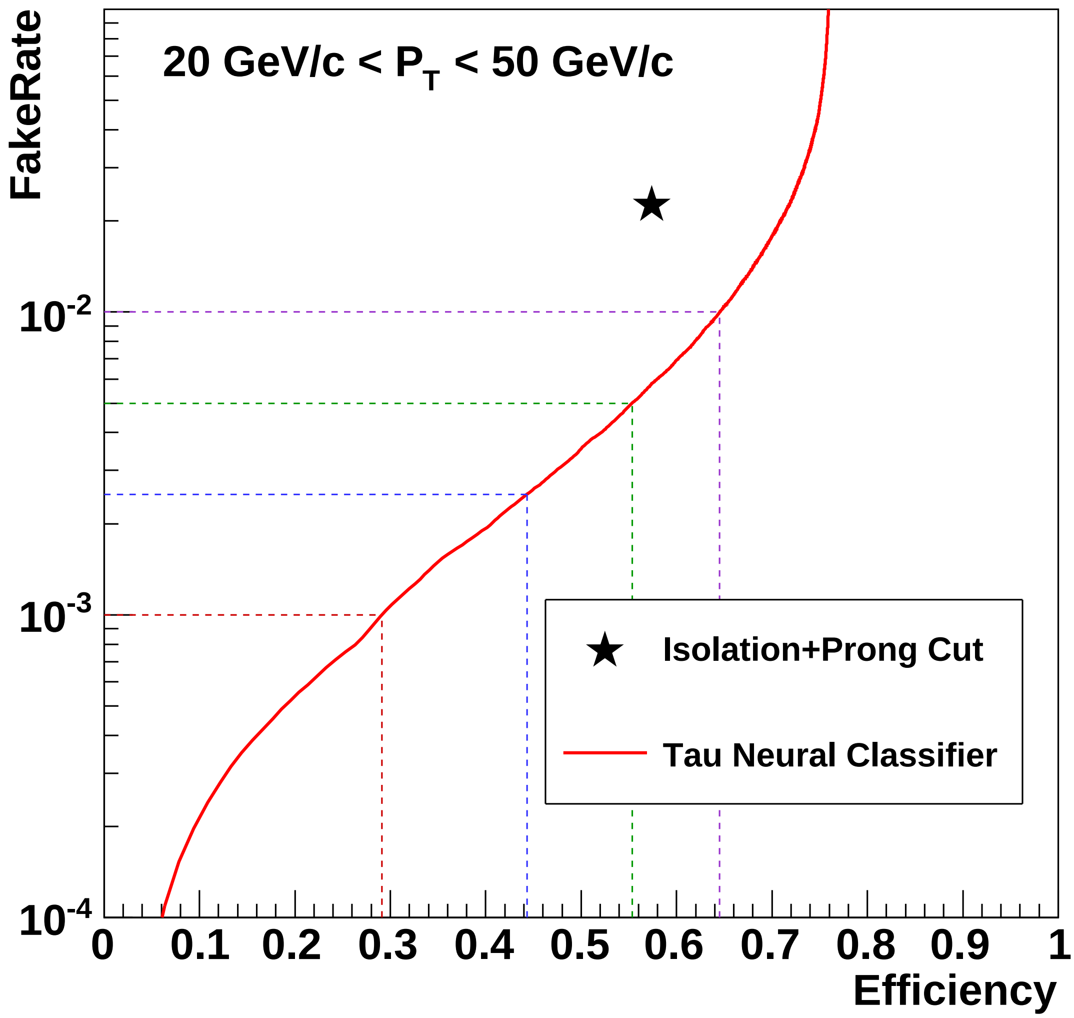


Figure 3.13: Performance curve (red) of the TaNC tau identification for various requirements on the output transformed according to equation ??. The horizontal axis is the efficiency for true taus with transverse momentum between 20 and 50 GeV/c to satisfy the tau identification requirements. The vertical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

⟨fig:finalPerfCurve⟩

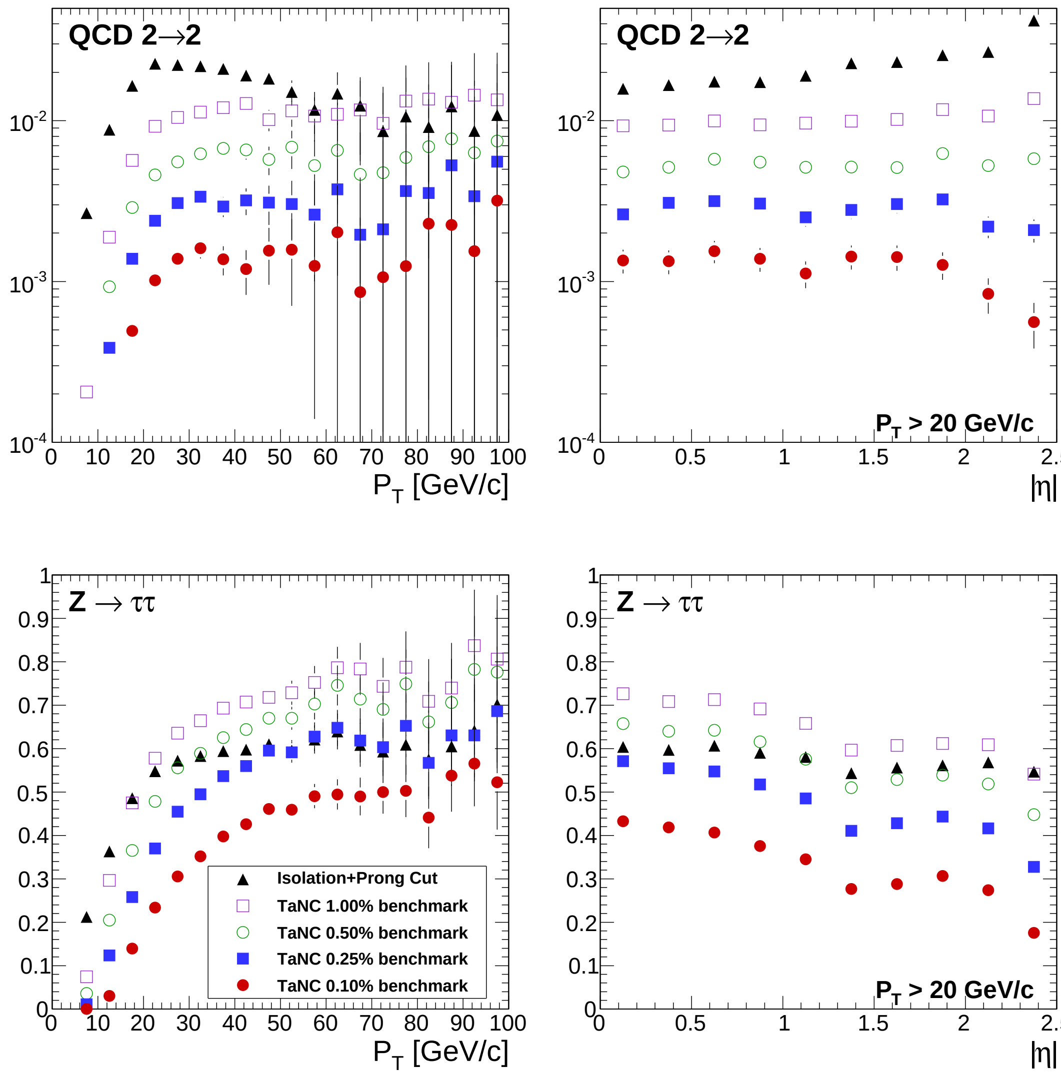


Figure 3.14: Comparison of the identification efficiency for hadronic tau decays from $Z \rightarrow \tau^+\tau^-$ decays (bottom row) and the misidentification rate for QCD di–jets (top row) versus tau–candidate transverse momentum (left) and pseudo-rapidity (right) for different tau identification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

§3.6 HPS+TaNC: A Hybrid Algorithm

⟨sec:TauId⟩ The techniques used in the TaNC have been hybridized with techniques used by the “Hadrons plus Strips” (HPS) algorithm. The combined algorithm is referred to “Hadrons plus Strips and Tau Neural Classifier” (HPS + TaNC) identification algorithm. The algorithm combines features of HPS [27] and 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 [28]. 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.

§3.6.1 Decay mode reconstruction

The decay mode reconstruction algorithm is seeded by particle–flow jets reconstructed by the anti–k_T algorithm [29]. In order to reconstruct the decay mode, the algorithm needs to merge photon candidates into candidate π^0 mesons. π^0 candidates are reconstructed by two algorithms which are executed concurrently. The “combinatoric” π^0 algorithm produces a π^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 π^0 candidate is determined according to the following categorical rankings:

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

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

m_{π^0} denotes the nominal neutral pion mass [16]. The choice of the invariant mass windows in the ECAL endcap and barrel regions is motivated by the resolution on the π^0 mass (illustrated in Figure 3.15) measured in particle–flow commissioning [30]. Multiple π^0 candidates in the same category are ranked in quality according to the difference of the reconstructed photon pair mass to the nominal π^0 mass. After the π^0 candidates are ranked, the highest ranked candidate is selected for the final collection. The photon constituents of the highest ranked candidate are removed from remaining π^0 candidates not yet selected for the final collection in order to prevent photons from entering more than one π^0 candidate. The rank of remaining π^0 candidates is reevaluated and the π^0 candidate with the next highest rank is selected for the output collection. The process is repeated until no more π^0 candidates are remaining.

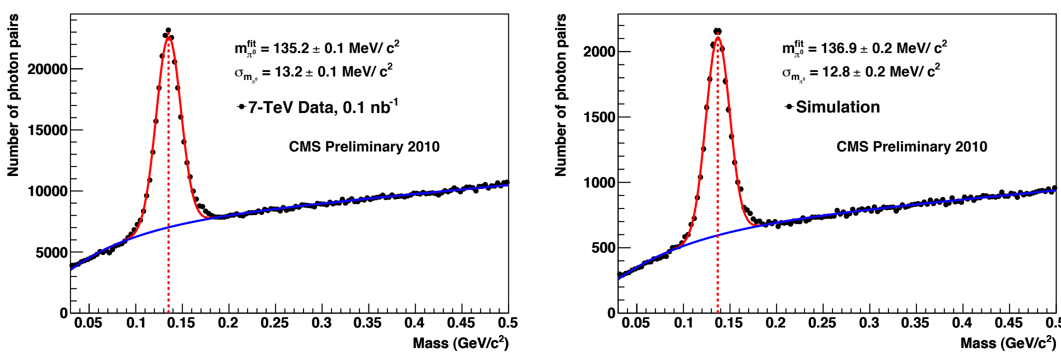


Figure 3.15: Invariant mass distribution of photon pairs reconstructed by the particle–flow in 2010 CMS minimum bias events (left), and predicted by the simulation (right). A clear resonant pick corresponding to the π_0 meson is visible above the combinatoric background. Reference: [30]

⟨fig:PFPiZeroRes⟩

Once the final collection of π^0 candidates is determined, tau reconstruction in the HPS + TaNC algorithm proceeds by building tau candidates from reconstructed π^0 candidates and charged hadrons reconstructed by the particle–flow algorithm. A combinatoric approach is again employed for the tau candidate building. A tau candidate hypothesis is built for every combination of jet constituents (π^0 candidates plus charged hadrons) which has a multiplicity consistent with a hadronic tau decay. The tau candidates are ranked analogous to the ranking utilized for the π^0 reconstruction, but with the following categorical rankings:

- In each decay mode category, the tau candidate with the highest neural network output is selected.
- The tau candidate has unit charge.
- The tau candidate passes the "lead pion" criteria, requiring that there is a photon or charged pion candidate with $p_{\mathrm{T}} > 5$ GeV/c.
- The tau candidate passes the HPS invariant mass and collimation[1] requirements.

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

§3.6.2 Hadronic tau discrimination

The final level of discrimination is performed by an ensemble of neural networks, with each neural network corresponding to a specific decay mode, analogously to the method used original TaNC algorithm (Section 3.4.2). The inputs of each neural network are different and correspond to the observables (invariant mass, Dalitz masses) available for its associated decay mode. The neural networks are trained on samples simulated $Z \to \tau^+\tau^-$ events ("signal") and QCD di–jet events selected in the 7 TeV data collected by CMS in 2010 ("background"). All of the tau hypothesis from a given jet reconstructed in data are used for training. The $Z \to \tau^+\tau^-$ signal sample is generated by PYTHIA [31] which has been interfaced TAUOLA [32] for the purpose of generating the tau decays and simulated passed through the "full" GEANT [33] based simulation of the CMS detector. Only tau candidates which have been reconstructed in a decay mode matching the true decay mode of the tau

[1] 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 ΔR between any two signal candidates to be less than $2.8/E_{\mathrm{T}}$, where E_{T} is the total transverse energy of the signal candidates. A full description is available in [27].

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 [34]. 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–only working point.

§3.7 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 [23]. 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 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.

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

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

⟨tab:nn˙var˙table⟩

Chapter 4

Mass Reconstruction: The Secondary Vertex Fit

?⟨ch:svfit⟩?

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

FiXme: *fuck*

§4.1 Existing mass reconstruction algorithms

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

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

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

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

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.

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

eq:CollinearAppr

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

§4.2 The Secondary Vertex fit

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

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

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

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

§4.3 Parametrization of tau decays

⟨sec:svParameterization⟩ The decay of a tau of visible four–momentum p_{vis} measured in the CMS detector ("laboratory") frame can be parametrized by three variables. The invisible (neutrino) momentum is fully determined by these parameters.

[1]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.

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

Given θ, $\overline{\phi}$ 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:

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

eq:restFrameMon

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

The opening angle $\overline{\theta}$ 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:

$$\begin{aligned} p_\perp^{vis} &= \overline{p}_\perp^{vis} \\ \Rightarrow \sin\overline{\theta} &= \frac{p^{vis}\sin\theta}{\overline{p}^{vis}} \end{aligned} \tag{4.4}$$

eq:labFrameOpe

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

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

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

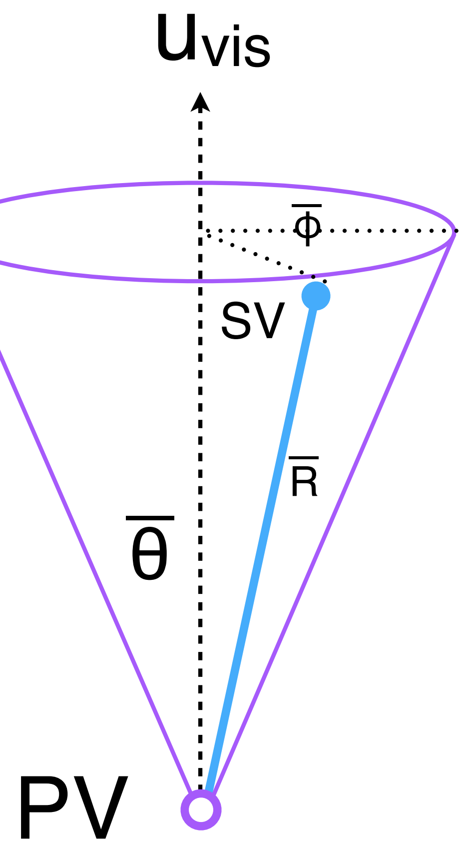


Figure 4.1: Illustration of the coordinate system used by the SVfit to describe the decays of tau leptons.

ig:svFitDecayParDiagram⟩

vector is not fully determined by θ and $m_{\nu\nu}$, but is constrained to lie on the surface of a cone of opening angle $\overline{\theta}$ (given by equation 4.4), the axis of which is given by the visible momentum vector. The tau lepton four–vector is fully determined by the addition of the third parameter $\overline{\phi}$, which describes the azimuthal angle of the tau lepton with respect to the visible momentum vector. The spatial coordinate system used is illustrated in Figure 4.1.

§4.4 Likelihood for tau decay

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

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

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

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

eq:pdfKineLepto

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

§4.4.1 Likelihood for reconstructed missing transverse momentum

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

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

$$\begin{aligned}
\sigma_{\|} &= \max\left(7.54\left(1 - 0.00542 \cdot q_T\right), 5.\right) \\
\mu_{\|} &= -0.96 \\
\sigma_{\perp} &= \max\left(6.85\left(1 - 0.00547 \cdot q_T\right), 5.\right) \\
\mu_{\perp} &= 0.0,
\end{aligned}$$

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

[2] 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 [?].

§4.4.2 Likelihood for tau lepton transverse momentum balance

⟨sec:ptBalance⟩ The tau lepton transverse momentum balance likelihood term represents the probability $p\left(p_{\mathrm{T}}^{\tau}|M_{\tau\tau}\right)$ for a tau to have a certain p_{T}, given 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 $\to \tau^{+}\tau^{-}$ 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 $p^{*}\left(p_{\mathrm{T}}|M\right)$, is derived by assuming an isotropic two–body decay, that is

$$\mathrm{d}p^{*} \propto \sin\theta \mathrm{d}\theta.$$

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

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

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 of scale parameter θ and shape parameter k, in order to account for tails in the p_{T} distribution of the tau lepton pair. The complete functional form is thus given by

$$p\left(p_{\mathrm{T}}|M\right) \propto \int_{0}^{\frac{M}{2}} p^{*}\left(p'_{\mathrm{T}}|M\right) e^{-\frac{\left(p_{\mathrm{T}}-p'_{\mathrm{T}}\right)^{2}}{2s^{2}}}\,\mathrm{d}p'_{\mathrm{T}} + a\Gamma\left(p_{\mathrm{T}},k,\theta\right). \tag{4.7}$$

eq:ptBalanceLike

Numerical values of the parameters s, θ and k are determined by fitting function 4.7 to the tau lepton p_{T} distribution in simulated Higgs $\to \tau^{+}\tau^{-}$ 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

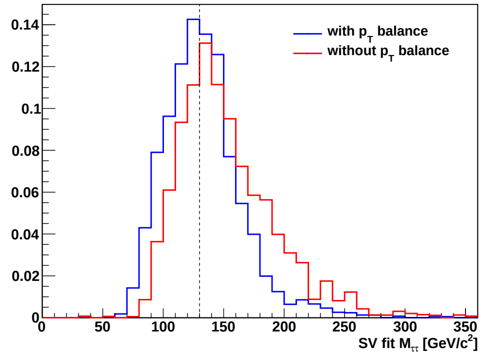


Figure 4.2: Distribution of di–tau invariant mass reconstructed by the SVfit algorithm in simulated Higgs events with $m_{A^0} = 130$ GeV/c^2. The SVfit algorithm is run in two configurations, with (blue) and without (red) the p_T–balance likelihood term included in the fit.

eImprovedMassResolution⟩

mass, we obtain the following numerical values for the parameters:

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

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

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

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

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

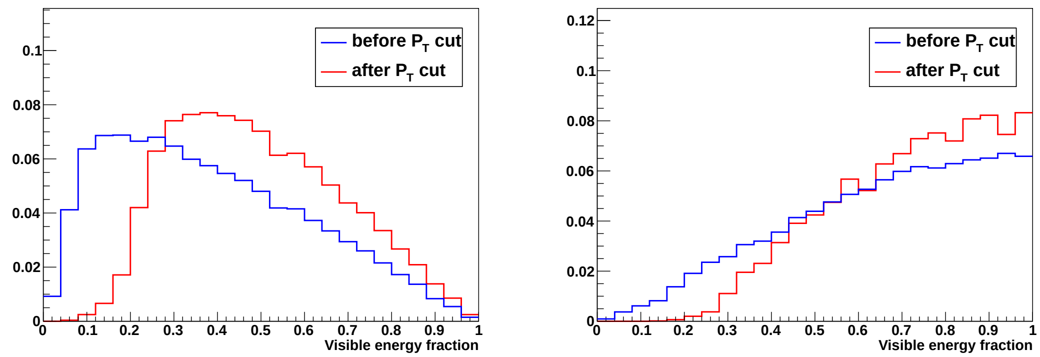


Figure 4.3: Normalized distributions of the fraction of total tau decay energy carried by the muon (left) and hadronic constituents (right) in simulated Higgs events with $m_{A^0} = 130$ GeV/c^2. The distribution is shown before (blue) and after (red) the requirement on the p_{T} of the visible decay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

⟨fig:ptBalancePtVisCuts⟩

§4.4.3 Secondary vertex information

The parametrization of the tau decay kinematics described in section 4.3 can be extended to describe the production and decay of the tau. As the flight direction of the tau is already fully determined by the parameters θ, $\overline{\phi}$ 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 $c\tau = 87$ μm is large enough to allow the displacement of the tau decay vertex from the primary event vertex to be resolved by the CMS tracking detector. The resolution provided by the CMS tracking detector is utilized to improve the resolution on the $\tau^+\tau^-$ 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:

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

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,

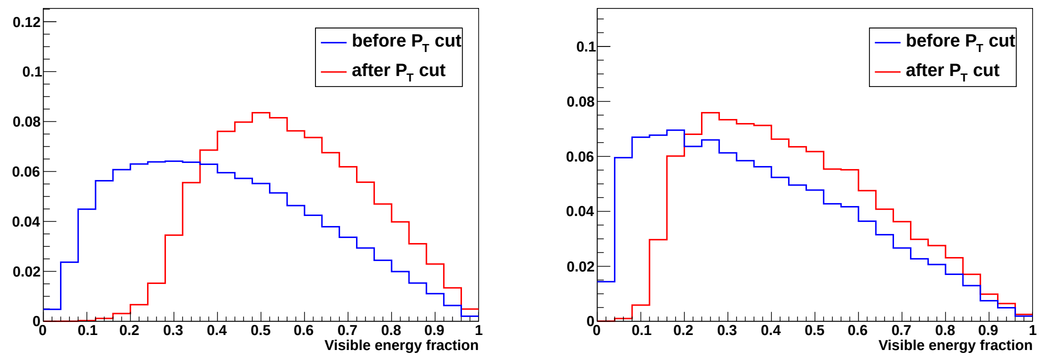


Figure 4.4: Normalized distributions of the fraction of total tau decay energy carried by the muon in simulated $Z \to \tau^+\tau^-$ (left) and Higgs events with $m_{A^0} = 200$ GeV/c^2 (right). The distribution is shown before (blue) and after (red) the requirement that the p_{T} of the muon be greater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

PtVisCutsCompareMasses⟩

however, because of systematic effects arising from tracker (mis–)alignment which are not yet fully understood.

§4.5 Performance

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

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

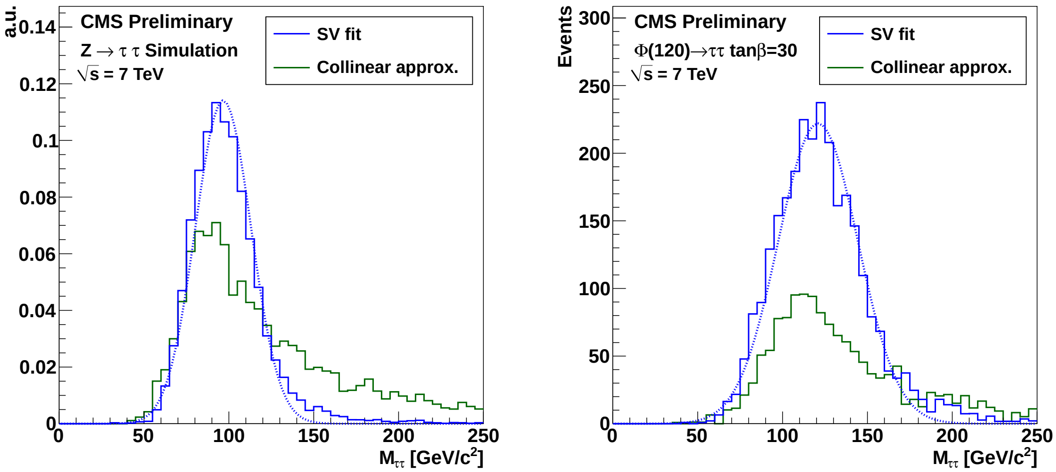


Figure 4.5: Comparison of the reconstructed tau pair mass spectrum in $Z \rightarrow \tau^+\tau^-$ (left) and MSSM $H(120) \rightarrow \tau^+\tau^-$ (right) events after the selections described in chapter 5. The mass spectrum reconstructed by the Secondary Vertex fit is shown in blue, the result of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

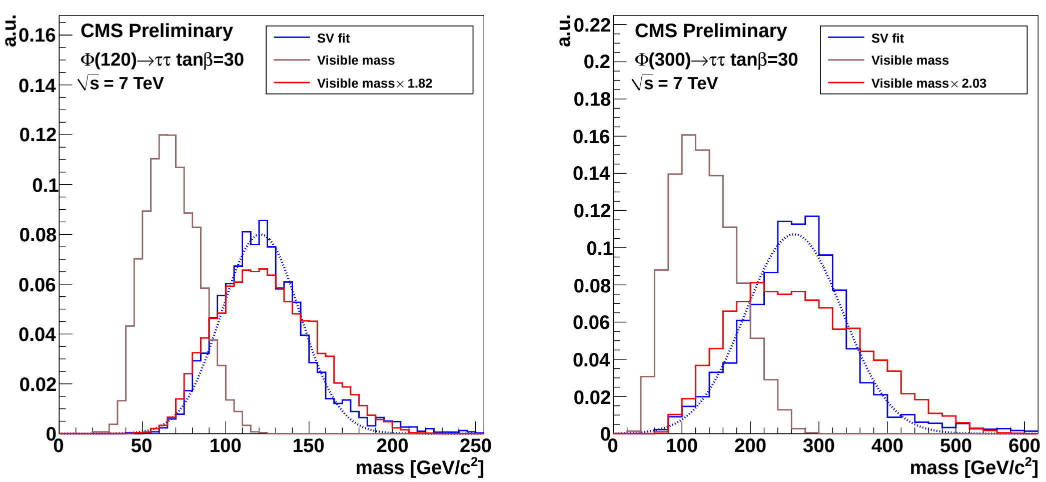


Figure 4.6: Comparison of the invariant mass of the muon and $\tau_{\rm jet}$ (the "visible mass") with the full $\tau^+\tau^-$ mass reconstructed by the SVfit. The spectrum is shown for two simulated MSSM Higgs samples, with $m_{A^0} = 120\,\mathrm{GeV}/c^2$ (left), and $m_{A^0} = 200\,\mathrm{GeV}/c^2$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

⟨fig:SVversusVis⟩

Chapter 5

Analysis Selections

⟨ch:selections⟩

§5.1 Particle Identification

§5.1.1 Muons

⟨sec:MuonId⟩ Muon candidates are required to be reconstructed as global and as tracker muons, meaning that a full track is reconstructed in the muon system and is well matched to a track in the silicon strip and pixel trackers. Additionally, they are required to pass the "Vector Boson Task Force" (VBTF) muon identification criteria developed for the $Z \rightarrow \mu^+\mu^-$ cross–section measurement [35]:

- ≥ 1 Pixel hits
- ≥ 10 hits in silicon Pixel + Strip detectors
- ≥ 1 hit(s) in muon system
- ≥ 2 matched segments
- $\chi^2/DoF < 10$ for global track fit
- transverse impact parameter of "inner" track $d_{\mathrm{IP}} < 2$ mm with respect to beam–spot

In order to reduce background contributions from muons originating from heavy quark decays in QCD multi–jet events, muons are required to be isolated. Isolation is computed as the p_{T} sum of charged and neutral hadrons plus photons reconstructed by the CMS particle–flow algorithm [24] within a cone of size $\Delta R_{iso} = 0.4$ around the muon direction. The innermost region of size $\Delta R_{veto} = 0.08$ (0.05) is excluded from the computation of the isolation p_{T} sum with respect to neutral hadrons (photons), in order to avoid energy deposits in the electromagnetic and hadronic calorimeters which are due to the muon to enter the sum. In order to reduce pile–up effects, particles entering the isolation p_{T} sum are required to have transverse momenta $p_{\mathrm{T}} > 1.0$ GeV/c. Charged particles are additionally required to originate from the same vertex as the muon. The muons are required to be isolated with respect to charged hadrons of $p_{\mathrm{T}} > 1.0$ GeV/c and photons of $p_{\mathrm{T}} > 1.5$ GeV/c as reconstructed by the particle–flow algorithm [24] in a cone of size $\Delta R = 0.4$ around the direction of the muon.

§5.1.2 Hadronic Taus

Hadronic decays of taus are identified by the HPS + TaNC hybrid algorithm described in Section 3.6. The "medium" working point is used, corresponding to an expected QCD fake–rate of about 1%. $Z \to \mu^+\mu^-$ background contributions are largely due to muons which failed to get reconstructed as global muons (thus failing the muon identification requirement) and are misidentified as tau–jet candidates. These muons are typically isolated and have a large chance to pass the hadronic tau ID discriminators. To reject these events, hadronic taus are additionally required to pass an anti–muon veto described in Section 3.7.

§5.1.3 Missing Transverse Energy

The missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$, in the event is reconstructed based on the vectorial momentum sum of particle candidates reconstructed by the particle–flow algorithm [24, ?]. In the ideal case, the $E_{\mathrm{T}}^{\mathrm{miss}}$ corresponds to the vector sum of the transverse components of all neutrinos in the event. The $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution in simulated $Z \to \mu^+\mu^-$ events is found to be smaller (better) than in the data. The reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ in the simulated events is "smeared" by a correction factor such that the data and simulation are in agreement. The "Z–recoil" $E_{\mathrm{T}}^{\mathrm{miss}}$ correction procedure is described in Section 7.2.

§5.2 Event Selections

The selections applied to the analysis are designed to reject large fractions of the background while maintaining a high efficiency for identifying signal (Higgs) events. The backgrounds

Background	Cross Section (pb)
QCD Heavy Flavor	84679[2]
$W \to \mu\nu + \text{jets}$	10435
$Z \to \mu\mu + \text{jets}$	1666
$t\bar{t} + \text{jets}$	158

Table 5.1: The different backgrounds to the analysis presented in this thesis that include misidentified hadronic taus.

⟨tab:FakeBackgrounds⟩

can be divided into two classifications: "fake" backgrounds, in which there is at least one misidentified hadronic tau decay, and the irreducible $Z \to \tau^+\tau^-$ background, which cannot[1] be distinguished from the potential presence of a Higgs boson of the same mass. Strategies for dealing with the irreducible Z background will be discussed in the Chapter 9. The different fake backgrounds, their cross section, and the basic removal strategies are outlined in Table 5.1.

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

The muon and tau–jet candidate are required to be of opposite charge, as the Higgs is neutral and charge is conserved. The muon is required to be pass the identification criteria described in Section 5.1.1. The tau-jet candidate is required to pass the "medium" TaNC tau identification discriminator.

FiXme: *is it really medium??*

Additional event selection criteria are applied to reduce contributions of specific background processes. In order to reject this background, a dedicated discriminator against muons is applied [23]. Remaining muon background is suppressed by rejecting events which

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

Requirement	
Trigger	HLT_Mu9 for MC *cf.* table 5.3 for Data
Vertex	reconstructed with beam–spot constraint: $-24 < z_{vtx} < +24$ cm, $\|\rho\| < 2$ cm, $N_{\mathrm{DOF}} > 4$
Muon	reconstructed as global Muon with: $p_{\mathrm{T}} > 15$ GeV, $\|\eta\| < 2.1$, VBTF Muon ID passed, isolated within $\Delta R = 0.4$ cone with respect to charged hadrons of $p_{\mathrm{T}} > 1.0$ GeV and neutral electromagnetic objects of $p_{\mathrm{T}} > 1.5$ GeV
Tau–jet Candidate	reconstructed by HPS + TaNC combined Tau ID algorithm TaNC "medium" Tau ID discriminator and discriminators against electrons and muons passed, calorimeter muon rejection passed
Muon + Tau–jet	charge(Muon) + charge(Tau–jet) = 0, ΔR(Muon, Tau–jet)> 0.5
Kinematics	M_T(Muon-MET)< 40 GeV $P_\zeta - 1.5 \cdot P_\zeta^{vis} > -20$ GeV

Table 5.2: Event selection criteria applied in the muon + tau–jet channel.

HtoMuTauEventSelection⟩

have a track of $p_{\mathrm{T}} > 15$ GeV and for which the sum of energy deposits in ECAL plus HCAL is below $0.25 \cdot P$ within a cylinder of of radius 15 cm(ECAL) and 25 cm(HCAL), respectively. The $t\bar{t}$ and $W+$jets backgrounds are suppressed by cuts on the transverse mass of the $\mu - -E_{\mathrm{T}}^{\mathrm{miss}}$ system and the $P_\zeta$variable. Contamination from $Z \to \tau^+\tau^-$ events in which the reconstructed tau–jet candidate is due to a $\tau \to e\nu\nu$ decay is reduced by applying a dedicated tau ID discriminator against electrons.

The complete set of event selection criteria applied in the muon + tau–jet channel are summarized in table 5.2.

The events are triggered by a combination of muon and muon + tau–jet "cross–channel" triggers. For the muon triggers, paths with lowest P_T thresholds are used as long as the path

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 5.3: Muon and muon + tau–jet "cross–channel" trigger paths utilized to trigger events in the muon + tau–jet channel in different data–taking periods.
⟨tab:AHtoMuTauTriggers⟩

remained unprescaled (see table 5.3). The muon + tau–jet "cross–channel" trigger paths increase the trigger efficiency for events containing muons of transverse momenta close to the $p_{\mathrm{T}}^{\mu} > 15$ GeV/c cut threshold. The trigger efficiency is measured in data via the tag–and–probe technique. Details of the muon trigger efficiency measurement are given in section 7.1 of the appendix. Monte Carlo simulated events are required to pass the HLT_Mu9 trigger path. Weights are applied to simulated events to account for the difference between the simulated HLT_Mu9 efficiency and the combined efficiency of the set HLT_Mu9, HLT_IsoMu9, HLT_Mu11, HLT_IsoMu13, HLT_Mu15, HLT_IsoMu9_PFTau15 and HLT_Mu11_PFTau15 used to trigger the data.

Chapter 6

Data–Driven Background Estimation

?⟨ch:backgrounds⟩?

§6.1 Introduction

§6.2 The Fake–rate Method

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

The results described in this note were obtained from Monte Carlo simulations of the $Z \rightarrow \tau^+\tau^-$ signal and different background processes for a centre–of–mass energy of $\sqrt{s} = 7$ TeV. Analysis of the $\sqrt{s} = 7$

TeV data recorded in 2009 [**?**] indicate that the probabilities for quark and gluon jets to fake the signatures of hadronic tau decays are well modeled by the Monte Carlo simulation. Once data–samples of sufficient event statistics are available at collision energies of $\sqrt{s} = 7$ TeV, fake–rates at the higher centre–of–mass energy will be measured in data and the values obtained from data will henceforth be used for the purpose of estimating background contributions via the fake–rate technique.

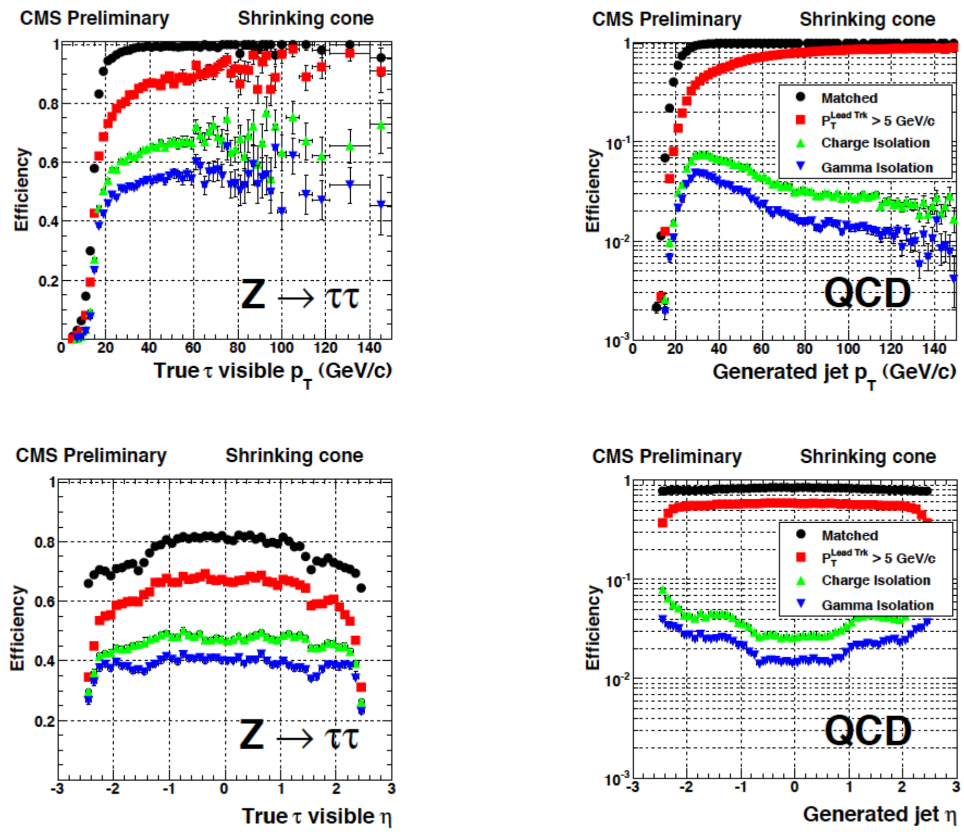


Figure 6.1: Cumulative efficiencies (left) and fake–rates (right) of successively applied tau identification cuts of the "shrinking signal cone" particle–flow based tau identification algorithm described in [?] as function of p_{T}^{jet} (top) and η^{jet} (bottom) of tau–jet candidates. The efficiencies/fake–rates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

EfficienciesAndFakeRates⟩

§6.2.1 Parameterization of fake–rates

Efficiencies and fake–rates of the tau identification algorithm based on requiring no tracks of $p_{\mathrm{T}} > 1$ GeV/c and ECAL energy deposits of $p_{\mathrm{T}} > 1.5\ GeV$ reconstructed within an "isolation cone" of size $dR_{iso} = 0.5$ and outside of a "shrinking signal cone" of size $dR_{sig} = 5.0/E_T$ as it is used in the $Z \to \tau^+\tau^- \to \mu + \tau$-jet analysis are displayed in figure 6.1. In order to account for the visible p_{T} and η dependence, we parametrize the fake–rates in bins of transverse momentum and pseudo–rapidity. As we will show in section **??**, the parametrization of the fake–rates by p_{T} and η makes it possible to not only estimate the total number of background events contributing to physics analyses, but to model the distributions of kinematic observables with a precision that is sufficient to extract information on the background shape.

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

$$R_{jet} = \sqrt{E\left(\eta^2\right) + E\left(\phi^2\right)} \tag{6.1}$$ {?}

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

§6.2.2 Measurement of fake–rates

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

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

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

Different sets of fake–rates are determined for the highest p_{T} and for the second highest p_{T} jet in QCD di–jet events, for jets in a QCD event sample enriched by the contribution of heavy quarks and gluons by requiring the presence of a muon reconstructed in the final state, and for jets in "electroweak" events selected by requiring a W boson in the final state.

§6.2.3 The Fake–rate method

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

Application of the fake–rate technique consists of two stages. The first stage consists of loosening the tau identification cuts and discriminators and applying only the preselection requirements defined by the denominator of equation 6.2, in order to obtain an event sample dominated by contributions of background processes, which are expected to increase by the inverse of the (average) fake–rate, typically by a factor $\mathcal{O}\,(100)$. In the second stage, weights are applied to all events in the background dominated control sample, according to the probabilities $P_{fr}\left(p_{\mathrm{T}}^{jet}, \eta_{jet}, R_{jet}\right)$ for jets to fake the signature of a hadronic tau decay. After application of the weights, an estimate for the total number of background events passing the tau identification cuts and discriminators and thus contributing to the final analysis sample is obtained.

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

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

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

“Simple” weight method

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

$$w_{jet}^{simple}\left(p_{\mathrm{T}}^{jet}, \eta_{jet}, R_{jet}\right) := P_{fr}\left(p_{\mathrm{T}}^{jet}, \eta_{jet}, R_{jet}\right) \tag{6.3}$$ eqBgEstFakeRat

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

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

$$W_{event}^{simple} := \Sigma w_{jet}^{simple} \tag{6.4}$$ eqBgEstFakeRat

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

$$w_{comp-part}^{simple}\left(p_{\mathrm{T}}^{jet}, \eta_{jet}, R_{jet}\right) := w_{jet}^{simple}\left(p_{\mathrm{T}}^{jet}, \eta_{jet}, R_{jet}\right) \tag{6.5}$$?eqBgEstFakeRat

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

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

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

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

"CDF–type" weights

eRate˙frCDFtypeWeights⟩

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

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

To be specific, the "CDF–type" weights assigned to tau–jet candidates are computed as (for a derivation of the formula, see section **??** of the appendix):

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

eqBgEstFakeRat

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

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

By definition of the tau identification efficiency $\varepsilon := \varepsilon\left(p_{\mathrm{T}}^{jet},\eta_{jet},R_{jet}\right)$ and fake–rate $f := f\left(p_{\mathrm{T}}^{jet},\eta_{jet},R_{jet}\right)$:

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

?eqBgEstFakeRat

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:

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

eqBgEstFakeRat

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:

$$w_{passed} = -\frac{1-\varepsilon}{\varepsilon} \cdot w_{failed}, \tag{6.9}$$ eqBgEstFakeRat

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

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

and

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

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

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

$$\begin{aligned} W^{CDF}_{event} &:= \Sigma w^{CDF}_{jet} \\ w^{CDF}_{comp-part}\left(p_{\mathrm{T}}^{jet}, \eta_{jet}, R_{jet}\right) &:= w^{CDF}_{jet}\left(p_{\mathrm{T}}^{jet}, \eta_{jet}, R_{jet}\right), \end{aligned} \tag{6.11}$$ eqBgEstFakeRat

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

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

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

Results obtained by the "CDF–type" fake–rate method are summarized in table **??**, in which the total number of background events estimated by equation 6.11 is compared to the true background contributions. The "best" estimate of the background contribution obtained from the "CDF–type" method is again taken as the arithmetic average of the estimates obtained by applying the fake–rate probabilities for the highest and second highest p_{T} jet in QCD di–jet events, jets in a muon enriched QCD sample and jets in W+jets events. Systematic uncertainties are taken from the difference between the computed average value and the minimum/maximum estimate. We obtain a value of $\mathcal{O}\,(15\text{–}20\%)$ for the systematic uncertainty of the "CDF–type" method, slightly higher than the systematic uncertainty obtained for the "simple" method. The small increase of systematic uncertainties is in agreement with our expectation for fluctuations of the jet–weights in case weights of negative and positive sign are used.

§6.2.4 k-Nearest Neighbor fake–rate calculation

For the fake–rate method to give correct results, care must be taken that the measured fake–rate is well defined in all of the regions of phase space where it will be used. Previous implementations [36] of the fake–rate method in CMS implemented the fake–rate parameterization by binning the numerator (tau ID passed) and denominator (tau ID passed and failed) distributions in the three dimensions of the parameterizations. This method has the disadvantage that the determination of the optimal binning is extremely difficult to determine. Furthermore, any bins with no entries in the denominator distribution caused the fake–rate to be undefined 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 (kNN). A kNN 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

$$p_{sig} = \frac{n_{sig}}{n_{sig} + n_{bkg}}, \tag{6.12}$$

eq:KNNEquation

where n_{sig},n_{bkg} are the observed number of signal and background events, respectively. By construction, $k = n_{sig} + n_{bkg}$. The principle of operation is illustrated in Figure 6.2

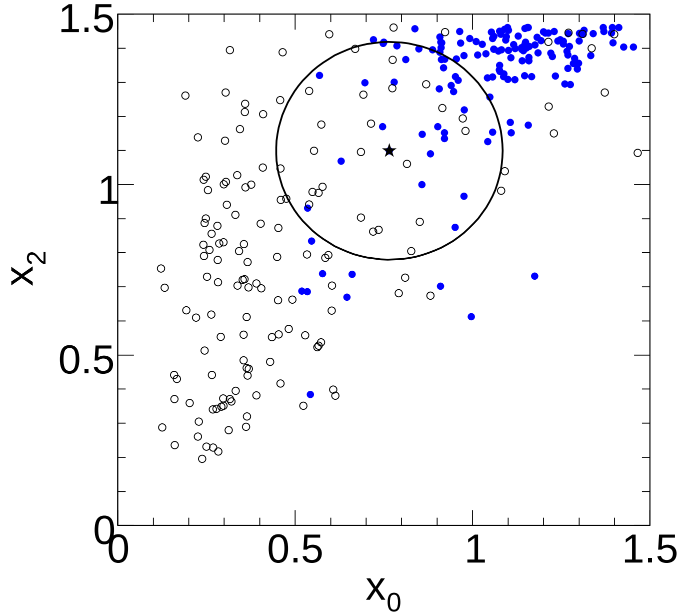


Figure 6.2: Example of the operation of a kNN classifier. The closest $k = 50$ neighbors (those inside the circle) to a test point (indicated by the star marker) are selected. The probability that the star marker is a signal event is given the number of signal neighbors (blue markers) in the circle divided by k. Image credit: [25]

⟨fig:KNN⟩

The classification feature of a kNN can be trivially adapted to parameterize the fake–rate such that it is defined everywhere. Examining the form of Equation 6.12, if one replaces n_{sig} with n_{passed} and n_{bkg} with n_{failed}, the equation is equivalent to the tau–fake rate. We thus train the kNN with tau–candidates which pass the tau–ID as signal events and those which fail as background events. The resulting classifier is a function which returns the expected fake–rate for any point in the space of the parameterization. The choice of k must be optimized. When k is low, the small number of neighbors causes large counting fluctuations in the fake rate. If k is too large, the kNN effectively averages over a large area of the space of the variables[1]. For the training statistics available in the 2010 data, $k = 20$ is found to be the optimal choice.

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

§6.2.5 Results of Background Estimation

An independent estimate of the background contributions is obtained by applying the fake–rate method described in [36]. Fake–rates in QCD multi–jet events (light quark enriched sample), QCD events containing muons (heavy quark and gluon enriched sample) and W + jet events are measured in data [**?**, **?**] and applied to events which pass all the event selection criteria listed in table 5.2 except for the requirement for tau–jet candidates to pass the "medium" tight TaNC discriminator and have unit charge.

No assumption is made on the composition of $Z \rightarrow \mu^+\mu^-$, W + jets, $t\bar{t} + jets$ and QCD backgrounds contributing to the event sample selected by the analysis. Differences between fake–rates obtained for QCD multi–jet, QCD muon enriched and W + jets background events are attributed as systematic uncertainties of the fake–rate method. Per jet and per event weights have been computed by the "simple" and "CDF-type" weights as described in [36] and the results are found to be compatible within statistical and systematic uncertainties. In the following, we present results for "CDF-type" weights. The "CDF-type" weights have the advantage that the background estimate obtained does not change, whether there is MSSM Higgs $\rightarrow \tau^+\tau^-$ signal present in the data or not.

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

The results of applying the fake–rate method to the mu + tau channel are summarized in table 6.1. The background prediction has been corrected for the expected contribution of $XX.X$ events from $Z \rightarrow \mu^+\mu^-$ background events in which the reconstructed tau–jet is

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

Table 6.1: Estimate for background contributions in mu + tau channel obtained by weighting events passing all selection criteria listed in table 5.2 except for the requirement for tau–jet candidates to pass the "medium" tight TaNC discriminator and have unit charge by fake–rates measured in QCD multi–jet, QCD muon enriched and W + jets data samples.

IuTauFakeRateResultsOS⟩

due to a misidentified muon. The obtained estimate is in good agreement with the Monte Carlo expectation.

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

The fake–rate method does not only allow to estimate the total number of background events, but allows to model the distributions of background processes as well. The capability to model distributions is illustrated in figure 6.3, which shows good agreement between the distributions observed in the like-sign data sample and the predictions obtained by the fake–rate method for the distributions of muon plus tau–jet visible mass and of the "full" invariant mass reconstructed by the SVfit algorithm.

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

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

MuTauFakeRateResultsSS⟩

Figure 6.3: 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⟩

Chapter 7

Monte Carlo Corrections

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

The corrections measured and used in this analysis can be divided into two categories, efficiency and scale corrections. Identification efficiency corrections scale the expected yield (due to a given identification selection) up or down. Scale corrections systematically scale the energy of a particle (or $E_{\mathrm{T}}^{\mathrm{miss}}$) up or down. In this analysis we apply efficiency corrections for the High Level Trigger, all stages of muon identification, and the hadronic tau identification. We apply a momentum scale correction to the muon and tau legs, and to the resolution of the $E_{\mathrm{T}}^{\mathrm{miss}}$. Finally, events are simulated with overlapping "pile–up"[1] events. The simulated events are weighted such that the number of pile–up events in the simulation matches that observed in the data.

§7.1 Muon Identification Efficiency

⟨sec:ZmumuTagAndProbe⟩ The identification efficiencies associated with the muon are measured in $Z \to \mu^+\mu^-$ events using the "tag and probe" technique [35]. $Z \to \mu^+\mu^-$ events are selected from the Muon

[1] 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.

7 TeV CMS 2010 datasets[2] by requiring that the events pass the "loose" Vector Boson Task Force (VBTF) event selections [35]. In the selected events, we define the "tag" muons as those that have transverse momentum greater than 15 GeV/c and pass the VBTF muon selection. The tag muons are further required to pass the "combined relative isolation" described in the VBTF paper. We finally require that the tag muon be matched to an HLT object corresponding to the run–dependent requirements listed in table 5.3. The trigger match requirement ensures that the event would be recorded independently of the probe muon. After the tag and probe muon pairs have been collected, we compare the muon identification performance in the probe collection in events selected in data to the performance in simulated $Z \to \mu^+\mu^-$ events. The selection of events and tag muon in the simulated sample is the same as the data sample, with the notable exception that the only HLT requirement applied in MC is that the tag muon is matched to an HLT_Mu9 object. Any difference in efficiency between the HLT_Mu9 path and the paths used to select the data (in the tag–probe measurement and in the analysis) will be considered implicitly in the correction faction.

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

- Efficiency of global probe muons to satisfy VBTF muon identification selections.
- Efficiency of global probe muons passing the VBTF muon identification selection to satisfy the isolation criteria described in Section 5.1.1.
- Efficiency of probe muons passing the offline analysis selection defined in Chapter 5 to pass the HLT selection.

In each case, the invariant mass spectrum of the tag–probe pair is fitted with a Crystal Ball function for the signal ($Z \to \mu^+\mu^-$) events and an exponential for the background. The fit is done for two cases; where the probe fails the selection and the where it passes. The method is illustrated in Figure **??**. The signal yield N is extracted from each fit and

[2] /Mu/Run2010A-Sep17ReReco2̌/RECO and /Mu/Run2010B-PromptReco-v2/RECO

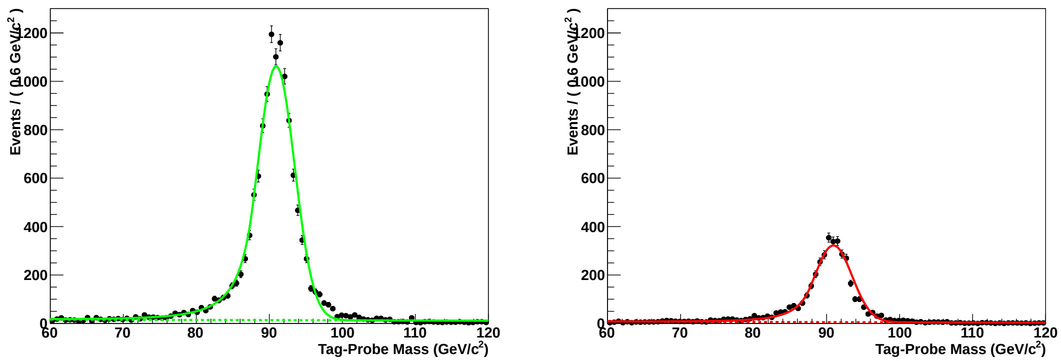


Figure 7.1: The tag–probe dimuon invariant mass spectrum in events in which the probe muon passed (left) and failed (right) the muon isolation requirement. The solid gives the result of the simultaneous fit of the signal (real $Z \rightarrow \mu^+\mu^-$ events) and background. The fitted background contribution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

g:MuonIsoCorrVersusPt⟩?

the efficiency is computed as $N_{pass}/(N_{pass} + N_{fail}$. Each efficiency is measured in both the data 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.

The correction for the trigger efficiency needs to take into account the differences in the HLT selections applied during different operating periods (see table 5.3). To determine the overall correction factor, we measure the trigger efficiency in data for each of the operating periods and compare it to the simulated efficiency of the HLT_Mu9 selection. The overall efficiency in data is taken as the average of the three periods, weighted by integrated luminosity.

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

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.

uonTagAndProbeResults⟩

the single muon triggers used in the run–range 147196 – 148058 (period B). The "cross–trigger" contribution is estimated as the difference between the efficiency in period B and the single–muon period C efficiency multiplied by a correction factor of $0.9 \pm 10\%$ to account for the τ leg efficiency. In the case that the measured single–muon period C efficiency is larger than the period B efficiency (due to statistical fluctuations and improvements in the trigger system), the period B efficiency is increased by 2%.

§7.2 Missing Transverse Energy Correction

⟨sec:ZRecoilCorr⟩

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

The $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution depends on the "recoil" of the Z boson. The reason for this effect is that for events where the Z is produced nearly at rest, the associated recoil products have very small transverse momentum and are produced at very high pseudorapidity. The $E_{\mathrm{T}}^{\mathrm{miss}}$ is corrected using a procedure called a "Z–recoil" correction, as described in [37]. The resolution of the $E_{\mathrm{T}}^{\mathrm{miss}}$ is measured in $Z \rightarrow \mu^+\mu^-$ events in simulation and data.

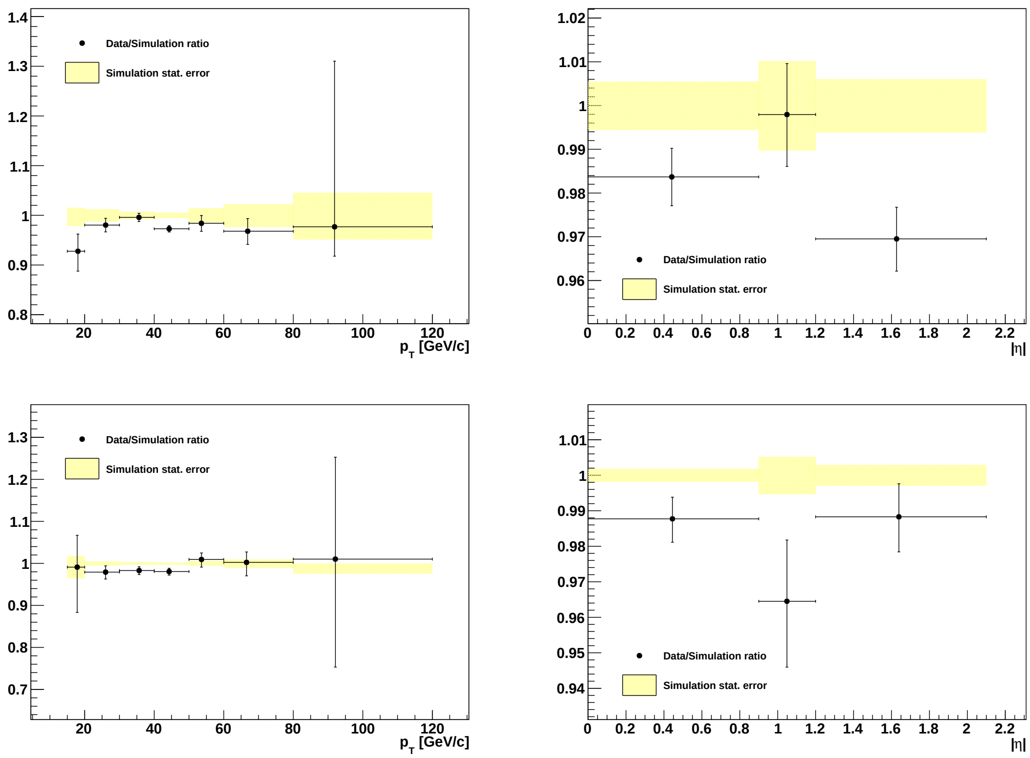


Figure 7.2: Ratio of muon isolation efficiency measured in data compared to simulated $Z \to \mu^+\mu^-$ events.

ig:MuonIsoCorrVersusPt⟩?

The difference in the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution in both samples is parameterized by the magnitude of the transverse momenta of the particles recoiling against the Z.[3] The reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ in the simulated $Z \to \tau^+\tau^-, Z \to \mu^+\mu^-$, and $W + \text{jets}$ samples is "smeared" by a random amount such that the resolution is the correct

Z–recoil corrections are determined as described in [37] and applied to simulated $Z \to \tau^+\tau^-$, $Z \to \mu^+\mu^-$ and $W + \text{jets}$ events, in order to correct for residual differences in $E_{\mathrm{T}}^{\mathrm{miss}}$ response and resolution between data and Monte Carlo simulation [38]. The corrections are obtained by an unbinned maximum likelihood fit (in data and simulation) of the transverse recoil vector $\vec{u}_T = -\left(\vec{q}_T + E_{\mathrm{T}}^{\mathrm{miss}}\right)$ as function of the transverse momentum $\vec{q}_T$ of the Z–boson in directions parallel and perpendicular to the Z–boson transverse momentum vector. The effect of the Z–recoil correction is illustrated in figure 7.3.

[3]The "recoil" particles are defined as all those not identified as Z decay produces. This definition is equivalent to the total decay product transverse momentum q_T added reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$.

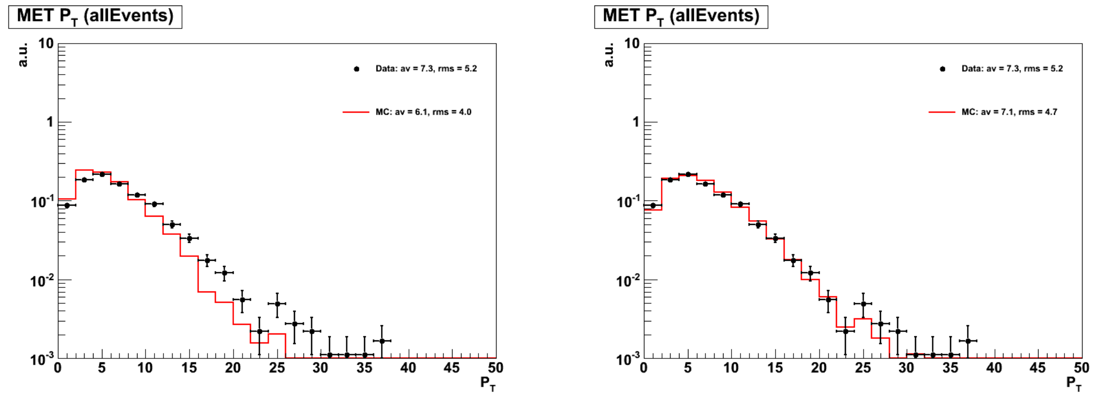


Figure 7.3: Missing transverse energy reconstructed in $Z \rightarrow \mu^+\mu^-$ events selected in data compared to $Z \rightarrow \mu^+\mu^-$ events in Monte Carlo simulation before (left) and after (right) the Z–recoil corrections to the E_T^{miss} resolution are applied.

⟨fig:ZrecoilCorrection⟩

§7.3 Pile–up Event Weighting

The average number of pile–up interactions in the event can effect almost all aspects of the analysis. In general, increasing pile–up lowers particle identification efficiencies and lowers E_T^{miss} resolution. It is therefore important that the distribution of pile–up events in the simulation matches the distribution found in the data. Differences in the number of pile–up interactions between the data (averaged over the analyzed run–range) and pile-up Monte Carlo samples produced for "BX156[4]" pile–up conditions are corrected for by reweighting Monte Carlo simulated events according to the number of reconstructed event vertices, in order to match the distribution measured in a $W \rightarrow \mu\nu$ dataset triggered by the HLT_Mu15 High Level Trigger path. Vertices considered for this purpose are required to pass $-24 < z_{vtx} < +24$ cm, $|\rho| < 2$ cm, nDoF > 4. In addition, the total transverse momenta of all tracks fitted to the vertex is required to exceed 10 GeV/c, assuming that "softer" vertices have little or no effect on the "hard" event to pass event selection criteria. The average vertex multiplicity distribution measured in data is compared to Monte Carlo simulation with "BX156" pile–up conditions in Figure 7.4. Both distributions are similar, resulting in Monte Carlo reweighting factors close to unity.

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

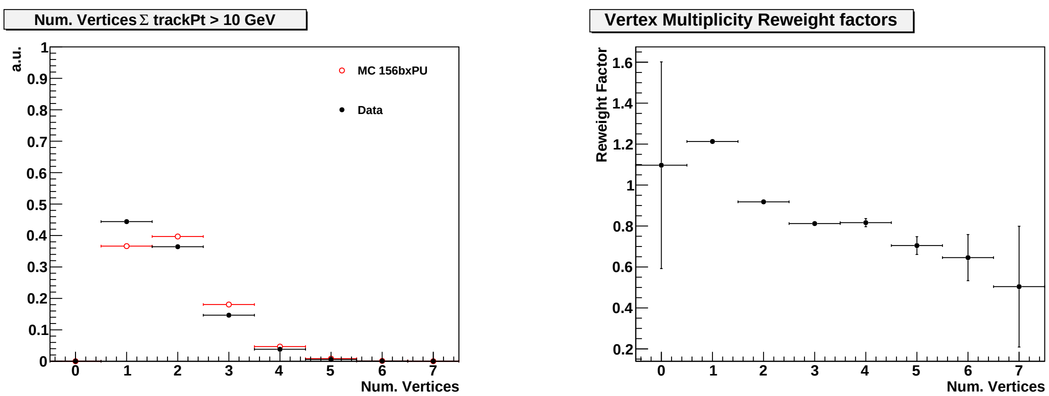


Figure 7.4: Vertex multiplicity distribution measured in the analyzed data–taking period compared to Monte Carlo simulation with "BX156" pile–up conditions (left) and resulting Monte Carlo reweighting factors (right).

⟨fig:pileUpReweighting⟩

Chapter 8

Systematics

Proper determination of systematic uncertainties is one of the most challenging and important components in performing a correct measurement. 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

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

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

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

It is interesting to examine Equation 8.2 in two scenarios. In the limit that N_{obs} is large compared to N_{bkg}, the effect of the error on the background estimate δN_{bkg} does not affect the final result. In 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 $\rightarrow \tau^+\tau^-$ signal extraction presented in this thesis are classified in three categories: normalization uncertainties on the signal, normalization uncertainties on background contributions, and shape uncertainties. Normalization uncertainties on the signal are due to lepton reconstruction, identification, isolation and trigger efficiencies. These terms are equivalent to the efficiency ϵ and acceptance terms $\mathcal{A}$ of Equation 8.2 and affect the expected yield of MSSM Higgs $\rightarrow \tau^+\tau^-$ signal and of $Z \rightarrow \tau^+\tau^-$ background events. They do not affect the *shapes* of visible and "full" invariant 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 are due to uncertainties on the momentum/energy scale of identified electrons, muons, tau and other jets in the event. The measurement of missing transverse energy represents another source of uncertainty specific to the "full" mass reconstruction by the SVfit algorithm. 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.

§8.1 Signal normalization uncertainties

The signal normalization uncertainties are due to imperfect knowledge of how improperly modeled effects in the detector affect our "acceptance" model, or the probability that a given signal event will pass one of the selections (detailed in Chapter 5). The general procedure to quantify these uncertainties is to measure the effect in some control region in both the data and Monte Carlo. The ratio of data to Monte Carlo then gives a correction factor which is applied to the simulation. An uncertainty on the measurement of the effect in control region (in data, simulation, or both) is then taken as the systematic uncertainties. The signal normalization uncertainties affecting this analysis on muon trigger, reconstruction, identification and isolation efficiencies are taken from the tag and probe analysis of $Z \rightarrow \mu^+\mu^-$ events presented in section 7.1. The uncertainty on the tau reconstruction and identification efficiency is taken to be 23%. The dependency of the Higgs signal extraction on the tau identification efficiency has been studied, the result being that uncertainties on the tau identification efficiency affect the limit on cross–section times branching ratio for MSSM Higgs $\rightarrow \tau^+\tau^-$ production by a few percent only. An uncertainty of 11% is attributed to the luminosity measurement.

§8.2 Background normalization uncertainties

Uncertainties on the normalization of background processes are obtained from the study of background enriched control regions presented in section ??. The main non–$Z \rightarrow \tau^+\tau^-$ background to the analysis is due to QCD multi–jet and W + jets events. These backgrounds are produced copiously enough for the backgrounds to be studied in control regions dominated by a single background process with a purity exceeding 90% and an event statistics exceeding the expected contribution of that background to the analysis by more than one order of magnitude. Both backgrounds are found to be well modeled by the Monte Carlo simulation. An uncertainty of 10% is attributed to the contribution of QCD and W + jet backgrounds to the analysis. The cross–section for $t\bar{t} + jets$ production makes it difficult to select a high purity sample of $t\bar{t}+jet$ events of high event statistics. From the study of the 19 events selected in the $t\bar{t}+jets$ background enriched control sample we assume an uncertainty on the $t\bar{t}+jets$ background contribution in the analysis of 30%. The $Z \rightarrow \mu^+\mu^-$ background has been studied with large statistical precision in two separate control regions, dominated by events in which the reconstructed tau–jet candidate is either due to a misidentified quark or gluon jet or due to a misidentified muon. Good agreement between data and Monte Carlo simulation is found in both cases. Sizeable uncertainties on the $Z \rightarrow \mu^+\mu^-$ background contribution arise due to the extrapolation from the background enriched control regions to the data sample considered in the analysis, however: the contribution of $Z \rightarrow \mu^+\mu^-$ background events to the analysis is due to events in which one of the two muons produced in the Z decay either escapes detection or fakes the signature of a hadronic tau decay. Both cases may be difficult to model precisely in the Monte Carlo simulation. The non–observation of a Z mass peak in the mu + tau visible mass distribution studied with the fake–rate method on the other hand sets a limit on possible contributions from $Z \rightarrow \mu^+\mu^-$ background events. Conservatively, we assume an uncertainty of 100% on both types of $Z \rightarrow \mu^+\mu^-$ background contributions.

§8.3 Shape uncertainties

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

The uncertainty on the muon momentum scale is taken from the analysis known di–muon resonances [**?**] and found to have a very small effect only. The uncertainty on the jet energy scale is determined from an analysis of the P_T balance between photons and jets in γ + jets events [**?**]. The jet energy scale uncertainties determined by the JetMET group are applied to tau–jets as well as other jets in the event. Work is ongoing in the Tau POG to determine energy scale correction factors and uncertainties specific to tau–jets [**?**].

The modelling of missing transverse energy in different types of background events has been studied in the background enriched control regions described in section **??**. No significant deviations between data and Monte Carlo simulation have been found (*cf.* control plots in the appendix). Uncertainties due to missing transverse energy are estimated by varying parameters of Z–recoil corrections within the uncertainties obtained when fitting the Z–recoil correction parameters in simulated $Z \rightarrow \mu^+\mu^-$ events versus $Z \rightarrow \mu^+\mu^-$ events selected in data.

§8.4 Theory uncertainties

The signal and background normalization as well as the shape uncertainties are all experimental uncertainties in nature. Additional theoretical uncertainties arise from imprecise knowledge of parton–distribution functions (PDFs) and of the exact dependency of signal cross–sections and branching ratios on $tan\beta$ and m_A.

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

[1] decreasing with m_A
[2] number quoted for $gg \to A/H$ and $b\bar{b} \to A/H$ sample as a whole;
in the subsample of events with b–tagged jets the effect of the JES uncertainty is 4%
[3] with small dependence on m_A

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

The uncertainties on the signal acceptance due to PDF uncertainties are estimated using tools developed by the EWK group [?]. The acceptance is computed with respect to MSSM Higgs $\to \tau^+\tau^-$ decays that have electrons of $P_T^e > 15$ GeV and $|\eta_e| < 2.1$, muons of $P_T^\mu > 15$ GeV and $|\eta_\mu| < 2.1$, jets produced in hadronic tau decays with visible $P_T^{vis} > 20$ GeV and $|\eta_{vis}| < 2.3$ on generator level, depending on the analysis channel considered. Acceptance values are computed for the central value and 44 eigenvectors of the CTEQ66 PDF set [?]. The systematic uncertainty on the signal acceptance is computed following the PDF4LHC recommendations [?, ?].

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

Chapter 9

Results

⟨ch:results⟩

Chapter 10

Conclusions

?⟨ch:conclusions⟩**?**

Bibliography

Goldstone:1961eq [1] J. Goldstone, "Field Theories with Superconductor Solutions", *Nuovo Cim.* **19** (1961) 154–164. doi:10.1007/BF02812722.

PhysRev.127.965 [2] 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 [3] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons", *Phys. Rev. Lett.* **13** (Aug, 1964) 321–323. doi:10.1103/PhysRevLett.13.321.

PhysRevLett.13.508 [4] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons", *Phys. Rev. Lett.* **13** (Oct, 1964) 508–509. doi:10.1103/PhysRevLett.13.508.

PhysRevLett.13.585 [5] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global Conservation Laws and Massless Particles", *Phys. Rev. Lett.* **13** (Nov, 1964) 585–587. doi:10.1103/PhysRevLett.13.585.

Glashow:1961tr [6] S. Glashow, "Partial Symmetries of Weak Interactions", *Nucl.Phys.* **22** (1961) 579–588. doi:10.1016/0029-5582(61)90469-2.

Weinberg:1967tq [7] S. Weinberg, "A Model of Leptons", *Phys.Rev.Lett.* **19** (1967) 1264–1266. doi:10.1103/PhysRevLett.19.1264.

Salam:1968rm [8] A. Salam, "Weak and Electromagnetic Interactions",. Originally printed in *Svartholm: Elementary Particle Theory, Proceedings Of The Nobel Symposium Held 1968 At Lerum, Sweden*, Stockholm 1968, 367-377.

Griffiths:IntroParticle [9] D. Griffiths, "Introduction to Elementary Particles". Wiley-VCH, 2004.

Morii:SMandBSM [10] S. M. T. Morii, C.S. Lim, "The Physics of the Standard Model and Beyond". World Scientific, 2004.

UA1WDiscovery [11] UA1 Collaboration, "Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540\,\mathrm{GeV}$", *Phys. Lett.* **B122** (1983) 103–116.

UA2WDiscovery [12] UA2 Collaboration, "Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN $\bar{p}p$ collider", *Phys. Lett.* **B122** (1983) 476–485. doi:10.1016/0370-2693(83)91605-2.

UA1ZDiscovery [13] 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. doi:10.1016/0370-2693(83)90188-0.

UA2ZDiscovery [14] UA2 Collaboration, "Evidence for $Z^0 \rightarrow e^+e^-$ at the CERN $\bar{p}p$ collider", *Phys. Lett.* **B129** (1983) 130–140. doi:10.1016/0370-2693(83)90744-X.

Martin:1997ium [15] S. P. Martin, "A Supersymmetry Primer", *arXiv* **hep-ph** (sep, 1997). 128 pages. Version 5 (December 2008) contains a change in convention that flips the signs of sigma and sigmabar matrices. It also contains a total of about 2 pages of updates, mostly on supersymmetry breaking issues. Errata and a version with larger type (12 pt, 142 pages) can be found at http://zippy.physics.niu.edu/primer.html.

PDG [16] Particle Data Group Collaboration, "Review of particle physics", *J. Phys.* **G37** (2010) 075021. doi:10.1088/0954-3899/37/7A/075021.

MSTWXSectionPlots [17]

LHCHiggsXSecGroup [18] LHC Higgs Cross Section Working Group Collaboration, "Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables", arXiv:1101.0593.

MHMaxBenchmark [19] M. Carena, S. Heinemeyer, C. Wagner et al., "MSSM Higgs boson searches at the Tevatron and the LHC: Impact of different benchmark scenarios", *The European Physical Journal C - Particles and Fields* **45** (2006) 797–814. 10.1140/epjc/s2005-02470-y.

CMSExperiment [20] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004.

CMS-PAS-TRK-10-004 [21] CMS Collaboration, "Measurement of Momentum Scale and Resolution using Low-mass Resonances and Cosmic-Ray Muons", *CMS PAS* **CMS-PAS-TRK-10-004** (2010).

CMS-PAS-TRK-10-005 [22] CMS Collaboration, "Tracking and Primary Vertex Results in First 7 TeV Collisions", *CMS PAS* **CMS-PAS-TRK-10-005** (2010).

CMS-PAS-PFT-08-001 [23] CMS Collaboration, "CMS Strategies for tau reconstruction and identification using particle-flow techniques", *CMS PAS* **CMS-PAS-PFT-08-001** (2008).

CMS-PAS-PFT-09-001 [24] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET", *CMS PAS* **CMS-PAS-PFT-09-001** (2009).

TMVA [25] A. Hoecker, P. Speckmayer, J. Stelzer et al., "TMVA - Toolkit for Multivariate Data Analysis", *arXiv* **physics.data-an** (mar, 2007). Published in: PoSACAT:040,2007 TMVA-v4 Users Guide: 135 pages, 19 figures, numerous code examples and references.

Kolmogorov [26] A. Kolmogorov, "On the representation of continuous functions of several variables by superposition of continuous functions of one variable and addition", *Doklady Akademiia Nauk SSSR* **114** (1957).

CMS˙AN˙2010-082 [27] 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 [28] 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 [29] 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 [30] CMS Collaboration, “Commissioning of the Particle–Flow reconstruction in Minimum–Bias and Jet Events from *pp* Collisions at 7 TeV”, *CMS PAS* **PFT-10-002** (2010).

pythia6.4 [31] S. M. T. Sjöstrand and P. Skands, “PYTHIA 6.4 Physics and Manual”, 2000.

tauola [32] S. Jadach, Z. Was, R. Decker et al., “The Tau Decay Library Tauola: Version 2.4”, *Comput. Phys. Commun.* **76** (1993) 361.

geant [33] 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.

CMS˙AN˙2010-099 [34] 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).

CMS-PAS-EWK-10-002 [35] 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˙AN˙2010-074 [36] J. Conway, E. Friis, and C. Veelken, “Estimation of Background contributions to Tau analyses via Fake–Rate technique”, *CMS Note* **2010/074** (2010).

CMS˙AN˙2010-332 [37] G. Bauer et al., “Modeling of $W \rightarrow \ell\nu$ MET with Boson Recoil”, *CMS Note* **2010/332** (2010).

CMS˙AN˙2010-460 [38] G. Cerati et al., “Search for MSSM neutral Higgs $\rightarrow \tau^+\tau^-$ Production using the TaNC Tau id. algorithm”, *CMS Note* **2010/460** (2010).