Search for Single-Top-Quark Production in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV

by

Sarah Kristina Wolinski

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Physics)
in The University of Michigan
2002

Doctoral Committee:

Professor Myron K. Campbell, Chair
Professor Ratindranath Akhoury
Professor Dante E. Amidei
Professor Robert D. Fogel
Associate Professor John K. Riles

The heavens are telling the glory of God;
and the firmament proclaims his handiwork.
Day to day pours forth speech,
and night to night declares knowledge.
There is no speech, nor are there words;
their voice is not heard;
yet their voice goes out through all the earth,
and their words to the end of the world.

Psalm 19

Acknowledgments

It seems that the longer one's time as a graduate student, the longer the list of the people one must acknowledge for their collaboration and encouragement over the years. At the top of this list is my advisor, Myron Campbell. From the days of Run II trigger studies all the way to the thesis-analysis phase, Myron has been an excellent advisor and a valued friend. Thanks to Myron's judicious choice of project and his help and insight along the way, the hardware phase of my graduate career—so often viewed as an obstacle—was a highlight. It was a privilege to learn electronics from "the great one." Whenever Myron would come into my office, sigh, and sit down, we would always talk about something interesting—and I will miss that.

Two other crucial players in the analysis phase have been Dan Amidei and Pierre Savard. I am grateful for the hours Dan took out of his packed schedule to talk about single top and for his advocacy in dealings with CDF godparents. And all levels of this analysis—from high-level strategies to minutiae—benefited immensely from my collaboration with Pierre Savard, a postdoc at the University of Toronto and a name to watch in HEP. It was a great experience to work with someone this good and this interested.

Other people who made this single-top analysis possible include the Rockefeller-Tsukuba team of Luc Demortier, Toshi Kikuchi, and Shinhong Kim, whose initially-competing and finally-complementary analysis provided guidance and inspiration. Thanks are also due to the CDF godparents of this analysis—Lina Galtieri, Richard Hughes, and David Waters. Their input was essential to improving the analysis and to engineering a paper for publication that included both the Rockefeller-Tsukuba separate Wg and W^* limits and the Michigan-Toronto combined $Wg + W^*$ limit.

I also want to thank Rob Roser, the Top-Group convener when I embarked on the single-top analysis five years ago, for suggesting that I take on the project and helping make data sets available. Stepping back even further, I owe thanks to Melissa Franklin of Harvard, who gave me a summer job at Fermilab and encouraged me to go to physics graduate school in the first place. I still remember a conversation in a grubby trailer office in which she brushed aside my self-doubts with an "Of course you're smart enough. The question is whether you're *crazy* enough."

There are others with whom I've enjoyed working in my graduate career. In my hardware days it was fun to be a part of the feminine hegemony on the Level 2 calorimeter trigger along with Monica Tecchio and Colleen Murphy. Dave Gerdes was always someone to whom I could take my hard physics questions. And finally, I want to thank Ken Bloom for actually reading almost all of my thesis and giving me comments, even during the year when I was away from Michigan and feeling distant from the field. It has been good to have Ken as an editor and friend.

Other friends made in my eight years in Ann Arbor added much to my graduate-school experience. Wine tastings and mushroom hunts with Rebecca Wilkin and Cédric Picard, twilit "panel discussions" with David Schwartz at the single-bench picnic table overlooking Sylvan Ave., civilized dinners with Renée and Franko Bayer—these are among my fondest memories of the graduate school years. Renée and Franko are also among those who helped me to the realization that one did not need to be weird to have Christ as the central focus of one's life. Thanks also to Peter Payne for the Science, Philosophy, and Christianity discussion group which made me realize I did not need to give up my scientific mindset either. But the real thanks go to God who sought out this skeptic in graduate school when I was not even looking and convinced me to put my trust in Him.

And how can I begin to thank my family? This thesis would not have been possible without the lifelong support and encouragement of my parents, Steve and Kristina Truitt. My earliest memories of getting interested in math and science date back two decades to discussions of Gödel's proof in the kitchen with my dad. While my sister Eliza Truitt often

provided a refreshing orthogonality to the physics lifestyle, she could also understand the joy and pain that are common to all forms of writing. I have to thank my daughter Kristina for sharing her mom with a thesis for her first year. Most of all, I thank my husband Dave Wolinski. From the days of problem sets to the days of LaTeX-ing our theses, Dave has been my co-worker, my support, and my mentor. In the past year he has uncomplainingly combined full-time work with half-time dad (nights, weekends) so I could finish this thesis. Dave's sharp mind and persistence in seeking solutions have been fantastic, if sometimes humbling, resources to have just one office door—or now, just a phone call—away. Over the years, I have found that whatever the problem might be, it always helps to talk to Dave.

Table of Contents

List of Tables

Table

List of Figures

Figure

List of Appendices

Appendix

Chapter 1

Introduction

The first top quarks ever observed, at Fermilab's Tevatron $p\bar{p}$ collider in 1995 [1], were produced in particle-antiparticle ($t\bar{t}$) pairs via a strong-nuclear-force interaction. Despite this historic role, $t\bar{t}$ pair production can provide only indirect information on some of the top quark's most interesting properties. The top quark's width, lifetime, mixing behavior, and weak couplings are all accessible only through electroweak interactions. But since $t\bar{t}$ production is mediated by the strong interaction, its rate conveys nothing about the top's weak interactions; any such information must instead be gleaned from the top's subsequent weak decay. However, $t\bar{t}$ production is not the only way to make top quarks at the Tevatron. In *electroweak single-top-quark production*, the production—not just the decay—of the top proceeds through a weak interaction. Therefore the rate of "single-top" production provides direct information on the top's weak couplings. Phenomenological calculations predict a total cross section for Standard-Model single top production at the Tevatron of approximately half that for $t\bar{t}$ production.

We present the results of a search for single-top-quark production in 106 pb^{-1} of data from $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV collected with the Collider Detector at Fermilab (CDF) during the 1992-1995 Run 1 of the Tevatron. This paper is organized as follows. After a review of the Standard Model of elementary particles and interactions in Chapter 2, we explain in Chapter 3 why single-top-quark production is especially well-suited to studying the top quark's properties, both within and beyond the Standard Model. In Chapter 4 we

describe the experimental apparatus. Chapter 5 describes the selection criteria by which we identify single-top (signal) events in the data and Chapter 6 discusses the types of non-single-top (background) processes that also satisfy these criteria. Chapter 6 concludes with an estimate of the numbers of signal and background events expected to remain in the data sample after the selection criteria have been applied. These predictions indicate that the CDF Run 1 data sample has insufficient statistical sensitivity to permit observation of Standard-Model single-top production.

Nevertheless, it is worthwhile to look for single top in data and set an upper limit on its production rate. Such an upper limit might constrain those models of new physics which predict larger single-top rates than would be expected in the Standard Model. Chapter 7 outlines a maximum-likelihood method for estimating the single-top content of data. Chapter 8 shows how to use the likelihood function to extract an upper limit on the single-top content of data. Chapter 9 explains how to incorporate systematic uncertainty in computing the upper limit and presents an *a priori* estimate of the limit we expect to be able to set in CDF Run 1 data. Chapter 10 presents the results obtained from real data. The upper limit is six times the Standard Model prediction—not a strong constraint, but a world record to date. It is hoped that this pioneering analysis will serve as a guide for future analyses in larger data sets.

Chapter 2

The Standard Model

All known matter in the universe, including the exotic forms present only at the Big Bang and in particle colliders, appears to be made of just a handful of different types of fundamental particles. Four fundamental forces account for all known interactions between those particles. The Standard Model is a description of all the matter particles and three of the four forces. Gravity, by far the weakest interaction of the four, does not enter the Standard Model.

2.1 Overview

A summary of the elementary particles and forces provides an overview of the Standard Model. Particles can be grouped into two categories according to their value of intrinsic angular momentum, or spin. The properties of the spin-$\frac{1}{2}$ *fermions* are summarized in Table 2.1. Fermions are matter constituents, though only three of the twelve shown (e, u, and d) appear in the atoms which comprise ordinary matter. Properties of the integral-spin *bosons* are summarized in Table 2.2. Bosons are force mediators, with the exception of the Higgs boson, which plays a special role that will be discussed later.

To understand the explanatory power of the Standard Model beyond the level of a list of particles, we must look at the mathematical structure of the theory. In the sections to follow we will see that the interactions between the particles are derived in the sense

	Particle	Symbol	Mass (MeV/c^2)	Electric Charge (Q)	Color (C)	Weak Isospin (T_W^3)
	Electron neutrino	ν_e	< 0.000003	0	0	$+\frac{1}{2}$
First	Electron	e	0.511	-1	0	$-\frac{1}{2}$
Generation	Up quark	u	$1-5$	$+\frac{2}{3}$	r,g,b	$+\frac{1}{2}$
	Down quark	d	$3-9$	$-\frac{1}{3}$	r,g,b	$-\frac{1}{2}$
	Muon neutrino	ν_μ	< 0.19	0	0	$+\frac{1}{2}$
Second	Muon	μ	106	-1	0	$-\frac{1}{2}$
Generation	Charm quark	c	$1150-1350$	$+\frac{2}{3}$	r,g,b	$+\frac{1}{2}$
	Strange quark	s	$75-170$	$-\frac{1}{3}$	r,g,b	$-\frac{1}{2}$
	Tau neutrino	ν_τ	< 18.2	0	0	$+\frac{1}{2}$
Third	Tau	τ	1777	-1	0	$-\frac{1}{2}$
Generation	Top quark	t	$174,300$	$+\frac{2}{3}$	r,g,b	$+\frac{1}{2}$
	Bottom quark	b	$4000-4400$	$-\frac{1}{3}$	r,g,b	$-\frac{1}{2}$

Table 2.1: Names, masses [2], and quantum numbers of the fundamental fermions. Weak isospin values pertain to the left-handed chiral state (e_L, u_L, d_L, *etc.*) only; right-handed fermions have $T_W^3 = 0$. The confidence intervals for the neutrino mass limits are as follows: "not without ambiguity", for m_{ν_e}; 90%, for m_{ν_μ}; 95%, for m_{ν_τ}.

that they spring from symmetry principles in the theory. We will examine the types of calculations that can be done in the Standard Model. And finally we will consider that the Standard Model only explains some of the properties listed in these tables. Other aspects of these tables remain empirical fact whose explanations await a deeper theory.

2.2 The Fermion Lagrangian

The Lagrangian $L = T - V$, or kinetic energy minus potential energy, is a succint statement of the dynamics of a system. In this section we outline the derivation of the Standard-Model fermion Lagrangian, which describes the interactions between the fermions and the gauge bosons [3].

	Particle	Symbol	Spin	Mass (GeV/c^2)	Electric Charge (Q)	Color (C)
Electromagnetic force	Photon	γ	1	0	0	0
Charged weak force	W boson	$W^{\pm}$	1	80.4	± 1	0
Neutral weak force	Z boson	Z^0	1	91.2	0	0
Strong force	Gluon	g	1	0	0	octet
Spontaneous symmetry breaking	Higgs boson	H^0	0	> 95.3	0	0

Table 2.2: Names, masses [2], and quantum numbers of the fundamental bosons. All have been observed by experiment except the Higgs. The confidence limit on the Higgs mass is 95%.

2.2.1 Gauge invariance and the covariant derivative

The Standard Model is a gauge theory, which means the interactions between particles are *determined* by the invariance of the theory under certain operations known as local gauge transformations. These transformations are equivalent to rotations in abstract, internal spaces analogous to "spin space" in which rotating a spin-up ($\uparrow$) particle changes it into a spin-down ($\downarrow$) particle. We now show that the principle of local gauge invariance requires the existence of gauge fields mediated by gauge bosons.

The simplest example is provided by the set of gauge transformations that form the mathematical group $U(1)$. Under a $U(1)$ gauge transformation, the wavefunction Ψ is multiplied by a local phase factor $e^{-i\chi(x)}$:

$$\Psi \rightarrow \Psi' = e^{-i\chi(x)}\Psi$$

Since physical observables depend only on the absolute square of the wavefunction, $|\Psi|^2$, it is reasonable to expect that this transformation would leave the theory invariant. However, the free-particle Schrödinger equation ("the theory"), $-\frac{1}{2m}(-\nabla)^2\Psi = i\frac{\partial}{\partial t}\Psi$, is *not* invariant under the substitution $\Psi \rightarrow \Psi'$, since the derivatives do not cancel, unless we simultaneously introduce a field A^μ which transforms according to

$$A^\mu \rightarrow A^{\mu\prime} = A^\mu - \frac{1}{e}\partial^\mu\chi$$

and replace the ordinary derivatives $\partial^\mu = (\frac{\partial}{\partial t}; -\nabla)$ in the Schrödinger equation by covariant derivatives

$$D^\mu = \partial^\mu - ieA^\mu.$$

The utility of the covariant derivative is that any equation of the theory written in terms of the covariant derivative will automatically be gauge-invariant [1].

The field A^μ introduced to preserve gauge invariance is known as a gauge field. Like all fields, it has an expansion in terms of creation and annihiliation operators and therefore has a particle associated with it. Moreover, since the field is a 4-vector, the associated particle must also be vector, or spin-1. This particle is the gauge boson. Thus the existence of a field A^μ with a corresponding gauge boson has been shown to follow from the requirement of $U(1)$ gauge invariance. The factor e will be seen to characterize the strength of the interaction between the gauge boson and other particles. The value of coupling strength e is arbitrary in the sense that the theory does not require it to have any particular value.

There are two other groups of transformations under which the theory is also invariant, $SU(2)$ and $SU(3)$. $SU(n)$ gauge transformations are more complicated than $U(1)$, because now the phase factor is a matrix operator which performs a rotation in the internal $SU(n)$ space. An $SU(n)$ gauge transformation has the form

$$\Psi \rightarrow \Psi' = e^{i\vec{\epsilon}(x)\cdot\vec{F}}\Psi$$

where $\vec{F}$ are the $n^2 - 1$ generators of group $SU(n)$ and $\vec{\epsilon}$ specifies the $n^2 - 1$ parameters of the rotation. Just as in the $U(1)$ case, the equations of the theory are not invariant under the substitution $\Psi \rightarrow \Psi'$ unless we replace ordinary derivatives with covariant derivatives containing new gauge fields, except that now $n^2 - 1$ gauge fields $\vec{G}^\mu$ are required in order to form a scalar product with the generators $\vec{F}$ of $SU(n)$:

$$D^\mu = \partial^\mu - ig_n\vec{F}\cdot\vec{G}^\mu$$

[1] This follows because it can be shown that if Ψ transforms as a wavefunction, *i.e.* $\Psi \rightarrow \Psi' = e^{-i\chi(x)}\Psi$, then so does $D^\mu\Psi$: $(D^\mu)'\Psi' = e^{-i\chi}D^\mu\Psi$. Since $D^\mu\Psi$ behaves like a wavefunction under gauge transformations, so does $D_\mu D^\mu\Psi$, and so on. An equation written in terms of D^μ's will, after a gauge transformation, permit the cancelling of factors of $e^{i\chi}$ from every term, leaving the original equation invariant.

Each field G^μ corresponds to a spin-1 gauge boson. Successive $SU(n)$ transformations do not commute. The term for a theory with local non-Abelian phase invariance is a Yang-Mills gauge theory.

Thus to achieve $SU(2)$ invariance (*i.e.* invariance under $\Psi \rightarrow \Psi' = e^{i\vec{\epsilon}(x)\cdot\vec{\tau}/2}\Psi$, also known as "weak isospin" invariance), where $\vec{\tau}$ are the $2^2 - 1 =$ three generators of $SU(2)$ (the Pauli matrices), three new gauge fields $\vec{W}^\mu$ are required. The covariant derivative associated with the $SU(2)$ sector is

$$D^\mu = \partial^\mu - ig_2\frac{\tau^i}{2}W^{\mu i}, \quad i = 1, \ldots, 3$$

Similarly, to preserve $SU(3)$ invariance, $3^2 - 1 =$ eight new gauge fields are required in the covariant derivative:

$$D^\mu = \partial^\mu - ig_3\frac{\lambda^a}{2}G^{\mu a}, \quad a = 1, \ldots, 8$$

Simultaneous $U(1)$, $SU(2)$, and $SU(3)$ invariance is achieved by adding terms in D^μ. Thus the covariant derivative expressing the full $SU(3) \times SU(2) \times U(1)$ symmetry of the Standard Model may be written as [2]

$$D^\mu = \partial^\mu - ig_1\frac{Y}{2}B^\mu - ig_2\frac{\vec{\tau}}{2}\cdot\vec{W}^\mu - ig_3\frac{\vec{\lambda}}{2}\cdot\vec{G}^\mu$$

Again, coupling strengths g_1, g_2, and g_3 are not fixed by any principle in the theory, but instead must be measured by experiment. A Lagrangian written in terms of this covariant derivative will be gauge-invariant. Thus we see that requiring the theory to be invariant under $SU(3) \times SU(2) \times U(1)$ gauge transformations has required the existence of eight, three, and one spin-1 gauge bosons corresponding to the fields $G^{\mu a}$, $W^{\mu i}$, and B^μ, respectively. In various guises, these are the spin-1 gauge bosons appearing in Table 2.2. All have been observed by experiment. Why nature is invariant under these groups of gauge transformations and not others is not known.

[2]The $U(1)$ portion of D^μ is rewritten in this expression (as $ig_1\frac{Y}{2}B_\mu$ rather than ieA_μ) for parallelism with the other two terms, and because the term A_μ is conventionally used to denote the electromagnetic field, but as we will see, the gauge field B^μ of the $U(1)$ sector does not correspond exactly to the electromagnetic field. The expression for D^μ is a matrix equation using the convention that each term only operates in the relevant space—*e.g.* the W^μ term is a 2×2 matrix in $SU(2)$ and a singlet in $U(1)$ and $SU(3)$.

2.2.2 *Transformation properties of the fermions*

We assign the various quarks and leptons into multiplets according to their transformation properties in these internal spaces. Particles only transform nontrivially in those spaces in which they carry nontrivial quantum numbers. The transformation properties of a given state must be determined experimentally. In the following discussion we will restrict our attention to the first generation of fermions, (ν_e, e^-, u, d), but the same remarks apply under the substitutions (ν_e, e^-, u, d) $\longrightarrow$ (ν_μ, μ^-, c, s) or (ν_τ, τ^-, t, b).

The quantum number associated with the $U(1)$ sector of the Standard Model is weak hypercharge Y. It can be shown from the requirement of chiral-anomaly cancellation [4] that for a given fermion, $Y = 2(Q - T_W^3)$. From Table 2.1 we see that all fermions have $Y \neq 0$, so all fermions couple to the weak-hypercharge boson B^μ [3]. However, since the $U(1)$ term in the covariant derivative is simply a number, not a matrix, it is not necessary to assign particles to multiplets depending on their $U(1)$ transformation properties. The $SU(2)$ and $SU(3)$ sectors are more complicated.

The quantum number associated with the $SU(2)$ sector of the Standard Model is weak isospin, T_W^3. It is observed that right-handed chiral states of quarks and leptons carry $T_W^3 = 0$ and so transform as singlets in $SU(2)$ weak-isospin space. Left-handed states, meanwhile, carry $T_W^3 = \frac{1}{2}$ and so transform as members of $SU(2)$ doublets:

$$\begin{array}{ll} e_R^-,\ u_R,\ d_R & SU(2)\ \text{singlets} \\ L = \begin{pmatrix} \nu_{eL} \\ e_L^- \end{pmatrix},\ Q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix} & SU(2)\ \text{doublets} \end{array}$$

Rotations in $SU(2)$ weak-isospin space leave the right-handed states invariant, but transform members of the left-handed lepton and quark doublets into one another. Thus when the lepton doublet points "up" in $SU(2)$ space ($T_W^3 = +\frac{1}{2}$), it is an electron neutrino;

[3] This would *not* be true of a right-handed neutrino, if it existed, because it would have both $Q = 0$ and $T_W^3 = 0$. In the Standard Model, right-handed neutrinos do not exist. But as experimental evidence mounts for neutrino oscillations [5], it seems that neutrinos do have small, nonzero masses. If this is true then their helicity is no longer a Lorentz-invariant quantity, and right-handed neutrinos must exist.

when it points "down" ($T_W^3 = -\frac{1}{2}$), it is an electron. Transitions between members of the electroweak $SU(2)$ doublets are caused by interactions with the gauge fields W^μ.

The quantum number associated with the $SU(3)$ sector of the Standard Model is color charge, C. Of the fermions, only quarks are observed to carry nonzero C. Quarks transform as $SU(3)$ triplets while leptons are colorless and so transform as $SU(3)$ singlets:

$$\begin{array}{cc} e^-,\ \nu_e & SU(3)\ \text{singlets} \\ \begin{pmatrix} u_r \\ u_b \\ u_g \end{pmatrix}, \begin{pmatrix} d_r \\ d_b \\ d_g \end{pmatrix} & SU(3)\ \text{triplets} \end{array}$$

where the r, b, g suffix refers to the value of the C quantum number for the particle (red, blue, green). Rotations in $SU(3)$ color space leave leptons invariant, but turn quark color states into each other. Transitions between color states are caused by interactions with the gauge fields G^μ. Gluons, the physical particles associated with the G^μ gauge fields, themselves carry color charge. *E.g.*, the gluon turning the state u_r (red u quark) into the state u_b (blue u quark) would carry color $b\bar{r}$ (blue/anti-red).

With these multiplet assignments, it is straightforward to write the fermion Lagrangian. For a fermion of wavefunction Ψ, we start from the free-particle Lagrangian $\mathcal{L} = \bar{\Psi}(i\gamma^\mu\partial_\mu - m)\Psi$ and make the kinetic-energy term $\bar{\Psi}i\gamma^\mu\partial_\mu\Psi$ gauge-invariant by the substitution $\partial_\mu \to D_\mu$, where D_μ is the full covariant derivative $D_\mu = \partial_\mu - ig_1\frac{Y}{2}B_\mu - ig_2\frac{\vec{\tau}}{2}\cdot \vec{W}_\mu - ig_3\frac{\vec{\lambda}}{2}\cdot\vec{G}_\mu$. A substitution of this form is made for each fermion or fermion multiplet (L, e_R, Q_L, u_R, d_R) and the terms are summed to yield the full fermion Lagrangian. Using the convention that a term in D_μ acting on the wrong kind of matrix form gives zero [4], $\mathcal{L}_{ferm}$ may be written as

$$\mathcal{L}_{ferm} = \sum_{f=L,e_R,Q_L,u_R,d_R} \bar{f}i\gamma^\mu D_\mu f$$

[4]For example, when it acts on a right-handed lepton, $D_\mu = \partial_\mu - ig_1\frac{Y}{2}B_\mu$, since the 2×2 $SU(2)$ matrix term $\tau^i W^i$ and the 3×3 $SU(3)$ matrix term $\lambda^a G^a$ both give 0 when acting on a state that is a singlet in the $SU(2)$ and $SU(3)$ spaces, respectively.

2.2.3 *Electroweak unification*

The expression for $\mathcal{L}_{ferm}$ can be rewritten in a more useful form. The physical particles associated with the $SU(2)$ gauge fields are not the Cartesian components $(W_\mu^1, W_\mu^2, W_\mu^3)$ appearing in the covariant derivative in the scalar product $\vec{\tau} \cdot \vec{W}_\mu$, but rather are the linear combinations $W_\pm$:

$$\begin{aligned} W_+ &= (-W_1 + iW_2)/\sqrt{2} \\ W_- &= (-W_1 - iW_2)/\sqrt{2} \\ W_0 &= W_3 \end{aligned}$$

Moreover, the neutral $SU(2)$ gauge boson W_0^μ itself does not correspond to a physical particle. Instead, it appears in nature in linear combinations with the neutral hypercharge gauge boson B^μ. Mixing between the $U(1)$ and $SU(2)$ neutral gauge bosons unifies the electromagnetic and weak sectors of the Standard Model.

To write $\mathcal{L}_{ferm}$ in a form that expresses electroweak unification, the compact expression $\mathcal{L}_{ferm} = \sum_f \bar{f} i\gamma^\mu D_\mu f$ is expanded and terms of the form $\bar{\nu}\gamma^\mu\nu$ are collected. These are found to be

$$-\mathcal{L}_{ferm}\,(U(1)\ \&\ SU(2),\ \text{neutrinos}) = (-\frac{g_1}{2} Y_{\nu_L} B_\mu - \frac{g_2}{2} W_\mu^0)\,\bar{\nu}_L\,\gamma^\mu\,\nu_L$$

Define field Z_μ to be proportional to the coefficient of the $\bar{\nu}$-ν interaction term: $Z_\mu \propto (g_1 Y_{\nu_L} B_\mu + g_2 W_\mu^0)$. Since there is no electromagnetic interaction between neutrinos, the electromagnetic field A_μ must be orthogonal to Z_μ: $A_\mu \propto (g_2 B_\mu - g_1 Y_{\nu_L} W_\mu^0)$. After normalization, the fields A^μ and Z^μ can be written as

$$\begin{aligned} A_\mu &= \frac{g_2 B_\mu - g_1 Y_{\nu_L} W_\mu^0}{\sqrt{g_2^2 + g_1^2 Y_{\nu_L}^2}} \\ Z_\mu &= \frac{g_1 Y_{\nu_L} B_\mu + g_2 W_\mu^0}{\sqrt{g_2^2 + g_1^2 Y_{\nu_L}^2}} \end{aligned}$$

After setting $Y_{\nu_L} = -1$ (recall $Y = 2(Q - T_W^3)$) and defining $\sin\theta_w = g_1/\sqrt{g_1^2 + g_2^2}$, $\cos\theta_w =$

$g_2/\sqrt{g_1^2+g_2^2}$, the fields A_μ and Z_μ can be expressed in terms of a mixing matrix:

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_w & \sin\theta_w \\ -\sin\theta_w & \cos\theta_w \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^0 \end{pmatrix}$$

Effectively, the weak mixing angle θ_w has been chosen so that A_μ only couples to charged particles [5]. A_μ is the gauge field of electromagnetism and Z_μ represents a new neutral-current weak force. The gauge bosons associated with these fields are the photon and the Z^0 boson, respectively.

All instances of B_μ and W_μ^0 in $\mathcal{L}_{ferm}$ can be rewritten in terms of the new fields A_μ and Z_μ. From the requirement that the electromagnetic interactions have the familiar form $\mathcal{L}_{EM} = eQ_f A_\mu[\bar{f}_L\gamma^\mu f_L + \bar{f}_R\gamma^\mu f_R]$, where Q_f is the electric charge of the fermion in units of e, the relation $e = \frac{g_1 g_2}{\sqrt{g_1^2+g_2^2}}$ can be deduced. After some manipulation, the fermion Lagrangian can be rewritten in the following useful form:

$$\begin{aligned}
\mathcal{L} &= \sum_{f=e,u,d} e\,Q_f\,(\bar{f}\,\gamma^\mu\,f)\,A_\mu \\
&+ \frac{g_2}{\cos\theta_w} \sum_{f=\nu_e,e,u,d} [\bar{f}_L\,\gamma^\mu\,f_L\,(T_f^3 - Q_f\sin^2\theta_w) \\
&\qquad + \bar{f}_R\,\gamma^\mu\,f_R\,(-Q_f\sin^2\theta_w)]\,Z_\mu \\
&+ \frac{g_2}{\sqrt{2}}\left[(\bar{u}_L\,\gamma^\mu\,d_L + \bar{\nu}_{eL}\,\gamma^\mu\,e_L)\,W_\mu^+ + (\bar{d}_L\,\gamma^\mu\,u_L + \bar{e}_L\,\gamma^\mu\,\nu_{eL})\,W_\mu^-\right] \\
&+ \frac{g_3}{2} \sum_{q=u,d} \bar{q}_\alpha\,\gamma^\mu\,\lambda^a_{\alpha\beta}\,q_\beta\,G^a_\mu
\end{aligned}$$

The second and third generations are included in $\mathcal{L}_{ferm}$ by adding in identical terms with the substitutions $(\nu_e,\ e^-,\ u,\ d) \longrightarrow (\nu_\mu,\ \mu^-,\ c,\ s)$ or $(\nu_\tau,\ \tau^-,\ t,\ b)$. This Lagrangian describes all known interactions [6] of the quarks and leptons. The A_μ term describes the electromagnetic interaction, which couples all charged fermions with strength eQ_f; the Z_μ term describes the weak-force neutral-current interaction, which couples all fermions having nonzero electric charge or weak isospin, with strength $\frac{g_2}{\cos\theta_w}(T_f^3 - Q_f\sin^2\theta_w)$ for left-handed fermions and strength $\frac{g_2}{\cos\theta_w}(-Q_f\sin^2\theta_w)$ for right-handed fermions; the W_μ terms describe

[5] Thanks to D. Carlson's thesis [6] for this phrasing.

[6] Excluding gravity, which is not described in the Standard Model, and the Higgs sector, which will be discussed in Section 2.4.

the weak-force charged-current interaction, which couples members of weak isodoublets with strength $\frac{g_2}{\sqrt{2}}$; and the G_μ term describes the strong-force interaction, which couples quarks with strength $\frac{g_3}{2}$. The *complete* Standard-Model Lagrangian would also describe the interactions of the gauge bosons among themselves, but that is beyond the scope of this paper.

2.2.3.1 Color confinement

Among the implications of the $SU(3)$ gauge structure of the strong interaction, not obvious from the form of the fermion Lagrangian but highly significant to hadron-collider physics, are the dual phenomena of *asymptotic freedom* and *color confinement*. Both ultimately spring from the fact that the gluons themselves carry color charge and so can self-interact. Asymptotic freedom refers to the property that the strong coupling $\alpha_s \equiv g_3^2/4\pi$ becomes smaller as higher and higher energies ($\Rightarrow$ smaller and smaller distance scales) are probed, with the effect that at sufficiently high energies, quarks behave like free particles. This property arises from antiscreening of color charge by gluons [3]. On the other end of the energy scale, the theory of the strong interactions predicts that the strong coupling should become nonperturbatively strong at a scale characterized by $\Lambda_{QCD} \simeq 200$ MeV ("QCD" is short for quantum chromodynamics, the term for the $SU(3)$-sector color-force interactions). In this nonperturbative regime, the potential energy of two colored particles increases linearly with the distance between them. This, again, is a consequence of the gluon self-interaction. The lines of force for a "QCD dipole" (*e.g.* a q-$\bar{q}$ pair) occupy a flux tube of constant cross-sectional area as separation r betwen the charges increases, due to the gluon self-coupling, so that the field energy grows as the volume of the tube (*i.e.*, as r [3]). Therefore a $q\bar{q}$ pair produced in a collision will move away from each other until at some separation it is energetically favorable to snap a particle-antiparticle pair out of the vacuum and form two color-singlet hadrons. Thus colored particles are forever confined in color-neutral hadrons. This implies that quarks produced in a collision will appear not as bare quarks, but as "jets" of hadrons.

2.2.4 *Quark mixing*

The statement that the fermion Lagrangian we wrote in Section 2.2.3 describes all known interactions of the quarks and leptons must be corrected in a subtle way. That Lagrangian assumes that the quark states of definite mass (eigenstates of the mass Hamiltonian) are the same as the quark states which couple to the $W^{\pm}$ bosons (eigenstates of the weak Hamiltonian). This is in fact not the case. As we have written them so far, the interaction terms between the W bosons and the quarks are

$$\mathcal{L}_{ferm}(CC,\ \text{quarks}) = \frac{g_2}{\sqrt{2}}[\bar{u}_L\gamma^\mu d_L W_\mu^+ + \bar{c}_L\gamma^\mu s_L W_\mu^+ + \bar{t}_L\gamma^\mu b_L W_\mu^+ + h.c.]$$

which we can write as $\frac{g_2}{\sqrt{2}}(J_{CC}^\mu W_\mu^+ + h.c.)$, where the charged current to which the W couples is

$$J_{CC}^\mu = \bar{u}\gamma^\mu P_L d + \bar{c}\gamma^\mu P_L s + \bar{t}\gamma^\mu P_L b$$

(Recall $\bar{\psi}_L\gamma^\mu\psi_L = \bar{\psi}\gamma^\mu\psi_L$). Using row and column vectors, we can rewrite J_{CC}^μ as

$$J_{CC}^\mu = (\bar{u}\ \ \bar{c}\ \ \bar{t})\,\gamma^\mu P_L \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad \text{(Slightly wrong.)}$$

The problem with this expression is that it is not mass eigenstates (d, s, b) which ought to appear in the charged-current coupling to the W, but weak eigenstates (d', s', b'):

$$J_{CC}^\mu = (\bar{u}\ \ \bar{c}\ \ \bar{t})\,\gamma^\mu P_L \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} \quad \text{(Correct.)}$$

Then the proper assignment of quark states to weak isodoublets would be $\binom{u}{d'}$, $\binom{c}{s'}$, $\binom{t}{b'}$.

We can rotate the mass eigenstates into the weak eigenstates by means of a unitary matrix V

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

and write the correct form of the charged current as follows:

$$J^{\mu}_{CC} = (\bar{u} \;\; \bar{c} \;\; \bar{t})\, \gamma^{\mu}\, P_L\, V \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

By convention, only the d-type quarks are rotated. However, since the product of two rotations is another rotation, this choice is completely general [7]. The unitary matrix V describing the rotation between the mass eigenstates and the weak eigenstates of the d-type quarks is called the Cabibbo-Kobayashi-Maskawa (CKM) matrix. It can be parameterized by three quark-mixing angles and one complex phase. The current experimentally-measured values [7] for the magnitudes of the CKM matrix elements are shown below.

$$V = \begin{pmatrix} 0.9742 \text{ to } 0.9757 & 0.219 \text{ to } 0.226 & 0.002 \text{ to } 0.005 \\ 0.219 \text{ to } 0.225 & 0.9734 \text{ to } 0.9749 & 0.037 \text{ to } 0.043 \\ 0.004 \text{ to } 0.014 & 0.035 \text{ to } 0.043 & 0.9990 \text{ to } 0.9993 \end{pmatrix}$$

The CKM matrix allows for generation-crossing charged-current weak couplings which, among other things, permit heavier quarks to decay into lighter ones. For example, if we expand J^{μ}_{ch}, we find there is a term $\bar{c}\,\gamma^{\mu}\, P_L\, V_{cb}\, b$ permitting the b quark to decay to a c and a virtual W. Such generation-crossing interactions are suppressed by a factor of an off-diagonal CKM matrix element, which are smaller than the diagonal elements V_{ud}, V_{cs}, V_{tb}. The CKM matrix does *not* permit flavor-changing neutral-current interactions, however. If

[7] Suppose we also replace the u-type quarks of definite mass $(\bar{u} \;\; \bar{c} \;\; \bar{t})$ in J^{μ}_{CC} by rotated versions $(\bar{u}' \;\; \bar{c}' \;\; \bar{t}')$, *i.e.* write

$$J^{\mu}_{CC} = \bar{U}'\, \gamma^{\mu} P_L D'$$

where $U = \begin{pmatrix} u \\ c \\ t \end{pmatrix}$ and $D = \begin{pmatrix} d \\ s \\ b \end{pmatrix}$ are the multiplets of mass eigenstates. Unitary matrices V_{up} and V_{down} rotate to weak eigenstates according to $U' = V_{up} U$, $D' = V_{down} D$. The factor $\bar{U}'$ appearing in J^{μ}_{CC} can be rewritten as $(\bar{u} \;\; \bar{c} \;\; \bar{t}) V^{\dagger}_{up}$, so

$$J^{\mu}_{CC} = (\bar{u} \;\; \bar{c} \;\; \bar{t})\, \gamma^{\mu}\, P_L\, V^{\dagger}_{up} V_{down} \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

Defining $V \equiv V^{\dagger}_{up} V_{down}$, we recover the situation with only the d-type quarks rotated.

we replace the mass eigenstates (d, s, b) in the neutral-current terms in the Lagrangian with weak eigenstates (d', s', b')—expanded in terms of mass eigenstates, mixing angles, and the complex phase—all flavor-changing terms cancel, leaving only flavor-diagonal interactions.

2.3 Feynman Diagrams and Matrix Elements

The Standard-Model fermion couplings can be taken directly from the fermion Lagrangian and expressed as Feynman-diagram vertices. These are shown in Figures 2.1–2.4.

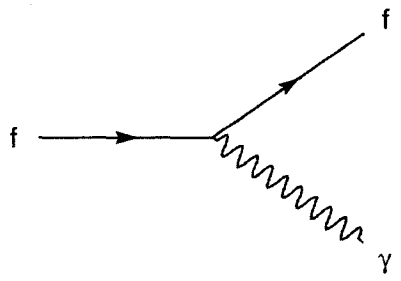


Figure 2.1: The electromagnetic vertex, from $\mathcal{L}_{EM} = eQ_f(\bar{f}\gamma^\mu f)A_\mu$. $f = e, \mu, \tau, u, d, s, c, t, b$. The interaction of charged fermions with the photon is also known as quantum electrodynamics, or QED.

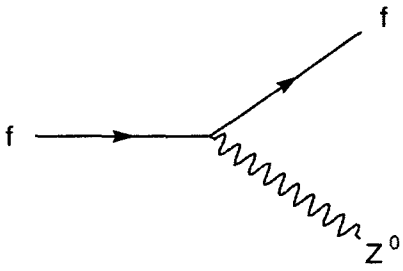


Figure 2.2: The neutral-current weak vertex, from $\mathcal{L}_{NC} = \frac{g_2}{\cos\theta_w}[\bar{f}_L\gamma^\mu f_L(T_f^3 - Q_f \sin^2\theta_w) + \bar{f}_R\gamma^\mu f_R(-Q_f \sin^2\theta_w)]Z_\mu$. $f = e, \mu, \tau, \nu_e, \nu_\mu, \nu_\tau, u, d, s, c, t, b$.

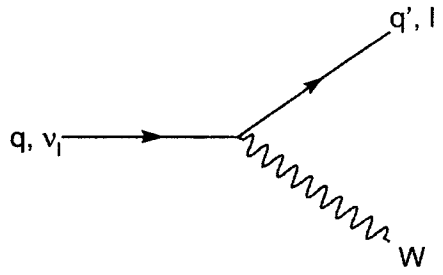


Figure 2.3:
The charged-current weak vertex, from $\mathcal{L}_{CC} = \frac{g_2}{\sqrt{2}}\left[(\bar{q}\gamma^\mu P_L V_{qq'} q' + \bar{\nu}_\ell \gamma^\mu P_L \ell) W_\mu^+ + h.c.\right]$. $\ell = e_L, \mu_L, \tau_L$; $q = u_L, c_L, t_L$; $q' = d_L, s_L, b_L$.

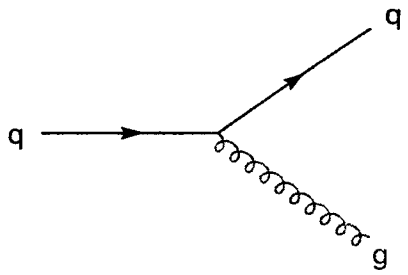


Figure 2.4: The strong vertex, from $\mathcal{L}_{QCD} = \frac{g_3}{2} \sum_q \bar{q}_\alpha \, \gamma^\mu \, \lambda^a_{\alpha\beta} \, q_\beta \, G^a_\mu$. $q = u, d, c, s, t, b$.

These vertices can be joined together to yield the diagrams for all possible Standard-Model interactions. A Feynman diagram is more than just a visualization of how a set of particles interact. Each diagram represents a transition-matrix element $\mathcal{M}_{fi}$ connecting the initial and final states of the diagram. This matrix element is a single term in an infinite series of all possible transitions connecting the initial and final states. The sum of all terms, called the S-matrix [8], contains all physical information for that scattering process.

Feynman rules tell how to write down the transition-matrix element $\mathcal{M}_{fi}$ represented by a particular diagram. These rules specify factors to be included in the transition matrix for each vertex, internal propagator line, or external wave function line in the diagram, along with rules for how to handle loops and other features. In a perturbative theory, each additional vertex introduces a multiplicative coupling-strength factor less than 1, so

that lowest-order terms in the series are the largest and most important in performing a calculation.

The Feynman rules of field theory are derived from the Lagrangian. For example, within a phase factor, the vertex factor for a given vertex is the relevant interaction term in the Lagrangian, with the external wavefunctions removed—*i.e.* $eQ_f\gamma_\mu$ for the electromagnetic interaction, $\frac{g_2}{\sqrt{2}}V_{qq'}\gamma_\mu P_L$ for the charged-current weak interaction between quarks, and so on. To paraphrase Ref. [3], the matrix element produced by the Feynman rules of the Standard Model is approximately $\mathcal{M}_{fi} \simeq \langle f|V|i\rangle$, where the potential V is equivalent to the interaction Lagrangian.

Observables such as cross sections and decay widths are in turn calculated from transition matrix element $\mathcal{M}_{fi}$. For example, the differential cross section for the general two-body scattering process $A + B \to C + D$ can be shown to be [3]

$$d\sigma = \frac{(2\pi)^4\delta^4(P_C + P_D - P_A - P_B)}{4\sqrt{(P_A \cdot P_B)^2 - m_A^2 m_B^2}} \frac{d^3P_C}{2E_C(2\pi)^3} \frac{d^3P_D}{2E_D(2\pi)^3} \overline{|\mathcal{M}_{fi}|^2}$$

The line over $\mathcal{M}_{fi}$ indicates that the matrix element has been summed or averaged over unobserved degrees of freedom such as spins or colors.

2.4 The Higgs Mechanism

For each massive fermion or gauge boson in a field theory, the Lagrangian ought to contain an appropriate mass term. For a fermion ψ, this term is of the form $m\bar{\psi}\psi$; for a neutral boson B_μ, the term is of the form $\frac{1}{2}m_B B^\mu B_\mu$; and for a charged boson W_μ^+, of the form $m_W W^{+\mu} W^+{}_\mu$. Experiment shows that all the quarks and charged leptons, along with the $W^\pm$ and Z^0 bosons, have nonzero mass. But the Standard-Model Lagrangian we have written so far contains no mass terms for either fermions or gauge bosons. Nor can such terms be simply added into the Lagrangian by hand without explicitly breaking desired invariances of the theory. The fermion mass term $m\bar{\psi}\psi$ can be rewritten as

$$m(\bar{\psi}_R\psi_L + \bar{\psi}_L\psi_R)$$

Since right-handed fermions are $SU(2)$ singlets while left-handed fermions occupy $SU(2)$ doublets, the products $\bar{\psi}_R\psi_L$ and $\bar{\psi}_L\psi_R$ are not themselves $SU(2)$ singlets. A Lagrangian containing them would no longer be $SU(2)$-invariant. Similarly, the product $B^\mu B_\mu$ appearing in a gauge-boson mass term is patently not invariant under a gauge transformation such as $B^\mu \to B^{\mu\prime} = B^\mu - \frac{1}{g}\partial^\mu\chi$. The only way to preserve these desired invariances is to set $m = 0$ for all fermions and gauge bosons, in contradiction to experiment.

The Higgs mechanism offers a solution to the problem of mass. This mechanism posits the existence of a fundamental scalar (spin-0) "Higgs field" pervading all space. This field carries nontrivial $SU(2)$ and $U(1)$ quantum numbers, yet is not orthogonal to the vacuum. This means that the vacuum does not share the $SU(2) \times U(1)$ symmetry of the Lagrangian. A symmetry possessed by the Lagrangian but not by the ground state of the system is said to be *hidden* or *spontaneously broken*. We will show in the remainder of this section that in the presence of the Higgs field, the fermions and gauge bosons no longer appear massless.

2.4.1 Masses for the gauge bosons

The Higgs field exists as an $SU(2)$ doublet of complex scalar fields:

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

where in terms of real components,

$$\begin{aligned} \phi^+ &= \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2) \\ \phi^0 &= \frac{1}{\sqrt{2}}(\phi_3 + i\phi_4) \end{aligned}$$

The quantum numbers of the two states ϕ^+, ϕ^0 are $Q = +1, 0$ and $T_W^3 = +\frac{1}{2}, -\frac{1}{2}$ respectively, from which we deduce that $Y = +1$ for both. The Higgs field is a singlet in $SU(3)$ color space. The $SU(2) \times U(1)$ gauge-invariant Lagrangian for the Higgs field is

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

where the covariant derivative D^μ includes the $SU(2)$ and $U(1)$ terms ($D^\mu = \partial^\mu - ig_1 \frac{Y}{2} B^\mu - ig_2 \frac{\vec{\tau}}{2} \cdot \vec{W}^\mu$) and the potential V describes cubic and quartic Higgs self-interactions:

$$V = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2$$

The minimum of potential V is obtained [8] for $\phi^\dagger \phi = 0, \frac{-\mu^2}{2\lambda}$. Parameter λ is necessarily ≥ 0 in order for V to be bounded from below, but μ^2 can be positive or negative. If $\mu^2 < 0$, the ground state of the system corresponds to

$$\phi^\dagger \phi = \frac{-\mu^2}{2\lambda} \equiv \frac{v^2}{2}$$

This means that the ground state of the system corresponds to a nonzero value of the Higgs field: the field has a nonzero *vacuum expectation value.* The surface of minimum potential $\phi^\dagger \phi = \frac{v^2}{2}$ corresponds to a 4-sphere in $(\phi_1, \phi_2, \phi_3, \phi_4)$ space:

$$\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2 = v^2$$

In order to study the spectrum of the system using perturbation theory, it is necessary to choose a vacuum around which to expand. The choice of any *particular* vacuum—*i.e.* any particular point on the 4-sphere—will break global symmetries of the system, leading to the presence of Goldstone bosons in the theory. We choose the vacuum to be

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}$$

or, in real coordinates, $(\phi_1, \phi_2, \phi_3, \phi_4) = (0, 0, v, 0)$. This choice breaks three global symmtries of the system, so we expect the spectrum to contain three Goldstone bosons. It was necessary to assign the vacuum expectation value only to the neutral member ϕ^0 of the Higgs doublet in order to preserve charge conservation, since the quantum numbers of ϕ are those of the vacuum. Had a vacuum expectation value also been given to ϕ^+, electric charge could disappear into the vacuum.

[8] With the definition $r^2 = \phi^\dagger \phi$, $V = \mu^2 r^2 + \lambda r^4$. The condition $\partial V / \partial r = 0$ is achieved for $r = 0$ and $r^2 = -\frac{\mu^2}{2\lambda}$.

The general form of ϕ expanded about the chosen vacuum is then

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} \eta_1(x) + i\eta_2(x) \\ v + H(x) + i\eta_3(x) \end{pmatrix}$$

When ϕ is substituted back into the Higgs Lagrangian, several things happen. As desired, mass terms for the gauge bosons appear in the Lagrangian. So does a mass term for the Higgs field $H(x)$. But undesired features appear as well: the fields η_1, η_2, and η_3 appear as three massless Goldstone bosons which are not seen in nature. Moreover, these fields have bilinear couplings to the gauge bosons, indicating that we have not correctly identified the fundamental particles of the system [9]. These couplings reflect the fact that the degrees of freedom represented by the Goldstone bosons and those represented by the longitudinal polarizations of the (now-massive) gauge bosons are not linearly independent. (If they were, the new Lagrangian would have too many degrees of freedom: three Goldstone bosons, one massive Higgs, and three gauge-boson longitudinal polarizations add up to more d.o.f. than the original Lagrangian's four scalars and no longitudinal polarizations.) To get rid of the Goldstone bosons, yet keep the mass terms for the gauge bosons, we exploit the fact that the Lagrangian is invariant under *local* $SU(2)$ gauge transformations [9]. Therefore we can always choose $\vec{\epsilon}(x)$ such that $\phi' = e^{i\vec{\epsilon}(x)\cdot\vec{\tau}/2}\phi$ is rotated into having zero components along the ϕ_1', ϕ_2', and ϕ_4' directions, *i.e.* is of the form

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}$$

(not bothering to rename the field H.) That is, we "gauge away" the three fields that would become the Goldstone bosons. The Goldstone bosons and their bilinear couplings disappear from the Lagrangian. These three degrees of freedom are given wholly to the longitudinal polarizations of the $SU(2) \times U(1)$ gauge bosons. Hence the saying that the gauge fields "eat" the Goldstone bosons and gain weight.

[9] Recall that $\mathcal{L} = (D^\mu\phi)^\dagger(D_\mu\phi) - V(\phi)$. The covariant-derivative term is obviously $SU(2)$-invariant; and since V depends only on the quantity $\phi^\dagger\phi$, V, too, is invariant under local $SU(2)$ transformations $\phi \to \phi' = e^{i\vec{\alpha}(x)\cdot\vec{\tau}/2}\phi$.

To see how the mass terms for the gauge bosons appear in the Lagrangian, it is necessary to give at least a cursory glance to the algebra of the Higgs mechanism. Starting from the original Lagrangian,

$$\begin{aligned}\mathcal{L} &= (D_\mu\phi)^\dagger(D^\mu\phi) - \dots \\ &= \phi^\dagger D_\mu^\dagger D^\mu\phi - \dots \\ &= \phi^\dagger(\partial_\mu - ig_1\frac{Y}{2}B_\mu - ig_2\frac{\vec{\tau}}{2}\cdot\vec{W}_\mu)^\dagger(\partial^\mu - ig_1\frac{Y}{2}B^\mu - ig_2\frac{\vec{\tau}}{2}\cdot\vec{W}^\mu)\phi - \dots\end{aligned}$$

Substituting in the vacuum-expanded expression for ϕ and focusing in on those terms which might give rise to mass terms for gauge fields B^μ and $\vec{W}^\mu$ [10],

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 0 & v \end{pmatrix}(ig_1\frac{Y}{2}B_\mu + ig_2\frac{\vec{\tau}}{2}\cdot\vec{W}_\mu)^\dagger(ig_1\frac{Y}{2}B^\mu + ig_2\frac{\vec{\tau}}{2}\cdot\vec{W}^\mu)\frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v \end{pmatrix}$$

When this matrix expression is fully expanded, it is found to be

$$\frac{1}{8}\left[v^2g_2^2(W_{1\mu}^2 + W_{2\mu}^2) + v^2(g_1B_\mu - g_2W_{3\mu})^2\right]$$

or, in terms of fields that correspond to particles of definite mass,

$$(\frac{1}{2}vg_2)^2W_\mu^+W^{\mu-} + \frac{1}{2}(\frac{1}{2}v\sqrt{g_1^2+g_2^2})^2Z_\mu Z^\mu + \frac{1}{2}(0)A_\mu A^\mu$$

from which we extract the following masses for the vector bosons:

$$\begin{aligned}M_W &= \frac{1}{2}vg_2 \\ M_Z &= \frac{1}{2}v\sqrt{g_1^2+g_2^2} \\ M_\gamma &= 0\end{aligned}$$

v, the vacuum expectation value of the Higgs field, is $\simeq 246$ GeV.

The absence of a mass term for the electromagnetic gauge boson is not a surprising result, but rather a necessary consequence of charge conservation, which required the vacuum state to be electrically neutral. This means that the charge operator Q gives 0 when

[10] *i.e.* neglecting the ∂_μ, cancelling the two minus signs, and neglecting the H term in ϕ

acting on the vacuum: $Q\phi_0 = (T_3 + \frac{Y}{2})\phi_0 = -\frac{1}{2}\phi_0 + \frac{1}{2}\phi_0 = 0$. Thus the vacuum is invariant under the $U'(1)$ gauge transformation $\phi_0 \rightarrow e^{i\alpha Q}\phi_0$:

$$\begin{aligned} Q\phi_0 \rightarrow \phi_0' &= e^{i\alpha(x)Q}\phi_0 \\ &= (1 + i\alpha Q - \frac{\alpha^2}{2}Q^2 + \ldots)\phi_0 \\ &= \phi_0 \end{aligned}$$

If the vacuum retains a symmetry under any subgroup of the original $SU(2)\times U(1)$ symmetry of the Lagrangian, the gauge boson associated with that group of transformations will be massless [3]. Therefore we expect the theory to have a massless gauge boson associated with the group of $U'(1)$ transformations whose generator is the electric charge operator. This massless gauge boson is the photon.

2.4.2 Masses for the fermions

We have seen how the nonzero vacuum expectation value v of the Higgs field leads to the presence of gauge-boson mass terms in the Lagrangian. We now show how the Higgs mechanism also gives mass to the fermions. We will see that while the Higgs mechanism provides a means to *accommodate* different, nonzero masses for the fermions, it does not derive the particular mass values.

The previously-noted difficulty with fermion mass terms $m\bar{\psi}\psi$ is that they can be rewritten as $m(\bar{\psi}_R\psi_L + \bar{\psi}_L\psi_R)$, which would explicitly break the $SU(2)$ invariance of the Lagrangian since right- and left-handed states are assigned to different multiplets. However, the Higgs doublet provides a means to construct new $SU(2)$ singlets to include in the Lagrangian. We can add in an interaction term for the first-generation fermions with the Higgs field:

$$\mathcal{L}_{int} = g_e(\bar{L}\phi e_R + \phi^\dagger \bar{e}_R L) + g_d(\bar{Q}_L\phi d_R + \phi^\dagger \bar{d}_R Q_L) + g_u(\bar{Q}_L\phi_c u_R + \phi_c^\dagger \bar{u}_R Q_L)$$

where $\phi_c = -i\tau_2\phi^* = \binom{-\phi^{0*}}{\phi^-}$ is also an $SU(2)$ doublet. The coupling strengths g_e, g_d, and g_u are completely arbitrary. To see how mass terms arise from this interaction Lagrangian,

we substitute the expression for the Higgs field, expanded about its vacuum expectation value—*i.e.* $\phi = \frac{1}{\sqrt{2}}\binom{0}{v+H(x)}$, $\phi_c = \frac{1}{\sqrt{2}}\binom{-v-H(x)}{0}$—into the above fermion-Higgs interaction Lagrangian, yielding

$$\begin{aligned}\mathcal{L}_{int} &= \frac{g_e v}{\sqrt{2}}(\bar{e}_L e_R + \bar{e}_R e_L) + \frac{g_e}{\sqrt{2}}(\bar{e}_L e_R + \bar{e}_R e_L)H \\ &+ \frac{g_d v}{\sqrt{2}}(\bar{d}_L d_R + \bar{d}_R d_L) + \frac{g_d}{\sqrt{2}}(\bar{d}_L d_R + \bar{d}_R d_L)H \\ &- \frac{g_u v}{\sqrt{2}}(\bar{u}_L u_R + \bar{u}_R u_L) - \frac{g_u}{\sqrt{2}}(\bar{u}_L u_R + \bar{u}_R u_L)H\end{aligned}$$

This Lagrangian contains a mass term for each fermion. We identify $m_e = \frac{g_e v}{\sqrt{2}}$, $m_d = \frac{g_d v}{\sqrt{2}}$, and $m_u = -\frac{g_u v}{\sqrt{2}}$. Unfortunately, since the coupling strengths g_e, g_d, g_u were arbitrary, the particular mass values m_e, m_d, m_u are not determined, but only accommodated. Rewriting the interaction Lagrangian in terms of the measureable masses m_f rather than the arbitrary coupling strengths g_f:

$$\mathcal{L}_{int} = m_e\bar{e}e + m_d\bar{d}d + m_u\bar{u}u + \frac{m_e}{v}\bar{e}eH + \frac{m_d}{v}\bar{d}dH + \frac{m_u}{v}\bar{u}uH$$

This Lagrangian also contains interaction terms for each fermion with the Higgs boson, which correspond to the Feynman vertex shown in Figure 2.5. The coupling of each fermion

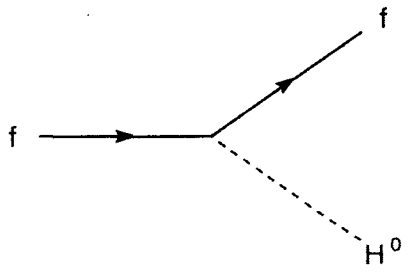


Figure 2.5: The fermion-Higgs vertex, from $\mathcal{L}_{FH} = \frac{m_f}{v}\bar{f}fH$. $f = e, \mu, \tau, u, d, c, s, t, b$.

to the Higgs field is m_f/v—*i.e.* proportional to the mass of the fermion.

2.5 Free Parameters and Unanswered Questions

Since the Standard Model of elementary particle physics was formulated in the 1970's, its predictions have been confirmed by experiment with spectacular precision. As just one example, the Standard-Model prediction for the anomalous magnetic moment of the muon $a_\mu = \frac{g_\mu - 2}{2}$ [10], the amount by which the magnetic moment of the muon departs from the value it would have in the absence of interactions with virtual particles [11], is

$$a_\mu^{\text{theory}} = 0.00116591596 \pm 0.00000000067 \tag{2.1}$$

while the world-averaged experimental measurement of this parameter is

$$a_\mu^{\text{exp}} = 0.00116592300 \pm 0.00000000840 \tag{2.2}$$

The Standard-Model prediction, and subsequent experimental discovery in 1979 and 1983, respectively, of the gauge bosons mediating the strong and weak nuclear forces, have rightly been ranked "among the major intellectual achievements of mankind" [3]. Yet for all its predictive power, the Standard Model leaves certain important questions unanswered.

The Standard Model requires nineteen input parameters [4]. Their presence suggests the need for a further theory which would explain why these parameters have the values they do. The input parameters to the Standard Model are the three gauge coupling constants g_1, g_2, g_3; the three quark-mixing angles and one complex phase needed to express the CKM matrix; the Higgs quartic coupling λ and mass $\sqrt{2\lambda v^2}$; the QCD vacuum angle; and the nine values of the fermion mass spectrum: the three charged-lepton masses m_e, m_μ, m_τ and the six quark masses m_u, m_d, m_c, m_s, m_t, m_b. Figure 2.6 shows the fermion mass spectrum.

The top quark mass, 175 GeV/c^2, is especially puzzling. Although the top quark is a noncomposite, zero-radius point particle, it is roughly as heavy as a gold atom and thirty-five times heavier than the next-heaviest quark. Moreover, the top mass is the only fermion mass comparable to $v \simeq 246$ GeV, the energy scale characterizing electroweak symmetry breaking. These facts make the top quark a particularly interesting object of study.

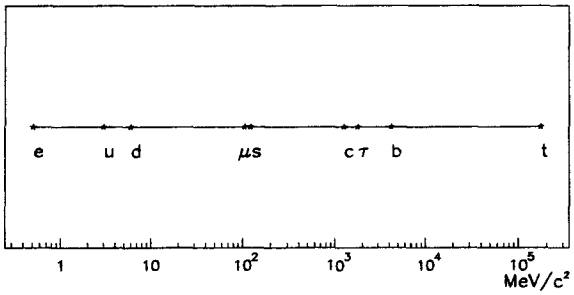


Figure 2.6: Experimentally-measured masses of the fundamental building blocks of matter, the fermions. (1 MeV/$c^2 \approx 1.8 \times 10^{-30}$ kg). The Standard Model says nothing about why these masses span five orders of magnitude.

Chapter 3

Electroweak Single-Top-Quark Production

In this chapter, we focus on electroweak single-top-quark production as a particularly powerful means of studying the top quark. We begin with a review of the various modes by which single top quarks can be produced at the Fermilab Tevatron. We next describe the Standard-Model parameters which can be measured in single-top events and the advantages of measuring them in single-top as opposed to $t\bar{t}$ events. We discuss how single top can be used to probe for new physics in the top sector. Finally, we review the Standard-Model-predicted rates for single-top production at the Tevatron.

3.1 Single-Top Production Modes at the Tevatron

Electroweak single-top-quark production can proceed through several different channels at the Tevatron. In "W-gluon fusion" (Wg) production (Figure 3.1) [12, 13, 14], a light quark emits a virtual t-channel W boson which fuses with a b quark from initial-state gluon splitting to yield a single top quark ($qg \to t\bar{b}q'$). Along with the top quark, the Wg final state also includes the light q' recoiling from W emission and the $\bar{b}$ left over from gluon splitting. In a second mode of single-top production, known as W^* production [15, 16], an initial-state q and $\bar{q}'$ annihilate to form a highly off-shell s-channel W boson—the term "W^*" emphasizes that it has $q^2 \gg M_W^2$—which subsequently decays to a t and $\bar{b}$ ($q\bar{q}' \to t\bar{b}$). Figure 3.2 shows a leading-order diagram for W^* production of single top. A third produc-

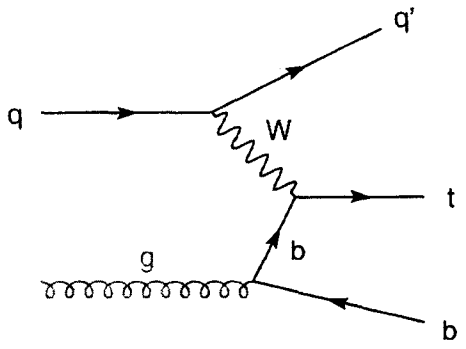


Figure 3.1: "W-gluon fusion" (Wg) single top.

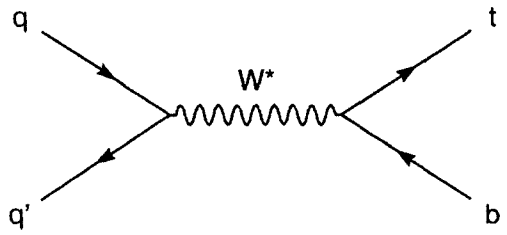


Figure 3.2: W^* single top.

tion mechanism, "associated production" of a single top quark and an on-shell W boson ($gb \rightarrow tW$) [17, 18], is shown in Figure 3.3. The heaviness of the final state [19] and the fact

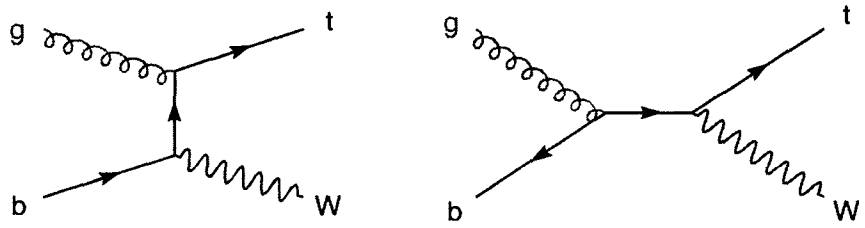


Figure 3.3: Associated production of single top and a real W (tW). Suppressed at Tevatron.

that it proceeds through a gluon-b interaction [20] makes this process much less important at the Tevatron energy ($\sqrt{s} = 1.8$ TeV) than the previous two.

It is not immediately obvious that any of these single-top production mechanisms would have appreciable cross sections at the Tevatron compared to the dominant pure-QCD

$t\bar{t}$ production mode (Figure 3.4). One of the first papers on Wg single top [12] describes

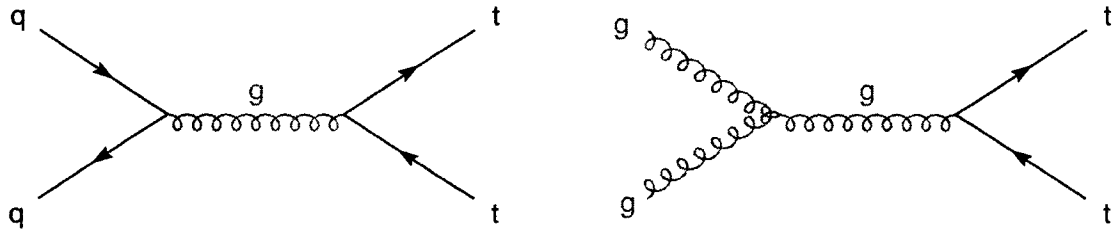


Figure 3.4: Top quark pair ($t\bar{t}$) production.

the result that Wg production of heavy u-type quarks can actually *exceed* strong-interaction production for sufficiently large mass splittings $m_u - m_d$ as "somewhat surprising", explaining

> Although W-gluon fusion is lower order in the strong interaction than the pure strong-interaction processes, it is second order in the weak interaction. W-gluon fusion also produces a three-body final state, and is therefore suppressed by a phase-space factor of about $(2\pi)^{-3} \times \frac{\pi}{4}$. We therefore expect W-gluon fusion to be suppressed relative to the strong processes by a factor $\alpha^2/4\pi\alpha_s \sim 10^{-4}$. [This includes a factor of 2 from the coupling of gluons to color-triplet quarks, which is $\alpha_s/2$ rather than α_s.] [12]

But Ref. [12] goes on to list a number of enhancements to Wg fusion which overcome this suppression factor. Some of these effects are more relevant at higher-energy colliders, but the effects apparently boost the Tevatron cross section as well. Among these effects are scaling behavior of the subprocess cross section at high energy ($\hat{\sigma} \sim 1/M_W^2$ for Wg due to the spacelike W propagator *vs.* $\hat{\sigma} \sim 1/\hat{s}$ for pure-strong processes); lack of color suppression for the Wg initial state relative to the gg; lower momentum fraction x in the structure functions due to the presence of only one heavy particle in the final state; and logarithmic enhancement associated with the collinear singularity if the b quark were massless. Ref. [21] also cites the presence of factors of $(m_{fermion}/M_W)$ in the couplings [1].

[1]Ref. [21] uses two approximations in calculating σ_{Wg}: the "effective-W" approximation, in which the

In sections to follow we will discuss in more detail the Standard-Model calculation of the single-top cross sections, along with the uncertainties on these cross sections. Here we just present a preview of the results in Table 3.1. Note that the sum of the cross sections

Process	Name	Theory Cross Section	Uncertainty
$q\bar{q}' \to t\bar{t}$	$t\bar{t}$	5.1 pb	18%
$qg \to q't\bar{b}$	Wg	1.70 pb	14%
$q\bar{q}' \to t\bar{b}$	W^*	0.73 pb	14%
$gb \to tW$	tW	< 0.1 pb	

Table 3.1: Theoretically-predicted cross sections for Standard-Model top production mechanisms at $\sqrt{s} = 1.8$ TeV [22, 20, 23]. Sources of theoretical uncertainty include uncertainties in parton distribution functions, uncertainties in m_t, and scale dependence.

for all the single-top processes is nearly one half the cross section for pair-produced top.

3.2 Single Top Within the Standard Model

Electroweak single-top-quark production is of theoretical interest because it provides direct information on the charged-current interaction of the top quark. Unlike the case of top-pair production, where the electroweak vertex Wtb plays a role only in the top quark's decay, in single top the *production* cross section contains information on the coupling of top to W and b. Specifically, the cross section is proportional to the square of the Wtb vertex factor. Therefore, as summarized by Ref. [20], "within the standard model, single-top-quark production offers a means to directly measure the... CKM matrix element V_{tb}. Beyond the standard model, it is sensitive to a nonstandard Wtb vertex, and to exotic single-top-quark production processes involving new particles." In this section we examine the first of these claims.

W boson is treated as a parton within the proton, and an effective Lagrangian describing the interactions between the longitudinally-polarized gauge bosons, the Higgs boson, and the fermions. They find that cross sections involving W_L's in the initial state are enhanced by powers of $(m_{fermion}/M_W)^2$.

3.2.1 Measuring V_{tb}

We recall from Chapter 2 that in the Standard Model, the unitary 3 × 3 Cabibbo-Kobayashi-Maskawa (CKM) matrix describes the rotation between the mass eigenstates and the weak eigenstates of the d-type quarks:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

Before discussing how single-top production can be used to measure V_{tb}, it is useful to review how well V_{tb} is known, and how well it can be measured through other avenues. (Since expressions for physical observables such as cross sections depend on the absolute square of V_{tb}, we really only measure the magnitude $|V_{tb}|$. But for convenience, we will often drop the absolute-value bars and refer to this term simply as V_{tb}.)

3.2.1.1 How well do we know V_{tb}?

How well we know V_{tb} depends on whether or not we assume that the three-generation Standard Model is the complete picture. If three-generation unitarity is required, then V_{tb} is extremely well constrained; if not, then V_{tb} is extremely poorly constrained. As Ref. [24] summarizes:

> If we assume unitarity and only three generations of quarks, then from direct and indirect measurements of the other elements, the magnitude of V_{tb} is very precisely known: $|V_{tb}| = 0.9991 \pm 0.0001$, without any direct measurements using the top quark system. However, once these two constraints are removed, then the allowed 90% confidence level (CL) limits [2] open up to leave essentially no bounds on V_{tb}:

$$\begin{pmatrix} — & — & — \\ — & — & — \\ 0.004 \text{ to } 0.014 & 0.034 \text{ to } 0.046 & 0.9989 \text{ to } 0.9993 \end{pmatrix}_{\text{3 gen.}} \Longrightarrow$$

$$\begin{pmatrix} — & — & — & \cdots \\ — & — & — & \cdots \\ 0 \text{ to } 0.11 & 0 \text{ to } 0.52 & 0 \text{ to } 0.9993 & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}_{\text{4+ gen.}}$$

[2] These values are from Particle Data Group, R. Barnett *et al.*, Phys. Rev. D **54**, 1 (1996). More-recently-updated values are available, but do not change the argument here. For example, the year-2000 PDG 90% C.L. limits [7] on $|V_{tb}|$ in the absence of three-generation unitarity are $0.07 - 0.993$.

A fourth generation of quarks $\binom{t'}{b'}$ could cause V_{tb} to be significantly less than 1 [3]. But the existence of a fourth generation cannot be ruled out by the above bounds on V_{tb}. A *direct* measurement of V_{tb}, by contrast, would provide model-independent bounds. A direct measurement of V_{tb} less than unity would be a major discovery [16].

3.2.1.2 Measuring V_{tb} in $t\bar{t}$ events

Since the $gt\bar{t}$ vertex is responsible for $t\bar{t}$ production, the cross section for $t\bar{t}$ production does not contain any information on V_{tb}. But information on V_{tb} can be deduced from the pair-produced t and $\bar{t}$'s subsequent weak decays. To measure V_{tb} in top decays, one looks at top events containing a W boson [4] and notes the fraction of those events which also contain b quarks:

$$B_b = \frac{t \to Wb}{t \to Wq} = \frac{|V_{tb}|^2}{|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2}$$

(Here "$t \to Wb$" is meant to indicate not the branching fraction to Wb, but the fraction of decays with W's which also contain b's [26].) The principle for extracting ratio B_b is simple. If a pure $t\bar{t}$ sample were available, the ratio would be

$$B_b = \frac{t \to Wb}{t \to Wq} = \frac{\text{No. tagged events}/\epsilon}{\text{No. untagged events} + \text{No. tagged events}/\epsilon}$$

where ϵ is the efficiency for identifying a b jet [24]. In practice, obtaining a pure $t\bar{t}$ sample is not feasible. Instead, various techniques—such as measuring the ratio of double-tagged to single-tagged jets in b-tagged "lepton+jets" events [5], or measuring the fraction of tagged jets in "dilepton" events [6]—are combined into a likelihood estimator to extract B_b [26]. The value for B_b thus obtained is 0.99 ± 0.29 [27].

If we assume three-generation unitarity, then $|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2 = 1$ and $B_b = |V_{tb}|^2$. This assumption yields $|V_{tb}^{3\ gen.}| > 0.80$ at 90% confidence level [27]. If, this assumption is dropped, then we can solve for $|V_{tb}|^2 = \frac{B_b}{1-B_b}\left[|V_{td}|^2 + |V_{ts}|^2\right]$. Clearly, some

[3] Even if the decay $t \to Wb'$ were kinematically disallowed, the element $V_{tb'}$ might still be large, which would force V_{tb} to be small [25].

[4] As opposed to a charged Higgs, for example.

[5] "Lepton+jets" events display the signature of one semileptonic t decay and one hadronic t decay, as if from $t\bar{t} \to \ell\nu_\ell bqq'b$.

[6] "Dilepton" events display the signature of two semileptonic t decays, as if from $t\bar{t} \to \ell_1\nu_{\ell_1}b\ell_2\nu_{\ell_2}b$.

assumptions must be made about CKM matrix elements V_{td} and V_{ts} in order to say anything about V_{tb} in the absence of three-generation unitarity. Requiring V_{td} and V_{ts} to have the mean values from the Standard-Model three-generation matrix ($V_{td} = 0.009$ and $V_{ts} = 0.040$) permits the following lower limit to be set on V_{tb} [27]:

$$|V_{tb}^{3\ gen.'}| > 0.055 \quad (90\%\ \text{C.L.})$$

Thus even with these large assumptions about V_{td} and V_{ts}, the lower limit on V_{tb} is weak. The "ironic" result that even excellent precision on B_b yields only a modest limit on V_{tb}, to paraphrase Ref. [26], lies in the relationship between B_b and V_{tb}. It can be shown that

$$\delta|V_{tb}|^2 \sim \frac{1}{(1-B_b)^2}\,\delta B_b$$

so that measuring B_b close to 1 results in large uncertainty on V_{tb}. Thus the $t\bar{t}$ decay method for measuring V_{tb} is model-dependent and not very powerful. Single top provides a better method.

3.2.1.3 Measuring V_{tb} in single-top events

At tree level, assuming pure Standard-Model couplings, the Wtb vertex factor may be written as

$$\frac{g_2}{\sqrt{2}}\,V_{tb}\,\bar{t}\,\gamma^\lambda\,P_L\,b\,\epsilon_\lambda$$

Since all the single-top production diagrams contain one instance of this vertex, any single-top cross section $\sigma \sim \overline{|\mathcal{M}|^2} \sim |V_{tb}|^2$. Therefore if the Standard Model holds, the single-top-quark production cross section directly measures V_{tb}. As Ref. [24] points out, the uncertainty on V_{tb} thus extracted is simply "the error on the measured single top cross section added in quadrature with the error on the theoretical calculation of that cross section, all divided by two." [7] Moreover, this measurement of V_{tb} is free from assumptions about other elements of the CKM matrix, which as we saw in the preceding section were a chief disadvantage of measuring V_{tb} through top-quark decays in $t\bar{t}$ events.

[7] Since $\sigma \propto |V_{tb}|^2$, it is straightforward to show by error propagation that $\delta|V_{tb}|/|V_{tb}| = \frac{1}{2}\delta\sigma/\sigma$.

3.2.2 *Deriving* $\Gamma(t \to Wb)$

Once V_{tb} has been measured, it can be used to normalize other top-quark observables, such as the partial decay width to W and b, through a relation such as [28]

$$\Gamma(t \to bW) = \frac{G_F\, m_t^3}{8\sqrt{2}\pi} |V_{tb}|^2 \left(\frac{2k}{m_t}\right) \left[\left(1 - \frac{m_b^2}{m_t^2}\right)^2 + \left(1 + \frac{m_b^2}{m_t^2}\right)\left(\frac{M_W^2}{m_t^2}\right) - 2\left(\frac{M_W^4}{m_t^4}\right)\right]$$

where k, the W momentum in the t rest frame, is

$$k = \left[m_t^2 - (M_W + m_b)^2\right]^{\frac{1}{2}} \left[m_t^2 - (M_W - m_b)^2\right]^{\frac{1}{2}} / 2m_t$$

For large m_t, this width takes on the asymptotic form

$$\Gamma(t \to bW) \simeq 180 \text{ MeV} \times |V_{tb}|^2 \left(\frac{m_t}{M_W}\right)^3$$

which for $m_t = 175$ GeV/c^2, corresponds to a partial width of ~ 1.8 GeV [26] (if we assume $V_{tb} \approx 1$). This is an important calculation because the top-quark width is difficult to measure directly at a hadron collider, where the typical experimental resolution with which a mass peak can be reconstructed is nearly an order of magnitude larger than this value [8] (Figure 3.5).

3.2.2.1 Top quark lifetime

The above-calculated value for the top-quark width $\Gamma \sim 1.8$ GeV corresponds to a lifetime $\tau = \hbar/\Gamma$ of $\sim 0.4 \times 10^{-24}$ sec. An interesting ramification of the shortness of this lifetime is that the top quark decays before the strong force can bind it into a hadron. Ref. [28] shows that the lifetime for an ultra-heavy quark Q that decays to Wq will, for sufficiently large quark mass m_Q, become shorter than the timescale $\Lambda_{QCD}^{-1} \sim 10^{-23}$ sec needed for $(Q\bar{q})$ or $(Q\bar{Q})$ hadrons to form. Therefore in many respects "such quarks behave

[8] The measured width of the top mass peak is effectively the quadrature sum of the intrinsic top width and the experimental resolution: $\Gamma_{meas} = \sqrt{\Gamma_{int}^2 + \Gamma_{exp}^2}$. Therefore the intrinsic top width can in principle be deduced: $\Gamma_{int} = \sqrt{\Gamma_{meas}^2 - \Gamma_{exp}^2}$. However, the uncertainty on the thus-derived intrinsic top width is $\delta\Gamma_{int} = \frac{1}{\Gamma_{int}}\sqrt{\Gamma_{meas}^2 \delta\Gamma_{meas}^2 + \Gamma_{exp}^2 \delta\Gamma_{exp}^2}$. We see in Figure 3.5 that Γ_{exp} is large, which means that Γ_{meas} will also be large. Thus to make $\delta\Gamma_{int}$ reasonably small, both $\delta\Gamma_{exp}$ and $\delta\Gamma_{meas}$ must be made VERY small. The latter requires large numbers of real top events in data.

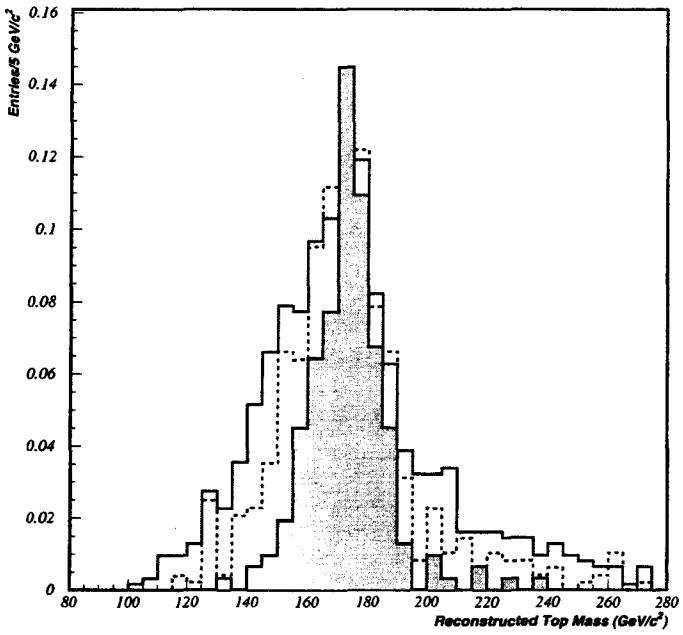


Figure 3.5: Distribution of reconstructed top-quark mass in $t\bar{t}$ events from HERWIG Monte Carlo with CDF detector simulation. Top mass assumed to be $m_t = 170$ GeV/c^2. The shaded histogram shows the mass distribution when the Monte-Carlo-level information is used to assign the correct identities of final-state jets. The distribution is approximately Gaussian, with mean of 170.0 GeV/c^2 and $\sigma = 11.0$ GeV/c^2 [26]. (Chapter 5 will provide more information on Monte Carlo event generators and detector simulation programs.)

like (quasi)-free particles, their dynamical properties determined only by electroweak and *perturbative* strong interactions." The reason top hadrons do not form is as follows:

> The binding force in a heavy quarkonium state is essentially coulombic. The revolution time of the $(Q\bar{Q})$ bound state is then estimated as $t_R \sim 9/4 m_Q \alpha_s^2$. If the lifetime of the $(Q\bar{Q})$ system becomes shorter than the revolution time t_R, then the quarkonium bound state cannot be formed any more. Setting $\alpha_s = 0.15$ to illustrate the point one finds this to happen for
>
> $$m_Q \gtrsim 125 \text{ GeV} \times |V_{Qq}|^{-\frac{2}{3}}$$
>
> ...
>
> A somewhat stronger limit applies to $(Q\bar{q})$ bound states. We find a mass limit of
>
> $$m_Q \gtrsim 100 \text{ GeV} \times |V_{Qq}|^{-\frac{2}{3}}$$
>
> above which no more open-flavor hadrons can exist, *i.e.* Q decays before it can form a meson by picking up a light quark q: $\tau_Q < t_{Had} \sim \Lambda_{QCD}^{-1} \sim 10^{-23}$ s [28].

With a mass of 175 GeV/c^2, the top quark is well above these thresholds for either $t\bar{q}$- or $t\bar{t}$-meson formation.

3.2.3 Does single top measure $\Gamma(t \to bW^+)$ directly?

Some single-top phenomenologists have argued that single-top-quark production provides a means to measure the top-quark width *directly*. For example, Ref. [19] makes the claim:

> [Wg fusion]...within the SM...provides a way to directly probe the partial width of the top quark, $\Gamma(t \to W^+b)$, through the effective-W approximation...valid at energies much larger than...m_W...in which the W boson is treated as a parton within the proton. Using this approximation, σ_{Wg} can be related to the width $\Gamma(t \to W^+b)$ by the equation [19]
>
> $$\sigma_{Wg} \simeq \sum_{\lambda=0,+,-} \int dx_1 dx_2 f_\lambda(x_1) b(x_2) \left[\frac{16\pi^2 m_t^2}{\hat{s}(m_t^2 - M_W^2)}\right] \Gamma(t \to W_\lambda^+ b)$$

However, Ref. [29] disagrees with this view, saying

> Partial widths are physical observables, in the same league with masses—in principle, one can measure them without recourse to an underlying theory. In the case of single top, the physical observable being measured is the cross section, not the partial width. Thus it is misleading to report the result of the measurement in terms of the partial width. It is better to report the result of the measurement in terms of an unphysical parameter, V_{tb}, which has no meaning outside of an underlying theory (the standard model) [29].

Ref. [29] speculates that the reason Wg fusion single top has been "misguidedly singled out as a measurement of the partial width" is an overly-literal interpretation of the effective-W approximation used in the earliest calculations of σ_{Wg} [30, 21]. In the effective-W approximation, one thinks of "a quark radiating a W boson, which then strikes a b quark in the proton sea, promoting it to a t quark". Ref. [29] does not dispute the validity of the effective-W method, but stresses that the radiated W boson is not on-shell even though it is treated as on-shell in the calculation. It is this difference in q^2 scale between the Wtb coupling involved in Wg single-top production *vs.* that in top decays which makes the identification of σ_{Wg} with $\Gamma(t \to Wb)$ invalid. As Ref. [29] explains,

> The W boson in t-channel single top production is far off shell; its virtuality is restricted by the kinematics to $q^2 < 0$. The W boson in top decay is on shell, $q^2 = M_W^2$. Thus the two processes are not directly related. To make my point, consider [W^* production]... where the W is again highly virtual, this time with $q^2 > (m_t + m_b)^2$. Nobody has proposed that one should consider this process to be a measurement of $\Gamma(t \to Wb)$, yet it is really no different than the t-channel process, in the sense that both processes have the t and b on shell and the W off shell.

Furthermore, Wg single-top production and top-quark decay "have significant QCD radiative corrections which are unrelated to each other. Thus any attempt to identify the two processes is at best valid at leading order." The correct way to proceed, according to Ref. [29], is to (1) measure σ_{Wg}; (2) compare to the theoretically-calculated σ_{Wg} (with QCD corrections); (3) extract V_{tb}; and (4) use V_{tb} to calculate $\Gamma(t \to Wb)$, again including QCD corrections. This approach "is correct to all orders in QCD, and makes clear that the partial width is not being measured, but rather is a derived quantity." The same procedure can be followed for W^* single top.

3.2.4 Polarization in single top

Before leaving the subject of Standard-Model measurements that can be made using single-top production, we mention that single-top events also provide an excellent opportunity to study top-quark polarization [31, 17, 20]. The weak charged-current interaction only couples left-handed chiral states, so the single top quarks are almost 100% left-handed polarized [17]. Moreover, we saw in Section 3.2.2.1 that the top decays before hadronization can obscure its spin state, so its polarization is observable in the distribution of its decay products [20]. Thus it has been said that in decaying, the single-top quark will "analyze its own polarization" [31].

In addition, the unique kinematics of single-top production permit the construction of a so-called "optimal basis" in which this polarization is especially easy to study [32]. Ref. [32] shows that in single-top-quark production, the spin of the top quark is highly polarized along the direction of the d-type quark in the event. In W^* production at the Tevatron, the source of the d-type quark is predominantly the antiproton. In Wg production,

the d-type quark ends up in the recoiling light-quark jet about three-quarters of the time. Ref. [20] explains why, in Wg single top, even in the case where the light-quark jet does *not* contain the d quark, the top is often still polarized along the light-quark jet's direction:

> Since W-gluon fusion proceeds via $ug \to dt\bar{b}$ about 77% of the time at the Tevatron ($\sqrt{s} =$ 2 TeV), the d quark is usually the light-quark jet. The other 23% of the events proceed via $\bar{d}g \to \bar{u}t\bar{b}$, in which case the $\bar{d}$ quark is moving along one of the beam directions. However, since the light-(anti)quark ($\bar{u}$) jet tends to move in the same direction as the $\bar{d}$ antiquark, the direction of the light-quark jet is still rather a good basis to analyze the spin in these events.

Ref. [32] shows that the top quark is 98% polarized along the direction of the antiproton in W^* single top and 96% polarized along the direction of the light-quark jet in Wg single top at the Tevatron.

Ref. [20] cites [33] in singling out the charged lepton from the semileptonic decay of the top quark—however produced—as a particularly sensitive indicator of the top spin. In the top rest frame, the angular distribution of the lepton is given by

$$\frac{1}{\Gamma}\frac{d\Gamma}{d\cos\theta} = \frac{1}{2}(1+\cos\theta)$$

where θ is the angle in the top rest frame between the direction of the lepton and the spin of the top. Since, for example, in Wg events the top spin is highly polarized along the direction of the light-quark jet, the angle between this jet and the charged lepton, in the top rest frame, would be expected to obey this distribution. Ref. [20] identifies the asymmetry in the angular distribution of ℓ as a simple test variable for polarization in Wg events:

$$A \equiv \frac{\sigma(-1 < \cos\theta < -0.1) - \sigma(-0.1 < \cos\theta < 0.8)}{\sigma(-1 < \cos\theta < -0.1) + \sigma(-0.1 < \cos\theta < 0.8)}$$

(The small-angle region is excluded since lepton and jet identification require a minimum angular separation between the two.) Statistically-significant nonzero A would constitute observation of polarization of the top quark in Wg events.

3.3 Single Top as a Probe for New Physics

If new physics is present in single-top production, then it cannot be used either to extract V_{tb} or derive $\Gamma(t \to Wb)$ [29]. But the prospect of discerning new physics beyond the Standard Model is far more interesting than that of just measuring another Standard-Model parameter.

It was observed over a decade ago [34] that the heaviness of the top quark makes the top sector an especially likely place to look for new physics. As the heaviest of all fermions, the top is most strongly coupled to the symmetry-breaking sector responsible for the generation of mass [35]. This fact, along with the tantalizing proximity of the top mass to the Fermi scale $v \simeq 246$ GeV characterizing electroweak symmetry breaking [9], have led to the suggestion that the top quark may participate in new dynamics underlying the generation of mass [37].

In the Standard Model, v is the vacuum expectation value of the complex scalar Higgs field which generates the masses of the gauge bosons and fermions. But some argue from "triviality and naturalness" [38] that the fundamental Higgs scalar of the Standard Model is only an effective theory. Instead, they propose dynamical models of symmetry breaking in which condensates of fermions play the role of the Higgs boson:

> In theories of dynamical electroweak symmetry breaking, the electroweak interactions are broken to electromagnetism by the vacuum expectation value of a fermion bilinear. These theories may thereby avoid the introduction of fundamental scalar particles, of which we have no examples in nature [35].

Dynamical symmetry breaking with preferential involvement of the third quark generation is attractive because it simultaneously avoids a fundamental scalar and explains the large top mass. Models in which top-quark condensates [39] alone are responsible for symmetry breaking are not favored since they predict a too-large top mass [40], but this problem can be avoided by combining top condensation with technicolor (another model of dynamical symmetry breaking) as is done in topcolor-assisted technicolor [41]. In all these models,

[9] Larios and Yuan cheekily write $m_t \simeq v/\sqrt{2}$ [36].

the large top mass is not an accident, but a necessary consequence: "the top mass is large because the strong forces needed to bind tops into Higgs result in a strong Higgs coupling to top." [37] Experimental study of the top quark is necessary to determine if these models are correct, *i.e.* in the words of Ref. [34], to determine "whether the top is just another quark or whether it holds the clue to the physics of the Fermi scale."

In this section we outline the main points covered in Ref. [37], an excellent review of the impact on single-top production of the various proposed types of new physics. We will see that the Wg and W^* single-top modes are sensitive to different kinds of new physics. Following this paper's example, we group the possible types of new physics that affect single top into two categories: those which involve the effects of a *new particle* that couples to the top quark, and those which involve the effect of a *modification of the Standard-Model couplings* between the top and other known particles.

3.3.1 New particles

Ref. [37] cites a wide variety of extensions to the Standard Model which call for an extra set of quarks $\binom{t'}{b'}$. For example, some models require additional fermions to participate in a seesaw mechanism to generate the top mass [42, 43]. These new quarks could affect single-top rates by mixing with the third generation. If V_{tb} were significantly less than unity, rates for the Wg and W^* diagrams shown in Figures 3.1 and 3.2 would be suppressed below Standard-Model expectations. However, other elements of the CKM matrix might also be affected by mixing in a way that counteracts the suppression. We saw in Section 3.2.1.1 that when the requirement of three-generation unitarity is removed, the constraints on V_{tb} and V_{ts} become much weaker [10]:

$$0.0 \leq |V_{ts}| \leq 0.55$$

$$0.06 \leq |V_{tb}| \leq 0.9993$$

[10]These values, taken from Ref. [37], are the 1998/1999 values available in C. Caso *et al.*, The European Physical Journal **C3**, 1 (1998) and 1999 off-year partial update for the 2000 edition available on the PDG WWW pages (URL: http://pdg.lbl.gov/). Since Ref. [37] was published, newer limits $V_{ts} = 0 - 0.16$ and $V_{tb} = 0.07 - 0.993$ have become available [7]. The tighter upper bound on V_{ts} makes Ref. [37]'s discussion of the large-V_{ts} scenario less likely, but we include it because the general principle still holds.

If V_{ts} took on the maximum allowed value of 0.55 and V_{tb} were as large as allowed by unitarity, *i.e.* $V_{tb} = 0.835$, since $0.55^2 + 0.835^2 = 1$, then Wg single top could proceed significantly through an initial-state gluon splitting to $s\bar{s}$ rather than $b\bar{b}$. The much larger parton distribution function for the s quark compared to the b would increase the Wg rate above the three-generation-unitarity Standard Model value (NLO cross section 4.07 pb *vs.* 2.44 pb [20] at Tevatron Run II.) The rate for W^* with final state $t\bar{b}$ would meanwhile be suppressed to ~ 0.7 of the Standard Model expectation. Of course, the rate for W^* with final state $t\bar{s}$ would be enhanced, but the absence of the long-lived $\bar{b}$ in the final state would make this mode distinguishable from $t\bar{b}$. Additionally, the new b' might be directly produced in an s-channel W^*-like interaction $q\bar{q}' \to t\bar{b}'$.

Other models of new physics posit the existence of additional gauge symmetries mediated by new gauge bosons. For example, a version of the top-flavor model proposes that the third-generation quarks transform under an additional $SU(2)_h$ symmetry that is related to the generation of the top mass [43]. This model requires three additional gauge bosons $W'^{\pm}$ and Z'. The W' boson could contribute to s-channel single-top production via $q\bar{q}' \to W' \to t\bar{b}$, with a possible resonant enhancement if enough energy were available to produce the W' near on-shell. Additionally, the W' process could interfere with the regular W^*-exchange diagrams, with the effect that the overall rate could be increased or decreased. The existence of a W' would have negligible impact on Wg fusion, however. There the virtual W boson is constrained by the kinematics to be spacelike, *i.e.* to have $q^2 < 0$, so that the contribution from t-channel exchange of a W' is suppressed by the heavy W' mass (by a propagator factor of $1/M^2_{W'}$). This type of argument is applicable to any new heavy particle's impact on single-top production: the s-channel process can receive resonant enhancement, while t-channel process is relatively insensitive.

A similar analysis holds for the case of additional scalar bosons. Such particles are called for in a wide variety of theories. For example, the top-color dynamical model calls for pion-like bound states of top and bottom quarks [44] ("top-pions" $\pi^{\pm}$, π^0) to play the role of the symmetry-breaking Higgs bosons. These composite scalars couple strongly to

the top, but through a right-handed interaction. They contribute to the s-channel mode of single-top through $c\bar{b} \to \pi^+ \to t\bar{b}$ due to a large $t_R - c_R$ mixing (not constrained by the CKM matrix, which refers only to left-handed states.) Since the interaction is right-handed, this process does not interfere with the Standard-Model W^* process. As in the case of the new gauge boson W', the s-channel process can receive a large resonant contribution from top-pion exchange, but any contribution to t-channel process is suppressed by $1/M_\pi^2$. Note that the single top quarks produced via an intermediate scalar particle will have different polarization properties than those produced through a vector boson. Specifically, the "optimal" polarization basis discussed in Section 3.2.4 no longer applies. Of course, if the interaction is right-handed, this also will show up in nonstandard polarization behavior.

3.3.2 Anomalous couplings

Ref. [37] uses a model-independent method of parameterizing the top's couplings to the known light particles called the "electroweak chiral Lagrangian". The EWCL approach involves writing an effective Lagrangian which applies at energies well below the scale Λ at which new physics becomes relevant. The first term in the EWCL is the usual Standard-Model Lagrangian. Subsequent terms represent deviations from the Standard Model in terms of operators in successively higher and higher mass dimensions, which operators are suppressed by successively higher and higher powers of Λ:

$$\mathcal{L}_{eff} = \mathcal{L}_{SM} + \mathcal{L}_4 + \mathcal{L}_5 + \ldots$$

Thus for energies well below Λ, the lowest-dimension operators have the largest effect. Dimension-4 operators which could affect single-top production include charged-current terms modifying the tWb coupling and flavor-changing neutral current (FCNC) terms involving the top's coupling to Z and c. (Terms with the u substituted for the c are also possible but are omitted here.)

$$\begin{aligned}\mathcal{L}_4 \;=\; & \frac{e}{\sqrt{2}\sin\theta_W} W_\mu^- (\kappa_{Wtb}^L\, e^{i\phi_{Wtb}^L}\, \bar{b}\,\gamma^\mu\, P_L\, t + \kappa_{Wtb}^R\, e^{i\phi_{Wtb}^R}\, \bar{b}\,\gamma^\mu\, P_R\, t) \\ & + \frac{e}{2\sin\theta_W\cos\theta_W} Z_\mu (\kappa_{Ztc}^L\, e^{i\phi_{Ztc}^L}\, \bar{c}\,\gamma^\mu\, P_L\, t + \kappa_{Ztc}^R\, e^{i\phi_{Ztc}^R}\, \bar{c}\,\gamma^\mu\, P_R\, t) + h.c.\end{aligned}$$

where $\phi^{L(R)}_{Wtb(Ztc)}$ are CP-violating phases. The dimension-5 operator contains terms that could also affect single-top production, through couplings of the top to the gluon and photon fields:

$$\begin{aligned}\mathcal{L}_5 &= \frac{g_3 G^a_{\mu\nu}}{\Lambda_{gtc}}(\kappa^L_{gtc}\, e^{i\phi^L_{gtc}}\, \bar{c}\, T^a\, \sigma^{\mu\nu}\, P_L\, t + \kappa^R_{gtc}\, e^{i\phi^R_{gtc}}\, \bar{c}\, T^a\, \sigma^{\mu\nu}\, P_R\, t) \\ &+ \frac{2eF_{\mu\nu}}{3\Lambda_{\gamma tc}}(\kappa^L_{\gamma tc}\, e^{i\phi^L_{\gamma tc}}\, \bar{c}\, \sigma^{\mu\nu}\, P_L\, t + \kappa^R_{\gamma tc}\, e^{i\phi^R_{\gamma tc}}\, \bar{c}\, \sigma^{\mu\nu}\, P_R\, t) + h.c.\end{aligned}$$

Here $i\sigma^{\mu\nu} = -\frac{1}{2}[\gamma^\mu, \gamma^\nu]$.

The terms in $\mathcal{L}_4$ that modify the Wtb coupling could affect single-top production. Form factor κ^R_{Wtb} is already strongly constrained by preexisting low-energy $b \to s\gamma$ data [45], but both the Wg and W^* modes are sensitive to κ^L_{Wtb}. Moreover, the $\mathcal{L}_4$ and $\mathcal{L}_5$ FCNC couplings Ztc, gtc, and γtc (and also with substitution $c \to u$) can have a large impact on single-top production. New modes of s-channel production such as $q\bar{q}' \to Z^*/g/\gamma \to t\bar{c}(\bar{u})$ would be open (Figure 3.6), but the absence of the long-lived $\bar{b}$ quark in the final state means they would probably not be confused with ordinary W^* production. Nor would these new s-channel modes interfere with Standard-Model W^*, since the final state is different. The

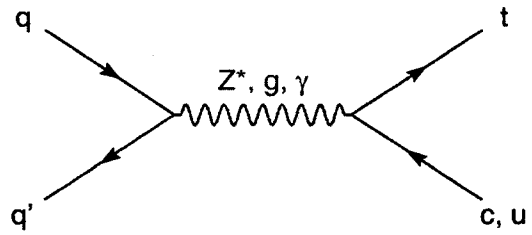


Figure 3.6: Exotic s-channel single-top production modes made possible by new FCNC couplings. The absence of the $\bar{b}$ in the final state makes these processes distinguishable from Standard-Model W^*.

t-channel mode of single top would be strongly affected by these new vertices, however, through diagrams in which initial-state b quarks are replaced by c's and the t-channel W by a $Z/g/\gamma$ (Figure 3.7.) As in the case of W^*, the final state contains a $\bar{c}$ or $\bar{u}$ instead of a $\bar{b}$. But as will be discussed in Chapter 5, the final-state $\bar{b}$ in Wg single top tends to be low-momentum and often missed anyway, so its absence here will not serve to distinguish this

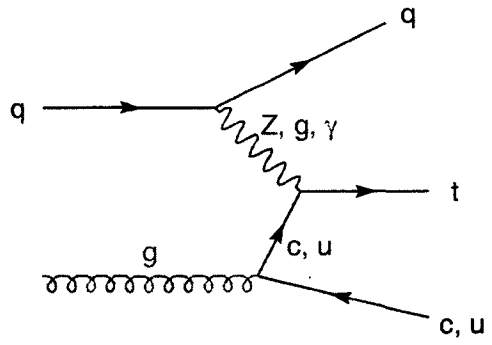


Figure 3.7: Sample exotic t-channel single-top production modes made possible by new FCNC couplings. The (absent) final-state $\bar{b}$ is not a significant part of the Wg signature, so these processes are *not* distinguishable from Standard-Model Wg single top.

final state from Standard-Model Wg. Moreover, the larger probability density for finding lighter c or u quarks in the parton sea could enhance this new t-channel rate with respect to ordinary Wg, an effect which might offset the presumably smaller FCNC coupling. These diagrams do not interfere with Standard-Model Wg fusion due to the different final state.

Measuring the rate of single-top production will help constrain the magnitudes of the interactions of these new FCNC operators, set by κ_{Ztc}, Λ_{gtc}, and $\Lambda_{\gamma tc}$. Some of these anomalous couplings are already constrained by low-energy data. For example, κ^L_{Ztc} and κ^R_{Ztc} are constrained by low-energy data to be less than 0.05 and 0.29, respectively [46]. But these constraints are indirect because they all rely on assumptions about the underlying theory. New physics could generate large cancellations to permit κ_{Ztc} to be as large as 1! Single top, by contrast, provides an opportunity to constrain these couplings directly, with fewer assumptions.

3.3.3 Summary of the impact of new physics on single top

Thus we see that the Wg and W^* modes of single-top production are sensitive to different kinds of new physics. While a modification of the Wtb coupling affects both modes, the presence of additional heavy particles affects primarily the s-channel mode and

anomalous couplings affect primarily the t-channel mode. Thus studying the correlation of the measured single-top rates in the t-channel *vs.* s-channel (σ_t-σ_s) plane can help discern what, if any, new physics is affecting the production. Figure 3.8 shows this plane for Tevatron Run II. The Standard Model point is indicated with the 3σ theoretical-uncertainty contour drawn around it. Points for several representative new-physics models are also shown. Figure 3.9 shows the analogous information for the CERN's Large Hadron Collider (LHC). Ref. [37] explains how these plots can be used:

> A deviation in σ_s that is not also reflected in σ_t is most likely due to the effect of nonstandard particles. A deviation in σ_t that is not also seen in σ_s is likely from a FCNC. A deviation that is comparable in both rates is most likely from a modification of the W-t-b interaction. In the very least, if the SM is a sufficient description of single top production, the fact that the two rates are consistent will allow one to use them to extract V_{tb} with confidence that new physics is not distorting the measurement.

Thus measuring single-top production is a win-win game: either we measure V_{tb}, or we find evidence of new physics.

3.4 The Standard-Model Calculations for σ_{Wg} and σ_{W^*}

Having established motivation for studying single top, we now take a closer look at the Standard Model calculations of the cross sections for the two main single-top processes at the Tevatron, the s-channel W^* process and the t-channel Wg process. We examine features which simplify or complicate each rate calculation and which affect the systematic uncertainty on the final answer.

3.4.1 The W^* cross section

Calculating the next-to-leading-order (NLO) cross section for $q\bar{q}' \rightarrow W^* \rightarrow t\bar{b}$ is relatively straightforward for several reasons. The energy scale of the process, characterized by the top mass ($m_t = 175$ GeV/c^2) is well away from the scale at which QCD becomes nonperturbative ($\Lambda_{QCD} \simeq 200$ MeV) so that the calculation can be performed in a regime

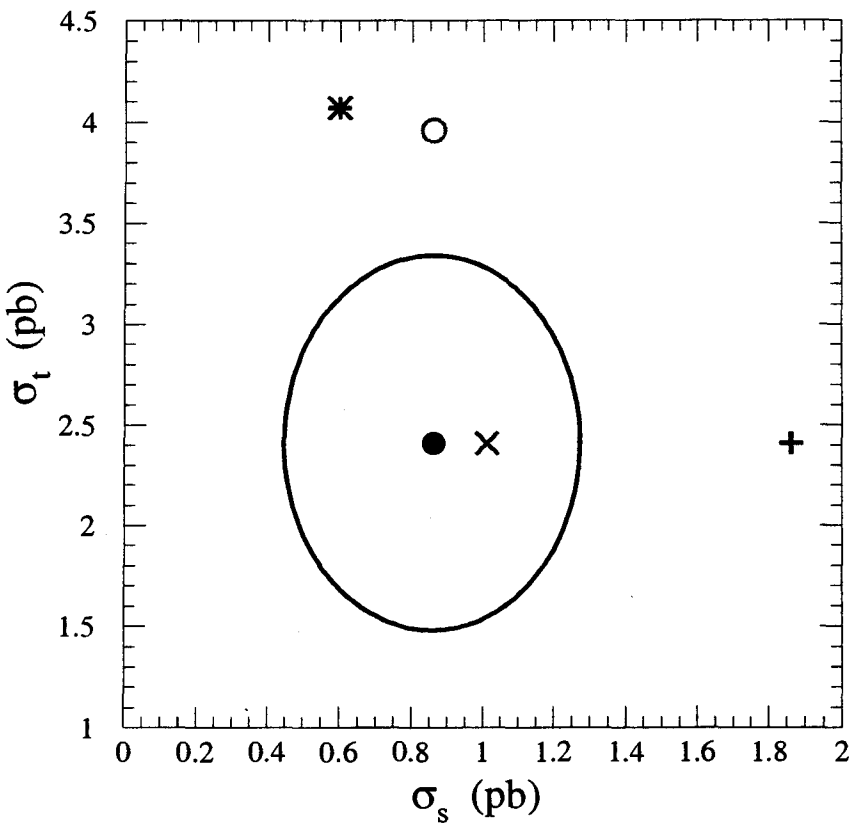


Figure 3.8: The location of the Tevatron SM point (the solid circle) in the σ_s-σ_t plane, and the 3σ theoretical deviation curve. Also shown are the points for the top-flavor model (with $M_{Z'} = 1$ TeV and $\sin^2\phi = 0.05$) as the $\times$, the FCNC Z-t-c vertex ($|\kappa_{Ztc}| = 1$) as the open circle, a model with a charged top-pion ($m_{\pi^\pm} = 250$ GeV and t_R-c_R mixing of $\sim 20\%$) as the cross, and a four quark generation scenario with $|V_{ts}| = 0.55$ and $|V_{tb}| = 0.835$ as the asterisk. All cross sections sum the t and $\bar{t}$ rates at $\sqrt{s} = 2$ TeV. Taken from [37].

where perturbative QCD is very reliable [23]. The q and $\bar{q}'$ parton distribution functions are evaluated at moderate values of momentum fraction x, where they are well-known [16]. Moreover, the process is similar to the well-studied Drell-Yan leptoproduction process $q\bar{q}' \to \ell\nu$. This confers several advantages. The initial-state QCD correction for W^* is identical

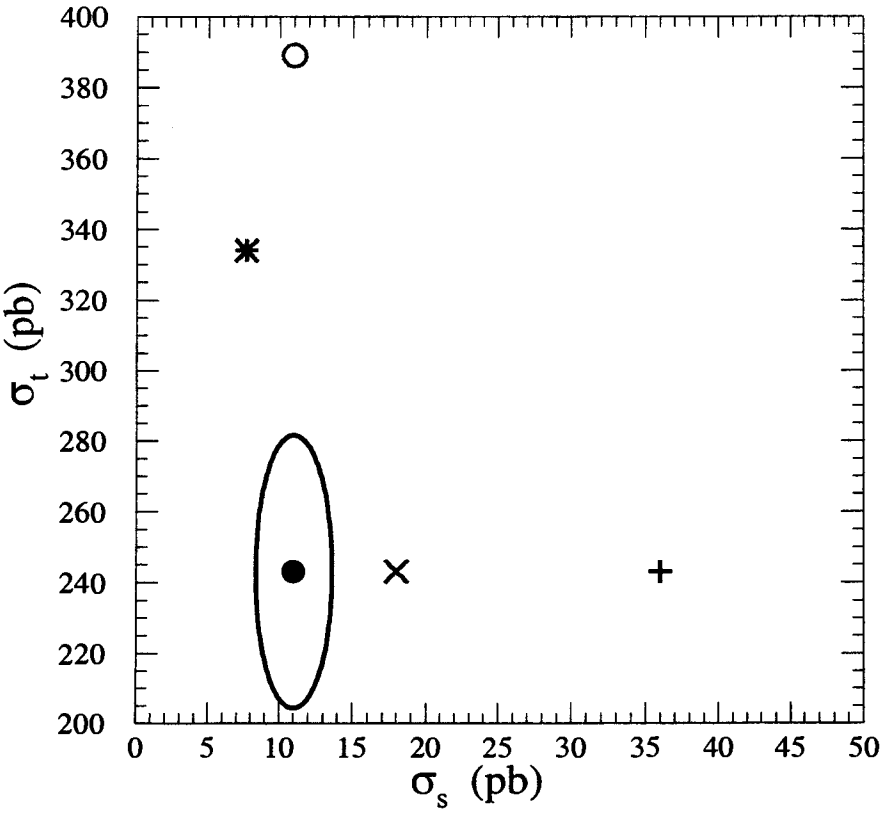


Figure 3.9: Analogous to Figure 3.8, except for the LHC (pp, $\sqrt{s} = 14$ TeV.) From [37].

to that for Drell-Yan, and the latter has been calculated to $\mathcal{O}(\alpha_s^2)$ [47]. This leaves only the final-state QCD corrections still to be done, and these themselves are straightforward since they do not contain any collinear or infrared singularities [16]. Furthermore, the initial- and final-state corrections do not interfere at NLO since the $t\bar{b}$ final state is still a color singlet when a gluon radiates from an initial-state quark but a color octet when the gluon radiates from a final-state quark [23]. Finally, in W^* the $q\bar{q}'$ initial state can be experimentally constrained by comparing to Drell-Yan data. Note that the $q\bar{q}'$ flux can only be *constrained* by measuring $q\bar{q}' \to \ell\nu$, not measured directly, because at a hadron

collider the p_z of the neutrino is not known, so the q^2 of the W boson cannot be fully determined [16]. Nevertheless, Drell-Yan data can be used to reduce systematic uncertainty on PDF's and QCD corrections to the initial state.

The most important corrections to the order-α_w^2 leading-order W^* cross section are the QCD correction of order α_s and the Yukawa correction of order $\alpha_w m_t^2/M_W^2$ [23]. We consider each in turn.

3.4.1.1 $\mathcal{O}(\alpha_s)$ corrections to σ_{W^*}

Diagrams contributing to the $\mathcal{O}(\alpha_s)$ QCD correction to the W^* process are shown in Figure 3.10 [48]. As discussed above, at this order the QCD corrections to the initial state

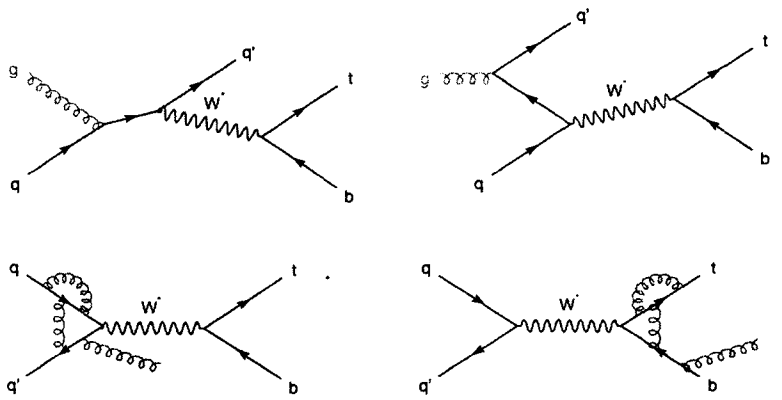


Figure 3.10: Diagrams contributing to the $\mathcal{O}(\alpha_s)$ correction to $q\bar{q}' \to tb$. From [23].

and final state do not interfere and may be considered separately. The $\mathcal{O}(\alpha_s)$ corrections to the production of a W^* of mass squared q^2 are the same as for the Drell-Yan process and are found to increase the LO cross section by +36% at the Tevatron. Final-state corrections to the subsequent propagation and decay of the W^* are obtained from the imaginary part of the self-energy of the W and are found to increase the cross section by +18% over LO. Thus the total QCD correction increases the LO cross section by a significant amount—over 50%.

3.4.1.2 $\mathcal{O}(\alpha_w m_t^2/M_W^2)$ Yukawa corrections to σ_{W^*}

Diagrams contributing to the $\mathcal{O}(\alpha_w m_t^2/M_W^2)$ Yukawa correction to the W^* process are shown in Figure 3.11. The Yukawa correction arises from loops of Higgs bosons and

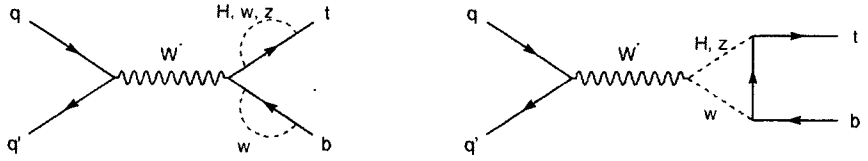


Figure 3.11: Diagrams contributing to the $\mathcal{O}(\alpha_w m_t^2/M_W^2)$ correction to $q\bar{q}' \to t\bar{b}$. The dashed lines indicate the Higgs boson and unphysical scalar W and Z bosons in the R_ξ gauge. The diagram on the left represents a wavefunction renormalization and the diagram on the right a vertex correction. From [23].

the scalar components of virtual vector bosons. (Note that a top-quark loop in the W propagator—which might be expected to contribute a term of Yukawa strength—has already been absorbed by the renormalized weak coupling constant $\alpha_w = g^2/4\pi = \sqrt{2}G_\mu M_W^2/\pi$.) The result for the Yukawa correction can be expressed as the fractional change in the total cross section, $\Delta\sigma_Y/\sigma_{LO}$, as a function of Higgs mass M_H. For Higgs mass values in the range 60 GeV/c^2–1 TeV/c^2, the absolute value of the correction is found to be never more than 1% of the leading-order cross section.

Furthermore, since the Yukawa correction is of $\mathcal{O}(\alpha_w m_t^2/M_W^2)$, it is expected to be at least as large as the ordinary $\mathcal{O}(\alpha_w)$ weak correction for the known value of the top mass 175 GeV/c^2. Thus since the Yukawa correction is found to be negligible, so too should be the weak correction (which as of 1996 could not yet be calculated explicitly due to the unavailability of weak-corrected parton distribution functions.) Since the W-gluon-fusion process also involves the $t\bar{b}$ weak charged current, the Yukawa correction to that process is probably also negligible.

3.4.1.3 Results for σ_{W^*}, with uncertainty

The central value of the NLO calculation of σ_{W^*} is 0.73 pb. Uncertainty on this value arises from several sources: scale dependence, uncertainty in the modeling of the parton distribution functions, and uncertainty in the measured top mass. Scale dependence represents uncertainty from uncalculated higher-order QCD corrections. It is estimated by noting the change in the cross section induced by varying the relevant scale—factorization scale μ_F for initial-state corrections and renormalization scale μ_R for final-state—between $\frac{1}{2}\sqrt{q^2}$ and $2\sqrt{q^2}$. The total uncertainty from scale dependence is found to be 4%. The central value for the cross section uses $\mu_F = \mu_R = \sqrt{q^2}$, the mass of the virtual W.

To minimize the effects of uncertainty in modeling of parton distribution functions, the cross section was calculated using two different sets of NLO PDF's—CTEQ3M and MRS(A')—and the central value was taken to be the average of the results. The difference between the results is a measure of the PDF uncertainty, which was taken to be 4%.

The calculated cross section for $q\bar{q}' \to t\bar{b}$ depends strongly on the top mass, so uncertainty on m_t leads to additional uncertainty on σ_{W^*}. Using the experimental result $m_t = 175 \pm 6$ GeV/c^2, 15% uncertainty due to m_t is obtained. The latest estimates of top-mass uncertainty ± 5.1 GeV/c^2 [49] reduce the σ_{W^*} uncertainty slightly, to 12.8%.

Thus the total uncertainty on the NLO σ_{W^*} calculation is $4\% \oplus 4\% \oplus 12.8\% = 14\%$.

$$\sigma_{W^*} = 0.73 \pm 0.10 \text{ pb}$$

3.4.2 The Wg cross section

In contrast to the case of W^*, the Wg cross section calculation is not at all straightforward. The crux of the difficulty is that the tree-level diagram for Wg fusion (Figure 3.1) contains what would be a collinear divergence if the b quark were massless [50]. Specifically, if $m_b = 0$, then the diagram is divergent when the initial-state gluon emits a collinear $\bar{b}$, *i.e.* splits into a real $b\bar{b}$ pair. Of course, the b is not massless; its mass regulates the singularity. But a memory of the collinear divergence remains in the total cross section as

logarithmically enhanced terms, *i.e.* terms proportional to

$$\ln\left[\frac{Q^2+m_t^2}{m_b^2}\right]$$

where $Q^2 \equiv -q^2$ is the virtuality of the W boson of four-momentum q. (Since Q^2 is usually less than or of order M_W^2, Ref. [50] abbreviates this logarithm in the text [11] as $\ln(m_t^2/m_b^2)$.) Ref. [50] summarizes the problem presented by these logarithmic terms as follows:

> The total cross section for W-gluon fusion contains these logarithmically-enhanced terms, of order $\alpha_s \ln(m_t^2/m_b^2)$, as well as terms of order α_s (both terms also carry a factor of α_w^2, which we suppress in the following discussion). Furthermore, logarithmically enhanced terms, of order $\alpha_s^n \ln^n(m_t^2/m_b^2)/n!$, appear at every order in the perturbative expansion in the strong coupling, due to collinear emission of gluons from the internal b-quark propagator. Since the logarithm is large, the perturbation series does not converge quickly, and it appears difficult to obtain a precise prediction for the total cross section.

The solution to this problem of large collinear logarithms lies in a paradigm shift in the treatment of the Wg process. Instead of the $2 \to 3$ diagram of Figure 3.1, the leading-order process is now taken to be the $2 \to 2$ diagram $qb \to q't$ (Figure 3.12) in which the initial-state b quark is described by a parton distribution function. Ref. [50] explains why

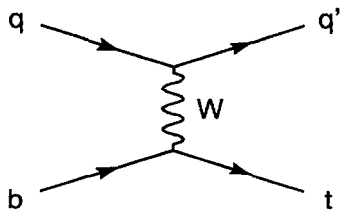


Figure 3.12: Leading-order diagram for Wg single top. This process is intrinsically of order $\alpha_s \ln(m_t^2/m_b^2)$ due to the b parton distribution function.

the b-quark PDF solves the problem of the large collinear logarithms as follows:

> The coefficient of the logarithmically-enhanced term is the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) splitting function P_{qg}, which describes the splitting of a gluon into a $b\bar{b}$ pair. One can sum the logarithms by introducing a b distribution function

[11] Though *not* the calculations, of course.

$b(x, \mu^2)$ and calculating its evolution with μ (from some initial condition) via the DGLAP equations. Thus the b distribution function can be regarded as a device to sum the collinear logarithms.

Moreover, since the b PDF "is calculated from the splitting of a gluon into a collinear $b\bar{b}$ pair, it is intrinsically of order $\alpha_s \ln(\mu^2/m_b^2)$" rather than "merely of order α_s". Ref. [50] shows analytically and graphically that

$$b(x, \mu^2) \sim \alpha_s(\mu^2) \ln(\mu^2/m_b^2) g(x, \mu^2)$$

where $g(x, \mu^2)$ is the gluon PDF. Thus the cross section for the $2 \to 2$ process is of order $\alpha_s \ln(m_t^2/m_b^2)$ due to the b PDF ($\mu \approx m_t$).

There are then two independent, numerically comparable corrections to this process that must be considered at NLO: one correction of order $1/\ln(m_t^2/m_b^2)$ and another of order α_s.

3.4.2.1 $\mathcal{O}(1/\ln(m_t^2/m_b^2))$ correction to σ_{Wg}

The $2 \to 3$ diagram of Figure 3.1 now represents a correction to the LO diagram from an initial gluon. The cross section contains terms of order $\alpha_s \ln(m_t^2/m_b^2)$ and order α_s. But the logarithmically-enhanced terms have *already* been included in the LO cross section by way of the b PDF. These collinear logarithms must be removed from the $2 \to 3$ diagram in order to avoid double counting them. This is done by subtracting the contribution of the diagram in Figure 3.13, which represents the collinear region of the $2 \to 3$ process where the gluon splits to a real $b\bar{b}$ pair. After the terms of order $\alpha_s \ln(m_t^2/m_b^2)$ have been subtracted from diagram of Figure 3.1 by the terms in the diagram of Figure 3.13, only terms of order α_s remain. Recalling that the LO process was of order $\alpha_s \ln(m_t^2/m_b^2)$, we see that since the diagrams of Figures 3.1 and 3.13, taken together, contribute a term of order α_s, this contribution is suppressed by a factor of $1/\ln(m_t^2/m_b^2)$ with respect to the LO process. We might naïvely have expected the diagram of Figure 3.1 to constitute a correction of order α_s to the LO process, but this is not the case. The existence of a correction of $\mathcal{O}(1/\ln(\mu^2/m_b^2))$ is "generic to any process involving a perturbatively derived heavy-quark

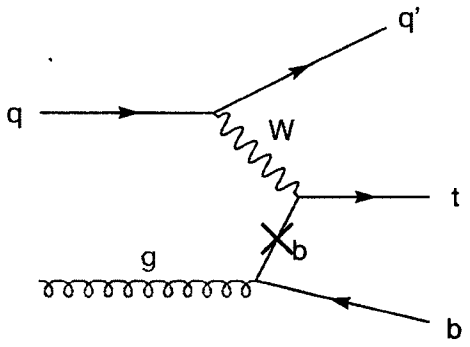


Figure 3.13: Diagram for the collinear region of the $2 \to 3$ process in which the gluon splits to a real $b\bar{b}$ pair. The cross section contains *only* the terms of $\mathcal{O}(\alpha_s \ln(m_t^2/m_b^2))$, *i.e.* the collinear-divergence logarithms. The $\times$ on the internal b quark line indicates that the quark is on-shell.

distribution function in the region $\mu^2 \gg m_Q^2$." The overall size of the $\mathcal{O}(1/\ln(m_t^2/m_b^2))$ correction at the Tevatron is found to be -20%.

3.4.2.2 $\mathcal{O}(\alpha_s)$ correction to σ_{Wg}

Diagrams contributing to $\mathcal{O}(\alpha_s)$ correction to the LO process $qb \to q't$ may be grouped into several categories: corrections to the light-quark vertex (Figure 3.14), a cor-

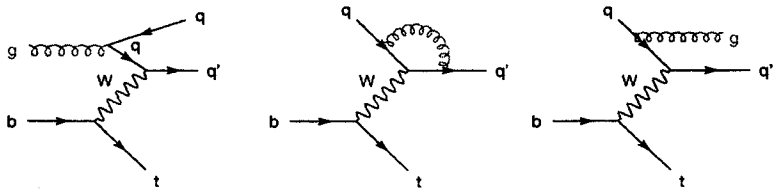


Figure 3.14: QCD corrections to the light-quark vertex of $qb \to q't$.

rection to the heavy-quark vertex which *does not* contain more collinear logarithms (Figure 3.15), and a correction to the heavy-quark vertex which *does* contain more collinear logarithms (Figure 3.16). The light-quark-vertex corrections do contain collinear logarithms of order $\ln(Q^2/m_q^2)$, where m_q is the light-quark mass, but these are "absorbed by the light-quark distribution functions in the usual way. Since the light-quark distribution functions

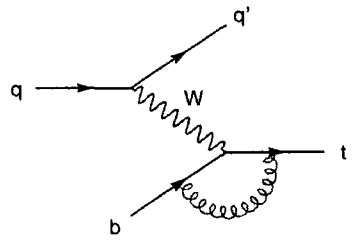


Figure 3.15: QCD correction to the heavy-quark vertex of $qb \to q't$.

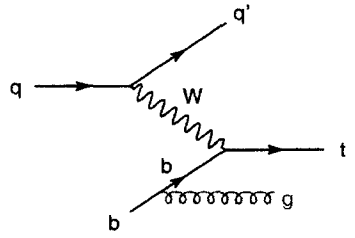


Figure 3.16: QCD correction to the heavy-quark vertex of $qb \to q't$. This diagram contains another power of the collinear logarithm in (m_t^2/m_b^2) corresponding to emission of a collinear gluon.

are intrinsically of zeroth order in α_s, the remaining corrections are of order α_s [with respect to the LO process]." The diagram of Figure 3.15, meanwhile, has a cross section of order $\alpha_s^2 \ln(m_t^2/m_b^2)$ (including the factor of $\alpha_s \ln(m_t^2/m_b^2)$ from the b PDF), so it is suppressed by a factor of α_s with respect to the LO process. The cross section for the diagram of Figure 3.16, however, contains terms of order $\alpha_s^2 \ln(m_t^2/m_b^2)$ and $\alpha_s^2 \ln^2(m_t^2/m_b^2)$. Collinear emission of the gluon, on top of the factor from the b PDF, yields the squared-logarithm term. Furthermore, "another power of this logarithm appears at every order in the strong coupling". To get the perturbative calculation to converge, it is necessary to modify the b PDF again. A term corresponding to gluon emission is added to the DGLAP evolution equation for the b PDF to sum these additional collinear logarithms. But once these logarithms have been absorbed into the b PDF (*i.e.* into the LO cross section), it is necessary not to double-count them by including the $\mathcal{O}(\alpha_s^2 \ln^2(m_t^2/m_b^2))$ terms from the diagram of Figure 3.16. The collinear region must be subtracted, leaving only the terms of order

$\mathcal{O}(\alpha_s^2 \ln(m_t^2/m_b^2))$, which correspond to non-collinear gluon emission and constitute a true $\mathcal{O}(\alpha_s)$ correction to the LO process.

An analogous approach can be used to include the contributions of still higher-order diagrams. As Ref. [50] summarizes,

> [A]ll collinear logarithms are ultimately summed into the b distribution function; no explicit collinear logarithms remain. The remaining terms are all of order α_s^n or, if the diagram has a b quark in the initial state, of order $\alpha_s^n \ln(m_t^2/m_b^2)$. These correspond to corrections of order $\alpha_s^{n-1} \times 1/\ln(m_t^2/m_b^2)$ or α_s^{n-1}, respectively, compared with the leading-order process.

The overall size of the $\mathcal{O}(\alpha_s)$ correction is found to be +13%. It is apparently a numerical accident that this correction tends to compensate the $\mathcal{O}(1/\ln(m_t^2/m_b^2))$ correction.

3.4.2.3 Results for σ_{Wg}, with uncertainty

The central value of the NLO calculation of σ_{Wg} is 1.70 pb [20]. This value was obtained using the CTEQ4M $\overline{MS}$ PDF's, $m_t = 175$ GeV/c^2, $V_{tb} = 1$, and scale $\mu^2 = Q^2$ at the light-quark vertex and $\mu^2 = Q^2 + m_t^2$ at the heavy-quark vertex. As in the case of W^* single top, uncertainty on the cross section arises from scale dependence, PDF uncertainty, and top mass uncertainty. Scale uncertainty, *i.e.* uncertainty from uncalculated higher-order QCD corrections, is estimated by varying the scale μ in the b-quark PDF between one-half and twice its central value of $\mu^2 = Q^2 + m_t^2$. The scale is also varied in the strong coupling [20]. But Ref. [50] argues that the scale in the light-quark PDF should not be varied, since there $\mu^2 = Q^2$. The total impact of scale variation on the calculated cross section is found to be $\pm 5\%$.

The cross section for Wg single-top production is not as strongly dependent on the top mass as was the cross section for W^* production. Although as m_t increases the PDF's do decrease, this effect is not augmented by an accompanying decrease in constituent cross section $\hat{\sigma}$, since $\hat{\sigma} \sim 1/M_W^2$ rather than $1/\hat{s}$. An uncertainty of ± 5.2 GeV/c^2 [12] corresponds to an uncertainty in the cross section of 9% due to m_t [20].

[12] Close enough to the official uncertainty ± 5.1 GeV/c^2 [49].

The worst source of uncertainty on σ_{Wg} is in the modeling of the gluon distribution function, which in turn affects the b PDF. Ref. [50] did not even venture to estimate this uncertainty. The more-recent paper [20] cites work [51] stating that this uncertainty appears to be less than 10% at the Tevatron and the LHC.

Thus the total uncertainty on the NLO σ_{Wg} calculation is $5\% \oplus 9\% \oplus 10\% = 14\%$.

$$\sigma_{W^*} = 1.70 \pm 0.24 \text{ pb}$$

Chapter 4

Experimental Apparatus

The data used in this analysis were collected by the Collider Detector at Fermilab (CDF) during the 1992-1995 run of Fermilab's Tevatron collider. The three-story-high, five-thousand-ton CDF detector surrounds the point at which the Tevatron's beams of 900-GeV protons and 900-GeV antiprotons collide. As summarized in a 1988 *NIM* article, the purpose of the CDF detector is "to measure the energy, momentum, and where possible the identity, of particles produced at the Tevatron collider over as large a fraction of the solid angle as practical." [52] The first section of this chapter reviews how the Tevatron's proton and antiproton beams are produced and accelerated. The second section describes how the CDF detector measures the properties of the particles arising from the collisions. The present tense will be used in all descriptions. It must be noted, however, that at the time of writing, significant changes have been made to both accelerator and detector in preparation for Run 2. This chapter describes the experimental apparatus used for CDF Run 1.

4.1 Colliding Beams at the Tevatron

The Tevatron's counter-rotating beams of protons (p) and antiprotons ($\bar{p}$) are produced in stages which are briefly described below. An overall view of the Fermilab accelerator complex is provided in Figure 4.1.

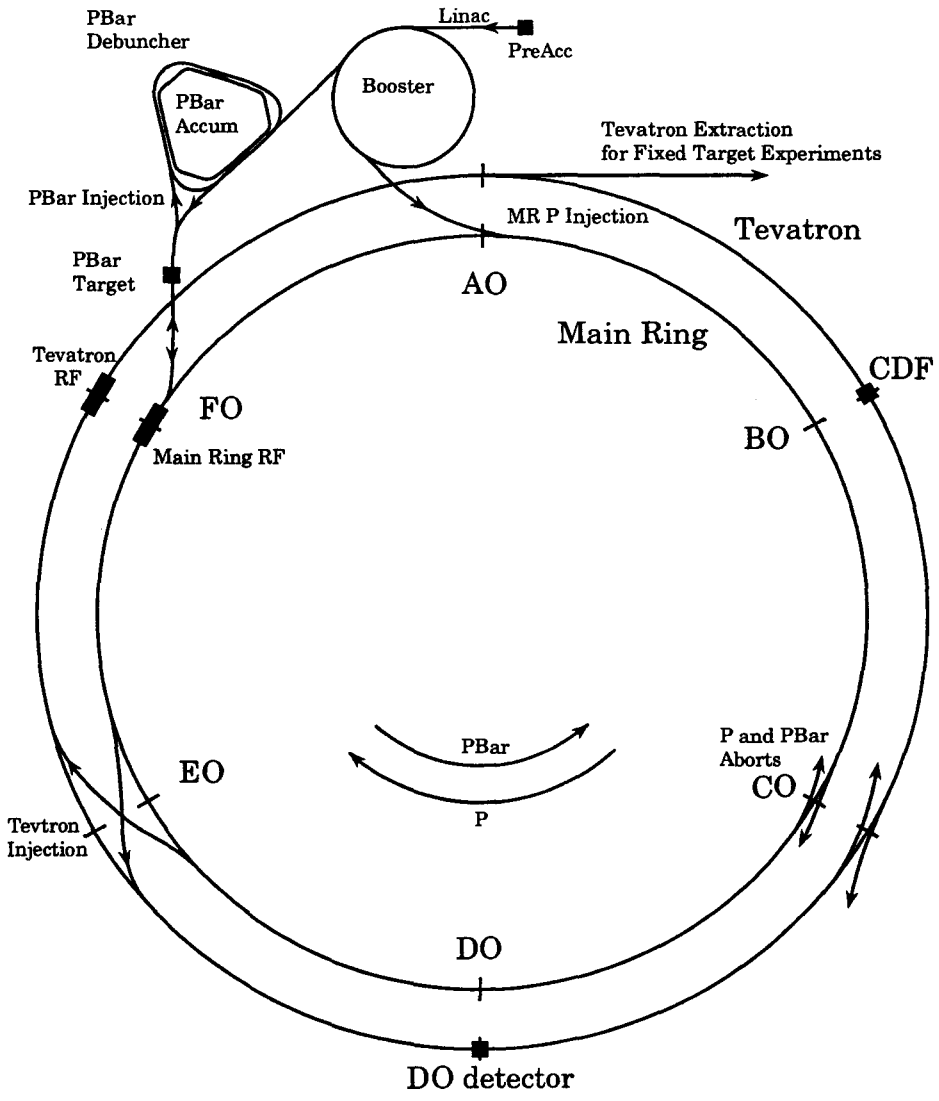


Figure 4.1: Fermilab accelerator complex. For clarity, the Main Ring and Tevatron are shown separated, while in reality they have the same radius.

4.1.1 Preaccelerators

The proton beam starts out as hydrogen gas added to the slender gap between anode and cathode in a magnetron surface-plasma source in the presence of constant electric and magnetic fields [53]. Some hydrogen atoms get adsorbed onto the cathode, while others dissociate into a plasma of electrons and H^+ ions. Several different reactions take place on the cathode surface to produce H^- ions, or atoms of hydrogen containing an extra electron. Plasma particles strike the cathode and sputter off H^- ions, or H^+ ions pick up two electrons from the cathode surface. The H^- ions are electromagnetically extracted through an aperture in the anode and steered into an electrostatic accelerating column. 750 kV of potential difference are maintained between the two ends of the column by a Cockcroft-Walton generator. The resulting 750-keV beam of H^- ions passes into a transport line where its size and focus are adjusted and it is formed into bunches with the same periodicity as that of the electromagnetic field of the linear accelerator (Linac) into which it will be injected.

4.1.2 Linac

The Linac uses a radiofrequency (RF) electromagnetic field to accelerate the H^- ions from 750 keV to 400 MeV. Approximately 150 m long, the Linac consists of a series of electrically resonant cylindrical metal tanks driven by klystron RF sources. A line of drift tubes, separated by gaps, runs down the axis of each tank. Spacing and length of drift tubes, and the frequency of the oscillating RF field, are arranged so that the ion bunches emerge into the gaps when the field is in the direction that will accelerate them, but hide shielded in the tubes when it is in the direction that would slow them down [54]. Furthermore, the phase of the RF field is adjusted so that slower particles receive more energy, and faster particles less, than synchronous particles. Transverse focusing of the beam is achieved by quadrupole magnets.

4.1.3 Booster

The next phase of acceleration takes place in the Booster, a 151-m-diameter synchrotron ring around which accelerating RF cavities are interspersed among orbital bending magnets. 400-MeV H^- ions from the Linac are brought tangent to the ring and sent through a carbon foil to strip off their electrons. The resulting bare protons are steered into the ring while electrons and unstripped ions are directed to a beam dump. After the Booster has been filled with roughly 3×10^{12} protons, the phase of the fields in its RF cavities is adjusted to form 84 stable phase regions known as "buckets". This captures the beam into 84 bunches which are then accelerated to 8 GeV. The acceleration mechanism is the same as that used in the Linac—bunches alternately hide in drift tubes or emerge into gaps and get boosted by the RF field—except the beam is constrained to a circular orbit, and so makes repeated passes through the accelerating cavities. Since the protons' velocity increases with each successive turn, the frequency of the accelerating RF field (and the resonance of the RF cavities) must also be increased with each turn [1]. So must the field strength of the magnets which confine the beam to a circular orbit. Synchronous increase of the RF frequency and magnet strength with particle momentum is the origin of the term "synchrotron". Quadrupole magnets that alternately focus in the x- and y-dimensions control the transverse beam profile.

4.1.4 Main Ring

Like the Booster, the Main Ring is also a proton synchrotron featuring tunable RF cavities and conventional bending and focusing magnets, except it is 2000 m in diameter. Protons are extracted from the Booster into the Main Ring for one of two purposes: either for acceleration to 150 GeV and injection into the Tevatron, or for acceleration to 120 GeV and injection to the Antiproton Source. The first step in either case is to capture the beam

[1] In order for the RF to be in phase with the particle, *i.e.* so that as the particle moves from bucket to bucket it always finds the RF at the same phase, the RF must oscillate at frequency $f = h/\tau$. Here h is the harmonic number of the ring (84) and τ is the revolution period of the particle $= 2\pi R_{ring}/v$. As v increases, so must f increase.

from the Booster and ramp it to the desired final energy ("flattop"). In Tevatron-injection mode, only a fraction of the 84 Booster bunches are used. After acceleration, these are all coalesced into one large bunch and deposited into the desired stable phase location, or RF bucket, in the Tevatron. In antiproton-producing ("pbar stack") mode, all possible Booster bunches are used. After flattop, the bunches are extracted from the Main Ring to the Antiproton Target. During normal Tevatron running, the Main Ring usually continually operates in pbar-stack mode whenever it is not needed for Tevatron injection.

4.1.5 Antiproton Source

Among the advantages of colliding p on $\bar{p}$ is that the same accelerators can be used for the two beams: their opposite charge means they can form separate counter-rotating beams within the same magnets and RF cavities. The disadvantage, however, is the cost and difficulty of making antiprotons. The Antiproton Source performs this task. It is comprised of a target where the particles are produced and two rings for cooling and collecting them into a beam. In the Antiproton Target, the 120-GeV Main Ring beam impinges on the edge of a nickel target disk. A wide variety of secondary particles, including antiprotons, is produced. The secondaries pass into a lithium lens, a cylinder of lithium carrying a large current, whose azimuthal magnetic field focuses the particles into a collimated beam. After the lens, beam particles with negative charge and momentum near 8 GeV are selected by a magnet and directed to the Debuncher ring. As they enter the Debuncher, the bunches of antiprotons are spread out in momenta. The Debuncher's RF fields rotate the bunches in phase space to be spread out in time rather than momentum, and then debunch the beam altogether into a continuous stream. Next the beam undergoes stochastic cooling to reduce its transverse (or betatron) oscillations by a factor of more than two. In stochastic cooling, pickups measure the beam's transverse size at a given point and send this information to a point across the ring so a corrective kick can be administered when the beam arrives at that point. The cooled batch of antiprotons enters the Accumulator ring where it is added to the tail of the accumulating stack of antiprotons and cooled further. When sufficient

antiprotons have accumulated in the stack core, they are extracted and injected to the Main Ring. There they undergo acceleration to 150 GeV, coalescing into one large bunch, and injection to the desired Tevatron RF bucket.

4.1.6 Tevatron

Like the Booster and Main Ring, the Tevatron is a synchrotron accelerator. Located directly under the Main Ring, in the same tunnel and at the same radius, the Tevatron has a very similar lattice of bending magnets as the Main Ring, except they are all superconducting and so can achieve higher field strengths at lower current. The magnets must be kept at liquid-helium temperatures (4.6 K). Under typical operating conditions, the Tevatron is injected with six bunches of protons (each containing $\sim 150 \times 10^9$ particles) followed by six counterrotating bunches of antiprotons (each containing $\sim 50 \times 10^9$ particles.) The bunches are roughly evenly spaced around the 1113 RF buckets of the ring and are separated by empty buckets. After both beams have been accelerated to their final energy of 900 GeV, they are permitted to continue circulating the ring for tens of hours for a colliding-beam session or "store". At two points around the ring, known as B0 and D0, the counterrotating beams undergo transverse squeezing by low-beta quadrupole magnets to $\sim 40\mu$m luminous spots and are directed head-on into each other. Bunch crossings occur at B0 and D0 approximately every 3.5 μsec. Thus are $p\bar{p}$ collisions at center-of-mass energy ($\sqrt{s}$) 1.8 TeV brought about.

4.2 The CDF Detector

A natural set of phase-space coordinates for describing the relativistic kinematics of $p\bar{p}$ collisions is provided by azimuthal angle ϕ, rapidity y, and the component of momentum transverse to the beamline, p_T [52, 55]. Ref. [55] explains the convenience of transverse momentum and rapidity for describing relativistic collisions as follows:

With these variables the configuration of particles in a collision process viewed in the laboratory frame differs only by a trivial shift of the origin of rapidity from the same process viewed in the center of mass frame.

Azimuthal angle ϕ is defined with respect to the x axis of CDF's coordinate system, as shown in Figure 4.2. Rapidity y is defined according to

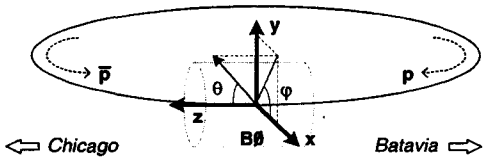


Figure 4.2: CDF coordinate system, with azimuthal (ϕ) and polar (θ) angles shown. The "Chicago" arrow corresponds to positive η, *i.e.* $\eta \to +\infty$, while the "Batavia" arrow corresponds to negative η, *i.e.* $\eta \to -\infty$.

$$y = \frac{1}{2} \ln \frac{E + p_{||}}{E - p_{||}}$$

where $p_{||}$ is the component of momentum lying along the beam direction. For convenience, the assumption of near-masslessness (valid at high energy) $E \simeq |\vec{p}|$ is often made at CDF, and the quantity pseudorapidity η substituted for rapidity:

$$\eta = \frac{1}{2} \ln \frac{p + p_{||}}{p - p_{||}}$$

Equivalently, pseudorapidity may be written as $-\ln(\tan\theta/2)$, where θ is the polar angle with respect to the proton-beam direction. The transverse component of momentum, p_T, is defined as $p \sin\theta$.

Reflecting this natural set of coordinates, the CDF detector surrounds the B0 $p\bar{p}$ collision point with approximate forward-backward and azimuthal symmetry. In structure the detector is akin to a cylindrical onion, composed of layered subsystems which perform the tasks of measuring the momentum, energy, and other properties of particles produced in the collisions. The first solid material encountered by a particle moving radially outward from the collision point is the wall of the 3.8-cm-diameter beampipe [56] which keeps the beam

volume under vacuum. To minimize multiple scattering [2], the beampipe walls are made of thin, ultra-light beryllium metal. Subsequently, the particle passes through successive layers of tracking, calorimetry, and muon subsystems. These subsystems are shown in one-quarter view in Figure 4.3. Each will be described in a separate section below.

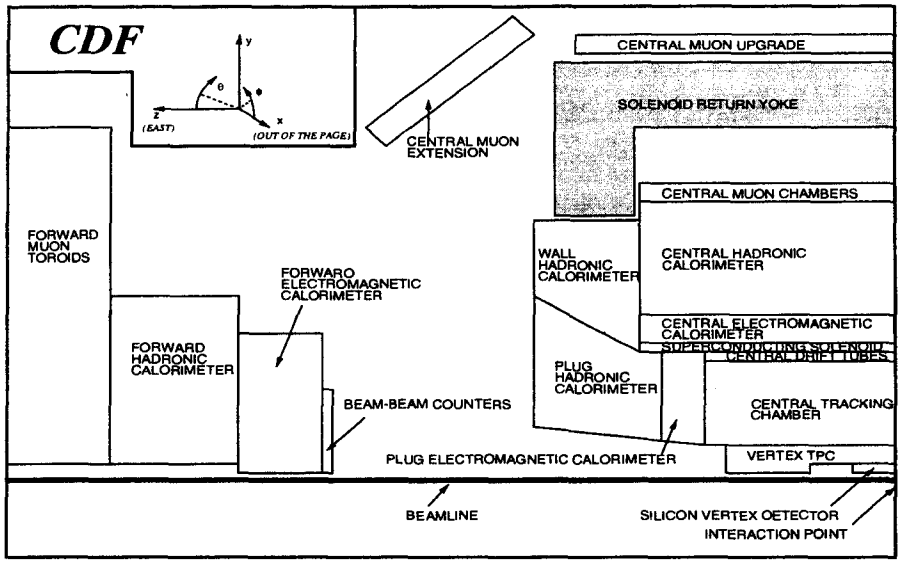


Figure 4.3: Quarter view showing major subsystems of the CDF detector. The detector is forward-backward symmetric across the transverse plane intersecting the collision point and azimuthally symmetric about the beam axis.

4.2.1 Tracking

After passing through the beampipe, a particle enters CDF's tracking systems. These consist of a precision silicon vertex detector, a vertex time projection chamber, and a large drift-chamber tracker immersed in a 1.4 T magnetic field. From measurements made in the tracking chambers, the three-dimensional trajectories of charged particles can be reconstructed. These trajectories are helical in geometry and are completely specified by five parameters: C (half-curvature), $\cot\theta$ (dip angle), z_0 (z-value of closest approach

[2] A charged particle traversing a medium will experience many small-angle deflections from Coulomb scattering off of the nuclei of the medium [57].

to beam axis), d_0 (transverse distance of closest approach, also called *impact parameter*), and ϕ_0 (azimuthal angle at point of closest approach). From reconstructed tracks, much information about the particles can be deduced, including momentum, charge, correlation with energy deposition in the calorimeters or hits in muon chambers, and the presence of *displaced vertices*, the hallmark of heavy-flavor quarks such as bottom and charm.

4.2.1.1 Silicon vertex detector

The first tracking system encountered by an outgoing particle is CDF's Silicon Vertex Detector (SVX) [58, 59]. The SVX provides ultra-precise tracking measurements in the $r - \phi$, or transverse, plane. The excellent resolution with which the SVX measures impact parameter d_0, in particular, is crucial to identifying tracks that reconstruct to displaced vertices. SVX information on secondary vertices is essential to the bottom-quark-identifying ("b-tagging") algorithms used in this analysis. Between CDF Run 1A and Run 1B, the original SVX was replaced by an improved model. The SVX′ differed from its predecessor in its use of a radiation-resistant readout chip, AC- rather than DC-coupled silicon strips, and a few minor geometrical improvements. In all other respects, the devices were similar. In subsequent discussion we refer to either of the devices as "SVX".

Centered on the nominal interaction point, the SVX consists of two identical 25.5-cm cylindrical "barrels" placed coaxial with the beam, separated from each other by 2.15 cm. The z coordinate of $p\bar{p}$ interaction vertices at CDF forms a Gaussian-like distribution around $z = 0$ with $\sigma = 30$ cm due to the length of the p and $\bar{p}$ bunches. The SVX's 51 cm-long active region was chosen as a compromise between need to cover this relatively long interaction region *vs.* the capacitive and leakage-current burden that long strips impose on the readout electronics. The chosen length covers approximately 60% of $p\bar{p}$ collision vertices and extends to $|\eta| < 1.9$. Each barrel consists of four concentric layers of silicon strip detectors. The layers, numbered $0 - 3$, range in radii from 2.9 to 7.9 cm. A layer, in turn, is composed of twelve flat rectangular panels known as "ladders" arranged parallel to the beam axis in a twelve-sided cylinder. An individual ladder is composed of three electrically-connected

silicon crystals held end-to-end in a lightweight frame. Ladder width increases with layer radius to keep ϕ-coverage constant. Each group of four ladders at constant azimuth forms a "wedge" which points back to the beamline. A diagram of one barrel is shown in Figure 4.4.

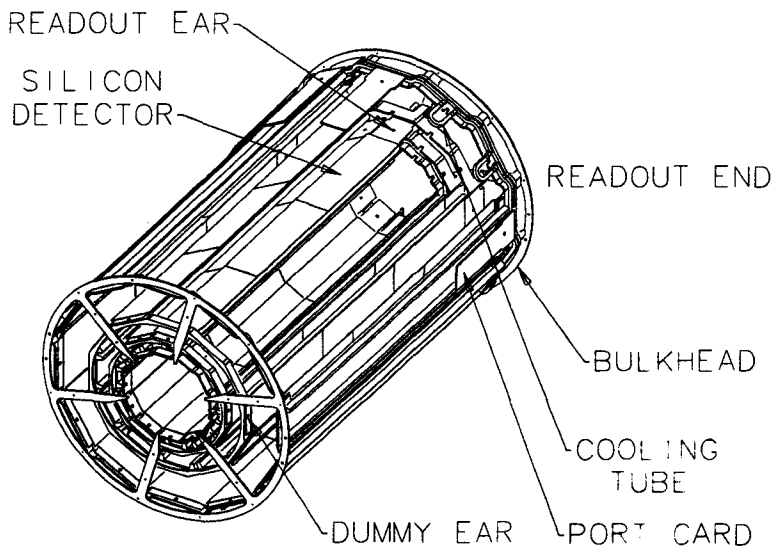


Figure 4.4: One barrel of the CDF silicon vertex detector (SVX). The four layers, each with its twelve ladders, can be seen.

The smallest subunit, or channel, of the SVX detector is a microstrip etched onto the three crystals of silicon in a ladder. Strips run parallel to the beam; the strip spacing, or pitch, is 60 μm on layers $0-2$ and 55 μm on layer 3. Each strip operates as a reverse-biased $p-n$ junction. When a charged particle passes through the depletion region of the silicon, it liberates many electron-hole pairs. These freed charge carriers move under the influence of the applied electric field and accumulate on the electrodes, where they are detected by a charge-integrating readout chip [60]. Under normal running conditions, the chips operate in "sparse" mode where only channels above a hardware threshold are read out. The SVX

readout chips are mounted on small circuit boards called readout hybrids located on the ends of each ladder.

Offline tracking algorithms associate raw charge levels of adjacent strips to *clusters*, which can be linked to track segments found in the outer tracking chamber. A cluster is defined as a contiguous group of strips whose pedestal-subtracted pulse height q is greater than a factor M times their noise σ, where M depends on the number of strips in the cluster. Cluster positions are defined as charge-weighted centroids: $x = \sum x_i q_i / \sum q_i$.

The SVX impact parameter resolution for high-momentum tracks is 13 μm [3].

4.2.1.2 Vertex time projection chamber

Event vertices at CDF can be located anywhere within the long luminous region, which was described above as centered at $z = 0$ but distributed in Gaussian fashion with $\sigma = 30$ cm. Accurate track reconstruction demands that the precise z-vertex of each event be determined. The vertex time projection chamber (VTX) provides the $r - z$ view tracking information necessary to this task.

Outside the SVX, extending in radius to 22 cm and in pseudorapidity from $-3.25 < |\eta| < 3.25$ [61], the VTX consists of twenty-eight time-projection chambers mounted end-to-end along the z-axis [62]. Like the SVX, the VTX is fabricated of lightweight materials to minimize multiple scattering [63]. Each time projection chamber or "module" of the VTX is a hollow octagonal cylinder, lying coaxial with the beam. At either end of the module are aluminum grids charged to a large negative voltage [64]. Separating each module into two drift regions is a central cathode plane charged to a lesser negative voltage. On either side of the central plane are strung sense wires at ground potential, followed by cathode grids. Figure 4.5 illustrates the chamber construction of the VTX.

All chambers are filled with a 50/50 mixture of argon and ethane gas. When a charged particle passes through a chamber, it ionizes the gas, liberating electrons. The electrons drift at constant speed away from the high-voltage grids and toward the central

[3] Compare this to the ~ 450 μm impact parameter resolution of the central tracking chamber alone [58].

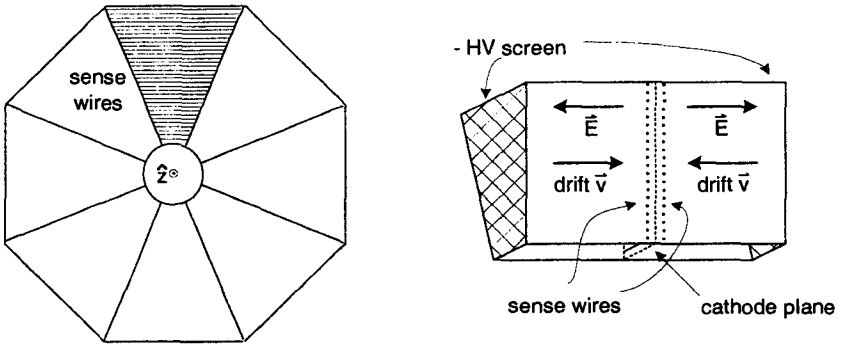


Figure 4.5: Cross-sectional view of the VTX; sketch of one sector of a time projection chamber.

sandwich of sense-wire and cathode planes, which acts as a multiwire proportional counter. As the drift electrons accelerate to the anode wires they ionize other gas molecules, resulting in an avalanche of electrons and a detectable pulse on the wire. A signal is also produced on pads etched onto the adjacent cathode plane. Signals from each pad and sense wire are separately amplified and read out. From the arrival times of electrons on the sense wires—and the known polynomial dependence of the drift velocity on time [65]—the $r-z$ (sideways) view of the particle's trajectory can be reconstructed. The z-vertex is measured with resolution 1 mm [61].

4.2.1.3 Central tracking chamber

While the SVX and VTX perform the complementary tasks of tracking in the $r-\phi$ and $r-z$ planes, CDF's Central Tracking Chamber (CTC) is capable of standalone three-dimensional track reconstruction [66]. Essential functions of the CTC for this analysis include measuring the momentum of charged particles, providing tracks to match calorimeter or muon-chamber information for lepton identification, and finding the tracks used for b tagging jets.

A particle of charge e moving in a uniform, static magnetic field $\vec{B}$ with velocity-component $v_{||}$ along the direction of $\vec{B}$ and $v_{\perp}$ perpendicular to $\vec{B}$ follows a helical path of radius $a = cp_{\perp}/eB$ and pitch angle $\alpha = \tan^{-1} v_{||}/v_{\perp}$ [67]. The relationship between

transverse momentum $p_\perp$ (or p_T as it is known on CDF [4]), radius of curvature a, and field strength B can be rewritten as

$$p_T \text{ (GeV/c)} = 3.00 \times 10^{-4} Ba \text{ (kG} - \text{cm)}^\dagger$$

Thus in a magnetic field $\vec{B} = B\hat{z}$, p_T can be directly derived from the radius of curvature a of the charged particle's track. To permit this p_T measurement, the CTC is surrounded by a superconducting solenoid of length 4.8 m and radius 1.5 m [56]. Inside the CTC volume, the solenoid provides uniform magnetic field $\vec{B} = 14.116$ kG $\hat{z}$. Outside the solenoid, the magnetic field is largely confined to a freestanding iron yoke which doubles as a flux-return path and a structural support member for other portions of the detector.

The CTC is a cylindrical drift chamber, strung with 84 layers of sense wires and filled with a 50/50 mix of argon/ethane gas with a small admixture of alcohol [68]. As in the VTX, an incoming charged particle ionizes the gas, liberating electrons, which then drift at approximately constant speed to the positively-charged sense wires and cause a detectable pulse. The CTC's active region extends from 31 cm to 132 cm in radius and 3.2 m in length, corresponding to the pseudorapidity range $|\eta| \leq 1$.

The CTC's 84 layers of sense wires are arranged in nine sub-groupings called "superlayers". In five superlayers, wires run parallel to the beam axis. Alternating with these *axial* superlayers are four *stereo* superlayers in which wires are canted at an angle of $+3°$ or $-3°$ with respect to the beam axis. Axial layers provide $r - \phi$ information about tracks, while stereo layers provide $r - z$. Within a superlayer, wires are arranged in parallel planes of twelve wires for an axial superlayer and parallel planes of six wires for a stereo superlayer. For both axial and stereo superlayers, the wire planes are tilted at an angle of approximately 45° with respect to the radial direction. A transverse view of the nine superlayers is shown in Figure 4.6. Field wire planes, not visible in the figure, are interleaved between every adjacent pair of sense wire planes.

[4]Recall that T stands for "transverse"—to the beamline, and to $\vec{B}$.

[4†]Jackson, p. 582, slightly modified. In terms of track-parameter "half-curvature" $C = 1/2a$, p_T (GeV/c) $= 0.00015B/C$ (kG $-$ cm). This is the origin of the formula in CDF tracking routines (c.f., for example, `QTRKUN.CDF`) `PT = 0.0001498980 * BMAGNT/ABS(TRKSFP(CUFP))`.

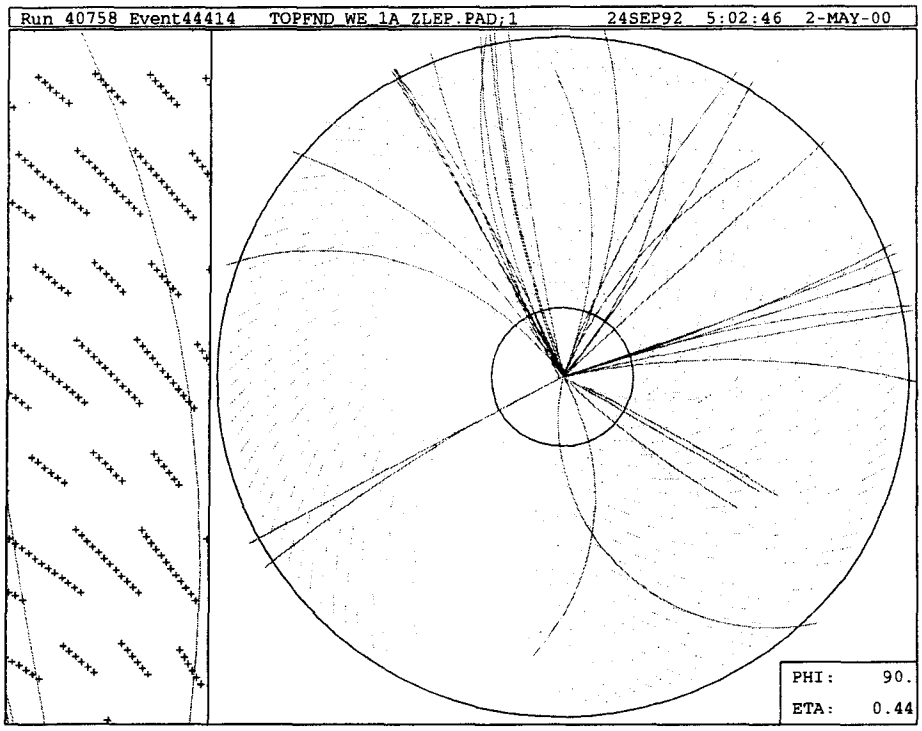


Figure 4.6: DF event display showing $r - \phi$ view of the CTC for a Run 1 event. Superlayer structure can be discerned. Close-up side window highlights the axial planes of twelve sense wires and the stereo planes of six sense wires (each appears as a small "+".)

The CTC's superlayered structure subdivides the tracking volume into multiwire, azimuthally-overlapping cells, where one cell is defined by a pair of field wire planes and their intervening sense wire plane. The number of cells per superlayer was chosen to provide approximately equal drift distances in all superlayers. The geometry of one drift cell is shown in more detail in Figure 4.7. The purpose of subdividing the tracking volume into drift cells is to limit the drift time required for electrons to reach a sense wire and to permit many measurements to be made for each track. The reason the cells are tilted with respect to the radial $\hat{r}$ is that in the presence of crossed electric and magnetic fields (drift field $\vec{E}$ and solenoid field $\vec{B}$), electrons drift at Lorentz angle β with respect to $\vec{E}$:

$$\tan\beta = \frac{v(E, B = 0)B}{kE} \quad \text{(MKS units)}$$

where $v(E, B = 0)$ would be the speed of the electrons if $\vec{B} = 0$. For ease in track recon-

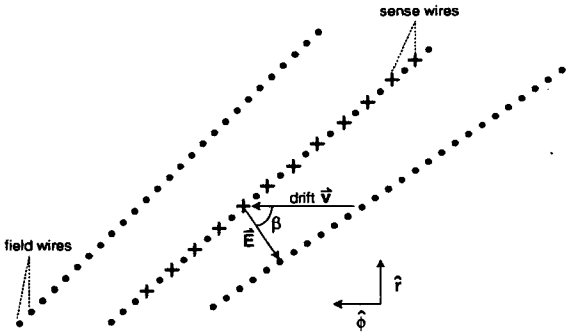


Figure 4.7: Geometry of one axial drift cell of the CTC. Note that the direction of the electric field is approximately 45° with respect to the radial and that the drift direction is roughly azimuthal.

struction, it is desirable that electrons should drift in the azimuthal ($\hat{\phi}$) direction. Tilting the cells by $(90° - \beta)$ with respect to $\hat{r}$ ensures that in spite of the Lorentz-angle effect, the electrons drift in the azimuthal direction $\hat{\phi}$ (see Figure 4.7.) Another advantage of tilting the drift cells is that it causes the cells to overlap in ϕ, guaranteeing that every high-p_T track must pass close to at least one sense wire per superlayer. Tilted drift cells also reduce ambiguity in hit-pattern recognition.

Each sense wire is connected directly to a preamplifier chip mounted on the chamber. From the preamplifier, the signals travel to an amplifier-shaper-discriminator (ASD) card mounted on the magnet yoke and then to a commercial LeCroy 1879 FASTBUS time-to-digital converter (TDC) in the CDF counting room.

Offline track reconstruction at CDF relies most heavily on CTC information. The first step is to look for segments, or patterns of hits, in the axial superlayers of the CTC. Next, segments are linked across superlayers to form two-dimensional $(r - \phi)$ tracks. The poorer resolution of the stereo superlayers (six sense wires *vs.* twelve) precludes segment-finding in the stereo layers. Instead, the two-dimensional tracks are made three-dimensional using z-vertex information from the VTX and attaching hits from the stereo superlayers [66]. The helix parameters of the resulting three-dimensional tracks are then extrapolated into the SVX, defining a "road" in which SVX hits (clusters) are searched for. Starting at the

outermost layer and moving inward, SVX hits are attached to the track layer by layer. At each layer, the helix parameters are recalculated using the new information. The track candidate with the largest number of clusters and the best χ^2 is selected [58].

In terms of performance, the CTC is highly efficient at reconstructing tracks from particles with momenta down to 300 MeV/c^2. Momentum resolution [5] is $\delta p_T/{p_T}^2 = 0.002$ $(\mathrm{GeV/c})^{-1}$. When beam-constraint or SVX information is included, the resolution improves to $\delta p_T/{p_T}^2 = 0.001$ $(\mathrm{GeV/c})^{-1}$ for high-p_T particles [66].

4.2.2 Calorimetry

After passing through the tracking volume, a radially-outgoing particle enters CDF's calorimeters, which measure its energy and position. CDF's calorimetry is divided into three separate η-regions: "central", "plug", and "forward", which together cover the range $-4.2 \leq \eta \leq 4.2$. Each η-region is separately instrumented with electromagnetic and hadronic calorimeters. The former are specialized for detecting electrons and photons and the latter for detecting strongly-interacting particles.

The task of the calorimetry is to reconstruct the energy flow in the largest possible fraction of the solid angle, with application to reconstruction of jets, identification of leptons, and reconstruction of the missing transverse energy associated with neutrinos or other noninteracting particles. "Transverse energy" E_T, in analogy with transverse momentum p_T, is defined as $E \sin\theta$. Conservation of momentum demands that the p_T of all particles in an event must sum to zero. On the assumption of near-masslessness, valid for highly relativistic particles, $E \simeq |\vec{p}|$, so $E_T \simeq p_T$, and the zero-sum relation is taken to hold for E_T as well. Thus the signature of an "invisible" particle such as a neutrino is an imbalance in E_T. Missing E_T ($\not\!\!E_T$) is defined as $-\sum_i E_{Ti}\hat{n}_i$, where the sum runs over calorimeter cells and $\hat{n}_i$ is a two-dimensional unit vector that points from the event vertex to the ith calorimeter cell [69].

[5] Note that the relative error on p_T, $\delta p_T/p_T = 0.002(\mathrm{GeV/c})^{-1} p_T$, increases linearly with p_T. For example, the relative error on a 5 GeV track would be $0.002(\mathrm{GeV/c})^{-1} \times 5$ GeV/c = 1%, while the relative error on a 50 GeV track would be $0.002(\mathrm{GeV/c})^{-1} \times 50$ GeV/c = 10%.

The different regions of CDF's calorimeters are used differently in this analysis. On account of their superior energy resolution and complete overlap with the tracking systems, the central calorimeters are used in jet reconstruction, lepton identification, and $\not\!\!E_T$ calculation. The endplug and forward-backward calorimeters miss the fiducial η-reach of the tracking, and so are not used for lepton identification, but are used to measure jets and $\not\!\!E_T$.

The principle of operation for either an electromagnetic or hadronic calorimeter is to cause the incident particle to undergo a cascade of interactions or *shower* in which it loses all its energy to the calorimeter material [70]. This energy deposition is then detected. All of CDF's calorimeters are *sampling* calorimeters, which means the tasks of passive particle absorption and active signal generation are separated. In all CDF's electromagnetic calorimeters, the passive absorber is lead; in the hadronic calorimeters, it is iron. In both cases, layers of the absorber are sandwiched between layers of active detection medium. In the central region, plastic scintillator is used as the active medium; in the plug and forward regions, gas proportional chambers are used. The term "sampling" reflects the fact that only the fraction of particle energy seen by the active detection material is measured.

In general, the energy resolution of a sampling calorimeter obeys the form $\sigma/E = \sigma_1/\sqrt{E} \oplus \sigma_2$, where the $\oplus$ symbol indicates that the terms are to be added in quadrature. Statistical effects, whether from sampling or intrinsic fluctuations in shower development, give rise to the $1/\sqrt{E}$ term, while the constant term comes from nonuniformities in the calorimeter. The $1/\sqrt{E}$ term is generally larger for a hadronic calorimeter. That energy resolution *improves* with increasing E reflects the statistical origin of the $1/\sqrt{E}$ term. Crudely speaking, the larger E, the larger the number N of shower particles that can be produced, and the less important the $\sqrt{N}$ statistical fluctuations in the production and detection of those shower particles [71]. By the same token, the smaller the mass of the shower particles, the larger the number N that can be produced for a shower of given E. This explains why hadronic calorimeters, in which the shower particles are primarily pions and nucleons, tend to have worse energy resolution than electromagnetic calorimeters, in

which the shower particles are primarily electrons. Since $m_\pi >> m_e$, for a fixed E the number N of particles that can be produced in a hadronic shower is much smaller than N for an equally energetic electromagnetic shower. Smaller N means worse statistical fluctuations in producing and detecting the showers, hence worse overall resolution [71].

In an electromagnetic shower, such as that produced by an electron (e) or photon (γ), the dominant mechanisms of energy loss are bremsstrahlung, in which an electron interacts with a nearby nucleus and emits a photon, and pair production ($\gamma\gamma \rightarrow e^+e^-$). The daughter particles produced in these processes are again e^+, e^-, and γ; the interactions occur repeatedly until the average energy per particle is low enough to stop further multiplication [6], a point known as *shower maximum*. Electromagnetic showers are typically compact and do not fluctuate widely in shape. A convenient metric for describing the longitudinal extent of an electromagnetic shower is the *radiation length* X_0. X_0 is defined to be "the mean distance over which a high-energy electron loses all but $1/e$ of its incident energy by bremsstrahlung" [72] and depends on the atomic number Z and atomic weight A of the calorimeter material. Thus the shower length scales as the radiation length of the material.

A hadronic shower proceeds by a succession of inelastic nuclear interactions including multiparticle production and emission from excited nuclei. Secondary particles produced are mostly pions and nucleons. The longitudinal extent of a hadronic shower scales with nuclear interaction length λ, which is defined as the mean free path of a particle before undergoing an inelastic collision. Hadronic showers tend to be much larger in extent than electromagnetic showers. They also display much wider fluctuations in shape. In addition to the relatively large mass of the shower particles, other factors contributing to the larger intrinsic fluctuations in hadronic showers include fluctuations in the fraction of neutral pions (π^0), since the π^0's decay to $\gamma\gamma$ without further nuclear interactions, and the fact that processes such as the excitation and breakup of nuclei can vary widely in the amount

[6] Radiative energy loss scales as E of the particle, while ionization energy loss scales as $\log E$. At the critical energy E_c the two are equal, and shower multiplication stops. Instead, the particles lose their remaining energy through ionization and excitation [72].

of detectable signal produced. The net effect of these fluctuations is to limit the energy resolution that can be achieved with a hadronic calorimeter.

The relative positioning and thickness of the electromagnetic and hadronic portions of the calorimeter are chosen to maximize discrimination between the two kinds of showers. The goal is to contain the electromagnetic showers over a short distance but to keep hadronic showers from developing until they reach the hadronic portion of the calorimeter. The figure of merit to maximize is the ratio (interaction length/radiation length) λ/X_0. To this end, the electromagnetic layers of the calorimeter are fabricated of high-Z materials (hence the choice of lead as absorber) and placed closer to the interaction point than the hadronic.

Reflecting the natural phase-space coordinates (η, ϕ, p_T) of the detector, CDF's calorimeters are segmented into $\eta - \phi$ towers that project back to the interaction point (Figure 4.8). Each tower spans approximately 0.11 units in η. Tower ϕ-segmentation is

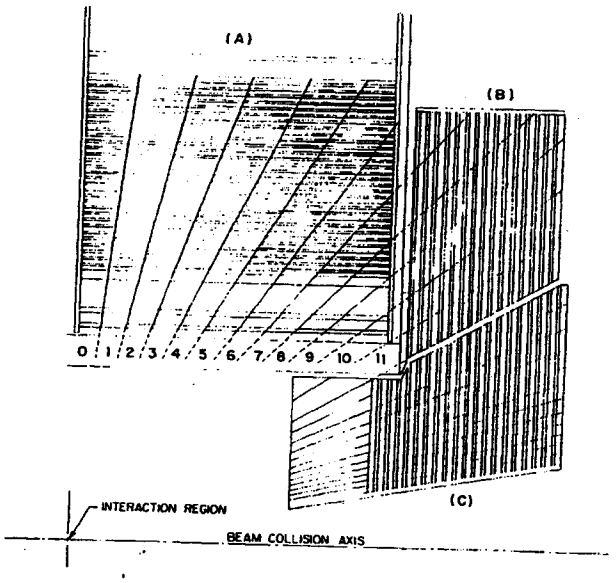


Figure 4.8: $r - \theta$ or "sideways" one-quarter view of the central (A), endwall (B), and plug (C) calorimeters illustrating the projective tower geometry. Note that towers can extend in η across modules.

15° in the central region and 5° in the plug and forward regions. Figure 4.9 illustrates the mapping of towers onto solid angle out to $|\eta| < 4.2$. Table 4.1 summarizes the properties

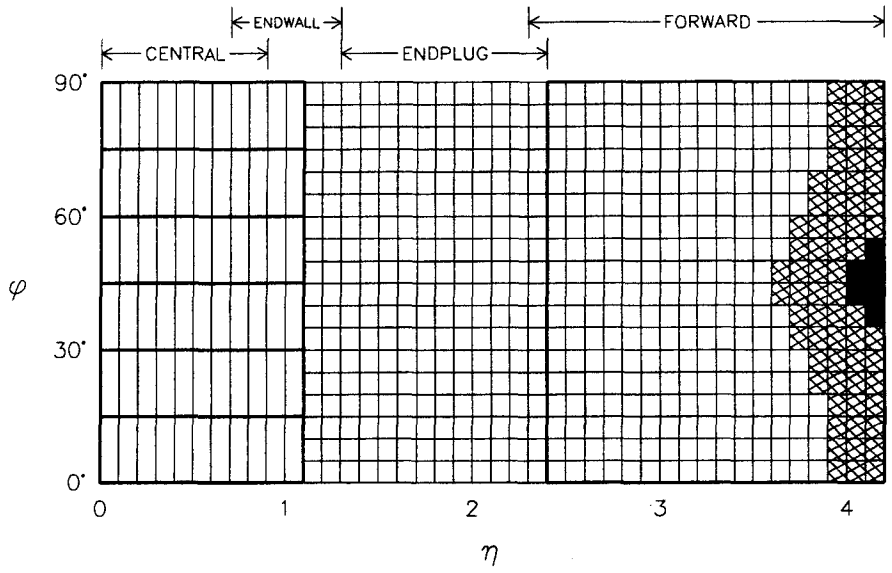


Figure 4.9: Hadron calorimeter tower segmentation in one of the eight identical $\eta - \phi$ quadrants of the detector ($0 < \phi < 90°, 0 < \eta < 4.2$). Shaded or blacked-out towers indicate regions of partial or no coverage. The corresponding electromagnetic towers have full coverage out to $\eta < 4.2$.

of CDF's calorimeters.

4.2.2.1 Central calorimetry

CDF's central calorimeters form a cylindrical annulus that can be separated into four freestanding C-shaped arches for maintenance. The inner surface of arches captures the tracking solenoid. Each arch is made up of twelve 15° wedges. The radially inner portion of each wedge is instrumented with electromagnetic calorimetry and the outer portion with hadronic. The electromagnetic layers of the wedge consist of thirty-one slabs of 5 mm thick polystyrene scintillator interleaved with thirty slabs of 1/8" thick lead [73]. As electromagnetic shower particles pass through the scintillator, it gives off blue light. Acrylic waveshifter sheets flanking the azimuthal sides of each tower collect the light and redirect it into a pair

System	"TLA"	η Range	Energy Resolution	Thickness
CENTRAL EM	CEM	$\lvert\eta\rvert < 1.1$	$13.7\%/\sqrt{E_T} \oplus 2\%$	$18X_0$
PLUG EM	PEM	$1.1 < \lvert\eta\rvert < 2.4$	$22\%/\sqrt{E} \oplus 2\%$	$18 - 21X_0$
FORWARD EM	FEM	$2.2 < \lvert\eta\rvert < 4.2$	$26\%/\sqrt{E} \oplus 2\%$	$25X_0$
CENTRAL HAD	CHA	$\lvert\eta\rvert < 0.9$	$50\%/\sqrt{E_T} \oplus 3\%$	$4.5\lambda_0$
ENDWALL HAD	WHA	$0.7 < \lvert\eta\rvert < 1.3$	$75\%/\sqrt{E} \oplus 4\%$	$4.5\lambda_0$
PLUG HAD	PHA	$1.3 < \lvert\eta\rvert < 2.4$	$106\%/\sqrt{E} \oplus 6\%$	$5.7\lambda_0$
FORWARD HAD	FHA	$2.4 < \lvert\eta\rvert < 4.2$	$137\%/\sqrt{E} \oplus 3\%$	$7.7\lambda_0$

Table 4.1: Summary of CDF calorimeter properties. Energy resolutions for the electromagnetic calorimeters are for incident electrons or photons, while resolutions for the hadronic calorimeters are for incident isolated pions. Energies are given in GeV. Thickness are given in radiation lengths X_0 for the electromagnetic calorimeters and interaction lengths λ_0 for the hadronic [56].

of acrylic-rod lightguides connected to photomultiplier tubes. The electromagnetic towers and light-collection apparatus for one wedge are shown in Figure 4.10.

In addition to the lead-scintillator layers, the central electromagnetic calorimeter also contains two layers of proportional chambers. Preshower wire chambers (CPR) mounted on the front of the wedges help identify photons and soft electrons that begin showering in the tracking material and solenoid coil even before they reach the lead [74]. Proportional strip chambers (CES) sandwiched between the eighth layer of lead and the ninth layer of scintillator in each wedge help determine shower position and transverse shower development. The location of the CES chambers in the lead-scintillator stack corresponds to the average depth of shower maximum. One CES chamber covers all the ten towers in the wedge. In each chamber, sixty-two parallel anode wires run along the azimuthal or x-dimension while 128 orthogonal cathode strips run along the polar or z-dimension. The CES achieves spatial resolution in x and z on the order of a few mm [75].

The central hadronic calorimeter consists of the hadronic layers of the wedges (CHA) plus additional endwall modules (WHA) mounted on the solenoid flux return yoke. As can be seen in Figure 4.8, in one CHA/WHA module pair, six towers are fully in the CHA, three are shared CHA/WHA, and three are fully contained in the WHA. The CHA consists

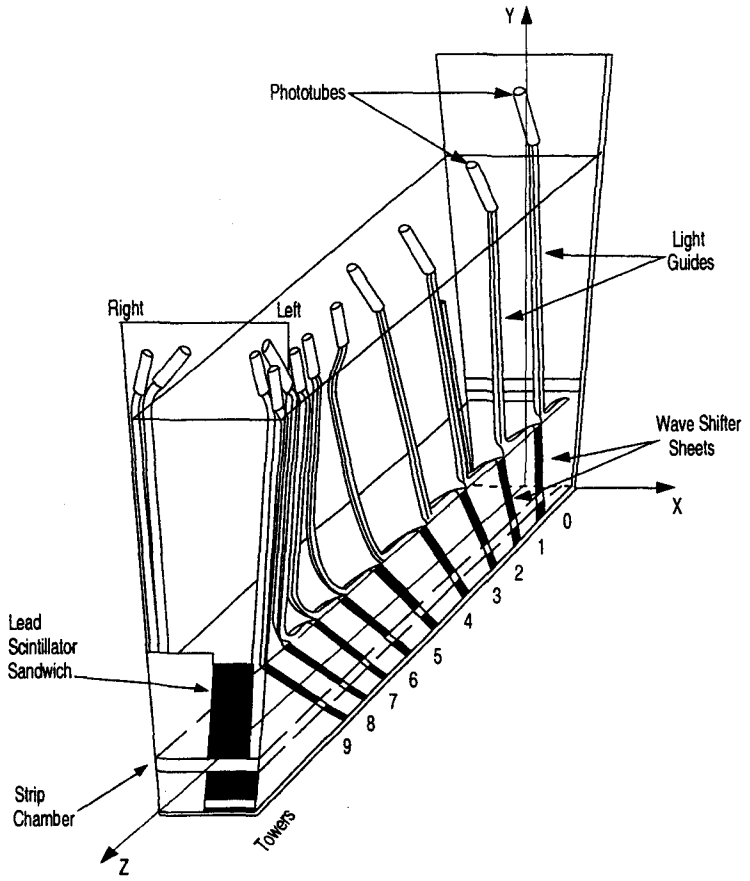


Figure 4.10: One wedge module of the central calorimeter. Electromagnetic instrumentation, including light collection apparatus, is shown. The EM portion of the instrumentation for one wedge weighs 2 tons; the hadronic portion weighs 13.2 tons.

of thirty-two interleaved layers of 2.5 cm steel and 1 cm scintillator [76]; the WHA consists of fifteen interleaved layers of 5 cm steel and 1 cm scintillator. Wavelength shifter strips lying alongside the long sides of each scintillator slab collect the light and redirect it to a pair of clear UVA light guides, one on either azimuthal side of the tower. The light guides from each sheet in a tower run up the ϕ-sides of the tower and into a pair of photomultiplier tubes.

Readout and digitization of the photomultiplier tubes for both the electromagnetic and hadronic central calorimeters are performed by a system of custom-designed circuit

boards and crates collectively known as the RABBIT system (Redundant Analog Bus Based Information Transfer).

4.2.2.2 Plug calorimetry

CDF's endplug calorimeters fit like corks into either end of the tube-shaped central calorimetry. Like the central calorimeter wedges, each endplug calorimeter consists of layers of active sampling medium sandwiched between layers of passive absorber. The stack of layers begins as a cylinder of diameter 2.8 m [77] and flares into a frustrum, as can be seen in Figure 4.8. A concentric conical hole of opening angle 10° penetrates the stack. The inner layers of the stack, *i.e.* the base of the cylinder, are instrumented for detection of electromagnetic showers. Thirty-four 2.7 mm layers of a lead alloy serve as absorber for the electromagnetic plug calorimeter (PEM). The outer layers of the cylindrical portion of the plug, along with the entire "frustrum" portion, are instrumented for hadronic showers. The absorber layers in the PHA are twenty 5 cm-thick iron plates [78]. The PEM covers polar angles between 36° and 10°; the PHA, between 30° and 10°, corresponding to the approximate η-range $1 < |\eta| < 2.4$.

For both the PEM and PHA, the active sampling medium was chosen to be gas proportional chambers with anode-wire and cathode-pad readout. The chief advantage of gas over scintillator in the $|\eta| > 1$ region is ease in constructing towers of constant η-segmentation. As $|\eta|$ increases, towers of constant $\Delta\eta$ subtend ever smaller polar angles; such small towers are easier to implement in stacks of cathode pads rather than stacks of scintillator with their associated light-collection apparatus. The radiation-hardness of gas proportional chambers is an additional advantage given the plug calorimetry's proximity to the beamline.

The proportional chambers themselves are high-resistivity conductive plastic tubes of 7 mm square cross section, strung with anode wires and filled with argon/ethane gas. In a given sampling layer, tubes are arranged side by side in a single flat layer and sandwiched between printed circuit boards onto which cathode pads have been etched. Avalanche signals

produced by shower particles are detectable both on the anode wires of the chambers and on the cathode-pad pickups abutting the tubes. Pads are segmented by constant $\Delta\eta \simeq 0.09$ and $\Delta\phi = 5°$ so that projective towers form when the layers are stacked. Figure 4.11 illustrates the pad patterns on one sampling layer of the PEM. Pads at the same (η, ϕ) are

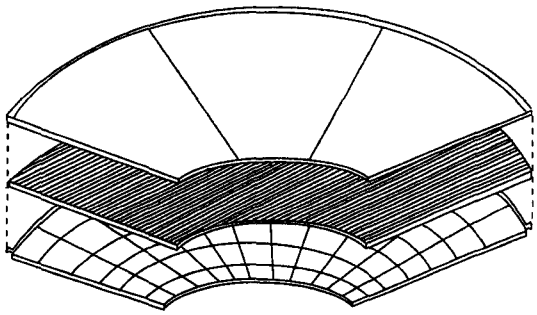


Figure 4.11: Exploded view of one quadrant of a sampling layer of the plug electromagnetic calorimeter. Gas proportional-tube layer is shown sandwiched between PC boards, one with etched pad-patterns and one serving as ground plane. Pad pattern is not to scale!

longitudinally ganged together to form projective $\eta - \phi$ towers. In the PEM, the thirty-four pads at a given (η, ϕ) are grouped to form three longitudinal segments of thickness five, twenty-four, and five layers respectively, whereas in the PHA all twenty layers are grouped into a single longitudinal segment. In addition to these cathode pads, the printed circuit boards on layers 6 through 15 of the PEM are also etched with cathode strips of finer segmentation. Even-numbered layers feature arc-shaped θ-strips ($\Delta\eta = 0.02$) while odd-numbered layers have spoke-shaped ϕ-strips ($\Delta\phi = 1°$). Anode wires, meanwhile, are ganged by plane and azimuthal sector (90° in the PEM, 30° in the PHA) to provide information about longitudinal shower development.

4.2.2.3 Forward calorimetry

CDF's forward/backward calorimeters are two identical systems placed fore and aft of the detector to cover the small-angle regions $2° < \theta < 10°$ and $170° < \theta < 178°$. (The term "forward" is used to refer to either the forward or backward system.) The forward

electromagnetic calorimeter (FEM) is located 6.5 m [79] from the nominal interaction point with the forward hadronic calorimeter (FHA) immediately behind it. Both the FEM and FHA are sandwich counters, the FEM using thirty layers of lead and the FHA twenty-seven layers of steel [80]. In both cases, single layers of Ar/Et gas proportional tubes are interleaved as the active sampling medium. The FEM and FHA are rectangular in shape, the FEM being roughly 3 m on a side and 1 m deep, while the FHA is 4.3 m on a side and 2 m deep. The alternating absorber sheets and proportional tube-layers of each calorimeter are arranged into rectangular quadrants notched at one corner to fit around the beampipe.

Instead of the conductive plastic tubes used in the plug, the forward gas calorimeters use tubes that are aluminum on three sides and fiberglass on the fourth. Each layer of proportional tubes is made from an aluminum "channel plate" which has the cross section of a comb with widely-spaced teeth. A single anode sense wire is stretched down the center of each channel and a fiberglass cover is epoxied on top of all the channels to form the fourth side of the tubes. The upper surface of this fiberglass panel is copper-clad with etched-on cathode pads of segmentation $\Delta\eta = 0.1$, $\Delta\phi = 5°$. Pads at constant (η, ϕ) are longitudinally ganged to form projective towers. In the FEM, two separate depth segments are maintained, each fifteen layers thick, while in the FHA all pads at a given (η, ϕ) are grouped into a single depth segment. Both pads and wires are read out.

All gas calorimeter signals—both plug and forward, pad and wire—are read out and processed by custom RABBIT hardware. Charge-integrating amplifiers preamplify the signals; sample-and-hold circuitry records the amplified signals before and after beam crossing; analog-to-digital converters digitize the difference between the "before" and "after" signal levels.

4.2.3 Muon systems

Muon detectors are the last CDF subsystem encountered by a particle moving radially outward from the beampipe. These are wire chambers mounted on the outermost surface of the detector, many inches of steel (hence many pion-interaction-lengths λ)

away from the interaction point. Since the cross section for radiative energy loss through bremsstrahlung drops with the inverse square of the mass of the incident particle [81], muons do not shower as electrons do. Nor do they engage significantly in the nuclear interactions responsible for hadron showers. Instead, muons pass through the many layers of lead and steel in the calorimeters largely unscathed. What energy they do deposit in the calorimeters is due to ionization, and even this is not large because their momentum lies in the *minimum-ionizing* [7] regime. Muons from $t \to Wb$ followed by $W \to \mu\nu_\mu$ are an essential part of the signal for this analysis. Muons are identified by reconstructing short segments ("stubs") from hits in the chambers and matching the stubs to tracks in the CTC.

CDF's central muon system has a certain "Rube Goldberg" flavor. It is composed of three subsystems which together cover the region $|\eta| < 1$, each with its own set of strengths and shortcomings:

- **Central Muon System (CMU)** Its barrel shape provides uniform $\eta - \phi$ coverage, but its location radially just outside the central calorimeter wedges means it has a low amount of shielding (it lies only $\sim 5.5\lambda$ [83] from the interaction region). Consequently the CMU is subject to a large "punch-through" background [8] from hadronic-shower pions which are not fully absorbed in the hadronic calorimeter but rather exit the back of the wedges and leave hits in the CMU. Moreover the CMU only covers to $|\eta| < 0.6$.

- **Central Muon Upgrade (CMP)** Shielded by 60 cm more steel than the CMU, the CMP enjoys greatly reduced background from hadronic punch-through. Its chief function is to confirm CMU hits. However, the CMP's box rather than barrel shape means it has uneven η coverage as a function of azimuth, extending at best to $|\eta| < 0.6$.

- **Central Muon Extension (CMX)** The CMX extends coverage from $0.6 < |\eta| < 1.0$ but has large azimuthal gaps (totalling 120° out of 360°) due to assorted mechanical obstructions.

[7] Ionization energy loss dE/dx reaches a broad minimum at $\gamma \approx 3$ [82].

[8] In CDF's 1989 run, when CMU was the only central muon system installed, the single muon trigger rate was dominated by punch-through [84].

Figure 4.12 shows the $\eta - \phi$ coverage of CDF's various central muon systems. Table 4.2

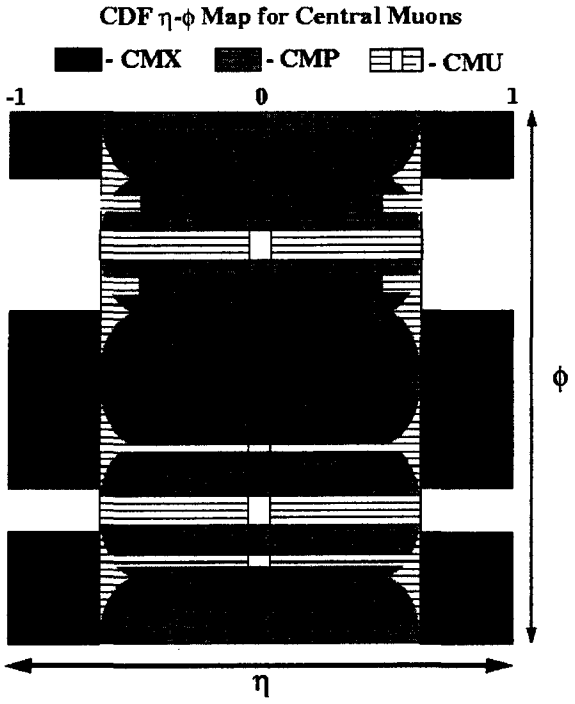


Figure 4.12: $\eta - \phi$ coverage of CDF's central muon systems.

summarizes some features of the central muon systems.

4.2.3.1 Central muon system

The central muon chambers are mounted parallel to the beamline, on the top of the central calorimeter wedges, as shown in Figure 4.13 [85]. Three chambers go on each wedge. Each chamber contains a 4×4 array of tube-shaped cells, as shown in Figure 4.14. The CMU chambers on a wedge do not reach all the way to the azimuthal edges of the wedge, so there is a 2.4° gap in muon coverage between each adjacent pair of wedges.

	CMU	CMP	CMX
η-coverage	$\lvert\eta\rvert \leq\sim 0.6$	$\lvert\eta\rvert \leq\sim 0.6$	$\sim 0.6 \leq \lvert\eta\rvert \leq\sim 1.0$
Drift-tube cross section	2.68×6.35 cm	2.5×15 cm	2.5×15 cm
Drift-tube length	226 cm	640 cm	180 cm
Maximum drift time	800 ns	1.4 μs	1.4 μs
N_{λ_π} from interaction region	5.5	7.8	6.2
Minimum detectable muon p_T	1.4 GeV/c	2.2 GeV/c	1.4 GeV/c

Table 4.2: Summary of CDF central muon system properties [83].

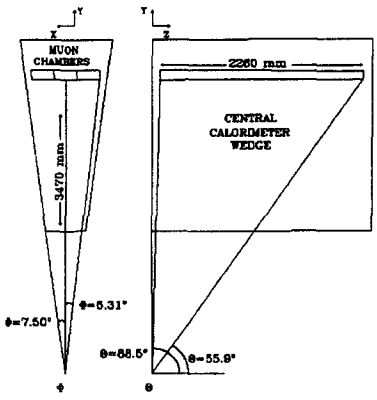


Figure 4.13: Location of central muon chambers on calorimeter wedge. In the head-on view (left) each of the three small boxes represents one CMU chamber (= sixteen tubes).

Each 4-tube $\times$ 4-tube CMU chamber is constructed of parallel aluminum sheets with glued-on aluminum I-beams and C-beams forming the separations between tubes. A 50 μm resistive sense wire is strung down the center of each tube. In a vertical stack of four tubes, known as a CMU tower, there are two pairs of wires. Wires in each pair lie on a radial line with the nominal interaction point. The two radially-aligned wire pairs are offset from each other by 2 mm. This offset causes a slight difference in arrival times at the two pairs, helping to resolve the ambiguity as to which side in ϕ a muon track passed. The operating voltages for the various parts of the chamber are +3150 V for the sense wires, -2500 V for the I-beam and C-beam tube walls, and 0 V for the aluminum sheets that form

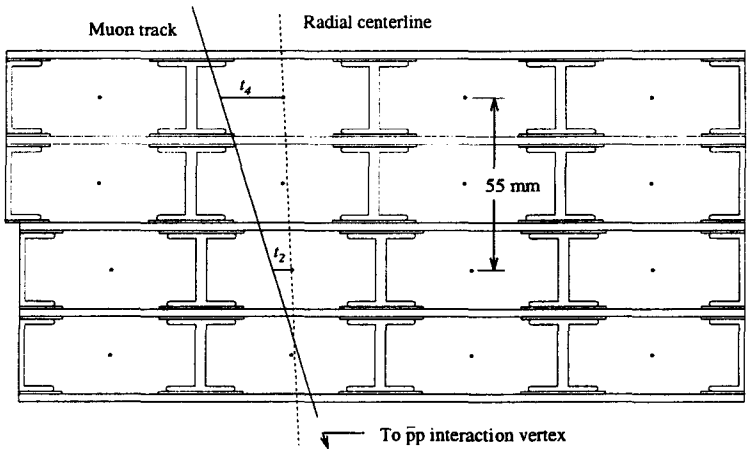


Figure 4.14: Cross section of a CMU chamber, showing the sixteen tubes. A radial line and a sample muon track are shown. Drift times t_2 and t_4 are used by the muon trigger to determine a muon momentum cutoff [52].

the floor and ceiling of each tube. Tubes are filled with 50/50 argon/ethane gas with a small amount of alcohol added. At these voltages, the tubes operate in *limited streamer mode*, in which the gas is close to breakdown and output pulses on the anode are much higher than in proportional mode [86].

Several types of information are recorded when a streamer pulse arrives on a sense wire anode. A TDC digitizes the drift time of the pulse; from this time measurement, plus knowledge of the drift velocity, the azimuthal or drift distance can be deduced. In addition to this transverse measurement, an analog-to-digital converter (ADC) measures the charge deposited on the wire to extract z-position. This is done using the method of *charge division*. Wires in alternate tubes in the same layer are electrically connected at one end and read out by a pair of ADC's at the other (Figure 4.15.) Charge deposited a distance z along the wire of length L sees resistance $R_1 = \rho(L - z)$ to the closer ADC1 and $R_2 = \rho(L + z)$ to the farther ADC2, where ρ = resistance/unit length of the wire. Charge Q_1 reaching ADC1 will be correspondingly larger than charge Q_2 reaching ADC2. It can easily be shown that $(Q_1 - Q_2)/(Q_1 + Q_2) = z/L$, so that z-position can be obtained from Q_1 and Q_2. The resolution achieved by the CMU for drift-distance measurements is 250 μm and for z-measurements 1.2 mm.

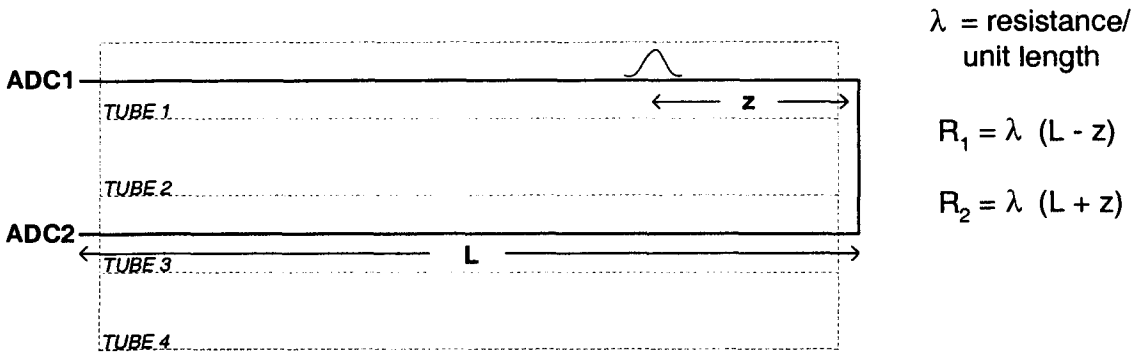


Figure 4.15: Diagram showing how charge division is used to extract z-position of a muon segment.

Another important feature of the CMU is its role in identifying muons for CDF's trigger system (c.f. Section 4.2.4.) The CMU uses drift-time differences between radially-aligned wires to perform a quick, crude measurement of muon p_T for use in the trigger. A muon's trajectory bends in the solenoidal B-field, then resumes a straight line. Because of the bending, the particle's path through the muon chambers deviates from a radial line by an angle α. The degree of bending—and consequently the magnitude of the angle α—depends inversely on the particle's p_T. The angle α can be measured from the drift-time difference Δs between a radially-aligned pair of wires in a CMU tower according to $\Delta s = H\alpha/v_{drift}$, where H is the spacing between the wire pair (55 mm; see Figure 4.14.) The smaller the drift-time difference Δs, the smaller the angle α, the less the track was bent in the solenoid, and the higher-p_T ("stiffer") the track. If Δs below a set threshold is detected, a trigger signal is generated.

4.2.3.2 Central muon upgrade

The CMP consists of 60 cm of steel located behind the CMU, followed by drift chambers [84]. Some of the steel is provided by the solenoid yoke; the remainder is provided by a pair of steel walls flanking the north and south sides of the detector. The steel adds an average of three pion interaction lengths to the central region, reducing hadronic

punch-through by a factor of twenty [9]. Chambers are arranged parallel to the beamline in a box-like geometry of fixed z. Mechanical obstructions limit the CMP's ϕ coverage to 76%; the rectangular as opposed to cylindrical geometry makes the CMP's η coverage a nonuniform function of ϕ, as shown in Figure 4.12.

CMP drift tubes are layered four deep in a half-cell-staggered configuration atop the steel shielding. The half-cell offset reduces left-right ambiguities. The chambers themselves are wide, flat tube-shaped aluminum extrusions. Cathode field-shaping strips glued onto the floor and ceiling of each tube help keep the drift field uniform. The central cathode strip is held at +3000 V; the other cathodes decrement in voltage in units of 375 V until they reach the (grounded) tube wall. A single anode sense wire at +5600 V is strung down the center of the tube. At these voltages, the chamber operates in *proportional* mode. The gas mixture is the usual argon/ethane with alcohol added.

Each CMP wire passes through a preamplifier and an ASD before being read out by a TDC. Drift-time measurements can be used to extract position measurements. The position resolution of CMP chambers is $\simeq 300$ μm. In addition to drift-time measurements, the CMP TDC's also send information to the trigger system on coincidences of chamber hits that would be consistent with a muon coming from the interaction point.

4.2.3.3 Central muon extension

If the CMX had no azimuthal gaps, it would form a pair of rings fore and aft of the detector, where the rings are not cylindrical but rather flared as if to lie on the surfaces of cones with apices pointing away from the interaction region. This configuration covers the pseudorapidity range to $0.6 < |\eta| < 1.0$. As it is, the CMX suffers from large gaps in azimuth due to obstruction from the Main Ring and solenoid refrigerator overhead, and the collision-hall floor underneath. The gaps total 120° in azimuth.

The chambers used in the CMX are identical in structure, resolution, and readout to those in the CMP, except for their shorter overall length (1.8 m *vs.* 6.4 m). The tubes

[9]Whereas the CMU-only trigger rate was dominated by punch-through, the CMU-CMP rate is dominated by single minimum-ionizing particles, as desired.

are stacked half-cell-staggered, eight layers deep. Since they fan out to form the surface of a frustrum, the tubes' edges do not touch, nor are their sense wires parallel to each other, though each tube's wire is coplanar with the beamline. Although the individual layers of tubes are non-hermetic on account of the tubes' having to fan out, alternating layers of tubes do compose a hermetic surface. The average muon passes though six layers of tubes on its trip through the CMX. Figure 4.16 illustrates the arrangement of tubes in the CMX.

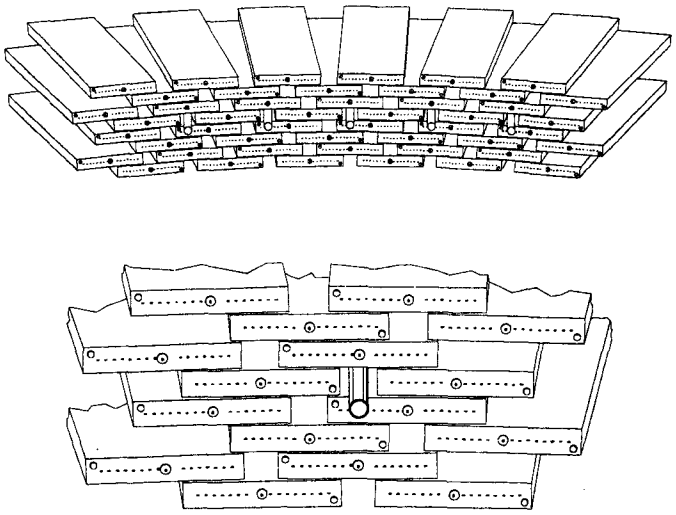


Figure 4.16: Tube layout in a portion of the CMX.

In addition to the drift tubes, the CMX features several layers of scintillation counters (CSX) mounted on both the inner and outer surfaces of the arches. The CSX system's fast response and excellent timing resolution helps reject the large CMX background coming from out-of-time interactions in the beampipe or the face of the forward calorimeter [83]. The CSX also provides fast information for the trigger, compensating for the long (1.4 μsec) maximum drift time of the CMX chambers. The signature of a real muon from the interaction region is then a coincidence between the CSX layers plus a stub in the CMX drift chambers.

Like the CMU, the CMX can also provide a crude meaurement of p_T for the trigger since it has projective (radially-aligned) wire pairs.

4.2.4 Trigger system

The total cross section for $p\bar{p}$ interactions at $\sqrt{s} = 1.8$ TeV has been measured to be $\sigma_{p\bar{p}} = 80.03 \pm 2.24$ mb [87]. Of this cross section, the portion visible to CDF's beam-beam counters (BBCs, discussed more fully in Section 4.2.5), is 51 mb [88]. Thus at a typical value of CDF Run 1 instantaneous luminosity $\mathcal{L} = 8 \times 10^{30}$ cm^{-2} s^{-1}, the rate of $p\bar{p}$ interactions [10] as measured by the BBCs is 0.41 MHz. Yet the rate at which events can be written to a storage medium such as disk or tape is at most a few Hz [90]. The task of the CDF trigger system is to choose the one out of roughly every 100,000 events that will be written to tape.

Happily, the vast majority of the events resulting from the 80-mb $p\bar{p}$ cross section are not very interesting. The single-top events relevant to this analysis typically contain physics objects such as electrons, muons, jets, or large $\not\!\!E_T$. Other events of interest at CDF might contain energetic photons, lepton pairs, *etc.* But most $p\bar{p}$ interactions lack any such features. Rather, the bulk of the cross section consists of low-momentum-transfer (low-q^2) interactions in which little trace is left in the detector, apart from the BBCs. Figure 4.17 illustrates that the cross sections for "interesting" processes at the Tevatron are many orders of magnitude smaller than $\sigma_{p\bar{p}} = 80$ mb. During datataking, CDF's trigger system performs a quick, on-the-fly analysis of a subset of the data from each event to determine if the event contains any interesting physics objects. The subset of data ("fast-outs") used by the trigger to make its decision, while not the full readout, does include information from the major subsystems of the detector including calorimetry, central tracker, and muon chambers. Events identified by the trigger as having interesting physics features are written to tape. Events rejected by the trigger are simply overwritten by the next incoming event.

CDF's trigger system is implemented as a three-level hierarchy. Each level reduces

[10] Dividing by the bunch crossing rate $1/3.5\mu$s = 0.286 MHz, the number of events per crossing is $\mathcal{O}(1)$. Note that the number of events per crossing will obey a Poisson distribution about this "average" number [89].

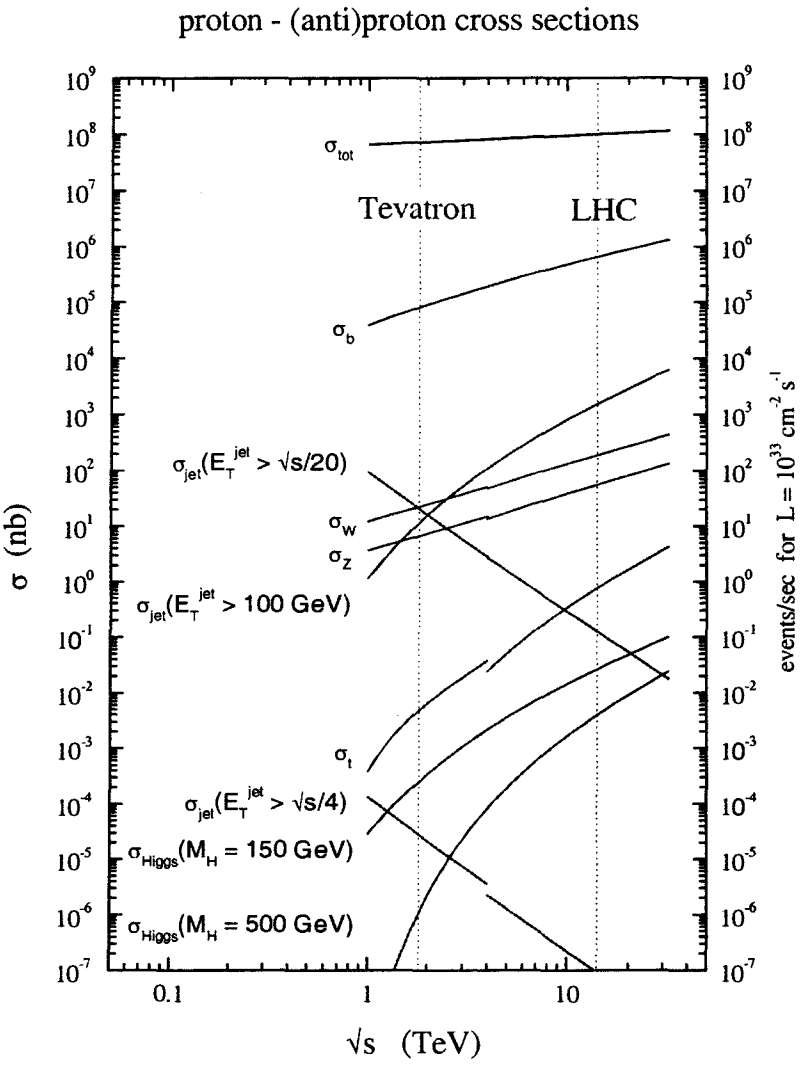


Figure 4.17: QCD predictions for hard scattering cross sections at the Tevatron and LHC [91].

the rate to a point where the next level can apply a more sophisticated (and more time-consuming) selection filter. Level 1 and Level 2 are implemented in dedicated hardware that uses fast, mostly analog inputs from the detector to make decisions. Level 1 discards roughly 99/100 of the events it sees, reducing the rate from the incoming once-per-bunch-crossing 286 kHz to a few kHz. Events passing this first cut are sent on to Level 2, which further reduces the rate to 100 Hz. Acceptance of an event at Level 2 initiates a full detector

readout, which takes 1 ms. During readout the detector is "dead", *i.e.* unable to take any new data. The 100 Hz rate of Level 2 Accepts was chosen to keep the readout deadtime at a manageable level (100 readouts/sec $\times$ 0.001 dead-sec/readout = 10% deadtime). Lastly, the full data for the event pass to the Level 3 trigger. Level 3 is implemented in software running on a group of processors. Here the rate is reduced by another factor of ten. The few Hz of events passing Level 3 are written to storage medium. Figure 4.18 summarizes the data- and decision-flow of the CDF trigger and data-acquisition (DAQ) system.

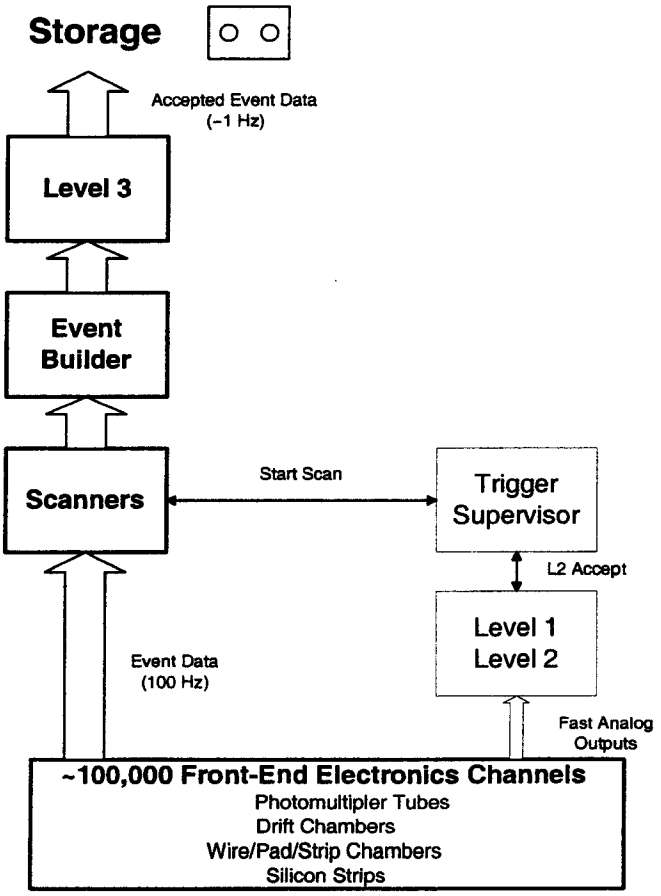


Figure 4.18: Overview of CDF trigger and data-acquisition system. The rate into Level 1 is 0.286 MHz and the rate out of Level 2 is 100 Hz.

Flexibility is built into the trigger system. A software table downloaded at run-time specifies all the filters to be used and cut values to be applied to each physics quantity

examined by each level of the trigger system [92]. (*Nomenclature note:* Such filters are themselves referred to as "triggers"—as in "This event passed a Level 2 electron trigger." This use should not be confused with "trigger" meaning the three-level system itself.) The trigger table can also specify rate limits or *prescales* to be applied to triggers that are expected to generate especially high rates of passing events. To prescale a trigger by a factor of N means to retain only one out of every N events passing the trigger.

4.2.4.1 Level 1 trigger

The Level 1 trigger is implemented in a set of custom-designed FASTBUS boards chiefly residing in crates in the CDF trigger room. Among the pieces of information calculated by Level 1 boards and used to make the Level 1 decision are the following:

- **Calorimetry information:** Physical calorimeter towers are summed into logical "trigger towers" of segmentation (15° in ϕ) $\times$ (0.2 in η). In the interest of speed, the energy information for each trigger tower is not digitized. Calorimeter trigger hardware computes the analog sum of E_T over all trigger towers above each of several programmable thresholds. Separate sums of electromagnetic, hadronic, and total (EM + had) energy are maintained. In addition to "sum E_T", the net imbalance in E_T $(E_T \cos\phi, E_T \sin\phi) = -\not\!\!E_T$ is calculated, and single towers above a high threshold are noted.

- **Tracking information:** A hardware track finder (Central Fast Tracker, or CFT) applies two coincidence gates to digitized "fast-out" signals from axial sense wires in the CTC to locate *prompt* and *delayed* hits corresponding to short and long drift-times, respectively, in a given drift cell [93]. While its chief function is to reconstruct $r - \phi$ tracks for Level 2 using both prompt and delayed hits, the CFT also performs a first pass using only prompt hits. Level 1 is notified if prompt hit patterns are found that correspond to tracks above a programmable threshold.

- **Muon information:** As discussed in Section 4.2.3.1, the angle α by which a muon stub departs from a radial line (a) depends inversely on the p_T of the particle, and (b) can be measured by the difference in drift time Δs between a radially-aligned pair of wires in the CMU or CMX. If Δs for a wire pair is detected below a particular threshold, signals are sent to Level 1.

- **Beam-beam counter information:** In addition to triggering on interesting physics objects, CDF also records "minimum bias" events at a fixed rate of approximately once every four seconds. The only requirement of a minimum-bias event is that it cause both the east and west BBC counters to fire within a 30 ns gate centered on the bunch-crossing time [94], signalling the fact that some sort of a beam-beam interaction has occurred. The BBCs are described in Section 4.2.5 where their role in calculating luminosity is discussed.

The Level 1 trigger decision is made by a special hardware board known as Fred (oddly, not an acronym). Fred receives a set of summary bits from all the subcomponents of Level 1, including calorimeter, tracking, muon, and BBC logic. Fred uses these bits as the address to a look-up table which tells whether to pass or fail the event. The entire Level 1 decision is made in the 3.5 μs between beam crossings, so the trigger incurs no deadtime. If an event is rejected at Level 1, all its data is promptly overwritten by the next incoming event.

4.2.4.2 Level 2 trigger

Like Level 1, Level 2 is implemented in custom FASTBUS hardware residing in the CDF trigger room; in fact, many of the same boards used for Level 1 also perform Level 2 functions. Also like Level 1, Level 2 uses the "fast-out" subset of data rather than the full detector readout. However, the Level 2 trigger is not deadtimeless. Instead, every time a Level 1 Accept is generated, the entire detector must hold onto its data and wait until a Level 2 decision is handed down. If Level 2 rejects the event, datataking resumes and the event is simply overwritten. If Level 2 accepts the event, the relatively lengthy 1 ms process

of readout (with its own associated deadtime) begins. The rate of events out of Level 1, *i.e.* into Level 2, is a few kHz; in order to keep Level 2-incurred deadtime at a $5-10\%$ level, the Level 2 decision time [11] must be kept to roughly 10 μs.

This is enough time to perform more sophisticated operations on the fast-out data than were done at Level 1. Among the physics quantities calculated at Level 2 are the following:

- **Calorimeter clusters:** Clusters of calorimeter energy deposition, such as would be left by electrons or jets, are identified by special hardware. The Cluster Finder board identifies "seed" towers over a high threshold and then looks around each seed for neighboring towers above a lower, "shoulder" threshold. Energy, position, width, and tower-multiplicity are among the quantities calculated for each cluster. In addition, clusters are matched to tracks identified by the CFT. This provides a first-order discrimination between es and π^0s, both of which shower in the EM calorimeter.

- **Tracking:** The CFT find $r-\phi$ tracks using both prompt and delayed hits. Momentum is measured with resolution $\delta p_T/p_T^2 \simeq 3\%$ [95]. While at Level 1 only the presence (not the direction) of high-p_T tracks was known, at Level 2 CFT the directions are known, so CFT tracks can be matched to calorimeter clusters or to muon stubs.

- **Muons:** "Golden muons" are stubs in the muon chambers that match high-p_T tracks found by the CFT. Calorimeter energy deposition by muon candidates is also identified.

- **Global $\sum E_T$; $\not\!\!E_T$:** Global sums of towers over threshold are redone at Level 2 since the analog signals have had more time to settle and are therefore more accurate.

A set of dedicated Level 2 processors assembles the above information, compares it to the filters and cut values specified in the software trigger table, and decides whether to accept or reject the event. The 100 Hz rate of Level 2 Accepts corresponds to a rejection factor of roughly $10-50$.

[11] Suppose the rate out of Level 1 is 5000 Hz. Then (5000 decisions/sec) $\times$ (10×10^{-5} dead-sec/decision) = 5% deadtime due to Level 2.

4.2.4.3 Level 3 trigger and data-acquisition system

A Level 2 Accept initiates a Start Scan broadcast to CDF's data acquisition system [92], prompting data from the front-end crates to be digitized and read by scanner modules [12]. The data for all channels is loaded into a specified buffer on all the scanners. The detector sits idle during this 1 ms readout process. Next, a FASTBUS object [13] known as the Event Builder pulls the data from that buffer on all the scanners and reformats it into subsections organized by detector component. These data-structure subsections are known as *banks* and are part of the YBOS memory management system [96]. The Event Builder writes the entire reformatted event to a single processor ("node") of the Level 3 farm.

Thus the Level 3 trigger makes its decision based on the full data for an event. Each node in the Level 3 processor farm executes FORTRAN filter algorithms which are a subset of the CDF event-reconstruction and offline analysis code [97]. By these sophisticated filters, Level 3 reduces the incoming data stream by another factor of $10 - 100$ down to a few Hz, the rate at which events can be written to disk or tape for offline analysis by individual physicists.

4.2.5 Luminosity monitoring

All cross sections measured at CDF, including the limit on the single-top cross section set by this analysis, are normalized relative to the *integrated luminosity* $\int \mathcal{L}dt$ of the CDF Run 1 dataset. This last section of the Experimental Apparatus chapter defines the quantity $\int \mathcal{L}dt$, describes how $\int \mathcal{L}dt$ is measured, and provides the value of $\int \mathcal{L}dt$ that will be used later in this analysis.

Instantaneous luminosity $\mathcal{L}$ is the product of incident beam flux (particles/second) with target density (particles/unit area) [98]. It is a measure of the intensity of a colliding beam machine, according to

$$\mathcal{L} = N_1 N_2 f W_x W_y$$

[12]The scanners are MX's for all systems read out by RABBIT, such as calorimetry and muon systems, and SSP's (later FRC's) for systems read out by FASTBUS, such as tracking.

[13]Implemented in a group of FASTBUS boards.

where N_1, N_2 are the number of particles in each bunch, f is the revolution frequency, and W_x, W_y are functions of the horizonal and vertical beam profiles, respectively, each with units (length^{-1}) [99]. The units of instantaneous luminosity are inverse cross-section per unit time, commonly expressed either as cm^{-2}/sec or b^{-1}/sec (1 barn = 10^{-24} cm^{-2}). The units of time-integrated luminosity $\int \mathcal{L}dt$ are inverse cross-section (b^{-1}). Luminosity integrated over the period in which a dataset was collected is a measure of the size of the dataset. Note that since the units of this measure are b^{-1}, the prefix decreases as the integrated luminosity increases (1 inverse femtobarn fb^{-1} > 1 inverse picobarn pb^{-1}).

Luminosity is the link between the cross section for a process and the rate of events from that process:

$$R = \sigma \mathcal{L}$$

or, integrating with respect to time, the number of events from that process:

$$N = \sigma \int \mathcal{L}dt$$

Thus, knowledge of the $\int \mathcal{L}dt$ used to collect the CDF dataset permits us to translate N observed from a given process to cross section for the process, according to $\sigma = \frac{N}{\int \mathcal{L}dt}$.

During running, CDF's luminosity is continuously monitored by a pair of beam-beam counters located ±5.82 m from the nominal interaction point [94]. Each BBC is a plane of scintillation counters with photomultiplier-tube readout, arranged in a rectangle around the beampipe. At this low beam angle (pseudorapidity range $3.24 < |\eta| < 5.9$) the counters detect a large fraction of the low-q^2 inelastic events that make up the bulk of the $p\bar{p}$ interactions [100]. A BBC "coincidence" is defined as hits in both the East and West counters in time with the bunch-crossing in the detector. The effective cross section for BBC coincidences has been measured to be 51.2 ± 1.7 mb in [88, 101]. The authors point out that σ_{BBC} is a *visible* cross section. That is, while σ_{BBC} is related to the physics of the $p\bar{p}$ interactions, it is also CDF-detector-specific, since it depends on the number of secondary interactions taking place in the material between the BBCs and the interaction point. σ_{BBC} is calculated by counting the respective numbers of elastic, inelastic, and BBC-visible events in special dedicated runs and applying the optical theorem [88].

From σ_{BBC}, integrated luminosity may be obtained using

$$\int \mathcal{L}dt = \frac{N_{BBC}}{\sigma_{BBC}}$$

The possibility of multiple interactions occurring in the same crossing means that N_{BBC} cannot directly be equated with the number of BBC coincidences. (A coincidence in the BBCs only indicates that *at least* one interaction has occurred.) Instead, we assume that the number of interactions per crossing is described by a Poisson distribution of mean μ, and write

$$N_{BBC} = \mu \cdot N_{BC}$$

where N_{BC} is the number of beam crossings, which is counted by a scaler. The mean number of interactions per crossing, μ, can be related to the Poisson probability of no interactions, P_0, by

$$P_0 = \exp(-\mu)$$

so that $\int \mathcal{L}dt = 1/\sigma_{BBC} \cdot -\ln P_0 \cdot N_{BC}$. With a little combinatorics and algebra, P_0 can be expressed as a function of number of hits in the West BBC only, East BBC only, and East-West coincidence, so that

$$\int \mathcal{L}dt = \frac{1}{\sigma_{BBC}} \cdot -\ln\left(1 - \frac{(N_{EW} - N_E \cdot N_W/N_{BC})}{(N_{BC} + N_{EW} - N_E - N_W)}\right) \cdot N_{BC}$$

The above formula assumes that μ remains constant, when in fact it drops with instantaneous $\mathcal{L}$ over the course of a run. In practice, integrated luminosity is calculated over many short intervals in the run, for each of which μ is assumed constant. The sum of $\int \mathcal{L}dt$ over all subintervals in the run, over all good runs in the data sample, constitutes the total $\int \mathcal{L}dt$ of the sample. The values of $\int \mathcal{L}dt$ for the CDF Run 1A and Run 1B samples are summarized in table 4.3.

Uncertainty on $\int \mathcal{L}dt$ derives chiefly from uncertainty on σ_{BBC}. Table 4.4 lists the source and magnitude of uncertainties on $\int \mathcal{L}dt$, along with the reference which discusses each source. To obtain the uncertainty on the total Run 1 $\int \mathcal{L}dt$, we add the contribution of

Run	Run Number Limits	Integrated Luminosity
Run 1A	40100 – 47834	19.65 pb^{-1}
Run 1B	55273 – 71023	86.34 pb^{-1}
Total Run 1		**106.0 pb^{-1}**

Table 4.3: Luminosity for CDF Run 1. From Ref. [100], Table 8, p. 33.

Uncertainty	Run 1A	Run 1B
Derivation of σ_{BBC} [102]		
Measurement uncertainty	3.13%	
Event definition	0.50%	
Beam pipe	1.00%	
Accidental correction [101]	1.00%	-
Correlated backgrounds [100]	1.00%	0.27%
Instantaneous $\mathcal{L}$ [100]	-	2.34%
Run-to-run [100]	-	0.36%
Total	3.61%	4.09%
TOTAL 1A + 1B uncertainty		**4.09 pb^{-1}**

Table 4.4: Uncertainty on integrated luminosity for Run 1A, Run 1B, and entire Run 1. Modeled after Ref. [100], Table 6, p. 31.

each row of table 4.4 in quadrature, *i.e.* assuming that rows (= sources of uncertainty) and columns (Run 1A, Run 1B) are uncorrelated [14]. This yields total uncertainty 4.09 pb^{-1}.

Thus the total Run 1 integrated luminosity value used in this analysis is **106.0 ± 4.1 pb^{-1}**.

4.2.5.1 Luminosity for sample collected by a prescaled trigger

Some of the events used in this analysis were selected by dynamically prescaled triggers. As discussed in Section 4.2.4, when a trigger is prescaled by a factor $p = 1/N$, only one out of every N events passing the trigger is retained. This effectively reduces

[14] Obviously the σ_{BBC}-related uncertainties are 100% correlated for Run 1A and Run 1B, which is why they are listed as a single column. The calculation would proceed according to (3.13%)(106.0 pb^{-1}) ⊕ (0.50%)(106.0 pb^{-1}) ⊕ (1.00%)(106.0 pb^{-1}) ⊕ (1.00%)(19.65 pb^{-1}) ⊕ (1.00%)(19.65 pb^{-1}) ⊕ (0.27%)(86.34 pb^{-1}) ⊕ (2.34%)(86.34 pb^{-1}) ⊕ (0.36%)(86.34 pb^{-1}) = 4.09 pb^{-1}.

the integrated luminosity for that trigger by a factor of $1/N$ ($\mathcal{L} = \frac{1}{N} R_{trigger}/\sigma_{trigger}$). "Dynamic" prescaling refers to the fact that the prescale factor p is not constant, but rather changes with instantaneous $\mathcal{L}$. If $p_i = 1/N_i$ is the prescale factor used during the ith segment of a run, then the prescaled (effective) integrated luminosity is

$$\left(\int \mathcal{L}dt\right)_{effective} = \sum p_i \left(\int \mathcal{L}dt\right)_i$$

In this analysis we finesse the tedious task of scaling each sub-run's worth of integrated luminosity by the appropriate dynamic prescale factor. Instead, we incorporate the effect of the prescales into our calculated acceptance using a trigger simulation. This will be discussed more fully in the section on signal acceptance (Section 5.4.2).

Chapter 5

Signal-Event Selection

As discussed in Chapter 3, the dominant Standard-Model modes of single-top production at the Tevatron are the t-channel process known as "W-gluon fusion" ($qg \to t\bar{b}q'$) and the s-channel process known as "W^*" ($q\bar{q}' \to t\bar{b}$). The theoretically-predicted single-top cross sections are 1.70 pb for Wg-fusion and 0.73 pb for W^*, for a total of 2.43 pb. Therefore CDF's entire Run 1 data set is expected to contain roughly

$$2.43 \text{ pb} \times 106.0 \text{ pb}^{-1} = 260 \text{ single top events}$$

These few hundred signal events are interspersed among millions upon millions of non-single-top (background) events. [1]

This chapter addresses the needle-in-a-haystack challenge of isolating the single-top signal. The goal is to devise selection criteria which, insofar as is possible, simultaneously retain the maximum number of signal events and reject the maximum number of background events. When applied to data, these selection criteria yield a signal-enriched sample from which information on the single-top cross section can be extracted. In this analysis, we treat both Standard-Model single-top production mechanisms (Wg fusion and W^*) as signal. That is, we design the selection criteria to retain both channels and we extract cross-section information as a combined result for both channels.

Section 5.1 surveys the main kinematic features of single-top signal events. Based

[1] If we assume data were taken at a rate of $\sim$1 Hz for 3 years, the CDF Run 1 data set contains roughly 100 million events.

on this information, Section 5.2 presents a combined strategy for selecting both Wg-fusion and W^* single-top events. Section 5.3 describes how this strategy is implemented in specific selection criteria or "cuts." Finally, Section 5.4 estimates the efficiencies of these cuts, *i.e.* how effective they are at retaining the desired signal events.

5.1 Monte-Carlo Studies of Signal Events

5.1.1 Simulating data with Monte Carlo techniques

> In 'real life', the machine produces interactions. These events are observed by detectors, and the interesting ones are written to tape by the data acquisition system. Afterwards the events may be reconstructed, *i.e.* the electronics signals (from wire chambers, calorimeters, and all the rest) may be translated into a deduced setup of charged tracks or neutral energy depositions ... In the Monte Carlo world, the rôle of the machine, namely to produce events, is taken by the event generators ... The behaviour of detectors—how particles produced by the event generator traverse the detector, spiral in magnetic fields, shower in calorimeters, or sneak out through cracks, *etc.*—is [done by] ... detector simulation [programs]. Ideally, the output of this simulation has exactly the same format as the real data recorded by the detector, and can therefore be put through the same event reconstruction and analysis chain ... [103]

To predict the kinematic features of single-top events, we use a Monte Carlo event generator. Coded into the Monte Carlo generator is the field-theory-derived expression for the total cross section for the process of interest. Denote this process by $A+B \to C+D$. The expression for the total cross section has the form of partonic cross section $\hat{\sigma}(A+B \to C+D)$, weighted by structure functions $F_{A/p}(x_1)$, $F_{B/\bar{p}}(x_2)$ expressing the probability to find initial-state partons A, B inside the beam particles p, $\bar{p}$ carrying fraction x_1, x_2 of the beam particle's momentum, integrated over momentum fractions. The constituent cross section $\hat{\sigma}$ itself has the form of a differential cross section integrated over final-state phase space. That is, the total cross section for process $A + B \to C + D$ has the form

$$\sigma(A + B \to C + D) = \int dx_1 dx_2 F_{A/p}(x_1) F_{B/\bar{p}}(x_2) \int dP_C dP_D \; d\sigma(A + B \to C + D)$$

This expression may be prohibitively difficult to evaluate. Mathematically, it is analogous

to

$$\bar{f} = \int \cdots \int f(x_1, \cdots, x_n) dx_1 \cdots dx_n$$

or "an integral of a weight function $f(x_1, \cdots, x_n)$ over variables $x_1, \cdots, x_n$ that parameterize the possible physical configurations." [104] The Monte Carlo generator randomly samples the integrand of this expression to achieve two purposes [105]: (a) to evaluate the integral [2], *i.e.* to estimate the value of the total cross section; and (b) to produce a set of events with the phase-space distributions that would be expected from the process of interest. Since our goal is to study the kinematic properties of single-top events, we are more interested in task (b).

The generator selects random events, or points $(x_1(i), \cdots, x_n(i))$ in phase space—each corresponding to a particular configuration of 4-momenta of initial- and final-state particles—and computes for each the value of the integrand $f(x_1(i), \cdots, x_n(i))$ in the expression for the total cross section. To achieve task (a), *i.e.* to estimate the total cross section, the generator treats the value $f(x_1(i), \cdots, x_n(i))$ as a random sample of the integrand. The average of N such samples constitutes an estimate of the integral [104]:

$$\bar{f} = \int \cdots \int f(x_1, \cdots, x_n) dx_1 \cdots dx_n \approx \frac{1}{N} \sum_{i=1}^{N} f(x_1(i), \cdots, x_n(i))$$

Task (b) is closely related. Here $f(x_1(i), \cdots, x_n(i))$ is treated as a *weight* representing the relative probability for phase-space point $(x_1(i), \cdots, x_n(i))$ to occur. To convert weighted events to unweighted, so that the denser regions of phase space are proportionately better represented just as in real data, the Acceptance-Rejection deweighting method is applied. Phase space is scanned and the maximum weight f_{max} is estimated. For subsequent phase space points the acceptance probability $R = f(x_1(i), \cdots, x_n(i))/f_{max}$ is computed, an auxiliary random number $y \in (0, 1)$ generated, and the point retained only if $y < R$ [104]. Unit-weight events thus generated will display "the same average behavior and same fluctuations as real data", except for the fact that the fluctuations arise from random sampling

[2]The advantage of estimating a d-dimensional integral using Monte Carlo methods is that the error on the estimate goes down as $1/\sqrt{N}$, where N is the number of sample points, regardless of d. By comparison, the error on the estimate of an integral obtained from numerical methods (the "trapezoid rule") only goes down as $1/N^{2/d}$ for N interpolation points [104].

rather than from the quantum mechanics of the underlying theory [103]. These fake events can then be used to study the kinematic properties of signal.

The primary Monte Carlo event generator we use in this analysis is PYTHIA version 5.6 [106]. In practice, PYTHIA's method of event generation departs in several respects from the above description. First, PYTHIA's phase-space sampling technique is more sophisticated than simply random; rather, those regions of phase space expected to be more dense are sampled with proportionately greater frequency. This leads to less fluctuation in weights and greater accuracy for a given number of sample points [107]. Moreover, the task of event generation is more complicated than simply simulating the underlying hard process whose Feynman diagram we draw. PYTHIA also incorporates the "bremsstrahlung"-type effects of initial- and final-state radiation (parton showers) and the hadronization effects of color confinement [103]. The final result is an event consisting not of simply the two or three final-state partons from the hard scattering, but also possibly hundreds of additional particles from the parton-showering and fragmentation. That is, the output of PYTHIA is an event "as detailed as could be observed by a perfect detector" [103].

Downstream of PYTHIA we use two additional Monte Carlo programs to refine the modeling of heavy quark and lepton decays. We use CLEOMC [108], a CDF adaptation of the CLEO Monte Carlo generator QQ version 9.0, to redecay B hadrons and TAUOLA version 2.5 [109] to redecay taus, all using updated decay tables.

Thus the output of the event generator for a single event is a long list of the identities and 4-momenta of particles arising from the hard scattering, showering, and fragmentation of the event. In CDF data-storage argot, such a list is known as a "GENP bank" (GENP being short for "generated particles"). The contents of a typical GENP bank for a PYTHIA Wg-fusion single-top event are shown in Figure 5.1.

The next step in generating fake signal events is to simulate the interaction of these particles with the CDF detector. A detector simulation program contains detailed information on the shape, spatial extent, and material composition of every section of the detector, from silicon layers through calorimeters through muon chambers. A "full" or

```
==== GENP printout for Run     1 Event      9 ====

GENP Particle list
 Ind  Name          Px         Py         Pz      Mass
   1 P            0.000      0.000    899.999  0.938
   2 AP           0.000      0.000   -899.999  0.938
   3 up          -0.440     -0.221    391.420  0.000E+00
   4 Gluon       -0.234     -0.045   -262.485  0.000E+00
   5 up           5.189     -3.912    369.318  0.000E+00
   6 bottom       4.111     -3.081    -40.670  0.000E+00
   7 down       -12.189    -16.287    175.406  9.900E-03
   8 top         21.489      9.295    153.242   175.
   9 W+           1.993     72.340    100.913   81.6
  10 bottom      17.817    -64.597     60.023   5.00
  11 W+           2.011     72.093    100.759   81.6
  12 bbar        -3.360     -1.295      0.360   5.00
  13 up          -0.527      1.967     -3.102  5.600E-03
  14 ubar         0.650      0.541     -1.355  5.600E-03
  15 Gluon       -0.318      0.198      0.253  0.000E+00
  16 Gluon       -0.836      1.657   -212.731  0.000E+00
  17 A(UD)        0.281      0.202   -376.419  0.771
  18 down       -12.387    -19.563     90.487  9.900E-03
  19 Gluon       -1.667     -2.974     17.610  0.000E+00
  20 Gluon        1.840      2.421     30.212  0.000E+00
  21 Gluon        0.121      1.320      9.303  0.000E+00
  22 Gluon        1.584      4.062     20.100  0.000E+00
  23 Gluon       -5.424      2.494      8.310  0.000E+00
  24 Gluon       -0.160      1.163      8.553  0.000E+00
  25 (UD)         0.440      0.221    399.811  0.771
```

Figure 5.1: Excerpt from the first twenty-five entries in the list of generated particles (GENP bank) for a PYTHIA Wg-fusion single-top signal event.

"detailed" detector simulation starts from each long-lived particle in the GENP bank and propagates it and its daughters through each successive layer of the detector, simulating its multiple-scattering, showering, and nuclear interactions along the way.

An alternative approach is to use a "fast" or "parameterized" detector simulation. A fast simulation, in contrast to a full simulation, does not "deriv[e] the detector response from first principles" [110]. Rather, properties of the long-lived GENP particles are input to parameterizations which yield high-level analysis quantities. Whereas a full simulation might return raw data in the form of photomultiplier-tube pulses or tracking-chamber hits, a fast simulation returns calorimeter-tower energies and tracking-chamber tracks, "eliminating the need to run the complete reconstruction package on the simulation output" [110]. The parameterizations are carefully tuned to ensure simulated data distributions reproduce real data distributions as closely as possible.

The detector simulation package we use, QFL′ [3], is a hybrid of these two kinds of detector simulation. The "prime" refers to the fact that the geometry of the Run 1B SVX′ is used rather than that of the Run 1A SVX. The interaction of particles with the silicon vertex

[3]The original meaning of this acronym is rumored to be "Quick, Fast, and Loose" (P. Savard).

detector is given a full simulation while the calorimeter interaction is parameterized. After QFL′ simulation, the Monte Carlo-generated signal events have exactly the same format as real CDF data and can be immediately subjected to the same event-reconstruction chain as real data would be. SVX hits are grouped into clusters; CTC tracks are reconstructed; calorimeter towers are clustered into jets; electrons, muons, and missing E_T *etc.* are all identified just as in real data. The event display for the simulated *Wg*-fusion single-top event whose GENP list appeared in Figure 5.1 is shown in Figure 5.2.

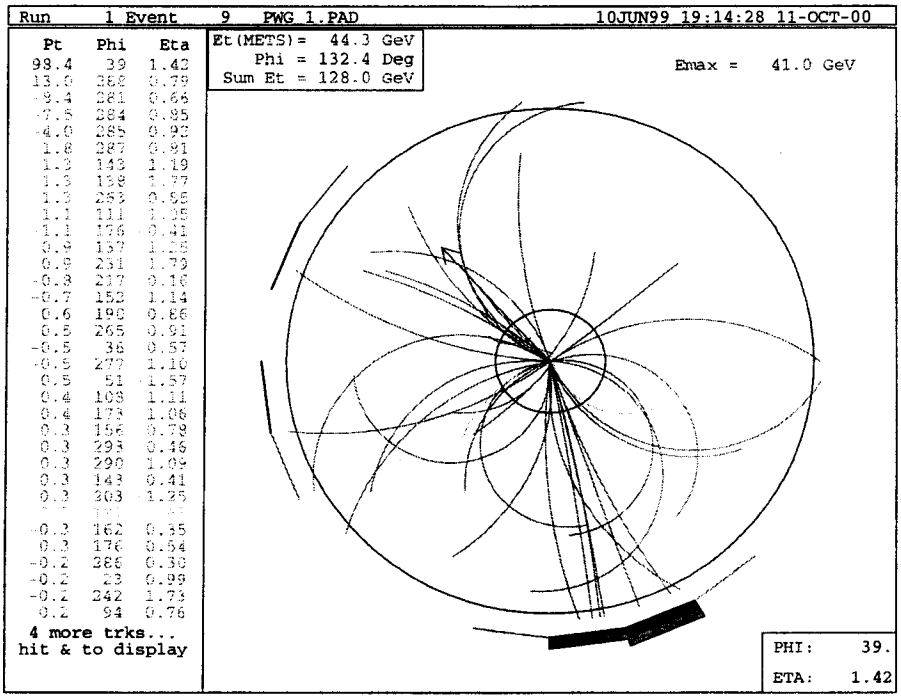


Figure 5.2: Event display, showing reconstructed CTC tracks and calorimeter energy deposition, after QFL′ detector simulation, for the same PYTHIA *Wg*-fusion single-top signal event whose GENP bank was shown in Figure 5.1.

5.1.2 Signal Monte Carlo samples

Tables 5.1 and 5.2 summarize the signal Monte Carlo samples generated for this analysis. Table 5.1 lists the standard samples used to study the properties of single-top

events and to calculate the *signal yield*, or the number of signal events expected to pass the analysis cuts. Table 5.2 lists the signal Monte Carlo samples used for studies of systematic uncertainty. These are included here for completeness.

Process	# evts generated	# evts passing PARTFILT	Filenames
Wg	152,912	42,500	`pwg_n.pad, n=1-8`
W^*	131,838	40,000	`pws_n.pad, n=1-8`

Table 5.1: Default signal Monte Carlo samples. All were generated with PYTHIA 5.6.

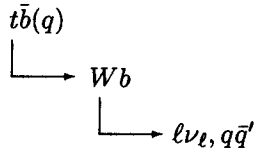
Effect	Process	# evts generated	# evts passing PARTFILT	Filenames
Initial-State	Wg	84,999	22,500	`pwg_noisr_n.pad, n=1-4`
Radiation	W^*	65,665	20,000	`pws_noisr_n.pad, n=1-4`
Parton Distrib.	Wg	85,227	23,366	`pwg_cteq3m_n.pad, n=1-5`
Function	W^*	65,544	20,000	`pws_cteq3m_n.pad, n=1-4`
Generator	Wg	75,578	20,000	`hwg_n.pad, n=1-4`
(HERWIG)	W^*	66,579	20,000	`hws_n.pad, n=1-4`
$M_{top} = 170$	Wg	72,958	20,000	`pwg_170_n.pad, n=1-4`
GeV/c^2	W^*	66,202	20,000	`pws_170_n.pad, n=1-4`
$M_{top} = 180$	Wg	71,586	20,000	`pwg_180_n.pad, n=1-4`
GeV/c^2	W^*	64,574	20,000	`pws_180_n.pad, n=1-4`

Table 5.2: Signal Monte Carlo samples used for studies of systematic uncertainty. Except where indicated, all samples were generated with PYTHIA 5.6.

5.1.3 Kinematic properties of Wg-fusion and W single-top signal events*

To recap from Chapter 3, a top quark decays after average lifetime $\sim 0.4 \times 10^{-24}$ seconds to a real W boson and a hard (*i.e.* high-momentum) b quark. The W, in turn, decays after average lifetime $\sim 0.3 \times 10^{-24}$ seconds to a lepton and neutrino (with branching fraction approximately $\frac{1}{3}$) or hadrons (branching fraction $\frac{2}{3}$). Thus the final-state partons in a Wg-fusion or W^* single-top event consist of the W-decay products plus two or three quarks: one b quark from the top decay, a second b quark from the tWb vertex in W^* events

or from the initial-state gluon splitting in Wg-fusion events, and, in the case of Wg only, a third quark recoiling from the W emission:



In the following series of plots we review some of the distinctive kinematic features of the final-state particles in single-top events. The W boson in these events was permitted to decay to any final state. However, because our analysis strategy will be to demand evidence of a leptonic W decay (a choice that will be explained in Section 5.2), Monte Carlo events are required to pass a generator-level filter (PARTFILT) requiring the presence of an energetic electron or muon in the fiducial region of the detector. Specifically, the PARTFILT cuts require an e or μ with:

- $p_T > 15$ GeV/c
- $|\eta| < 3.0$

5.1.3.1 Parton-level properties

Figure 5.3 shows the p_T spectra of the final-state quarks in Wg-fusion and W^* single-top events. The η distributions of the same partons are shown in Figure 5.4. From these plots, it may be observed that the final state of both processes features a high-p_T, low-$|\eta|$ b quark from top decay. The kinematic properties of the second b quark in the event, however, differ markedly between the two processes. In W^* this second b quark is just as hard and central as the b quark from top decay. In Wg fusion, however, the second b quark has much lower p_T and is not confined to the central region of the detector. A further difference between the final states of the two processes is the presence, in Wg fusion alone, of the additional light quark recoiling from the W emission. This quark is produced at high p_T and high $|\eta|$.

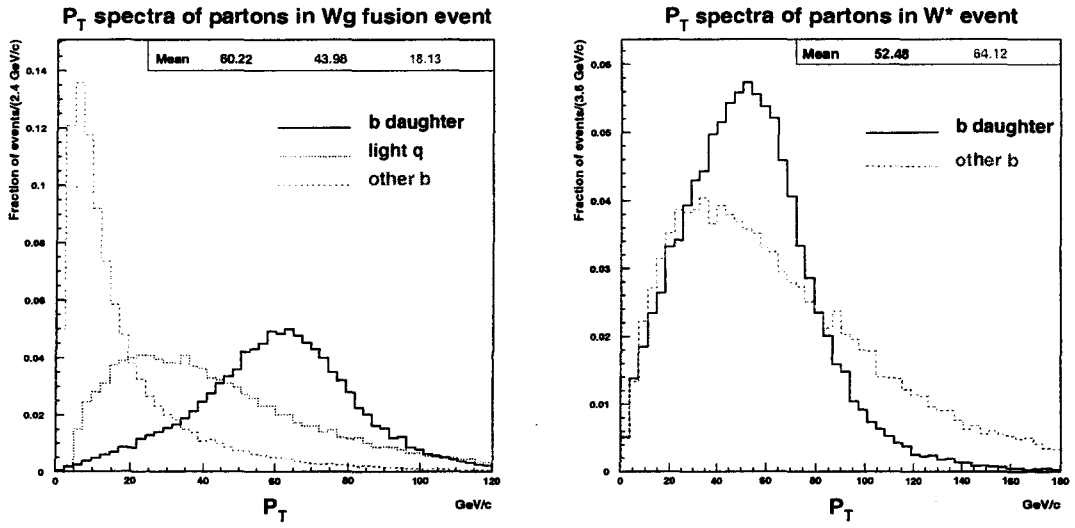


Figure 5.3: p_T of final-state quarks in PYTHIA Wg-fusion and W^* single-top events. All histograms normalized to 1.

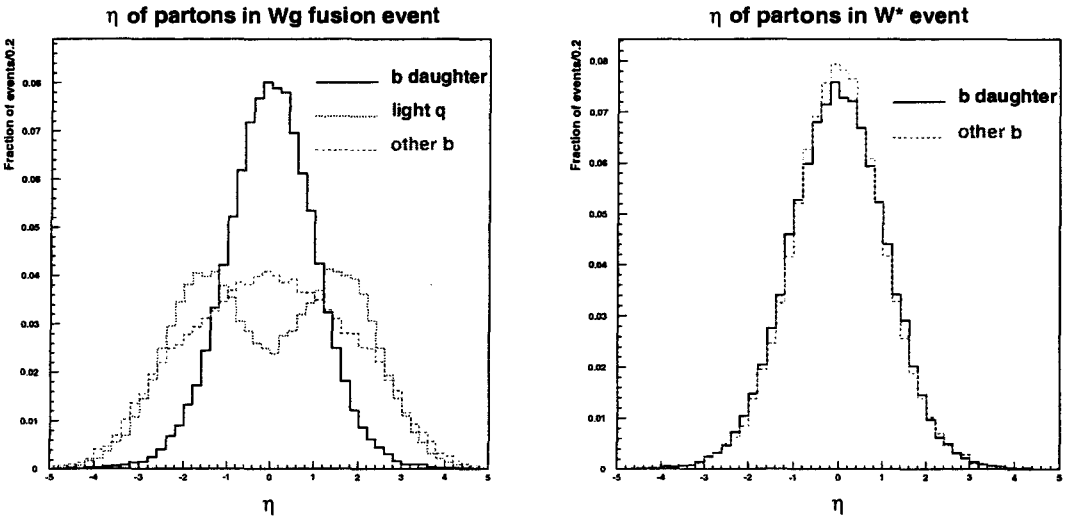


Figure 5.4: η of final-state quarks in PYTHIA Wg-fusion and W^* single-top events. All histograms normalized to 1.

A curious feature of this recoiling light quark in Wg-fusion events is that its η value displays a characteristic correlation with the charge of the lepton from top decay. This can be seen as follows. The recoiling light quark tends to keep going in the direction of the

initial-state hadron which emitted both it and the virtual W boson. That is, if a q in the proton emits $q'W$, the q' tends to move along $\eta > 0$, while if a $\bar{q}$ in the antiproton emits $\bar{q}'W$, the $\bar{q}'$ tends to move along $\eta < 0$. From the uud valence makeup of the proton we expect it to emit dW^+ $\frac{2}{3}$ of the time and uW^- $\frac{1}{3}$ of the time. Similarly, the $(\bar{u}\bar{u}\bar{d})$ composition of the antiproton means it ought to emit $\bar{d}W^-$ $\frac{2}{3}$ of the time and $\bar{u}W^+$ $\frac{1}{3}$ of the time. From these facts we deduce that no matter which hadron emits the virtual W, the sign of the W boson and the η of the recoiling light quark or antiquark ought to be the same $\frac{2}{3}$ of the time. Since the sign of the virtual W is carried into the sign of the lepton from top decay, Wg-fusion events display a pronounced bias toward $Q_\ell \times \eta_q > 0$. This phenomenon is shown in Figure 5.5. While we include this plot for completeness, this particular kinematic feature

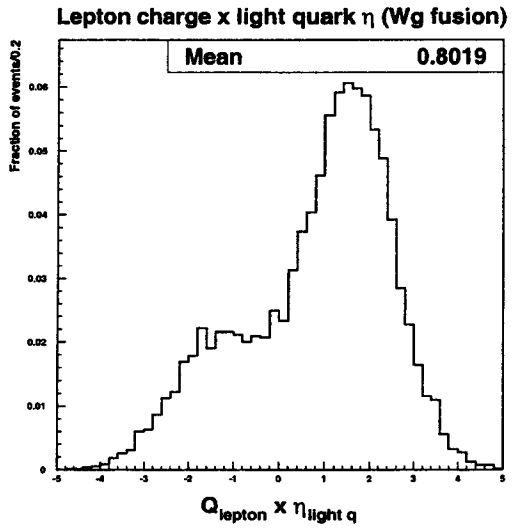


Figure 5.5: Product of charge of lepton (from W decay, from top decay) and η of recoiling light quark in PYTHIA Wg-fusion events is preferentially > 0.

will not be used in this analysis since it is not common to both single-top processes. In higher-statistics data samples the $Q \times \eta$ correlation of Wg-fusion single top will be a useful tool for separating Wg and W^*.

5.1.3.2 Validity check of signal model

While on the subject of parton-level quantities for Wg-fusion and W^* single-top events, we address the question of how well PYTHIA 5.6 is expected to model both processes. The choice of PYTHIA is uncontroversial in the case of W^*, where the similarity to the precisely-calculated Drell-Yan process ($q\bar{q} \to l^+l^-$) makes the uncertainties in the theoretical calculation of the production cross section relatively small. However, the same is not true for the Wg-fusion channel, where the treatment of the soft b quark in particular can differ from Monte Carlo generator to generator. (Even within a generator—PYTHIA—the handling of the soft b quark can be dramatically altered by the setting of switches; turning off initial-state radiation causes this b to disappear from the GENP bank.) To assess how well our chosen generator models Wg-fusion single top, we compare kinematic distributions from PYTHIA with corresponding distributions from ONETOP [6], a generator that performs a full helicity-amplitude matrix-element calculation. Since ONETOP outputs 4-vectors, these comparisons were made at parton level. (Had ONETOP's matrix calculation been spliced to PYTHIA's parton-showering and fragmentation, a truly superior generator would have resulted, but this option was not available in time for this analysis.)

Nevertheless, the comparisons, shown in Figures 5.6–5.8 for the p_T of the second b, the p_T of the recoiling light q, and the scalar sum E_T over all final-state partons and W daughters, reveal little discrepancy between the two generators. This bolsters our confidence in using PYTHIA to model Wg signal. Also shown in these plots are corresponding distributions for a third generator, HERWIG 5.6 [111]. HERWIG's distribution for the second b quark is noticeably different from that of either PYTHIA or ONETOP but even this effect does not have a large impact on the overall $\sum E_T$ distribution, for which all three generators agree well.

Thus we conclude that PYTHIA appears to model signal (*i.e.* agree with gold-standard ONETOP) reasonably well. In Chapter 9 we quantify this assertion by assigning a systematic uncertainty associated with choice of Monte Carlo generator. We estimate this

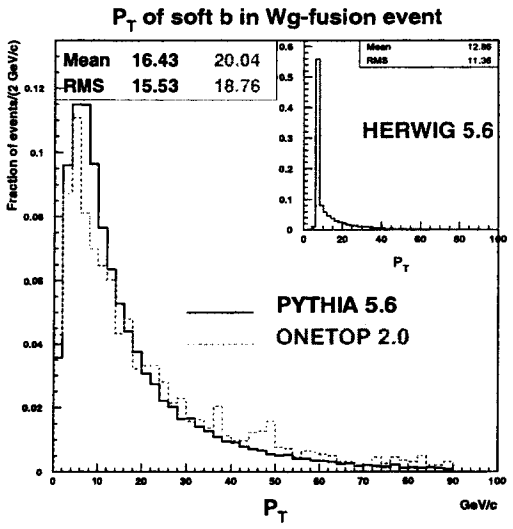


Figure 5.6: Comparison of p_T spectrum of second b quark in Wg-fusion events between three generators. Note acceptable agreement between PYTHIA and ONETOP but strange distribution for HERWIG. All histograms normalized to 1. Cuts applied: E_T of W daughters above 15 GeV; E_T of second b quark below 90 GeV.

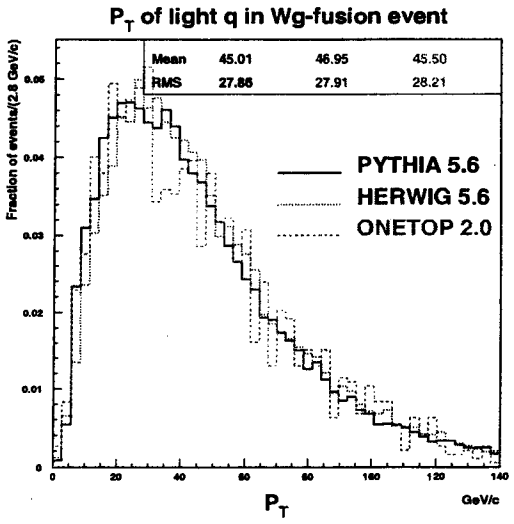


Figure 5.7: Comparison of p_T spectrum of light quark in Wg-fusion events between the three generators. All histograms normalized to 1. Same cuts as in Figure 5.6.

"generator systematic" uncertainty by assessing the impact on our fitted result of switching between HERWIG and PYTHIA.

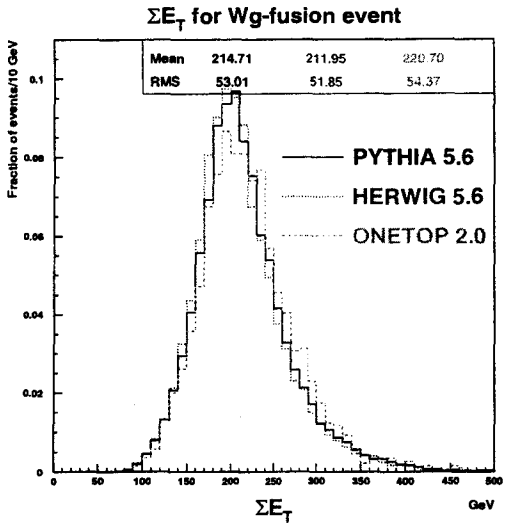


Figure 5.8: Comparison of scalar sum E_T for the three final-state quarks plus the two W daughters, for the three generators. All histograms normalized to 1. Same cuts as in Figure 5.6.

5.1.3.3 Detector-level properties

Thus far we have only examined parton-level kinematic properties of single-top events. To study the properties such events will display in the CDF detector, we examine the same Monte Carlo events after QFL′ detector simulation. In the following plots, events were required to contain evidence of a leptonic W decay: an energetic electron or muon plus the large missing energy suggestive of the lepton's undetected neutrino counterpart. The motivation behind the requirement of leptonic W decay will be discussed in Section 5.2.

Figure 5.9 shows the type of the W daughters in events passing the leptonic-W selection cuts. Clearly not all such events arise from $W \to e\nu_e$ or $W \to \mu\nu_\mu$; there is a substantial (~6%) contribution from $W \to \tau\nu_\tau$ followed by leptonic decay of the tau.

Of course, the final-state partons from a Wg-fusion or W^* single-top event appear in the detector not as bare quarks, but as jets of hadrons. We define "jet multiplicity" to be the number of energetic ($E_T > 15$ GeV), central ($|\eta| < 2.0$) jets in the event. Figure 5.10 shows the jet multiplicity in single-top events after W selection, both before and after application of an additional requirement that at least one jet be tagged as likely to contain a b quark.

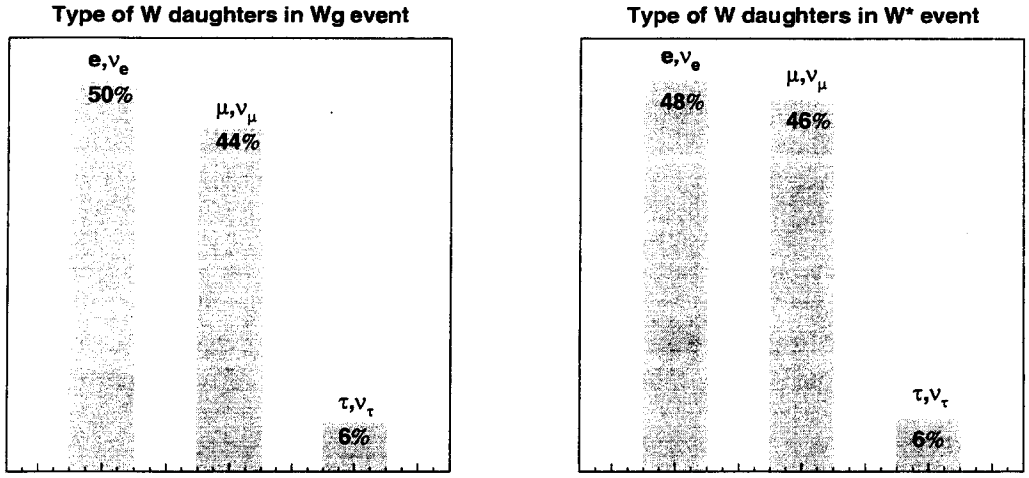


Figure 5.9: Identity of lepton from W decay in Wg-fusion and W^* single-top events passing the leptonic-W-selection analysis cuts. Note that though cuts require the presence of an electron or muon, 6% of the events are from $W \to \tau\nu_\tau$.

The algorithms used to define jets and to tag b-quark jets will be described in Section 5.3. Note that most of the signal resides in the one-, two-, and three-jet multiplicity bins. This is

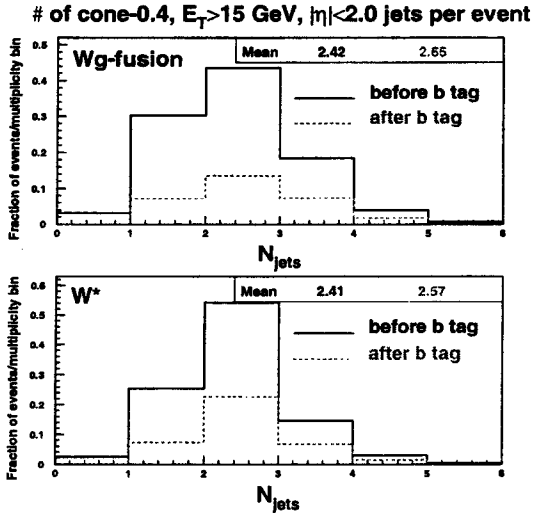


Figure 5.10: Jet multiplicity in single-top events after W selection. Solid histograms normalized to 1; dashed histograms normalized to the respective fraction of events passing b-jet identification algorithm (SECVTX tag).

true of both Wg and W^*. The reason both processes are spread out over several multiplicity bins rather than being confined to the bin corresponding to the expected number of final-state partons—*i.e.* three in the case of Wg $(b, \bar{b}, q)$ or two in the case of W^* $(b, \bar{b})$—is that some events will have final-state jets that fail the $E_T > 15$ GeV, $|\eta| < 2.0$ binning requirements, while others will have extra jets due to gluon radiation.

Lastly, we examine two kinematic variables that will be crucial to isolating the single-top signal. The first of these is the "reconstructed top mass", or the invariant mass of the lepton, neutrino, and highest-E_T b-tagged jet in the event. The specific techique used to construct $M_{\ell\nu b}$ will be discussed in Section 5.3. The second key kinematic variable for

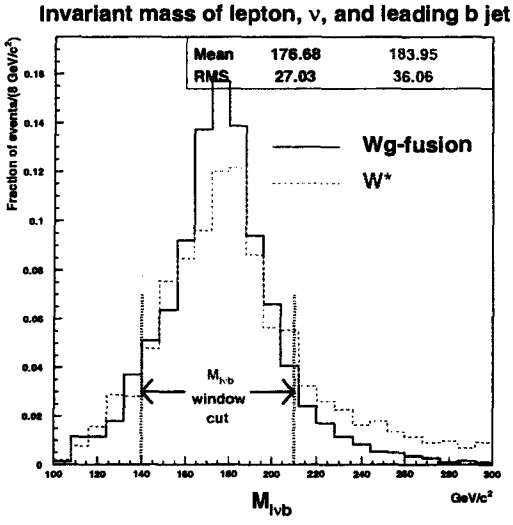


Figure 5.11: Reconstructed mass of lepton, neutrino, and leading b jet in single-top events. The cut window $140 < M_{\ell\nu b} < 210$ GeV/c^2 (*i.e.* $M_{\ell\nu b} = 175 \pm 35$ GeV/c^2) is shown. This cut is designed to improve signal strength and will be explained in Section 5.2. Histograms normalized to 1.

this analysis is $H_T \equiv \sum E_T$ {over all jets, lepton, and missing energy in the event}. The remarkable similarity in the H_T spectra of Wg-fusion and W^* single-top events (Figure 5.12) motivates our choice to perform a combined $Wg + W^*$ search using this variable.

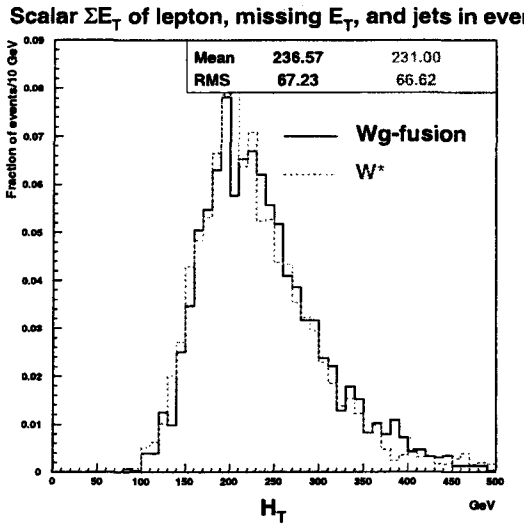


Figure 5.12: H_T, or scalar $\sum E_T$, in single-top events after $M_{\ell\nu b}$ cut. The virtual indistinguishability of the H_T distributions for the Wg-fusion and W^* processes (Kolmogorov-Smirnov probability of 1.0) makes H_T an ideal variable to use in a combined $Wg + W^*$ single-top search. Histograms normalized to 1.

5.2 Search Strategy

To review from the previous section, the final state of a single-top event consists of W-decay products ($q\bar{q}'$ or $\ell\nu_\ell$) plus one b-quark jet from $t \to Wb$, a second b-quark jet from initial-state gluon splitting (in Wg) or the Wtb production vertex (in W^*), and, in Wg only, a third jet from the light quark recoiling from W emission.

If the W decays hadronically ($W \to q\bar{q}'$), the final state consists of four or five energetic jets. This final state is difficult to distinguish from the vastly more common background of pure-QCD events which produce multiple jets in the detector. Sample Feynman diagrams for QCD multijet production are shown in Figure 5.13. While the unique kinematic properties of single-top events could be brought to bear in a search for "all-hadronic" single top (for example, requiring two jets to reconstruct near the W mass, *etc.*) we do not consider the all-hadronic channel any further in this analysis.

Instead, in order to isolate the single-top signal from the QCD multijet background, we demand evidence of a leptonic W decay. The first requirement of leptonic W selection is

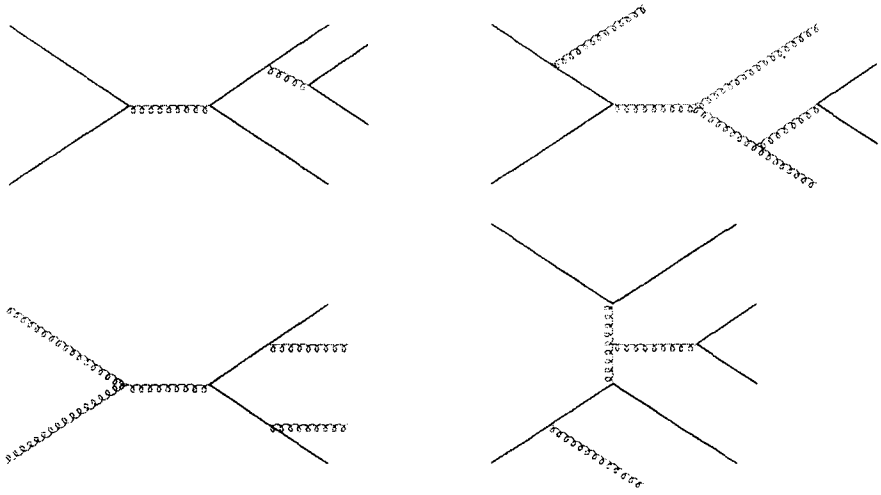


Figure 5.13: Representative Feynman diagrams for QCD multijet production.

the presence of an energetic electron or muon candidate in the central region of the detector. The lepton from W decay is expected to be *isolated*, *i.e.* free from excess additional energy in its immediate vicinity. The specific cuts used to identify leptons from W decay are described and motivated in Section 5.3. The reason we ask for an electron or muon candidate but not a tau, despite the fact that the W is roughly equally likely to decay to any of the three lepton channels e, μ, τ, is that while electron and muon identification are straightforward at CDF, tau identification is less so. Nevertheless, as we saw in Figure 5.9, the $W \to \tau\nu_\tau$ channel does contribute some additional events passing the leptonic-W-selection cuts. The second requirement of W-selection is the presence of large missing energy ($\not\!\!E_T$, defined in Section 4.2.2) such as would be left behind by the noninteracting neutrino from $W \to \ell\nu_\ell$. In letting the W boson decay to any final state in our Monte Carlo samples, we make sure to include the extra events from τ decays of the W in our kinematic distributions and acceptance calculations.

From Figure 5.10, we saw that if after W selection we group events by jet multiplicity, or number of central, high-E_T jets in the event, the bulk of the single-top signal resides in the one-, two-, and three-jet bins. This is true for both Wg and W^* single top. Thus we

narrow our search to events which pass W-selection cuts and which have one, two, or three jets.

While the leptonic-W-selection requirements greatly reduce background from QCD multijet processes, large backgrounds still remain from QCD or electroweak processes which produce a real W boson along with jets. However, these QCD-produced jets do not preferentially contain heavy-flavor quarks [4]. Therefore an effective way to isolate single-top events from this "W+jets" background is to demand evidence of at least one b quark in the final state. The relatively large mass and long lifetime of b quarks permit the use of "b-tagging" algorithms which distinguish b jets from light-quark jets with reasonable efficiency. The b-tagging algorithm used in this analysis, known as SECVTX, is described in Section 5.3. We require events to contain at least one b-tagged jet.

Even after requiring events to pass $W+1,2,3$-jet cuts with b tag, background remains from processes which produce a W boson along with true heavy-flavor (b- or c-quark) jets. Examples of true W + heavy flavor processes include $Wb\bar{b}$ production ($p\bar{p} \rightarrow Wg$, followed by $g \rightarrow b\bar{b}$) or electroweak Wc production ($p\bar{p} \rightarrow Wc$). In addition to these non-top backgrounds, top pair production ($t\bar{t}$) itself constitutes a background to single top. All our lepton, $\not\!\!E_T$, and b-tag cuts are by construction efficient at retaining the $t\bar{t}$ background! To reject these remaining backgrounds, we apply an additional cut on kinematic variable $M_{\ell\nu b}$.

Thus we conduct our combined $Wg+W^*$ single-top search in the channel $W+1,2,3$ jets plus b tag plus additional $M_{\ell\nu b}$ cut. In the next section we explain in more detail how these cuts are implemented.

[4]Since QCD interactions are flavor-independent, if we neglect phase-space considerations, the QCD-produced jets must be distributed equally across the five lighter quark flavors. In reality, the relative heaviness (4.2 GeV/c^2) of the b quark means it will be even less prevalent than $\frac{1}{5}$ in the QCD-produced jets.

5.3 Implementation: Analysis Cuts

5.3.1 W selection

In this section we describe the cuts used to identify the lepton candidate and large $\not\!\!E_T$ which we demand as evidence of a leptonic W decay. Our reason for choosing these particular cuts is that they are the same as those developed for CDF's $t\bar{t}$ analyses. The kinematics of W daughters from top decay ought to be very similar for the single-top and $t\bar{t}$ processes. Ref. [112] studied the efficiencies of the lepton-identification cuts by applying the cuts to samples of $Z \to e^+e^-$ and $Z \to \mu^+\mu^-$ events in CDF data and noting the respective numbers of events in which both leptons *vs.* only one lepton passed the cuts.

5.3.1.1 Electron cuts

An electron candidate is defined as a cluster of calorimeter towers associated with a three-dimensional CTC track [113]. Electron clustering starts from an array of "seed" towers with electromagnetic transverse energy $E_T > 3$ GeV. The E_T of a tower is defined as $E \sin\theta$, where E is the tower energy and θ is the polar angle defined by the center of the tower and the event *primary vertex*, or $p\bar{p}$ interaction point [5]. Towers adjacent to the seed are added to the cluster, up to the maximum cluster size of three towers in pseudorapidity ($\Delta\eta \sim 0.3$) by one tower in azimuth ($\Delta\phi = 15°$). Table 5.3 lists the identification cuts we apply to electron candidates. Explanation of each cut variable follows.

- **Calorimeter**

  Only electron candidates which shower in the Central Electromagnetic Calorimeter are considered.

[5] Recall from Section 4.2.1.1 that event vertices at CDF are distributed along the luminous interaction region according to a Gaussian centered at $z = 0$ of width ~30 cm. The primary vertex (P.V.) is identified event-by-event by software module VXPRIM [114]. The code starts from a seed, or guess, of the P.V. and adds tracks to the vertex using a constrained fit. Results of this fit are then used as inputs to an unconstrained vertex fit. The process is repeated, each time discarding the track with the largest residual above a cut until no such track remains. The coordinates of the final fit are deemed the P.V.

Quantity	Cut Value
Calorimeter	CEM
E_T	≥ 20 GeV
Had/EM (3×3)	≤ 0.05
E/P	≤ 1.8
Track-Strip Match	$\lvert\Delta x\rvert \leq 1.5$ cm
	$\lvert\Delta z\rvert \leq 3.0$ cm
L_{shr}	≤ 0.2
Strip χ^2	≤ 10.0
$\lvert Z_{vertex} - Z_0^e\rvert$	≤ 5.0 cm
$\lvert Z_{vertex}\rvert$	≤ 60.0 cm
Fiducial	.true.
Conversion	.false.
Isolation ($E_T{}^{0.4}/E_T$)	≤ 0.1

Table 5.3: Electron candidate selection cuts.

- E_T

  Total transverse energy of towers in the cluster.

- **Had/EM** (3×3)

  Ratio of hadronic to electromagnetic E_T for the towers in the cluster.

- E/P

  Ratio of the energy of the cluster to the magnitude of the momentum of the CTC track.

- **Track-Strip Match**

  Electron position at shower maximum is obtained from a fit to the shower shapes measured by the CES, using a nominal electron profile from test-beam data. This shower-max position is required to match the extrapolated position of the CTC track in two different views: polar (z) and azimuthal (x). x-coordinate matching is done between the track and the largest-pulse-height CES wire cluster; z-coordinate matching is done between the track and the largest-pulse-height CES strip cluster [115].

- L_{shr}

  A measure of the lateral development of an electron shower, L_{shr} compares the electron candidate's observed lateral shower profile to the expected profile as calculated from test-beam data. Defined as

$$L_{shr} = 0.14 \sum_i \frac{E_i^{adj} - E_i^{prob}}{\sqrt{(0.14\sqrt{E})^2 + (\Delta E_i^{prob})^2}}$$

  where E_i^{adj} is the measured energy in the ith tower; E_i^{prob} is the expected energy in that tower calculated using a shower profile parameterization from test-beam data; E is the EM energy of the cluster; ΔE_i^{prob} is the uncertainty on E_i^{prob}, and the sum runs over the towers adjacent to the seed tower in the same azimuthal wedge. Note that $0.14\sqrt{E}$ is the energy resolution of the CEM [6] so that L_{shr} has the form of (energy deviation)/(characteristic fluctuation) [116].

- **Strip χ^2**

  χ^2 of the fit to shower shapes used to determine the shower-max position.

- $|Z_{vertex} - Z_0^e|$

  Distance between z coordinate of event primary vertex and parameter z_0 of the electron candidate's CTC track. If the event has multiple primary vertices (due to multiple interactions), the distance to the closest vertex is used [117].

- $|Z_{vertex}|$

  z coordinate of the event primary vertex, *i.e.* z distance from the nominal collision point $(0, 0, 0)$.

- **Fiducial**

  Electron shower must fall within fiducial region of the CEM, *i.e.* avoid cracks, tower boundaries, and other uninstrumented portions of the CEM volume. The fiducial fraction of the solid angle for $|\eta| < 1.0$ is 84% [56].

[6] Thanks to J. Guimarães's thesis for pointing this out.

- **Conversion**

  Electron candidate must not come from photon conversion ($\gamma(N) \rightarrow e^+e^-$). Conversion electrons are identified by their low VTX occupancy and by being accompanied by an oppositely-charged track within a close angular separation and a characteristic radial distribution [118].

- **Isolation ($E_T{}^{0.4}/E_T$)**

  A measure of the presence of other energetic particles near the electron candidate. Defined as

  $$I = \frac{E_T^{cone} - E_T^{cluster}}{E_T^{cluster}}.$$

  E_T^{cone} is the E_T of all towers in a cone of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.4$ in $\eta - \phi$ space centered on the electron cluster; $E_T^{cluster}$ is the E_T of only those towers included in the cluster. Isolation is the ratio of excess E_T in the cone to E_T of the cluster.

5.3.1.2 Muon cuts

A muon candidate is defined as a track segment ("stub") in a muon chamber associated with a track in the CTC [56]. Table 5.4 lists the identification cuts we apply to muon candidates. Explanation of each cut variable follows.

- **Detector**

  Only muon candidates which left hits in the CMU, CMP, or CMX chambers are considered.

- P_T

  The CTC-measured transverse momentum of the muon track.

- E_{em}

  Energy deposited in the CEM tower traversed by the muon track. Required to be characteristic of a minimum-ionizing particle, *i.e.* small.

Quantity	Cut Value
Detector	CMU, CMP, CMX
P_T	≥ 20 GeV/c
E_{em}	≤ 2.0 GeV
E_{had}	≤ 6.0 GeV
$E_{em} + E_{had}$	≥ 0.1 GeV
$\|\Delta x\|$	$\|\Delta x\|_{CMU} \leq 2.0$ cm .or.
	$\|\Delta x\|_{CMP} \leq 5.0$ cm
	$\|\Delta x\|_{CMX} \leq 5.0$ cm
d_0	≤ 0.3 cm
$\|Z_{vertex} - Z_0^{\mu}\|$	≤ 5.0 cm
$\|Z_{vertex}\|$	≤ 60.0 cm
Isolation ${E_T}^{0.4}/P_T$	≤ 0.1

Table 5.4: Muon candidate selection cuts.

- E_{had}

  Energy deposited in the CHA tower traversed by the muon track.

- $E_{em} + E_{had}$

  Sum of CEM and CHA tower energies. In requiring this sum to be above a small minimum, this cut is designed to remove tracks which point at empty calorimeter towers due to track reconstruction problems [56].

- $|\Delta x|$

  Distance (cm) in $r - \phi$ plane between the muon stub and the extrapolated CTC track [117].

- d_0

  Impact parameter, or distance of closest approach as measured in the $r - \phi$ plane, between the CTC track and the beam position [117].

- $|Z_{vertex} - Z_0^{\mu}|$

  Distance along z axis between event primary vertex and muon candidate's CTC track.

If the event has multiple primary vertices (due to multiple interactions), the distance to the closest vertex is used [117].

- $|Z_{vertex}|$

  z distance from the primary vertex to the nominal interaction point $(0,0,0)$.

- **Isolation** ${E_T}^{0.4}/P_T$

  E_T deposited in a cone of 0.4 centered on the muon track, less the E_T of the tower actually traversed by the track, divided by the p_T of the track.

In addition to these quality cuts, the electron or muon candidate was required to have passed the appropriate Level 2 hardware trigger for Run 1A [119] or Run 1B [120, 121]. These triggers are summarized in Table 5.5 below [7].

5.3.1.3 Missing E_T

$\not\!\!E_T$ was defined in Section 4.2.2 as the negative of the vector sum of transverse energy over all calorimeter towers, or $-\sum_i E_i \sin\theta_i \hat{n}_i$, where E_i is the energy deposited in the ith calorimeter tower, θ_i is the polar angle of the ith tower, and $\hat{n}_i$ is a unit vector in the azimuthal plane pointing from the event vertex to the ith tower. The sum runs over all towers from $-3.5 < \eta < 3.5$ [122]. The reason a portion of the forward calorimetry is omitted in this sum—recall that it extends to $|\eta| = 4.2$—is that the low-β quadrupole magnets obscure part of the FHA. To be included in the sum, towers must have energies above certain detector-dependent thresholds on the order of several hundred MeV [69].

Since muons do not shower in the calorimeter, their energy does not get included in the $-\not\!\!E_T$ sum. Therefore, we correct $\not\!\!E_T$ to account for the presence of muons, according to [123]

$$({\not\!\!E_T}^{corr})_x = ({\not\!\!E_T}^{raw})_x - (p_\mu)_x + ({E_T}_\mu^{depos.})_x$$

[7] This task was nontrivial in the case of Run 1A events, since the trigger masks in `topfnd.cdf` are valid only for 1B events. We fixed this problem by deriving a correct set of Run 1A masks from the list of Run 1A triggers described in CDF Note 3855 [119], page 4. We implemented the new masks in a private version of `topfnd.cdf`.

Lepton candidate	Run 1A	Run 1B
HPTE		
	CEM_9*	CEM_16_CFT_12*
	MET_20*	MET_20*
CMU or CMU-CMP		
	CMUP_CFT_9_2*	CMUP_CFT_12_5DEG_V*
	CMNP_CFT_9_2*	CMNP_CFT_12_5DEG_V*
	CMU_CMP_CFT_9_2*	CMUP_JET*
	MET_35_TWO*	CMNP_JET*
	MET_35_TEX*	CMU_CMP_JET*
		CMUP_CFT_12_5DEG_M*
		CMNP_CFT_12_5DEG_M*
		MET_35_TWO*
		MET_35_TEX*
CMP-only		
	MET_35_TWO*	MET_35_TWO*
	MET_35_TEX*	MET_35_TEX*
CMX		
	CMX_CFT_9_2_ET*	CMX_CFT_12_5DEG_V*
	CMX_CFT_9_2_V5*	CMX_CFT_12_5DEG_M*
	CMX_CFT_9_2	CMX_CFT_12_5DEG_E*
	MET_35_TWO*	CMX_JET*
	MET_35_TEX*	MET_35_TWO*
		MET_35_TEX*

Table 5.5: Hardware triggers for an electron or muon candidate.

$$(\not\!\!E_T{}^{corr})_y = (\not\!\!E_T{}^{raw})_y - (p_\mu)_y + (E_T{}_\mu^{depos.})_y$$

That is, the CTC-measured muon p_T is taken to be an approximation of the muon E_T and vectorially added to $\not\!\!E_T$ (recall $\not\!\!E_T$ is a *negative* sum) while whatever small amount of energy the muon did deposit in the calorimeter is vectorially subtracted [56]. The above correction is done not just for muon candidates which pass the quality cuts of Section 5.3.1.2, but also for muon candidates which pass a looser set of cuts and for "minimum-ionizing objects" which display the stiff track and low energy deposition of a muon candidate but which traverse an uninstrumented region of CDF's muon chambers. In CDF jargon, $\not\!\!E_T$ after these muon corrections is known as "version 4" $\not\!\!E_T$.

5.3.2 *Jets*

We identify jets using a cone-clustering algorithm that sums the energy of calorimeter towers lying within an $\eta - \phi$ cone of fixed size $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.4$. A full description of the jet-reconstruction algorithm may be found in [124]. We require events to have one, two, or three jets which are both energetic ($E_T > 15$ GeV) and centrally-located ($|\eta| < 2.0$).

For the purposes of counting jets with $E_T > 15$ GeV, we use regular calorimeter-measured transverse energy. However, this measurement is known to underestimate the actual energy of the original parton which gave rise to the jet. It is this original parton energy that we wish to use in reconstructing the kinematic variable $M_{\ell\nu b}$, discussed in Section 5.3.5 below. Jet-energy mismeasurement arises from a variety of effects, some due to detector limitations, others due to unavoidable features of jet physics [124]. The former category includes calorimeter nonlinearities, $\vec{B}$-field bending of low-p_T charged particles, and reduced calorimeter response at tower and module boundaries. The latter category includes extraneous energy contributions from the underlying event [8], out-of-cone losses, and undetected energy carried off by muons or neutrinos [56]. Figure 5.14 shows that jet corrections raise the E_T of a typical light-quark jet (in this case, the jet arising from the recoiling light q in Wg-fusion events) by an factor on the order of 1.4.

The jet corrections we use are as follows. To the p_T of every jet in the event we apply a set of "flavor-independent" corrections of the form [125]

$$p_T^{corr} = (p_T^{raw} \times f_{rel} - UEM) \times f_{abs} - UE + OOC$$

where

- f_{rel}, the relative jet energy scale, corrects for nonuniformities in η response;
- UEM accounts for energy due to multiple interactions in the event;
- f_{abs} is the absolute jet energy scale;

[8] The underlying event is defined as the "ambient energy produced in hadron collisions associated with the soft interactions of spectator partons" [124].

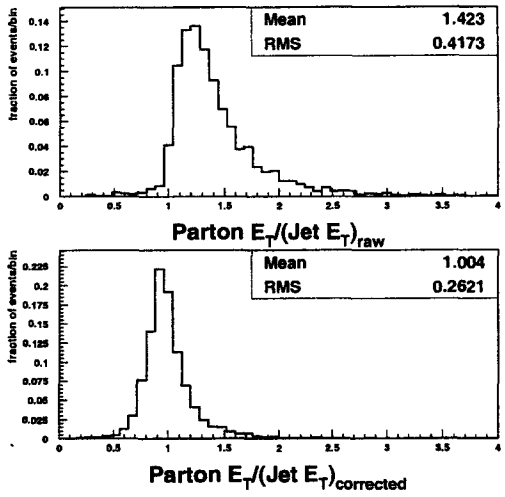


Figure 5.14: Effect of jet energy corrections: ratio of {original parton E_T from the GENP bank} to {jet E_T}, both before and after corrections, for the jet arising from the recoiling light q in Wg-fusion events. The corrections increase the E_T of the average jet by a factor of 1.4.

- *UE* accounts for energy due to the underlying event;
- *OOC* corrects for out-of-cone losses.

We use CDF routine JTC96S (by way of driver-routine JETOPS) to implement these corrections.

In addition, we apply a set of $t\bar{t}$-specific corrections (known in CDF jargon as "AA corrections") to the four most energetic jets in the event. In a $t\bar{t}$ lepton + jets event, these four jets presumably arise from the four hadronic-top daughters (b, $\bar{b}$, q, $\bar{q}'$); the purpose of these additional corrections is to obtain a refined estimate of these original partons' momenta [125] above and beyond the "flavor-independent" corrections described above [9]. Different corrections are applied to each jet's energy depending on its b-tagged status. Four categories of jet are distinguished: (a) not SVX-tagged; (b) SVX-tagged but not soft-lepton-tagged; (c) SVX-tagged and soft-electron-tagged; (d) SVX-tagged and soft-muon-tagged.

[9] For example, among other things, the $t\bar{t}$-specific corrections account for the energy lost through semileptonic decays of b and c hadrons [125].

(See Section 5.3.3 below for an explanantion of SVX tagging and soft-lepton tagging.) While these $t\bar{t}$-specific corrections were tailored to the final state of $t\bar{t}$ rather than single top, the similarity is close enough to warrant their application.

5.3.3 b tagging

The long lifetime of the b quark—$\mathcal{O}(10^{-12})$ s—means that a B hadron formed at the interaction point of a single-top event may travel a significant distance before decaying into a collection of lighter hadrons. In principle, the tracks from these decay hadrons can all be traced back to the spot where the B hadron decayed. This so-called secondary vertex will be displaced from the $p\bar{p}$ interaction point, or primary vertex, by the decay length [10]. The main method of identifying b jets ("b tagging") used in this analysis looks for the presence of secondary vertices within jets.

The b-tagging algorithm we use, called SECVTX [126], applies a loose set of quality cuts to all the SVX tracks [11] lying in a cone of $\Delta R = 0.4$ about the jet axis, then searches the list of passing tracks for sets of three or more tracks that form a vertex. If this attempt fails, the sub-list of tracks passing a set of tighter cuts is compiled to see if the jet contains at least two such higher-quality tracks. If either attempt is successful, the list of candidate vertex tracks is then revertexed afresh in three dimensions. This process of revertexing the candidate track list is iterated, each time discarding the track contributing the most to the vertex χ^2, until the worst track contributes less than a χ^2 cut value. (If this condition is still not met by the time the track list has dropped to two, the search for a secondary vertex is deemed a failure.) If a secondary vertex is found, its transverse distance L_{xy} from the primary vertex (P.V.) of the event is calculated, along with the error on L_{xy}. If the L_{xy} *significance* satisfies the condition

$$|L_{xy}/\sigma_{L_{xy}}| \geq 3$$

[10] Another factor contributing to long average decay length for B hadrons in single-top events is that the p_T of the b quarks (Figure 5.3) are large compared to the scale set by the b-quark mass (Ref. [56] explains this for the case of $t\bar{t}$.)

[11] A track having at least two hits in the SVX.

(along with a few other cuts), the jet is termed "tagged".

The sign of L_{xy} indicates the position along the jet axis of the secondary vertex with respect to the primary vertex. Positive-L_{xy} secondary vertices, located in front of the P.V., are consistent with having been produced by a hadron that traveled a distance away from the P.V. before decaying. Negative-L_{xy} vertices, located behind the P.V., are likely to be products of coincidence. Figure 5.15 illustrates the difference between positive-L_{xy} and negative-L_{xy} secondary vertices. The utility of L_{xy} for b tagging is that coincidence is

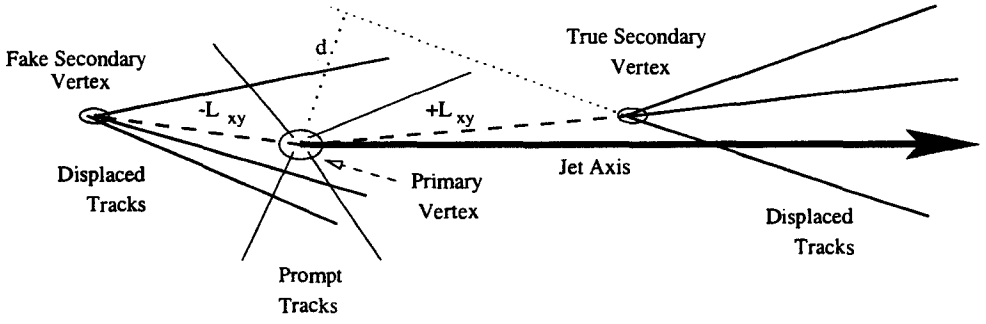


Figure 5.15: Diagram showing real-tag ($L_{xy} > 0$) and fake-tag ($L_{xy} < 0$) reconstructed secondary vertices.

equally likely to produce positive or negative tags in light-quark jets, which contain no true secondary vertices, while heavy-quark jets show a preponderance of positive tags.

We also use another b-tagging algorithm [127] which looks for medium-energy ("soft") electrons or muons, such as would be expected from semileptonic b decay ($b \to c\ell\nu_\ell$). The soft-lepton-tagging (SLT) algorithm has lower efficiency and higher fake rates than the SECVTX algorithm [125]. The only role it plays in this analysis is in identifying soft-lepton-tagged jets for extra jet energy corrections (see Section 5.3.2 above.)

5.3.4 Removals

We remove from our sample events that fall into the following categories:

- **Bad runs:** CDF maintains a run-by-run record

  `run_sum$library:master_exp_run.list`

of which, if any, detector subsystems were nonfunctional while that run's data were collected. We exclude runs in which any subsystem was known to have been nonfunctional.

- Z **candidates:** We remove events which satisfy a loose set of Z-candidate criteria, namely the presence of a second lepton which forms an invariant mass with the primary e or μ in the range $75 < M_Z < 105$ GeV/c^2.

- **Dilepton events:** We employ a two-stage process to cut out background from $t\bar{t}$ events in which both W bosons decay to leptons. First, we remove events that pass the cuts designed to select events for CDF's $t\bar{t}$ dilepton analysis. Next we reduce background from $t\bar{t}$ dilepton events where one lepton fails the lepton-identification cuts by rejecting events with an isolated, high-momentum track with charge opposite to that of the primary lepton (after CDF Note 4112 [128]). The p_T criterion for the tracks is 15 GeV/c and the isolation requirement is $\sum p_T < 2.0$ GeV/c in a cone of radius 0.4 around the candidate track.

To summarize, using CDF top-analysis-code jargon, we apply the following cuts on "TOPFND" variables: BR_CODE = 15 or -1; ZFLAG = 0; DILCD = 0.

5.3.5 Kinematic variables

Finally, we explain the specific techiques used to construct kinematic variables $M_{\ell\nu b}$ and H_T.

The challenge in constructing $M_{\ell\nu b}$, or the invariant mass of the lepton, neutrino, and highest-E_T ("leading") b-tagged jet in the event, *i.e.*

$$M_{\ell\nu b} = \sqrt{(E_\ell + E_\nu + E_b)^2 - (\vec{p_\ell} + \vec{p_\nu} + \vec{p_b})^2}$$

lies in the fact that we need all three components of the neutrino momentum, yet we only have information on *transverse* missing energy ($\not\!\!E_T$).

To get around this problem, we constrain the lepton-neutrino invariant mass to the W-boson mass, 80.22 GeV/c^2 [129] [12]:

$$(E_\ell + E_\nu)^2 - (\vec{p_\ell} + \vec{p_\nu})^2 = M_W^2$$

Since we have measurements for E_ℓ, $\vec{p_\ell}$, $p_{\nu x}$, and $p_{\nu y}$, we can solve the resulting quadratic equation for the unknown $p_{\nu z}$. For the purposes of computing $p_{\nu z}$, we do not equate $(p_{\nu x}, p_{\nu y})$ with the "version 4" muon-corrected $(\not\!\!E_{Tx}, \not\!\!E_{Ty})$ described in Section 5.3.1.3. Instead, we perform a much more careful computation of $(p_{\nu x}, p_{\nu y})$, according to

$$p_{\nu x} = -((E_{lepton}^{corr})_x + \sum_{jets}(E_{jet}^{corr})_x + 1.6 \cdot (E_{uce})_x)$$

Electron energies are corrected with CEMFIX. The sum over jets includes all jets with raw $E_T > 8$ GeV and $|\eta| < 2.4$. Jet E_T's are corrected as described in Section 5.3.2 above. Unclustered energy E_{uce} is defined as the negative sum of {raw lepton E_T, raw E_T of all jets with $E_T > 8$ GeV, and lepton-corrected $\not\!\!E_T$}. It is multiplied by K-factor 1.6. The origin of this factor is explained in [125].

In general, this process yields two solutions for p_z of the neutrino. We pick the solution with the smallest $|p_z|$ value, or, in the case of an imaginary solution, we retain Re(p_z). Once we have p_z, it is straightforward to compute $M_{\ell\nu b}$.

The second key kinematic variable for this analysis, H_T, is constructed by summing the following quantities:

- Primary electron(muon) $E_T(p_T)$;
- Muon-corrected ("version 4") $\not\!\!E_T$;
- Corrected [13] E_T's for all jets with raw $E_T > 8$ GeV and $|\eta| < 2.4$.

[12] This value is a bit out of date. The most recently-available value for the W mass is 80.419 GeV/c^2 [2]. This miniscule inaccuracy does not affect the analysis. A general principle to keep in mind when discovering any sub-optimal feature of one's cuts after the fact is that "cuts are neither right nor wrong, they're smart or dumb." (K. Bloom)

[13] Corrected with flavor-independent jet corrections (JTC96S).

5.4 Signal Acceptance and Yield

Having chosen a set of specific cuts to implement our analysis strategy, we must now determine the numbers of signal events expected to pass the cuts. To do this, we apply the cuts to the Monte Carlo signal samples described in Section 5.1.2 and note the ratio of the number of events passing all cuts to the number of events generated. We define *total acceptance* A_{tot} to be this fraction, $N_{pass}/N_{generated}$. Folded into A_{tot} are all the individual acceptances for the various criteria the event must satisfy. That is, we can factorize A_{tot} as

$$A_{tot} = BR(\text{single top} \to \ell X) \cdot \epsilon_{p_T\ell>20} \times \epsilon_{lep} \times \epsilon_{MET} \times \epsilon_{b-tag} \times \epsilon_{removals} \times \epsilon_{M_{\ell\nu b}}$$

where

- $BR(\text{single top} \to \ell X) \cdot \epsilon_{p_T\ell>20}$

  represents the branching ratio of single top to a final state containing an electron or muon, which includes $BR(W \to e\nu)$, $BR(W \to \mu\nu)$, and extra efficiency [14] from $W \to \tau\nu$ followed by $\tau \to \ell\nu_\ell\nu_\tau$, times the efficiency for the lepton to have detector-measured $p_T > 20$ GeV/c;

  $\equiv N_{p_T\ell>20}/N_{generated}$

- ϵ_{lep}

  represents the efficiency of all the lepton-identification cuts, including geometric, kinematic, quality, and trigger requirements;

  $\equiv N_{p_T\ell>20,\ passing\ lepton\ cuts}/N_{p_T\ell>20}$

- ϵ_{MET}

  represents the efficiency of the requirement $\not\!\!E_T > 20$ GeV;

  $\equiv N_{p_T\ell>20,\ passing\ lepton\ cuts,\ miss.\ E_T>20}/N_{p_T\ell>20,\ passing\ lepton\ cuts}$

[14] The terms *efficiency* and *acceptance* are essentially interchangeable. Both refer to the fraction of events passing a particular cut.

- ϵ_{b-tag}

  represents the efficiency of the requirement that the event have one, two, or three jets with $E_T > 15$ GeV and $|\eta| < 2.0$ and that at least one of these jets be SECVTX-tagged;

  $\equiv N_{\text{... and } \geq 1 \text{ SVX tag}} / N_{\text{... and miss. } E_T > 20}$

- $\epsilon_{removals}$

  represents the efficiency of the Z and dilepton removal requirements;

  $\equiv N_{\text{... and removals}} / N_{\text{... and } \geq 1 \text{ SVX tag}}$

- $\epsilon_{M_{\ell\nu b}}$

  represents the efficiency of the requirement that the event have $140 < M_{\ell\nu b} < 210$ GeV/c^2;

  $\equiv N_{\text{... and } M_{\ell\nu b} \text{ cut}} / N_{\text{... and removals}}$

  $= N_{\text{pass all cuts}} / N_{\text{generated}}$.

With this factorization, it is evident that $A_{tot} = N_{pass}/N_{generated}$.

5.4.1 Data-to-Monte Carlo scale factors

Since no Monte Carlo generator perfectly models every aspect of a given physics process, acceptances measured in a Monte Carlo sample are not exactly correct for real events. For example, in the cleaner Monte-Carlo environment, uncontaminated by multiple interactions, lepton-identification cuts might perform better than they would in a real event. Or else the Monte Carlo might overestimate the track multiplicity in b jets, making the Monte-Carlo-measured b-tagging rate appear too high. We compensate for such known shortcomings of our Monte Carlo by correcting our Monte-Carlo-measured efficiencies by the linear scale factors summarized in Table 5.6 below. With these corrections,

$$\begin{aligned} A_{tot} \;=\; & BR(\text{single top} \to \ell X) \cdot \epsilon_{p_{T\ell}>20} \times \\ & (sf_{lep} \cdot \epsilon_{lep}) \times \epsilon_{MET} \times (sf_{b-tag} \cdot sf_{1A} \cdot \epsilon_{b-tag}) \times \epsilon_{removals} \times \epsilon_{M_{\ell\nu b}} \end{aligned}$$

Effect	Scale factor
Lepton trigger & ID efficiency	$sf_{lep} = 0.94 \pm 10\%$
b-tagging efficiency	$sf_{b-tag} = 1.0 \pm 10\%$
Run 1A SVX correction	$sf_{1A} = 0.97 \pm 1\%$

Table 5.6: Data-to-Monte-Carlo scale factors used to correct acceptances.

5.4.1.1 sf_{lep}: Lepton-trigger and ID-efficiency scale factor

Lepton triggering and identification is more efficient in Monte Carlo than in data. To compensate for this effect, we degrade the Monte-Carlo-measured acceptance ϵ_{lep} by lepton-trigger and ID-efficiency scale factor $sf_{lep} = 0.94$. This factor was derived in CDF Note 3855 [119] by comparing the lepton-ID efficiency measured in $Z \rightarrow ll$ Run 1B data to that measured in a HERWIG $W+ \geq 1$ jet Monte Carlo simulation. The comparison was done separately for electrons and muons. The resulting ratios of efficiencies $\epsilon_{data}/\epsilon_{W+\geq 1j\ MC}$ were averaged over the expected abundances of electrons and muons in a $t\bar{t}$-pair-production sample to yield average scale factor ${sf_{lep}}^{1B}$. (The relative abundances of electrons and muons in a single-top sample ought to be similar enough that the same scale factor applies.) This factor was further degraded to account for differences in W rates between Run 1A and Run 1B to yield the average ${sf_{lep}}^{1A+1B} = 0.94$.

Ref. [119] cites two sources of systematic uncertainty on sf_{lep}: a 9% uncertainty taken from CDF Note 3403 [130] and a 10% uncertainty taken from CDF Note 3481 [131]. In quadrature these yield 13.4%, or ± 0.125. The second, or 10%, uncertainty is ascribed to uncertainty on the trigger efficiency used in the simulation. We retain this 10% uncertainty. However, we neglect the former, or 9%, uncertainty, for the following reasons:

The 9% uncertainty taken from [130] represents the total impact of sources of systematic uncertainty such as jet E_T scale, initial-state radiation (ISR), M_{top}, *etc.* on $t\bar{t}$ acceptance $A_{t\bar{t}}$. If we make the assumptions (a) that ϵ_{data} in $\{sf_{lep} = \epsilon_{data}/\epsilon_{W+\geq 1j\ MC}\}$ is fixed and (b) that $\epsilon_{W+\geq 1j\ MC} \approx \epsilon_{t\bar{t}\ MC}$, then it is reasonable to ascribe this 9% un-

certainty to sf_{lep} [15]. However, in this analysis, we will take into account the single-top acceptance uncertainty introduced by precisely the same list of systematic effects (jet E_T scale, initial-state radiation, M_{top}, *etc.*: see Section 9.2.) To do this *and* include the 9% error on sf_{lep} would be to double-count these uncertainties [16]. Thus we neglect the 9% and use $sf_{lep} = 0.94 \pm 10\%$ as our lepton-trigger and ID-efficiency scale factor.

5.4.1.2 sf_{b-tag}: b-tag efficiency scale factor

The method used to estimate the scale factor for correcting the Monte-Carlo-measured b-tag efficiency changed over the course of CDF Run 1B. An initial study estimated the scale factor to be 0.87 *per tagged event.* However, a later study using improved methods derived the value 1.25 *per tagged jet.* The differences in method and philosophy between these two derivations of the b-tag-efficiency scale factor are summarized in CDF's official internal statement on the matter, CDF Note 4939, *Godparents Report on the New $t\bar{t}$ Cross Section Measurement in the Lepton + Jets Channel Using SECVTX Tags* [132], Section 9. A major difference between these estimates is that the later study assumes the use of "track-degraded" Monte Carlo, in which track-reconstruction performance is deliberately reduced to better mimic tracking in real data. Adopting the new official sf_{b-tag} for this analysis was complicated by the fact that we use *non*-track-degraded Monte Carlo to model both signal and background. The informal recommendation of CDF Top Group experts under these circumstances has been to use b-tag-efficiency scale factor $1.0 \pm 10\%$. In this section we attempt to put this choice on a firmer footing.

CDF Note 4939 defines the data-to-Monte-Carlo b-tagging-efficiency scale factor as

[15] Error propagation gives $\sigma_{sf_{lep}} = \epsilon_{data}/\epsilon_{MC} \times (\sigma_{\epsilon_{MC}}/\epsilon_{MC}) = 9\% sf_{lep}$

[16] The most correct strategy would have been to generate new samples of HERWIG $W+ \geq 1$ jet Monte Carlo, each with a separate systematic effect shifted, to recompute sf_{lep} for each, and to correct our systematically-shifted signal acceptances each by a separate scale factor. Since the systematically-shifted sf_{lep}'s would tend to display the opposite fluctuations as the systematically-shifted signal acceptances, the net impact would be to *reduce* overall systematic fluctuations in signal acceptance. For example, suppose turning off ISR improves lepton-ID efficiency for HERWIG $W+ \geq 1$ jet Monte Carlo. The same is likely true of signal Monte Carlo. Thus both ϵ_{MC} and A_{signal} will fluctuate up, so the product $sf_{lep} \times A_{signal} = (\epsilon_{data}/\epsilon_{MC}) \times A_{signal}$ will fluctuate *less* than if we had used a fixed sf_{lep}. Thus we conclude that our strategy of using a fixed $sf_{lep} = 0.94$ and incorporating systematic effects after the fact is, if anything, conservative.

the quantity in parentheses in the following expression:

$$\epsilon_{data}^{t\bar{t}} = \epsilon_{MC}^{t\bar{t}} \left(\frac{\epsilon_{data}^{ele}}{\epsilon_{MC}^{ele}} \right)$$

$\epsilon_{data}^{t\bar{t}}$, $\epsilon_{MC}^{t\bar{t}}$ are the efficiencies for tagging a b jet in a $t\bar{t}$ event for data and Monte Carlo, respectively. ϵ_{data}^{ele}, ϵ_{MC}^{ele} are the efficiencies for tagging a b jet in a low-p_T electron sample for data and Monte Carlo, respectively. The text explains how to use the scale factor $sf_{b-tag} = (\epsilon_{data}^{ele}/\epsilon_{MC}^{ele})$:

> Using track degraded MC... this scale factor [is] 1.25 ± 0.13 per b-jet for Run 1B... Note that [this] scale factor [is] specifically for the MC package... which includes track degradation. One can use [this] scale factor[ ] if track degradation is used by either the explicit use of the routine that removes tracks from the MC, or by taking account of the effect of track degrading by multiplying by 0.87 ± 0.07... All the error in the b-tagging efficiency is in the above scale factor[ ]. That is, if one had previously used the 0.87 factor and taken the b-tagging efficiency error to be 0.07, then one should now take no error in the 0.87 and use instead the error in the ... [1.25 ± 0.13] scale factor...

This passage tells us to use *per-jet* scale factor 1.25. In addition, we must apply *per-event* scale factor 0.87 since our Monte Carlo is non-track-degraded. Any uncertainty we ascribe to b-tag efficiency should derive solely from that on the 1.25 ± 0.13 scale factor, *i.e.* should be 10.4%.

The first step in using new scale factor 1.25 is to convert this *per-jet* value into a *per-event* scale factor. This must be done separately for $t\bar{t}$, Wg, and W^*, since the tag efficiency will not necessarily be the same for the various b jets in each process. Each per-event scale factor must then be multiplied by 0.87. If the informal Top Group prescription is correct, the final overall per-event scale factors thus obtained ought to be consistent with $1.0 \pm 10\%$.

To derive the per-event scale factor for a particular process, we start from a sample of Monte Carlo events and measure the efficiency to tag each b jet in the event. Suppose the efficiency to tag the first b jet is e_1 and the efficiency to tag the second b jet is e_2. Suppose moreover that the efficiency to tag all other (non-b) jets in the event is zero. The efficiency to tag the entire Monte Carlo *event* is then

$$\text{event tag efficiency} \quad = \quad \text{probability that event has at least one tag}$$

$$
\begin{aligned}
&= \text{single-tag efficiency} + \text{double-tag efficiency} \\
&= e_1(1-e_2) + e_2(1-e_1) + e_1e_2 \\
&= e_1 + e_2 - e_1e_2
\end{aligned}
$$

The desired per-event scale factor is just the event tag efficiency for data divided by the event tag efficiency for Monte Carlo:

$$\text{event scale factor} = \frac{\text{tag efficiency for data event}}{\text{tag efficiency for MC event}}$$

To obtain the tag efficiency for a data event, we start from the expression for the tag efficiency for a Monte Carlo event and correct the tag efficiency for each *jet* (e_i) by the per-jet scale factor 1.25, *i.e.*

$$\text{event scale factor} = \frac{1.25e_1 + 1.25e_2 + (1.25)^2e_1e_2}{e_1 + e_2 - e_1e_2}$$

The overall event scale factors thus derived, along with the Monte Carlo tag efficiencies used to derive them, are recorded in Table 5.7 for $t\bar{t}$, Wg, and W^*. Event scale factors in the third column of the table have been multiplied by the necessary track-degradation factor 0.87.

Process	Tag efficiency for b jet 1 Tag efficiency for b jet 2	Event s.f. *including* $\times 0.87$
$t\bar{t}$	$27.2 \pm 0.2\%$ $26.7 \pm 0.2\%$	1.04 ± 0.09
Wg	$24.2 \pm 0.2\%$ $6.5 \pm 0.1\%$	1.07 ± 0.10
W^*	$24.8 \pm 0.2\%$ $20.5 \pm 0.2\%$	1.05 ± 0.10

Table 5.7: Event b tag scale factors for top processes.

We conclude that the event scale-factor values in Table 5.7 are consistent enough with the informal Top Group prescription of $1.0 \pm 10\%$ that we should continue to use this

value and uncertainty. In doing so we remain consistent with other published CDF analyses that use non-track-degraded Monte Carlo [17].

5.4.1.3 sf_{1A}: Run 1A correction

Roughly 20% of our data was collected with the Run 1A SVX. Radiation damage suffered by this device reduced its b-tagging efficiency to

$$\epsilon^{1A}_{b-tag} = (0.855 \pm 0.044)\epsilon^{1B}_{b-tag} \quad [130]$$

To account for the reduced tagging efficiency in Run 1A data, we perform detector simulation on all Monte Carlo events using the parameters of the Run 1B SVX′, then reduce the tagging efficiency by the luminosity-weighted average of ϵ^{1A}_{b-tag} and ϵ^{1B}_{b-tag}, namely $(0.97 \pm 0.01)\epsilon^{1B}_{b-tag}$ [130].

The 10% uncertainty on the b-tagging-efficiency scale factor, in quadrature with the 1% uncertainty on the Run 1A correction factor, gives overall b-tag-efficiency uncertainty 10%.

5.4.2 Signal acceptance and yield

Scale-factor-corrected efficiencies $(sf_{lep} \cdot \epsilon_{lep})$, ϵ_{MET}, $(sf_{b-tag} \cdot sf_{1A} \cdot \epsilon_{b-tag})$, $\epsilon_{removals}$, and $\epsilon_{M_{l\nu b}}$, along with total corrected acceptance A_{tot}, are given in Tables 5.8 and 5.9. The values for A_{tot} can be used to calculate expected yields in Run 1 data using the relation

$$\text{Yield} = N_{expected} = \sigma_{SM} \cdot \int \mathcal{L}dt \cdot A_{tot}$$

We use dataset size $\int \mathcal{L}dt = 106.0$ pb^{-1} and cross sections σ_{SM} from theory. In the case of electron events, it is valid to scale the product $\sigma_{SM} \cdot A_{total}$ by the full Run 1 luminosity because these events are selected by a hardware trigger [18] which is ~100% efficient [133]. Muon events, however, are selected by an assortment of prescaled triggers which are *not*

[17] We note that, of course, the most-correct approach would be to re-do the entire analysis with track-degraded Monte Carlo and to use the per-event scale factors from Table 5.7 *without* the 0.87 degradation factor.

[18] A combination of CEM_16_CFT_12 and MET_20

100% efficient. We incorporate the reduced effective luminosity for muon events by applying a muon trigger simulation `sim_mutrig` [134] to all Monte Carlo muon events and only counting events which pass this trigger simulation in our measured muon acceptance. The entire expected signal yield in Run 1 data is $N_s = 4.3$ events.

Wg	Acceptance	Yield (# events)
Produced	-	180
Lepton	8.7%	15.6
Lepton+$\not\!\!E_T$	7.5%	13.6
Lepton+$\not\!\!E_T$+b tag	2.2%	4.0
...and Z, dilepton removal	2.2%	3.9
...and mass window cut	1.8%	3.2
...W + 1jet bin	0.44%	0.80
...W + 2jet bin	0.81%	1.45
...W + 3jet bin	0.41%	0.75
TOTAL W+1,2,3 jet bin	**1.7%**	**3.00**

Table 5.8: Acceptance and yield for Wg-fusion single top ($\sigma_{Wg} = 1.7$ pb)

W^*	Acceptance	Yield (# events)
Produced	-	77
Lepton	7.8%	6.0
Lepton+$\not\!\!E_T$	6.8%	5.3
Lepton+$\not\!\!E_T$+b tag	2.5%	2.0
...and Z, dilepton removal	2.5%	1.9
...and mass window cut	1.7%	1.4
...W + 1jet bin	0.36%	0.28
...W + 2jet bin	1.0%	0.79
...W + 3jet bin	0.29%	0.23
TOTAL W+1,2,3 jet bin	**1.7%**	**1.29**

Table 5.9: Acceptance and yield for W^* single top ($\sigma_{W^*} = 0.73$ pb)

The Wg-fusion and W^* signal yields quoted above have both statistical and systematic uncertainties. These uncertainties are studied in Chapter 9 and are found to be 18.5% for μ_{Wg} and 18.6% for μ_{W^*}. (Section 9.2.2.) Thus $\mu_{Wg} = 3.00 \pm 0.56$ and $\mu_{W^*} = 1.29 \pm 0.24$.

Chapter 6

Background Composition and Modeling

Certain categories of non-single-top events are expected to pass the selection cuts described in the previous chapter. We group these surviving background processes under the headings "$t\bar{t}$" and "non-top", where the former category refers to $t\bar{t}$ pair production and the latter to all other QCD and/or electroweak processes that can mimic the single-top signature. In this chapter we describe the Monte Carlo samples we use to model the kinematics of the $t\bar{t}$ and non-top backgrounds and the techniques we use to estimate how many events of each type of background will pass our selection cuts. We conclude the chapter with a table summarizing the respective numbers of signal and background events expected to remain in the final sample after all cuts.

6.1 $t\bar{t}$ Background to Single Top

Since the lepton, $\not\!\!E_T$, and b-tag cuts we use in this analysis were developed for CDF's $t\bar{t}$ search, it is not surprising that these cuts are efficient at retaining $t\bar{t}$ events. We will show in Section 6.2.2 how the additional cuts we apply to jet multiplicity and $M_{\ell\nu b}$ help reject $t\bar{t}$ events. Dilepton removal, *i.e.* removing events with evidence of a second lepton such as would have come from $t\bar{t} \to \ell_1\nu_{\ell_1} b \ell_2 \nu_{\ell_2} \bar{b}$, further reduces background from $t\bar{t}$. Nevertheless $t\bar{t}$ remains a major background to single top, particularly in the two- and three-jet samples.

6.1.1 *Modeling the $t\bar{t}$ background*

To model the kinematic features of the $t\bar{t}$ background, we use a sample of $t\bar{t}$ Monte Carlo events generated by the CDF Top Group using HERWIG 5.6 (Table 6.1). Table 6.2 lists the additional $t\bar{t}$ Monte Carlo samples we use for studies of systematic uncertainty.

Process	# evts generated	Filenames
$t\bar{t}$	210,526	`htt_wsel_175_n.pad, n=a,b`

Table 6.1: Default $t\bar{t}$ background Monte Carlo sample. Generated with HERWIG 5.6.

Process	Effect	# evts generated	Filenames
$t\bar{t}$	$M_{top} = 170$	80,000	`htt_wsel_170.pad`
	$M_{top} = 180$	80,000	`htt_wsel_180.pad`

Table 6.2: $t\bar{t}$ background Monte Carlo samples used for studies of systematic uncertainty. Generated with HERWIG 5.6.

In all these samples the W boson is permitted to decay to any final state, just as was the case for signal Monte Carlo, in order to account for extra acceptance from events with a final-state electron or muon but which did not come from $W \to e\nu_e$ or $W \to \mu\nu_\mu$ (which came from, *e.g.*, $W \to \tau\nu_\tau$.)

6.1.2 *$t\bar{t}$ background prediction*

To determine the number of $t\bar{t}$ background events expected to pass the selection cuts described in Chapter 5, we perform an acceptance calculation exactly as was done for signal in Section 5.4. We apply the cuts to the $t\bar{t}$ Monte Carlo sample, correct the acceptances by the appropriate data-to-Monte Carlo scale factors, multiply by the size of the CDF Run 1 dataset, and normalize to the theoretically-calculated $t\bar{t}$ production cross section, according to

$$N_{t\bar{t},\ expected} = \sigma_{t\bar{t}} \cdot \int \mathcal{L}dt \cdot A_{tot}$$

We use theory cross section $\sigma_{t\bar{t}} = 5.1$ pb [22]. The reason we normalize to a theoretically-calculated value rather than the CDF-measured $t\bar{t}$ cross section $\sigma_{t\bar{t}} = 6.5^{+1.7}_{-1.4}$ pb [135] is because implicit in the CDF measurement are assumptions about the single-top content of data, which is what we seek to measure in this analysis. The issue of $t\bar{t}$ background normalization is discussed more fully in Section 6.2.3.2.

Table 6.3 lists the scale-factor-corrected efficiencies and yields for the $t\bar{t}$ background. We expect a total of 8.4 events from this background.

6.1.3 Uncertainty on $t\bar{t}$ background prediction

Uncertainty on the above $t\bar{t}$ background prediction arises from uncertainty on $\sigma_{t\bar{t}}$, $\int \mathcal{L}dt$, and A_{tot}. With regard to uncertainty on $\sigma_{t\bar{t}}$, Ref. [22] states the result $\sigma_{t\bar{t}} = 5.06^{+0.13}_{-0.36}$ pb at $\sqrt{s} = 1.8$ TeV and $M_{top} = 175$ GeV/c^2, *i.e.* that there is 7% uncertainty on $\sigma_{t\bar{t}}$, which is ascribed to scale dependence. However, the calculated value for $\sigma_{t\bar{t}}$ is a function of top quark mass, as is shown in Figure 6.1. The Run 1 Tevatron measurement for the top quark mass is $M_{top} = 174.3 \pm 5.1$ GeV/c^2 [49]. From the plot in Figure 6.1 we estimate that this ±5-GeV uncertainty on M_{top} introduces an additional uncertainty on $\sigma_{t\bar{t}}$ of ~17% [1]. In quadrature with the 7% this gives ~18%.

Next we must include uncertainty due to $\int \mathcal{L}dt$ and A_{total}. Since our analysis cuts closely parallel those of CDF's standard $t\bar{t}$ analysis [136], we use the same acceptance uncertainty derived there (13.5% [2]), except rounded up to 20% to account for integrated-luminosity uncertainty. Lastly, we take into account the fact that the standard CDF $t\bar{t}$ analysis does not include a cut on $M_{\ell\nu b}$ as we do in this analysis. To estimate the additional uncertainty introduced by this cut, we recompute the $t\bar{t}$ background prediction from HERWIG samples (listed in Table 6.2) in which $M_{top} = 170$, 180 GeV/c^2 respectively. We compare the shifted-M_{top} background predictions to the nominal and take the larger per-

[1] To estimate the additional uncertainty introduced by δm_t, we note the values of $\sigma_{t\bar{t}}$ that correspond to $174 + \delta m_t$ and $174 - \delta m_t$ and take a generous interpretation of the average difference from the nominal $\sigma_{t\bar{t}}$, *i.e.* ~17%.

[2] Table II, p. 12, of [136] cites $\epsilon_{total} = 0.037 \pm 0.005$ in the SVX-tagged lepton+jets channel, or 13.5%.

$t\bar{t}$	Acceptance	Yield (# events)
Produced	-	540.6
Lepton+$\not\!\!E_T$	**14.9%**	**80.68**
... W + 0jet bin	0.03%	0.15
... W + 1jet bin	0.66%	3.59
... W + 2jet bin	3.2%	17.48
... W + 3jet bin	5.2%	28.22
... $W+ \geq$ 4jet bin	5.8%	31.23
Lepton+$\not\!\!E_T$+b-tag	**6.6%**	**35.76**
... W + 1jet bin	0.18%	0.97
... W + 2jet bin	1.3%	7.04
... W + 3jet bin	2.3%	12.54
... $W+ \geq$ 4jet bin	2.8%	15.21
After Z removal	**6.5%**	**35.06**
... W + 1jet bin	0.17%	0.94
... W + 2jet bin	1.3%	6.84
... W + 3jet bin	2.3%	12.25
... $W+ \geq$ 4jet bin	2.8%	15.04
After dilepton removal	**5.5%**	**29.91**
... W + 1jet bin	0.11%	0.60
... W + 2jet bin	0.86%	4.62
... W + 3jet bin	1.9%	10.37
... $W+ \geq$ 4jet bin	2.6%	14.32
After mass window cut	**3.2%**	**17.37**
... W + 1jet bin	0.04%	0.21
... W + 2jet bin	0.42%	2.27
... W + 3jet bin	1.1%	5.90
TOTAL W+1,2,3 jet bin	**1.6%**	**8.38**

Table 6.3: $t\bar{t}$ background prediction.

centage difference (17%) as the additional $t\bar{t}$-acceptance uncertainty due to the $M_{\ell\nu b}$ cut. Adding these contributions in quadrature (18% $\oplus$ 20% $\oplus$ 17%), we assign total uncertainty 32% to our $t\bar{t}$ background prediction.

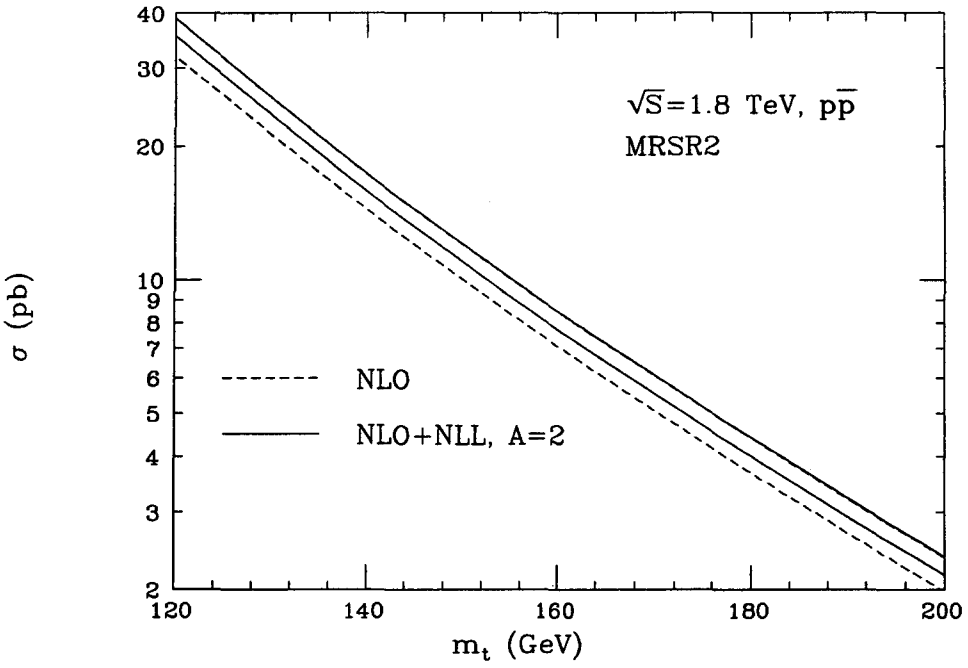


Figure 6.1: Total $t\bar{t}$ production cross-section in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV, as a function of the top-quark mass. Dashed lines: NLO result; solid lines: NLO+NLL result. Upper lines: $\mu = m_t/2$; lower lines: $\mu = 2m_t$. Figure and above caption text taken from [22]. We use this plot to estimate the additional uncertainty on $\sigma_{t\bar{t}}$ introduced by top-mass uncertainty.

6.2 Non-Top Background to Single Top

Unlike the $t\bar{t}$ background, which consists of a single well-defined process (strong-interaction $t\bar{t}$ pair production), the background we call "non-top" is a collection of disparate processes whose only common feature is that they all possess the signature of—though perhaps not the actual presence of—a leptonically-decaying W boson and a b jet. Such events can survive our lepton, $\not\!\!E_T$, and b-tag requirements and contaminate the final sample.

The largest contributor to the non-top background is direct production, through QCD and electroweak vertices, of a real W boson along with heavy-flavor quarks. Tree-level diagrams for direct W+heavy flavor production, which we call "$Wb\bar{b}/Wc\bar{c}$" and "Wc", are shown in Figures 6.2 and 6.3. Analogous processes in which a leptonically-decaying

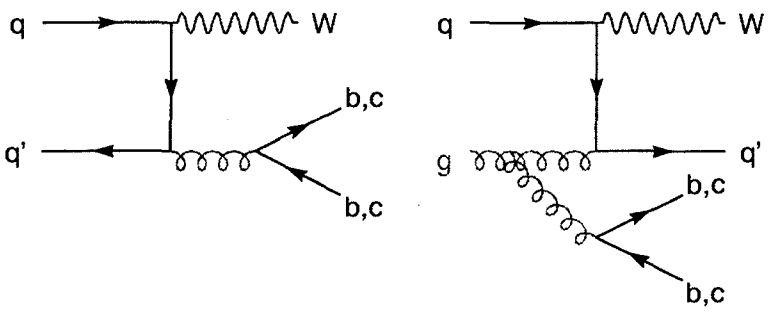


Figure 6.2: Sample tree-level diagrams for $Wb\bar{b}/Wc\bar{c}$ production.

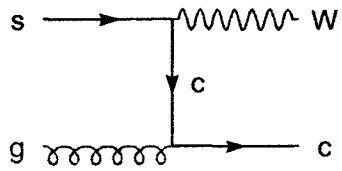


Figure 6.3: Diagram for Wc production.

Z boson is substituted for the W (and one decay lepton is lost) also contribute, though to a much smaller degree. In addition to direct W/Z+heavy flavor production, small contributions come from "diboson" production (WW, WZ, ZZ) where one boson decays to a lepton and the other to heavy-flavor quarks (Figure 6.4) and from $q\bar{q} \rightarrow Z \rightarrow \tau\tau$, since the τ lifetime can lead to a secondary vertex.

The remaining contributions to the non-top background come from events which fake

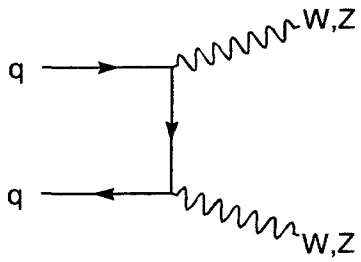


Figure 6.4: Diagram for diboson production.

the signature of either an electroweak boson or a heavy-flavor quark. The fake-electroweak-boson type of event, known as "non-W" (where we really mean "non-W or Z") contains a lepton and missing E_T from some other source, such as from direct $b\bar{b}$ production followed by semileptonic decay of one b. While it would be unusual for this process to give rise to a lepton sufficiently energetic and isolated to pass our cuts, $b\bar{b}$ production is copious at a hadron collider. Finally, the fake-b-quark category of events, known as "mistags", are simply W+light flavor events (substitute light quarks for the b's or c's in Figures 6.2 and 6.3) in which a light-quark jet is mistakenly identified as a b by the SECVTX b-tagging algorithm. So abundant is the production of W+light-quark jets that even a small SECVTX fake rate makes for a significant background from mistags.

6.2.1 Modeling the non-top background

To model the kinematic features of the non-top background, we use a VECBOS [137]-generated sample (Table 6.4) of W + 2-jet events in which the two quarks were forced to be $b\bar{b}$. Events were sent through HERPRT to model parton showering and fragmentation. In this sample the W boson was forced to decay to $e\nu_e$. However, since—in contrast to the

Process	# evts generated	Filenames
$Wb\bar{b}$	107,973	`w2bmrgbb_vec_n.hqfl, n=01,...,50`

Table 6.4: $Wb\bar{b}$ Monte Carlo sample (Courtesy of F. Ptohos). Generated by VECBOS with force $b\bar{b}$ option; used to model kinematic features of *all* sources of non-top background.

cases of signal and $t\bar{t}$ Monte Carlo—we will not use this sample to calculate an acceptance for the non-top background but only to model its kinematic behavior, we need not worry about accounting for the $W \to \mu\nu_\mu$ channel or extra acceptance from $W \to \tau\nu_\tau$. Instead we will normalize to a standard CDF non-top background estimate known as the "Method 2 background calculation." The subject of non-top background normalization is addressed in Section 6.2.3.

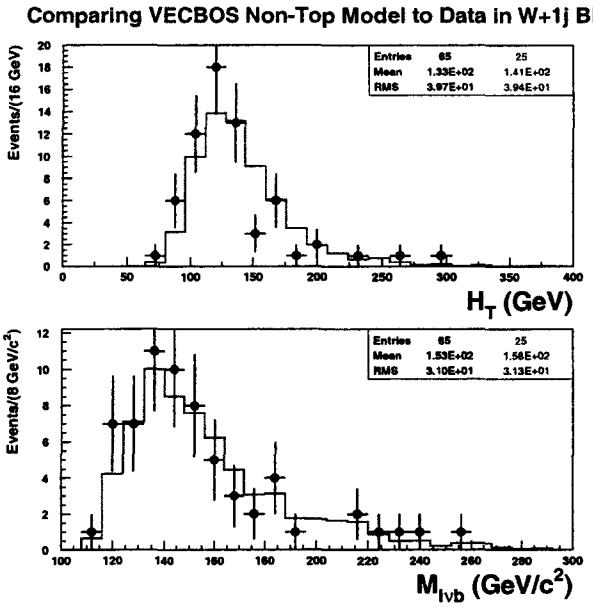


Figure 6.5: Check of VECBOS non-top background model in the $W + 1j$ bin after all cuts except $M_{\ell\nu b}$. Points are CDF data, line is VECBOS. VECBOS distributions normalized to CDF Method 2 prediction (65 events.) K-S probability of agreement for H_T distributions: 2%; for $M_{\ell\nu b}$ distributions, 29%.

6.2.1.1 Validity checks of non-top background modeling

The choice to use VECBOS $Wb\bar{b}$ Monte Carlo to model all sources of non-top background is reasonable, since $Wb\bar{b}/Wc\bar{c}$ production is the largest single contributor to the non-top background [3]. However, we cannot necessarily expect $Wb\bar{b}$ to reproduce accurately the kinematics of such diverse processes as WW, Wc, and direct $b\bar{b}$. To check that VECBOS $Wb\bar{b}$ adequately models the mix of processes that make up the non-top background in data, we compare two relevant kinematic distributions predicted by VECBOS to the corresponding distributions in data after all cuts except $M_{\ell\nu b}$ (Figure 6.5.) We restrict our attention to the $W + 1j$ bin, where the top content of data is expected to be very small. The good agreement we observe between data and Monte Carlo model leads us to conclude that our use of VECBOS to model all non-top backgrounds is acceptable.

[3]Section 6.2.3, Table 6.6 lists the bin-by-bin predictions for all processes contributing to the Method 2 background calculation.

In Chapter 9 we assign a systematic uncertainty associated with using VECBOS $Wb\bar{b}$ to model all non-top backgrounds. We estimate this "Background Model" systematic by assessing the impact on our measured $\sigma_{single\ top}$ of switching between an all-VECBOS-$Wb\bar{b}$ background model and a model that combines VECBOS $Wb\bar{b}$ and PYTHIA Wc. Table 6.5 lists the Wc Monte Carlo files that were generated for this purpose.

Process	# evts generated	# passing filters	Filenames
Wc	4,210,372	30,000	`pwc_n.pad, n=1,...,4`

Table 6.5: Wc Monte Carlo sample used for studies of systematic error on non-top background model. Generated using PYTHIA 5.6 process $fg \to f'W$. Two GENP filters applied: the usual requirement of a 15-GeV lepton with $|\eta| \leq 3.0$ plus the additional requirement of a 10-GeV c or $\bar{c}$ quark with $|\eta| \leq 2.0$.

6.2.2 Background kinematics

Figures 6.6 and 6.7 demonstrate the usefulness of the jet-multiplicity cut in rejecting $t\bar{t}$ and the $M_{\ell\nu b}$-window cut in rejecting both $t\bar{t}$ and non-top background events. Figure 6.6 shows the jet multiplicity in signal and background events after W selection and b tagging. Note that while signal resides chiefly in the one-, two-, and three-jet multiplicity bins, the $t\bar{t}$ background prefers the three-jet and higher multiplicity bins.

Figure 6.7 shows the invariant mass $M_{\ell\nu b}$ of the lepton, neutrino, and leading b-tagged jet for signal and background events after W selection and b tagging. The cut window $140 < M_{\ell\nu b} < 210$ GeV/c^2 clearly discriminates against the non-top background. Note that it is also effective against the $t\bar{t}$ background, despite the fact that real top quarks are present in $t\bar{t}$ and single-top events alike. This is because the more-complicated $t\bar{t}$ final state (nominally $\ell\nu_\ell b\bar{b}qq'$) reduces the chances that the ℓ, ν_ℓ, and leading b-tagged jet will have all come from the same t quark, broadening the distribution of reconstructed $M_{\ell\nu b}$ compared to single top. This effect was first anticipated in Ref. [13].

Figure 6.8 shows the distributions of the H_T variable for signal and background

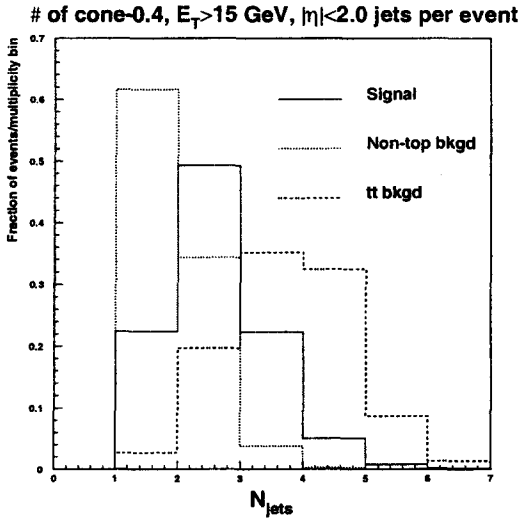


Figure 6.6: Jet multiplicity in signal and background events. Solid line is single-top signal (weighted sum of Wg and W^*); dashed is $t\bar{t}$; dotted is non-top. All histograms normalized to 1. The jet multiplicity ≤ 3 cut helps reject $t\bar{t}$.

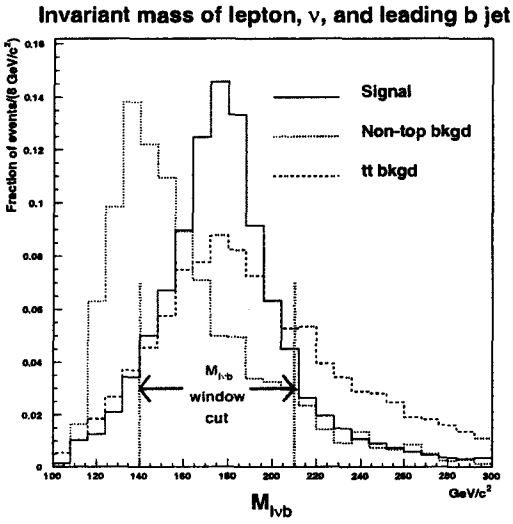


Figure 6.7: Reconstructed mass of lepton, neutrino, and leading b jet in signal and background events. Solid line is single-top signal (weighted sum of Wg and W^*); dashed is $t\bar{t}$; dotted is non-top. All histograms normalized to 1. Note that the $M_{\ell\nu b}$ cut window discriminates against both $t\bar{t}$ and non-top backgrounds.

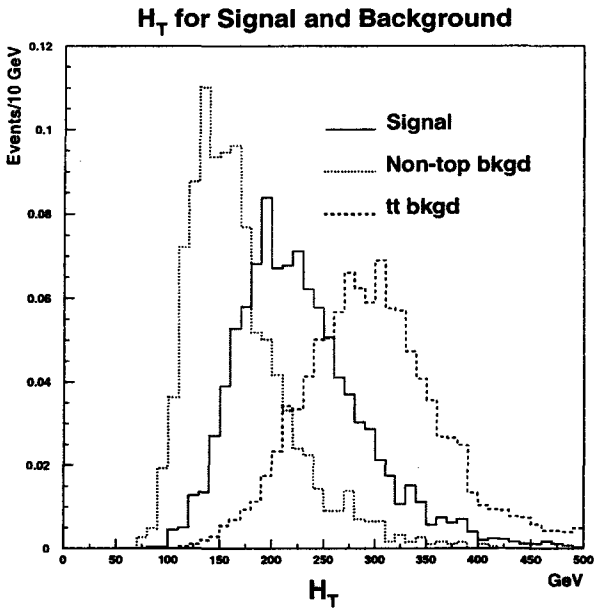


Figure 6.8: H_T distributions for signal and background after all analysis cuts. Solid line is single-top signal (weighted sum of Wg and W^*); dashed is $t\bar{t}$; dotted is non-top. All histograms normalized to 1.

after all analysis cuts. This plot underscores the usefulness of H_T for a combined single-top search. We saw in Figure 5.12 that the distributions for Wg and W^* signal are virtually indistinguishable; here we see that signal H_T is also reasonably separated from backgrounds.

6.2.3 Non-top background prediction

We now address the topic of normalization of the non-top background. While for signal and for the $t\bar{t}$ background we normalized Monte-Carlo-calculated acceptances to theoretically-predicted cross sections, this is not feasible for the diverse collection of processes we call "non-top". (While, *e.g.*, the theoretical cross section for diboson production is known, the concept of cross section for the mistag background is not well-defined.) Instead, to estimate the non-$t\bar{t}$ background to single top, we rely on the background calculation

from another CDF analysis which had the same lepton, $\not\!\!E_T$, b-tag, Z, and dilepton cuts we use [128].

6.2.3.1 CDF "Method 2" calculation

This calculation, first described in CDF internal notes [138] and [139] and known as the "Method 2" background calculation, relies on a mixture of techniques to estimate the various contributions to the non-top background. Contributions from processes with well-defined cross sections, such as diboson production, are calculated from Monte Carlo normalized to theory predictions just as is done in this analysis for signal and for $t\bar{t}$ background. The $Z \to \tau\tau$ contribution is estimated from Monte Carlo normalized to the CDF-measured Z cross section. Contributions from processes with less-well-known (in the case of W+heavy flavor) or ill-defined (in the case of the mistag and non-W backgrounds) cross sections are estimated with reference to the data.

The direct W/Z+heavy flavor rates are obtained from Monte Carlo normalized to data. In the case of $Wb\bar{b}$, Monte Carlo is used to derive two ratios, $F1_{Wb\bar{b}}^{[Njet]}$ and $F2_{Wb\bar{b}}^{[Njet]}$, representing the fraction of direct W+jets events in the N_{jet} bin containing, respectively, exactly one or two final-state jets with (a) $E_T > 15$ GeV and (b) a GENP-level b hadron within a conesize of $\Delta R = 0.4$ about the jet axis. Since the available Monte Carlo models have complementary strengths and weaknesses, different approaches—and different Monte Carlo programs—are used to calculate the two ratios: HERWIG W+1 jet with parton shower to derive $F1$, and VECBOS with forced $b\bar{b}$ to derive $F2$. Tagging efficiencies $\epsilon 1_{Wb\bar{b}}^{[Njet]}$ and $\epsilon 2_{Wb\bar{b}}^{[Njet]}$ are also derived from Monte Carlo for the one- and two-b-jet classes of event. Then these ratios are normalized to the number of W+jets events in data to obtain the expected number of tagged $Wb\bar{b}$ events in each jet-multiplicity bin, according to

$$N_{Wb\bar{b}}^{[Njet]} = N_W^{[Njet]} \times \left\{ F1_{Wb\bar{b}}^{[Njet]} \times \epsilon 1_{Wb\bar{b}}^{[Njet]} + F2_{Wb\bar{b}}^{[Njet]} \times \epsilon 2_{Wb\bar{b}}^{[Njet]} \right\}$$

In the above expression $N_W^{[Njet]}$ is the number of events from direct W+jets production in the $Njet$ data bin. Direct W+jets is not the only source of W candidates in data, however. To calculate N_W, we count the total number N_{data} of W candidates in data and subtract the

contributions from all sources other than direct W+jets production. These other sources include non-W, diboson, $Z \to \tau\tau$, and top:

$$N_W^{[Njet]} = N_{data}^{[Njet]}(1 - F_{non-W}^{[Njet]}) - N_{diboson}^{[Njet]} - N_{Z\to\tau\tau}^{[Njet]} - N_{top}^{[Njet]}\ ^4$$

$F_{non-W}^{[Njet]}$ is the fraction of W candidates that are not real W events. Estimation of the non-W background is discussed below.

The calculation of the $Wc\bar{c}$ background, $N_{Wc\bar{c}}^{[Njet]}$, is exactly analogous. The Wc calculation is also similar, namely

$$N_{Wc}^{[Njet]} = N_W^{[Njet]} \times \left\{ F_{Wc}^{[Njet]} \times \epsilon_{tag}^{[Njet]} \right\}$$

where F_{Wc} is the fraction of W+jet events from Wc, estimated from HERWIG W+1 jet with parton shower. Calculation of the analogous direct Z+heavy flavor rates is essentially similar, though with a few additional complications [139]. The crucial point regarding all these calculations is that while Monte Carlo simulation is trusted to provide the *fraction* of the $W + Njet$ bin that each type of event occupies, the overall *normalization* must be obtained from data.

For the backgrounds we call mistag and non-W, no adequate Monte Carlo model even exists, so these rates are estimated entirely from data. Mistags are estimated by parameterizing, as a function of E_T and track multiplicity, the probability that a light-quark jet will be misidentified as a b. This parameterization is derived from the $-L_{xy}$ tagging rate in generic-jet events. (Of course, only vertices with $+L_{xy}$ are counted as true b tags. But the rate of $+L_{xy}$ tags in generic jets is not a good measure of the mistag rate, because it consists of both mistags and real tags from the real heavy flavor in the sample. $-L_{xy}$ tags, however, are all fake since they come from a coincidence of tracks rather than a real secondary vertex. Light-quark jets are equally likely to have $-L_{xy}$ as $+L_{xy}$ tags. Thus

[4]This expression is a simplification—the LHS should really be $N_W + N_Z$, where N_Z is the number of Z+jets events where the Z decays to $\ell\ell$ and is mistaken for a W. The equation then becomes

$$N_W = \frac{N_{data}(1 - F_{non-W}) - N_{diboson} - N_{Z\to\tau\tau} - N_{top}}{1 + N_Z/N_W}$$

where the ratio N_Z/N_W is obtained from Monte Carlo [140].

the $-L_{xy}$ rate in generic jets is an approximation of the $+L_{xy}$ rate for light-quark jets, *i.e.* the mistag rate.) Thus to obtain the mistag contribution for a given set of data events, this mistag probability is computed for each jet in each event and summed over all events in the sample.

Finally, the non-W background is estimated using a technique known as the "Isolation *vs.* Missing E_T" method [141]. The scatterplot for pretagged events of isolation of the primary lepton *vs.* $\not\!\!E_T$ of the event is divided into four regions:

- Region A: $\not\!\!E_T \leq 15$ GeV and iso ≤ 0.1
- Region B: $\not\!\!E_T \leq 15$ GeV and iso ≥ 0.2
- Region C: $\not\!\!E_T \geq 20$ GeV and iso ≥ 0.2
- Region D: $\not\!\!E_T \geq 20$ GeV and iso ≤ 0.1 (W signal region)

The assumptions are made (1) that real-W events are confined to Region D and (2) that non-W events are distributed among all four regions with no correlation between isolation and $\not\!\!E_T$. With these assumptions, an estimate of the number $N_{D,non-W}$ of non-W background events in the W-signal region is provided by $(N_A \times N_C)/N_B$. The non-W fraction in this region F_{non-W} (which is needed to calculate N_W) is then $N_{D,non-W}/N_D$, or $F_{non-W} = \frac{(N_A \times N_C)/N_B}{N_D}$. Next the tagging rate, defined as the ratio of tagged jets to taggable jets, is computed for Region A, the region most analogous to the signal region. By counting the number of taggable jets in Region D and multiplying by the tagging rate and F_{non-W}, the number of tagged non-W events in Region D is estimated [5].

The Method 2 calculations of [128] for the contributions of the various sources of non-top background in the tagged W+jets sample after Z and dilepton removal are given in Table 6.6.

[5] Since (taggable jets)×(tagging rate) is in units of *tagged jets* and F_{non-W} is in units of *events*, Ref. [141] seems to have left out a piece of information on the number of tagged jets per tagged event (perhaps assumed to be 1?), but the general idea can be discerned.

Source	W+1jet	W+2jet	W+3jet
Mistags	18 ± 7	6.3 ± 2.6	1.6 ± 0.6
$Wb\bar{b}/Wc\bar{c}$	18 ± 7	9.0 ± 2.5	1.7 ± 0.5
Wc	17 ± 6	4.3 ± 1.5	0.6 ± 0.2
$Zb\bar{b}/Zc\bar{c}$, Zc	1.5 ± 0.5	0.7 ± 0.3	0.2 ± 0.1
Diboson, $Z \to \tau\tau$	1 ± 0.4	1.4 ± 0.5	0.3 ± 0.1
Non-W	9 ± 4	2.3 ± 1.0	0.8 ± 0.4
$\geq$ 1-SVX-tag events	65 ± 13	24 ± 4.5	5.2 ± 1.1
After mass window cut	37.4	13.9	2.7
TOTAL W+1,2,3j	**54 ± 12 events**		

Table 6.6: Numbers of tagged non-top background events in CDF Run 1 data predicted by Method 2 calculation of [128]. The mass window cut is discussed in Section 6.2.3.3. The ± 12 uncertainty cited in the last line of the table is explained in Section 6.2.4.

6.2.3.2 Dependence of Method 2 numbers on assumptions about single-top content of data

It must be noted that CDF's Method 2 background calculation is dependent on assumptions about the single-top content of data for the same reason CDF's $t\bar{t}$ cross section measurement was. However, we will show in this section that while the single-top assumptions could not be neglected in the case of the $t\bar{t}$ prediction, they can in the case of the non-top prediction, so that it is still valid to use the Method 2 numbers despite their dependence on these assumptions.

CDF measures the $t\bar{t}$ cross section using an iterative procedure [142] inextricably linked to the Method 2 background calculation. First, a Method 2 calculation is performed in which $\sigma_{t\bar{t}}$ is taken to be 0.0 pb, *i.e.* $t\bar{t}$ production is assumed to contribute zero events to N_{data}, the total number of W candidates in data. Therefore to calculate N_W (the number of events from direct W+jet production, which appears in the expressions for $N_{Wb\bar{b}}$, $N_{Wc\bar{c}}$, and N_{Wc}), we count N_{data} and subtract the contribution from all sources other than direct W+jets production—non-W, diboson, $Z \to \tau\tau$, single top, *but not* $t\bar{t}$: $N_W = N_{data}(1 - F_{non-W}) - N_{diboson} - N_{Z\to\tau\tau} - N_{single\ top}$. From this N_W, all the various Method 2 backgrounds ($N_{Wb\bar{b}}$, $N_{Wc\bar{c}}$, *etc.*) are calculated. The excess of tagged data events

over this predicted background in the $W+ \geq 3$-jet bin is used to estimate $\sigma_{t\bar{t}}$. The Method 2 calculation is redone using this new, nonzero value of $\sigma_{t\bar{t}}$; and the process is iterated until a stable value of $\sigma_{t\bar{t}}$ is reached.

Since $N_{single\ top}$ is among the non-W+jets contributions that must be subtracted from N_{data} in the above calculation of N_W, both the Method 2 background and the measured $t\bar{t}$ cross section depend on assumptions about single top. Typically, $N_{single\ top}$ is obtained from Monte Carlo normalized to a theory prediction for $\sigma_{single\ top}$. This poses a problem for an analysis attempting to measure $\sigma_{single\ top}$. However, it can be shown with some algebra that while the predicted number of $t\bar{t}$ events $N_{t\bar{t}}$ depends fairly sensitively on the single-top cross section, this is not true of the predicted number of non-top events N_{nt}. Appendix A of Ref. [143] proves this for the W+2-jet bin after tagging. Defining x to be the ratio between the single-top cross section in the data to that predicted by theory, Ref. [143] uses information on the pre- and post-tag composition of the $W+2$-jet and $W+ \geq 3$-jet samples to derive expressions for the $t\bar{t}$ and non-top normalizations in the tagged $W+2$-jet sample as a function of x. Ref. [143] shows [6] that in the tagged W+2-jet bin,

$$N_{t\bar{t},2,SVX} = 4.12 - 0.15 \times (x-1)$$

$$N_{nt,2,SVX} = 25.2 - 0.11 \times (x-1)$$

[6] Ref. [143] defines six variables:

- $N_{WZj,\geq 3}$, number of W,Z+jets events in the pretag $W+ \geq 3$ sample;
- $N_{t\bar{t},\geq 3}$, number of $t\bar{t}$ events in pretag $W+ \geq 3$;
- $N_{WZj,2}$, number of W,Z+jets events in pretag W+2;
- $N_{t\bar{t},2,SVX}$, number of $t\bar{t}$ events in tagged W+2;
- $N_{nt,2,SVX}$, number of non-top events in tagged W+2;
- x, the ratio of single-top cross section in data to theory;

and writes five equations using these variables, for the following quantities:

- number of events in pretag $W+ \geq 3$;
- number of events in tagged $W+ \geq 3$;
- number of events in pretag $W+2$;
- number of $t\bar{t}$ events in tagged $W+2$;
- number of non-top events in tagged $W+2$;

using information from the Method 2 pre- and post-tag tables of Ref. [142]. Next, Ref. [143] solves for $N_{t\bar{t},2,SVX}$ and $N_{nt,2,SVX}$ in terms of x.

If we increase x from 1 to 10, *i.e.* if the data contains ten times more single top than expected, $N_{t\bar{t}}$ changes by 33% while N_{nt} changes by only 4%. Clearly it is acceptable to neglect the non-top prediction's dependence on an assumed single-top cross section but not the $t\bar{t}$ prediction's. While [143] makes this argument for the W+2-jet bin only, we generalize the procedure of using pre- and post-tag information from bins N and $N+1$ to derive expressions for the $t\bar{t}$ and non-top normalizations in the tagged bin N as a function of x to the cases of $N=1$ and $N=3$, *i.e.* for the $W+1$-jet and $W+3$-jet bins [7]. Summing the results across all three bins, we show that the same conclusion holds for the SVX-tagged 1, 2, or 3-jet bin: that it is safe to neglect the non-top prediction's single-top dependence but not the $t\bar{t}$ prediction's.

6.2.3.3 Effect of $M_{\ell\nu b}$ cut

The only remaining obstacle to using the Method 2 numbers of [128] is the fact that that analysis does not apply a cut on reconstructed top mass $M_{\ell\nu b}$ as we do. To estimate the amount of non-top background that survives this additional cut, we measure the efficiency of this cut on our VECBOS $Wb\bar{b}$ sample. We measure the efficiency separately for each jet-multiplicity bin. These efficiencies are summarized in Table 6.7 and applied in the second-to-last line of Table 6.6. That is, we scale the total prediction in each bin

[7] In the $W+1$-jet bin, we obtain the equations

$$\begin{aligned} N_{t\bar{t},\,1,\,SVX} &= 0.95 - 1.03 \times (x-1) \\ N_{nt,\,1,\,SVX} &= 63.4 - 0.064 \times (x-1) \end{aligned}$$

and in the $W+3$-jet bin,

$$\begin{aligned} N_{t\bar{t},\,3,\,SVX} &= 9.71 - 0.118 \times (x-1) \\ N_{nt,\,3,\,SVX} &= 5.74 - 0.032 \times (x-1) \end{aligned}$$

Summing the results across all three bins, we obtain

$$\begin{aligned} N_{t\bar{t},\,123,\,SVX} &= 14.75 - 1.41 \times (x-1) \\ N_{nt,\,123,\,SVX} &= 94.3 - 0.21 \times (x-1) \end{aligned}$$

so that a tenfold increase in the single-top cross section reduces the $t\bar{t}$ prediction for the $W+$ 1, 2, or 3-jet bin by 86% but the non-top prediction by only 2%.

Bin	$\epsilon_{M_{\ell\nu b}\ cut}$
W + 1jet	57.6%
W + 2jet	57.9%
W + 3jet	52.0%

Table 6.7: Bin-by-bin efficiency of $M_{\ell\nu b}$ cut for non-top background.

(third-to-last line of Table 6.6) by the appropriate efficiency from Table 6.7:

$$N_{M2\ after} = \epsilon_{M_{\ell\nu b}\ cut} \cdot N_{M2\ before}$$

We expect a total of 54 events from non-top backgrounds.

6.2.4 Uncertainty on non-top background prediction

To obtain the total non-top background prediction, we add across the jet-multiplicity bins of Table 6.6. This yields 94.2 events before the cut on $M_{\ell\nu b}$. To estimate the uncertainty on this value, we must consider the impact of bin-to-bin correlations. While the correlation matrix for the background uncertainties found in [128] is not available, the correlation matrix for the uncertainties of a very similar Method 2 background table may be found in [139]. A portion of this correlation matrix is reproduced below (Table 6.8). Assuming

W + 1jet	W + 2jet	W + 3jet
1.00000	0.85748	0.78414
0.85748	1.00000	0.93922
0.78414	0.93922	1.00000

Table 6.8: Correlation matrix for Method 2 background uncertainties. From [139].

these correlations are approximately valid for the Method 2 calculation of [128], we derive [8] that the uncertainty on the sum of the three jet-multiplicity bins is ±18.0, or 19.1%. (Note

[8] We use the fact that for $f(x_1, x_2, x_3) = x_1 + x_2 + x_3$, $\delta f^2 = \sigma_{x_1}^2 + \sigma_{x_2}^2 + \sigma_{x_3}^2 + 2\rho_{x_1x_2}\sigma_{x_1}\sigma_{x_2} + 2\rho_{x_1x_3}\sigma_{x_1}\sigma_{x_3} + 2\rho_{x_2x_3}\sigma_{x_2}\sigma_{x_3}$.

that if we neglected bin-to-bin correlations, we would have found the uncertainty on the total value to be ±13.8, or 14.7%.)

We must also include uncertainty associated with the additional $M_{\ell\nu b}$ cut. We measure the efficiency of this cut using VECBOS $Wb\bar{b}$ and assume that this efficiency is valid for all sources of non-top background. As an estimate of the additional uncertainty this introduces, we test the effect on the efficiencies in Table 6.7 of shifting the $M_{\ell\nu b}$ 35-GeV/c^2 cut window, nominally centered at 175 GeV/c^2, up or down by 5 GeV/c^2. We find this introduces an additional uncertainty on the 54 events of ±6 events, or 11%. Adding this in quadrature with the previous 19% uncertainty, we assign 22% uncertainty to our non-top background prediction.

6.3 Summary: S/B

The predicted amounts of signal and background, along with their uncertainties, are summarized in Table 6.9. We conclude that after applying our selection cuts to CDF

Process	$W+1$ jet	$W+2$ jet	$W+3$ jet	$W+1,2,3$ jet
Wg	0.80	1.45	0.75	3.00 ± 0.56
W^*	0.28	0.79	0.23	1.29 ± 0.24
$t\bar{t}$	0.21	2.27	5.90	8.38 ± 2.68
non-top	37.4	13.9	2.7	54 ± 12
Total	38.7	18.4	9.6	66.7

Table 6.9: Predicted numbers of signal and background events after analysis cuts.

Run 1 data, we expect roughly 4.3 single top signal events to remain, on a background of 62 non-top and $t\bar{t}$ events.

These predictions for S and B can then be used to determine *a priori* whether we expect to have sufficient statistical sensitivity to observe single top, and if so, then with what accuracy we might be able to determine a benchmark Standard-Model parameter such as V_{tb}. As Ref. [20] points out, the figure of merit is not the same for these two goals:

> The statistics for discovering a signal are different from those for measuring its cross section. To claim a discovery, one needs to demonstrate that the signal is not consistent with a fluctuation in the background. The discovery significance is therefore governed by... $S/\sqrt{B}$. On the other hand, the accuracy with which a cross section can be measured is limited by the fluctuation in the total number of expected events in the signal region, $S + B$. Thus the fractional uncertainty on the measured cross section is $\sqrt{S + B}/S$.

The *a priori* numbers we obtain for S and B yield a signal-to-background ratio of approximately $1 : 15$; $S/\sqrt{B} = 0.5$. Clearly, the CDF Run 1 data sample does not have sufficient sensitivity to observe Standard-Model single-top production. Therefore in this analysis we will set an upper limit on the cross section.

Chapter 7

Fit Method

In previous chapters we discussed a kinematic variable, H_T, which is identical for Wg and W^* signal yet well-separated from backgrounds. In this chapter we describe how the H_T distribution can be used to measure the signal content of data. We construct a distribution which is a superposition of the single-top, $t\bar{t}$, and non-top H_T shapes and find the single-top and background normalizations such that if the N events observed in data were chosen from respective single-top and background Poisson distributions having those means, the likelihood of observing these particular data would be maximized. This technique is called "maximum likelihood fitting" and we speak of "fitting the data" to extract the single-top content. After describing the maximum-likelihood fit technique, we describe checks of fit performance. Finally, we note that when we apply this fit to CDF Run 1 data, we expect from the S/B calculations of the previous chapter that the statistical precision on the signal normalization returned by the fit will be poor. In a later chapter we show how to use the likelihood-fit technique to obtain an upper limit on the single-top content.

7.1 Derivation of the Likelihood Function

Our data is of the form of measured values of a variable—call it x for simplicity—for each of N observed events. We assume that each event is either from signal or background, *i.e.* $N = N_s + N_b$, and that N_s and N_b are themselves distributed according to Poisson

distributions about unknown means n_s, n_b. We further assume that we have an independent estimate μ_b of the mean background content n_b of our sample. In this section we derive [144] the general form of a function $\mathcal{L}$ that expresses the joint probability of observing these N data events at their respective values $\{x_i\}$ of variable x. $\mathcal{L}$, which we call the unbinned likelihood function, is a function of unknown Poisson means n_s, n_b; if we permit n_s and n_b to vary, those values n_s^{MAX}, n_b^{MAX} at which $\mathcal{L}$ achieves its maximum are our best estimate of the signal and background content N_s, N_b of the sample. Finally we apply this general treatment to the case of extracting a measurement of single-top content from the H_T distribution of CDF data.

7.1.1 General form of $\mathcal{L}$

The likelihood function $\mathcal{L}$ consists of two factors: $\mathcal{L} = \mathcal{L}_{Poisson/shape} \times \mathcal{L}_{Gaussian}$. The first factor, $\mathcal{L}_{Poisson/shape}$, embodies the shape information from the x-distributions for signal and background (which distributions we call *templates*) and the Poisson behavior of the respective amounts of signal and background about unknown means n_s, n_b. The second factor, $\mathcal{L}_{Gaussian}$, incorporates the additional information from the independent background calculation.

7.1.1.1 Constructing $\mathcal{L}_{Nm\delta_k\{x_k\}}$: Known m, Known Arrangement

We construct $\mathcal{L}_{Poisson/shape}$ in a series of steps, starting with simpler artificial cases. First we consider the probability of observing a particular arrangement of N events—say *sbbsb...ssb*—in which m events are signal, $N - m$ are background, and each event has its respective observed x value. Let $\delta_k = 1$ if the kth event is from signal, 0 if from background. Let the respective density functions (templates) for signal and background be $f_s(x)$, $f_b(x)$. The probability of observing this arrangement can be written as

$\mathcal{L}_{Nm\delta_k\{x_k\}}$ = (Poisson probability of observing $m = \sum_{k=1}^{N} \delta_k$ signal events in any order) $\times$
(Poisson probability of observing $N - m$ background events in any order) $\div$
(binomial coefficient $\binom{N}{m}$ expressing number of ways in which m signal can be

chosen from N total, since we seek P for *one particular* arrangement) $\times$

(probability that event 1 has x-value between x_1, $x_1 + dx_1$) $\times \cdots \times$

(probability that event N has x-value between x_N, $x_N + dx_N$)

That is,

$$\begin{aligned}\mathcal{L}_{Nm\delta_k\{x_k\}} &= \frac{e^{-n_s}{n_s}^m}{m!}\frac{e^{-n_b}{n_b}^{N-m}}{(N-m)!}\frac{1}{\binom{N}{m}}\left[\left({f_s}^{\delta_1}(x_1){f_b}^{1-\delta_1}(x_1)dx_1\right)\cdots\left({f_s}^{\delta_N}(x_N){f_b}^{1-\delta_N}(x_N)dx_N\right)\right] \\ &= \frac{e^{-(n_s+n_b)}}{N!}{n_s}^m{n_b}^{N-m}\prod_{k=1}^{N}{f_s}^{\delta_k}(x_k){f_b}^{1-\delta_k}(x_k)dx_k\end{aligned}$$

7.1.1.2 Constructing $\mathcal{L}_{Nm\{x_k\}}$: Known m, Unknown Arrangement

Of course in a real data sample, the arrangement of which events are s and which are b is not known. To construct the probability $\mathcal{L}_{Nm\{x_k\}}$ of observing any arrangement of N events in which m are signal and all the x-values are those observed, we sum $\mathcal{L}_{Nm\delta_k\{x_k\}}$ over all possible sequences which give m signal events, *i.e.* over sequences for which $\delta_1 + \ldots + \delta_N = m$. That is,

$$\begin{aligned}\mathcal{L}_{Nm\{x_k\}} &= \sum_{\delta_1+\ldots+\delta_N=m}\mathcal{L}_{Nm\delta_k\{x_k\}} \\ &= \frac{e^{-(n_s+n_b)}}{N!}{n_s}^m{n_b}^{N-m}\sum_{\delta_1+\ldots+\delta_N=m}\prod_{k=1}^{N}{f_s}^{\delta_k}(x_k){f_b}^{1-\delta_k}(x_k)dx_k\end{aligned}$$

7.1.1.3 Constructing $\mathcal{L}_{N\{x_k\}}$: Unknown m, Unknown Arrangement

Finally we are in a position to write down the likelihood function for the situation we face in real data, in which only the total number of events N and the individual x-values $\{x_k\}$ are known. Both the arrangement of signal and background events and the number of signal m are unknown. This probability $\mathcal{L}_{N\{x_k\}}$ is the sum of $\mathcal{L}_{Nm\{x_k\}}$ over all possible values of m, *i.e.*

$$\begin{aligned}\mathcal{L}_{N\{x_k\}} &= \sum_{m=0}^{N}\mathcal{L}_{Nm\{x_k\}} \\ &= \frac{e^{-(n_s+n_b)}}{N!}\left[\sum_{m=0}^{N}{n_s}^m{n_b}^{N-m}\sum_{\delta_1+\ldots+\delta_N=m}\prod_{k=1}^{N}{f_s}^{\delta_k}(x_k){f_b}^{1-\delta_k}(x_k)\right]dx_1\cdots dx_N\end{aligned}$$

The bracketed portion [ ] of the above equation can be expanded, and terms regrouped, to give

$$[\,] = (n_s f_s(x_1) + n_b f_b(x_1))\,(n_s f_s(x_2) + n_b f_b(x_2)) \cdots (n_s f_s(x_N) + n_b f_b(x_N))$$

so that

$$\begin{aligned} \mathcal{L}_{N\{x_k\}} &= \frac{e^{-(n_s+n_b)}}{N!}\left[\prod_{k=1}^{N} n_s f_s(x_k) + n_b f_b(x_k) dx_k\right] \\ &= \frac{e^{-(n_s+n_b)}(n_s+n_b)^N}{N!}\prod_{k=1}^{N}\frac{n_s f_s(x_k) + n_b f_b(x_k)}{n_s+n_b} dx_k \end{aligned}$$

By construction, the function $\mathcal{L}_{N\{x_k\}}$ embodies the shape information from density functions f_s, f_b and the Poisson behavior of the numbers of signal and background about unknown means n_s, n_b. This is the desired $\mathcal{L}_{Poisson/shape}$.

7.1.1.4 $\mathcal{L}_{Gaussian}$: Incorporating an independent background calculation

If we have an independent estimate of the amount of background present in our sample, we ought to incorporate that information into the likelihood function. Specifically, suppose we have an estimate μ_b with uncertainty σ_{μ_b} of the unknown Poisson mean n_b. To add this information to our likelihood function, we introduce an extra term $\mathcal{L}_{Gaussian}$:

$$\mathcal{L}_{Gaussian} = \frac{1}{\sqrt{2\pi{\sigma_{\mu_b}}^2}} e^{-(n_b-\mu_b)^2/2{\sigma_{\mu_b}}^2}$$

The term $\mathcal{L}_{Gaussian}$ introduces a Gaussian penalty for unknown Poisson mean n_b to go too far from expected Poisson mean μ_b, within the uncertainty σ_{μ_b} on that estimate of the expected Poisson mean. The complete likelihood function is then $\mathcal{L} = \mathcal{L}_{Poisson/shape} \times \mathcal{L}_{Gaussian}$.

7.1.1.5 Corollary: Interpretation of n_s^{MAX}, n_b^{MAX} returned by a likelihood fit

It is useful to review the implications of the above derivation of the unbinned likelihood function $\mathcal{L}$ for our understanding of the meaning of fit parameters n_s, n_b. Specifically,

we note that n_s, n_b appearing in $\mathcal{L}$ are not N_s, N_b—the *actual* numbers of signal and background in our particular sample—but are rather the *Poisson means* about which N_s and N_b are assumed to be distributed. We permit these unknown Poisson means n_s, n_b to vary and we find the values n_s^{MAX}, n_b^{MAX} at which $\mathcal{L}$ achieves its maximum.

Therefore the correct interpretation of parameters n_s^{MAX}, n_b^{MAX} returned by the fit is "those Poisson means such that if the N events observed in this experiment were chosen from respective signal and background Poisson distributions having those means, the likelihood of observing these particular data would be maximized." A *wrong* interpretation of parameters n_s^{MAX}, n_b^{MAX} returned by fit is "the actual amounts N_s and N_b of signal and background in this experiment." All we can know about N_s and N_b are n_s^{MAX} and n_b^{MAX}. In practice, we assume

$$\begin{aligned} n_s^{MAX} &\rightarrow \text{best guess at } N_s \\ n_b^{MAX} &\rightarrow \text{best guess at } N_b \end{aligned}$$

as our best estimate of the signal and background content of the sample. This is valid because the most-probable value of a Poisson-distributed variable is the Poisson mean.

These observations have implications for applying background constraints. As discussed above, we Gaussian-constrain fit parameter n_b to μ_b within uncertainty σ_{μ_b} in order to incorporate the result of an independent background calculation $\mu_b \pm \sigma_{\mu_b}$ into our likelihood function. Since n_b is a mean—not a fluctuation about a mean—it obeys Gaussian statistics by the Central Limit Theorem and can be constrained to an expected mean μ_b. Had we made the "$n_b = N_b$" misinterpretation, we might have worried that Gaussian-constraining n_b would "kill" the Poisson fluctuations of N_b, since we expect N_b to fluctuate about μ_b with uncertainty $(\sigma_{\mu_b}^2 + (\sqrt{\mu_b})^2)^{\frac{1}{2}}$. It is true that the desideratum N_b does fluctuate according to Poisson statistics. But fit parameter n_b does not.

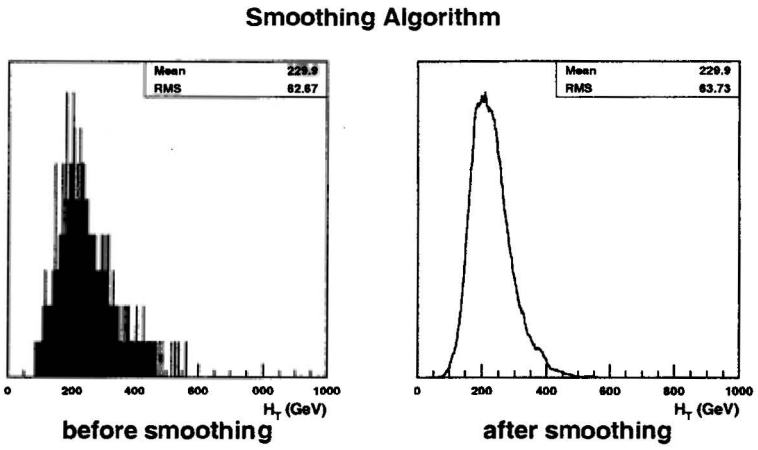


Figure 7.1: H_T distributions for Wg signal before and after smoothing.

7.1.2 $\mathcal{L}$ for single top

We can now write the likelihood function for the specific case of our single-top search. The x variable is of course H_T. Since we have two signal sources and two backgrounds there are four fit parameters (n_{Wg}, n_{W^*}, $n_{t\bar{t}}$, n_{nt}) and four H_T templates (f_{Wg}, f_{W^*}, $f_{t\bar{t}}$, f_{nt}) to incorporate into the Poisson/shape factor $\mathcal{L}_{Poisson/shape}$. With the shorthand $n = n_{Wg} + n_{W^*} + n_{t\bar{t}} + n_{nt}$, $\mathcal{L}_{Poisson/shape}$ may be written as

$$\mathcal{L}_{Poisson/shape} = \frac{e^{-n} n^N}{N!} \prod_{i=1}^{N} \frac{n_{Wg} f_{Wg}(H_{Ti}) + n_{W^*} f_{W^*}(H_{Ti}) + n_{t\bar{t}} f_{t\bar{t}}(H_{Ti}) + n_{nt} f_{nt}(H_{Ti})}{n}$$

Templates $\{f_j(H_T)\}$ are simply smoothed versions of the signal and background H_T distributions from Monte Carlo that were shown in Figure 6.8. We smooth the templates in order to remove statistical fluctuations. To do the smoothing, we bin H_T histograms with extremely fine granularity (10,000 bins) and average each bin with the 200 bins on either side of it. As an example, the H_T distribution for Wg signal before and after smoothing is shown in Figure 7.1. Smoothed distributions for all four processes are shown in Figure 7.2. The N events over which the product $\prod$ runs are the CDF Run 1 data events that survive our selection cuts.

Meanwhile Gaussian-constraint term $\mathcal{L}_{Gaussian}$ has two separate factors, one for each

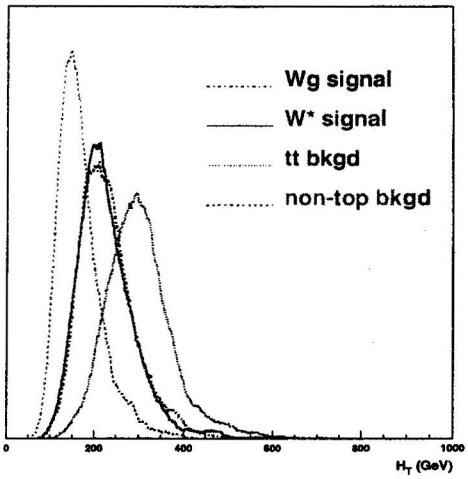


Figure 7.2: Smoothed H_T distributions (templates) for signal and background.

of the two background estimates:

$$\mathcal{L}_{Gaussian} = \frac{1}{\sqrt{2\pi\sigma_{\mu_{t\bar{t}}}{}^2}} e^{-(n_{t\bar{t}}-\mu_{t\bar{t}})^2/2\sigma_{\mu_{t\bar{t}}}{}^2} \times \frac{1}{\sqrt{2\pi\sigma_{\mu_{nt}}{}^2}} e^{-(n_{nt}-\mu_{nt})^2/2\sigma_{\mu_{nt}}{}^2}$$

The background estimates $\mu_{t\bar{t}} \pm \sigma_{\mu_{t\bar{t}}}$ and $\mu_{nt} \pm \sigma_{\mu_{nt}}$ appearing in $\mathcal{L}_{Gaussian}$ are simply the $t\bar{t}$ and non-top predictions from Table 6.9, ($\mu_{t\bar{t}} = 8.38$; $\mu_{nt} = 54.0$) with their respective uncertainties ($\sigma_{\mu_{t\bar{t}}} = 2.68$; $\sigma_{\mu_{nt}} = 12$).

Thus we arrive at the final form of $\mathcal{L} = \mathcal{L}_{Poisson/shape} \times \mathcal{L}_{Gaussian}$. This likelihood function expresses the joint probability of observing these N data events at their respective values $\{H_{Ti}\}$. As the four fit parameters n_{Wg}, n_{W^*}, $n_{t\bar{t}}$, n_{nt} are permitted to vary, $\mathcal{L}_{Gaussian}$ constrains $n_{t\bar{t}}$ and n_{nt} to stay close to predictions $\mu_{t\bar{t}}$ and μ_{nt} with respective Gaussian penalties. Those values n_{Wg}^{MAX}, $n_{W^*}^{MAX}$, $n_{t\bar{t}}^{MAX}$, n_{nt}^{MAX} for which $\mathcal{L}$ is maximized are our best estimate of the signal and background content N_{Wg}, N_{W^*}, $N_{t\bar{t}}$, N_{nt} of the sample.

7.2 Maximizing $\mathcal{L}$

In earlier chapters we emphasized that the H_T distributions for Wg and W^* are virtually indistinguishable, making H_T an ideal variable to fit for the combined $Wg + W^*$

signal content of data. To treat the two processes as a single signal, we ask the fit to tell us the combined $Wg + W^*$ content of data rather than the separate amounts of Wg and W^*. We do this by fixing the relative normalizations of templates f_{Wg} and f_{W^*} at the ratio of the Standard-Model cross sections. That is, we fix

$$n_{W^*} = n_{Wg}\left(\frac{\sigma_{W^*}}{\sigma_{Wg}}\right)$$

By forcing n_{W^*} to covary with n_{Wg} in the fit, we have effectively reduced our fit parameters to three: $n_{t\bar{t}}$, n_{nt}, and "n_{signal}", where $n_{signal} = n_{Wg} + n_{W^*} = (1 + \frac{\sigma_{W^*}}{\sigma_{Wg}})n_{Wg}$.

The goal is then to maximize $\mathcal{L}$ with respect to the three fit parameters $n_{t\bar{t}}$, n_{nt}, and n_{signal}. This problem is equivalent to maximizing $\log\mathcal{L}$, or minimizing $-\log\mathcal{L}$. We use the MINUIT [145] software package to perform this task. MINUIT subroutine MIGRAD performs a gradient search for the minimum value of a user-supplied multiparameter function using the Davidon-Fletcher-Powell algorithm. In addition to finding the parameter values corresponding to the function minimum, MINUIT finds the parameter errors. The one-standard-deviation error on a parameter corresponds to the amount by which the parameter would need to change in order to increase $-\log\mathcal{L}$ by 0.5. This error is statistical only.

7.3 Checks of Maximum Likelihood Fit Method

To test our maximum likelihood fit method, we apply it to many *pseudoexperiments*, or samples of fake data such as would be expected from repeats of the CDF experiment. We examine the resulting distributions of fit parameters.

7.3.1 Method for generating pseudoexperiments

A pseudoexperiment consists of an ensemble of H_T values randomly generated according to the expected H_T distribution for CDF data. We expect this distribution to have contributions from the two signal (Wg, W^*) and two background ($t\bar{t}$, non-top) processes, in the respective amounts predicted in Table 6.9. The smoothed H_T distributions for each of

these four processes were shown in Figure 7.2. To generate a pseudoexperiment, we draw random H_T values from each of the four smoothed distributions in turn.

To choose the number of times N_i to draw from the ith distribution, we take a Poisson fluctuation about predicted mean number of events μ_i from the ith process. To perform the actual random drawing according to the ith distribution, we use CERNLIB routines HISPRE and HISRAN. HISPRE takes distribution $g(x)$ and forms its cumulative distribution $G(x) = \int_0^x g(x')dx'$. HISRAN generates a random number $\alpha \in (0,1)$ and inverts the cumulative distribution to find x such that $\alpha = G(x)$. x values thus generated are distributed according to $g(x)$ [1]. After sampling N_{Wg} times from the Wg distribution, N_{W^*} times from the W^* distribution, and so on for all four distributions, we arrive at a random ensemble of H_T values chosen according to the expected distribution for CDF data. Of course, the ensemble of H_T values will be different for each pseudoexperiment.

Finally, we address the fact that the background means $\mu_{t\bar{t}}$, μ_{nt} appearing in the Gaussian-constraint term of our likelihood function $\mathcal{L}$ are known with finite precision. For each pseudoexperiment, we Gaussian-fluctuate these means within their uncertainties $\sigma_{\mu_{t\bar{t}}}$, $\sigma_{\mu_{nt}}$ to obtain values $\mu_{t\bar{t}fluctuated}$, $\mu_{ntfluctuated}$. We use these fluctuated means in place of the default means $\mu_{t\bar{t}}$, μ_{nt} in the Gaussian-constraint term in $\mathcal{L}$. By dynamically generating background constraints in this way for each pseudoexperiment, we account for the fact that we only know $\mu_{t\bar{t}}$, μ_{nt} to within $\sigma_{\mu_{t\bar{t}}}$, $\sigma_{\mu_{nt}}$.

7.3.2 Results of fitting pseudoexperiments

Next we fit this fake sample of "data" for a given pseudoexperiment, *i.e.* ensemble of H_T values, using the pair of background constraints $\{\mu_{t\bar{t}fluctuated}, \mu_{ntfluctuated}\}$ and the maximum likelihood method. MINUIT returns fitted values for the three parameters n_{signal}, $n_{t\bar{t}}$, n_{nt} along with errors $\sigma_{n_{signal}}$, $\sigma_{n_{t\bar{t}}}$, $\sigma_{n_{nt}}$ on these parameters. We repeat the process of generating and fitting pseudoexperiments a large number of times and examine the resulting distributions of fit parameters and errors. We also examine the distribution of *pulls*, where

[1] $\frac{dN}{dx} = \frac{dN}{d\alpha}\frac{d\alpha}{dx} = 1 \cdot g(x)$

pull is defined as the difference between fitted *vs.* predicted number of events, divided by the error on the fit parameter:

$$\text{pull (for fit parameter } n) = \frac{n - \mu}{\sigma_n}$$

Here σ_n is the parabolic error returned by MINUIT. If the fit is performing correctly, we expect to see pull distributions centered at 0 with standard deviation 1.

Figure 7.3 shows the distributions of fit parameters, fit-parameter errors, and pulls for each of 10,000 pseudoexperiments. While we see that overall the fitter appears to be performing adequately (fit-parameter distributions centered roughly at their input values; pull distributions centered at 0 with standard deviation 1), there are several non-ideal features. The distribution of fitted n_{signal} is biased a bit high ($\langle n_{signal} \rangle = 5.4$ *vs.* input value 4.3) while the distributions of fitted numbers of background are biased a bit low ($\langle n_{t\bar{t}} \rangle = 8.1$ *vs.* input value 8.4; $\langle n_{nt} \rangle = 52.6$ *vs.* input value 54). Also, the pull distribution for the $t\bar{t}$ background has a distinctly non-Gaussian low tail. However, it is reassuring for the purposes of a limit-setting analysis to note that the bias is on the conservative side, *i.e.* tends to *over*estimate the signal content.

To rule out the possibility that the issues mentioned above indicate the presence of bugs in our fitting software, we check how well the fitter performs on pseudoexperiments with less-challenging statistics. These plots are shown in Figure 7.4. With signal fraction enhanced so that $S/\sqrt{B} = 5$ and background uncertainties reduced to 10% each, the biases disappear and the pull distributions look fine. We conclude that the fitter biases are an artifact of the low statistics (4.3 on a background of 62) rather than a reflection of a problem with the fitting method.

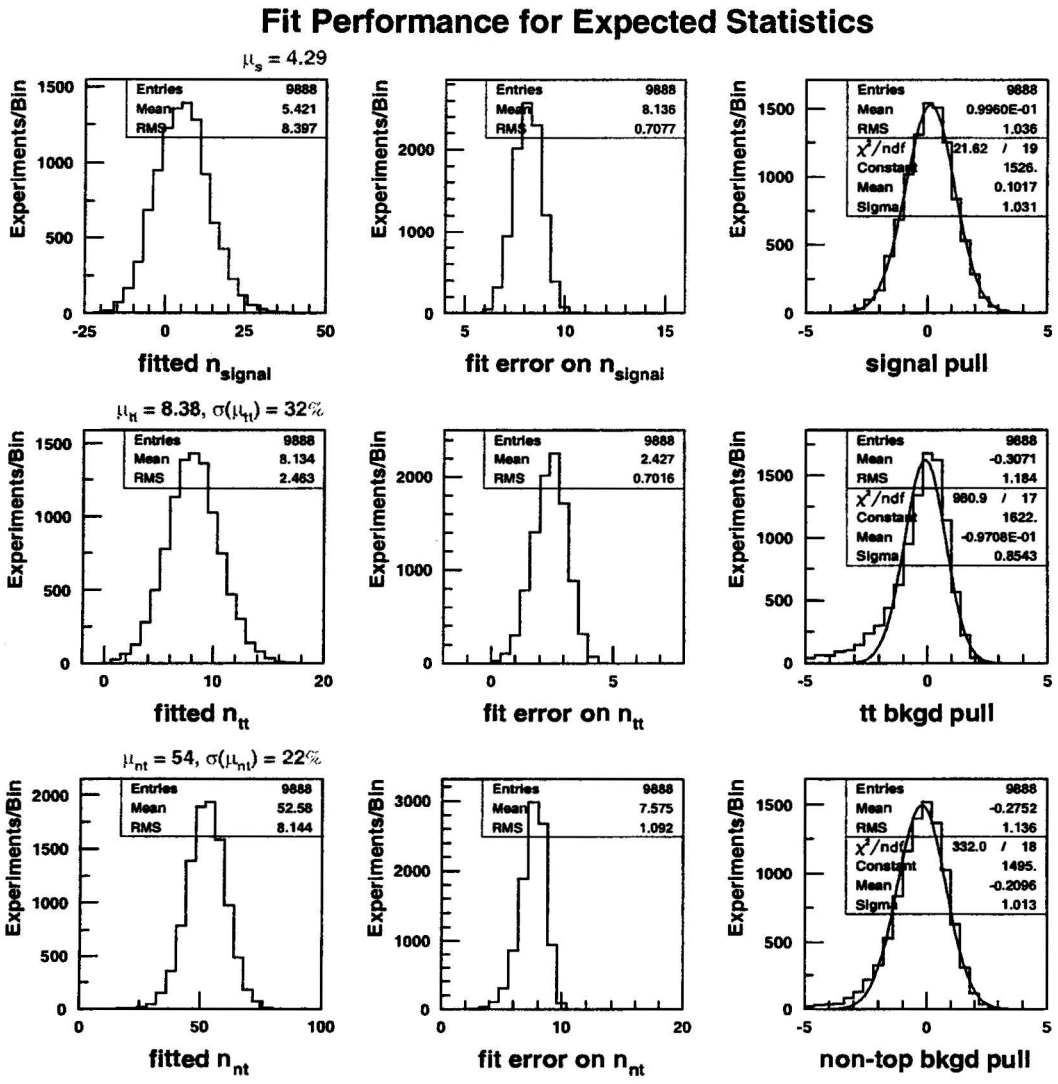


Figure 7.3: Distributions of fit parameters, fit-parameter errors, and pulls for 10,000 pseudoexperiments performed using the expected amounts of signal and background from Table 6.9.

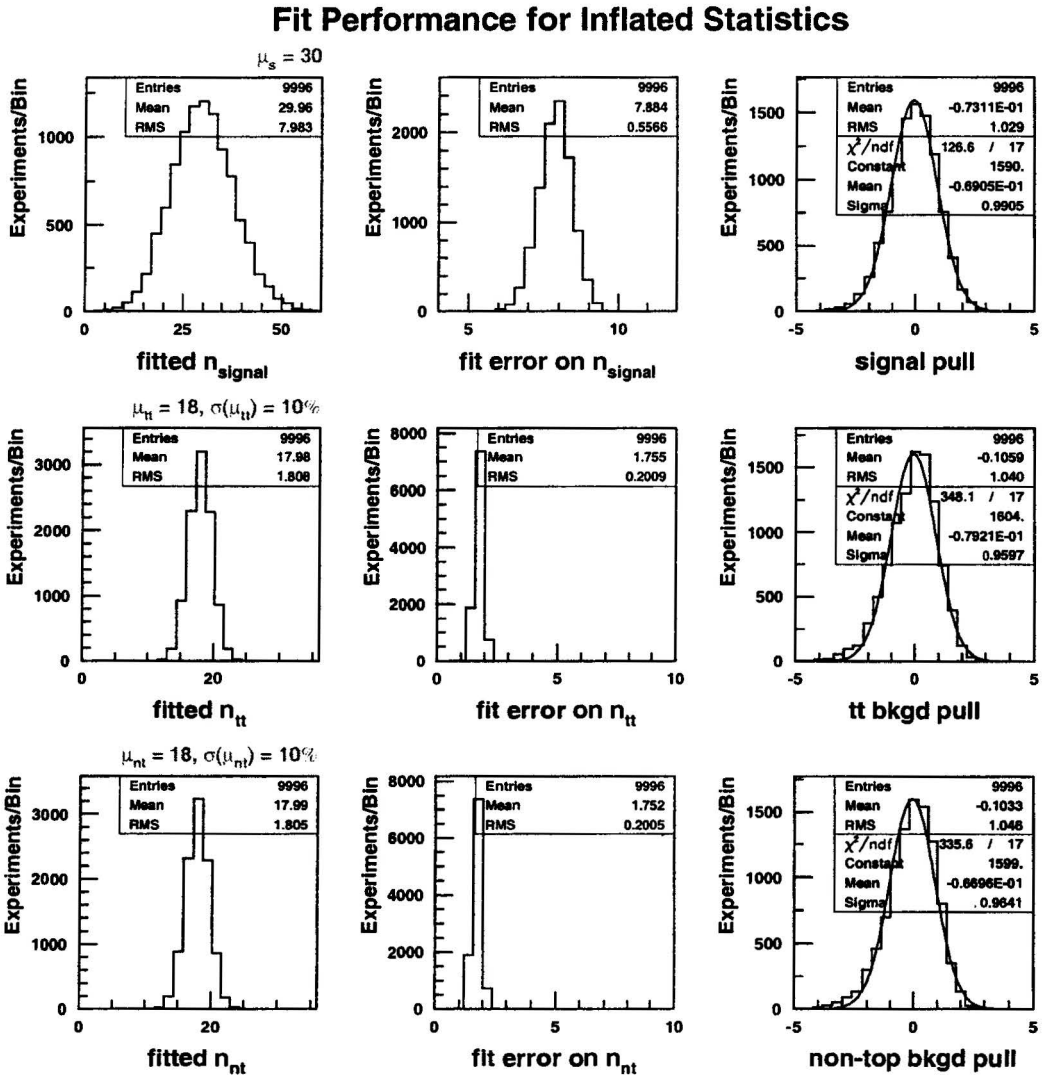


Figure 7.4: Distributions of fit parameters, fit-parameter errors, and pulls for 10,000 pseudo-experiments performed using artificially enhanced signal fraction ($\mu_{signal} = 30$; $\mu_{t\bar{t}} = \mu_{nt} = 18$; $S/\sqrt{B} = 5$) and artificially reduced uncertainty on backgrounds ($\sigma_{\mu_{t\bar{t}}} = \sigma_{\mu_{nt}} = 10\%$). Note the improved performance of the fitter with these conditions.

Chapter 8

Setting an Upper Limit

In Chapter 6 we calculated that we expect CDF Run 1 data to contain roughly 4.3 signal events on a background of 62, which corresponds to $S/\sqrt{B} = 0.54$. In Figure 7.3 we saw the impact of these challenging statistics on the performance of our likelihood fitter: in default pseudoexperiments using the expected amounts of signal and background, the mean value of fit parameter n_{signal} is 5.4 but the mean value of the *error* on that fit parameter is 8.1. In other words, we expect the signal normalization returned by the fit to have poor statistical precision. We shall instead use the likelihood-fit technique to obtain an upper limit on single-top production. In this chapter we begin by defining the upper limit at a given confidence level. We consider both Bayesian and frequentist methods of constructing an upper limit. The method we choose for this analysis is the Bayesian, but we explain how to check our result against the frequentist standard as well. Next we apply the upper-limit-contruction method to many pseudoexperiments in order to arrive at an *a priori* estimate of the limit we expect to be able to set in real data. We compare this *a priori* limit with the limit we would expect from a pure counting experiment to show the power of including shape information. Finally, we check if there is *a priori* motivation for narrowing our search to a subset of the jet-multiplicity bins $W+1$, $W+2$, or $W+3$. We conclude the answer is no.

8.1 Definition of Upper Limit at 95% Confidence Level

8.1.1 Introduction to β, or the relative amount of signal

It is convenient to rephrase our likelihood function $\mathcal{L}$ in terms of fit parameters and density functions with a different normalization. We define the following terms:

- μ_j: Predicted number of events for process j.
- $F_j(H_T)$: Template for process j, normalized to the total number of events predicted for process j, *i.e.*, $\int F_j(H_T)dH_T = \mu_j$.
- β_j: Fit parameter representing the relative number of events from process j, *i.e.* $\beta_j = \frac{n_j}{\mu_j}$.

β_j is interpreted as a deviation from the expected number of events from process j. For example, the fit returns $\beta_{signal} < 1$ if it prefers less than the predicted μ_{signal}, $\beta_{signal} > 1$ if it prefers more, and $\beta_{signal} = 1$ if it prefers exactly the predicted amount. With this nomenclature, and with the previously-defined shorthand $n = n_{W_g} + n_{W^*} + n_{t\bar{t}} + n_{nt} = \beta_{W_g}\mu_{W_g} + \beta_{W^*}\mu_{W^*} + \beta_{t\bar{t}}\mu_{t\bar{t}} + \beta_{nt}\mu_{nt}$, we rewrite the Poisson/shape factor as

$$\mathcal{L}_{Poisson/shape} = \frac{e^{-n}n^N}{N!}\prod_{i=1}^{N}\frac{\beta_{W_g}F_{W_g}(H_{Ti}) + \beta_{W^*}F_{W^*}(H_{Ti}) + \beta_{t\bar{t}}F_{t\bar{t}}(H_{Ti}) + \beta_{nt}F_{nt}(H_{Ti})}{n}$$

Recall that the product Π runs over the N data events passing all cuts. Recall also that β_{W_g} and β_{W^*} are forced to covary in the fit, so $\beta_{W_g} = \beta_{W^*} \equiv \beta_{signal}$. Rewriting the Gaussian factor in terms of β's,

$$\mathcal{L}_{Gaussian} = \frac{1}{\sqrt{2\pi\sigma_{\mu_{t\bar{t}}}{}^2}}e^{-(\beta_{t\bar{t}}\mu_{t\bar{t}}-\mu_{t\bar{t}})^2/2\sigma_{\mu_{t\bar{t}}}{}^2} \times \frac{1}{\sqrt{2\pi\sigma_{\mu_{nt}}{}^2}}e^{-(\beta_{nt}\mu_{nt}-\mu_{nt})^2/2\sigma_{\mu_{nt}}{}^2}$$

Thus we arrive at the final form of $\mathcal{L}(\beta_{signal}, \beta_{t\bar{t}}, \beta_{nt}) = \mathcal{L}_{Poisson/shape} \times \mathcal{L}_{Gaussian}$. Subsequent discussion will refer to setting an upper limit on β_{signal}, or "β_s" for short, rather than on n_{signal}. Next we define what we mean by an upper limit on β_s. We begin with the frequentist definition.

8.1.2 *Frequentist definition of upper limit at 95% confidence level*

Measurement of a physical constant β results in an estimator $\hat{\beta}$ [146, 147]. For a given value of β, measurements $\hat{\beta}$ will fall in a probability distribution function $f(\hat{\beta};\beta)$ in which the true value β appears as a parameter. We can construct γ such that repeated measurements produce results ($\hat{\beta} > \gamma$) 95% of the time by integrating $f(\hat{\beta};\beta)$ up to $\hat{\beta} = \gamma$ such that 5% of the area under the curve lies in the interval $(-\infty, \gamma)$: $0.05 = \int_{-\infty}^{\gamma} f(\hat{\beta};\beta)d\hat{\beta}$. (Figure 8.1). Note that γ depends on the unknown true value β. Repeating this procedure

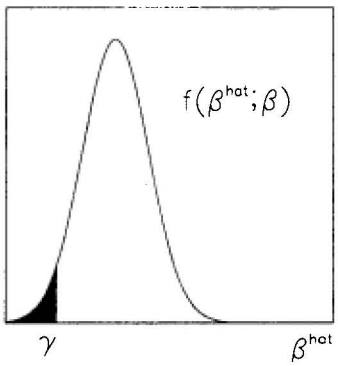


Figure 8.1: Cartoon showing the distribution of measurements $\hat{\beta}$ for a given unknown true value β. Also shown is γ, the value of $\hat{\beta}$ such that 95% of measurements are larger than this value.

for different true values of β, we construct curve $\gamma(\beta)$. Then, when we make a measurement using real data (in which $\beta = \beta_{actual}$) and observe $\hat{\beta} = \beta_{exp}$, we note that β_{exp} intersects the curve $\gamma(\beta)$ at $\beta = C$. Since by definition $\beta_{exp} > \gamma(\beta_{actual})$ 95% of the time, by construction $C > \beta_{actual}$ 95% of the time. (Figure 8.2.) Call this C "β_{95}". The β_{95} thus constructed has a 95% probability of being an upper bound on β_{actual}; therefore we call it "the upper limit on β_{actual} at 95% confidence level."

In practice, we obtain upper limit β_{95} from measured value β_{exp} by a series of iterations. Since $\gamma(\beta_{95}) = \beta_{exp}$, or $0.05 = \int_{-\infty}^{\beta_{exp}} f(\hat{\beta};\beta_{95})d\hat{\beta}$, we know that pseudoexperiments conducted with $\beta = \beta_{95}$ will yield measured values $\hat{\beta} \leq \beta_{exp}$ 5% of the time. Thus we generate ("throw") pseudoexperiments with input β equal to guessed value $\beta_{95\,guess}$ and integrate the resulting distribution of measured values $f(\hat{\beta};\beta_{95\,guess})$ up to β_{exp}. If this

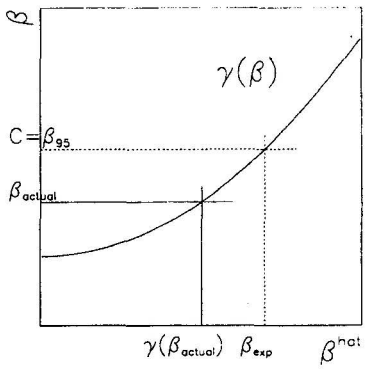


Figure 8.2: Cartoon showing how upper limit β_{95} is constructed: $\beta_{95} = \gamma^{-1}(\beta_{exp})$.

integral is <(>) 5%, we know β_{exp} <(>) $\gamma(\beta_{95_{guess}})$, so we must decrease(increase) $\beta_{95_{guess}}$ (Figure 8.3).

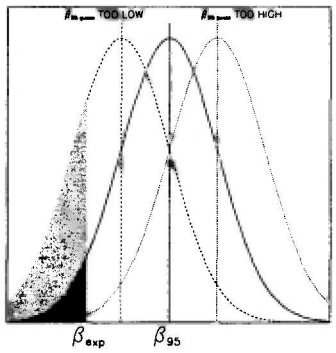


Figure 8.3: Cartoon showing the iterative procedure for deducing β_{95}. If the integral of $f(\hat{\beta}; \beta_{95_{guess}})$ up to β_{exp} is $< 5\%$, $\beta_{95_{guess}}$ must be too high, but if it is $> 5\%$, $\beta_{95_{guess}}$ must be too low.

8.1.3 Bayesian definition of upper limit at 95% confidence level

In a Bayesian treatment [148, 149], we treat the unknown physical constant β as a statistical variable rather than a parameter. We replace the parameterized p.d.f. $f(\hat{\beta}; \beta)$ by the conditional p.d.f. $f(\hat{\beta}|\beta)$ and rephrase the confidence-limit question as follows. Our measurements sample $f(\hat{\beta}|\beta)$, the distribution of measured values $\hat{\beta}$ for fixed, unknown β, when what we really seek is $g(\beta|\hat{\beta})$, which tells us that, given our measurement $\hat{\beta}$, the

"true answer" β lies between β and $\beta + d\beta$ with probability $g(\beta|\hat{\beta})d\beta$. This viewpoint provides another way to define upper limit at 95% confidence level: we seek β_{95} such that $\int_{-\infty}^{\beta_{95}} g(\beta|\hat{\beta})d\beta = 0.95$, *i.e.* such that the probability (given the data) that $\beta \leq \beta_{95}$ is 95%.

The connection between the data-derived $f(\hat{\beta}|\beta)$ and the desired $g(\beta|\hat{\beta})$ is provided by Bayes's Theorem [1]:

$$g(\beta|\hat{\beta}) = \frac{f(\hat{\beta}|\beta)P(\beta)}{\int f(\hat{\beta}|\beta)P(\beta)d\beta}$$

In a Bayesian interpretation, we take $P(\beta)$ appearing on the RHS to be the "prior" p.d.f. for β, reflecting our beliefs about the true value of the parameter before conducting the experiment. In similar fashion we take $g(\beta|\hat{\beta})$ appearing on the LHS to be the "posterior" p.d.f. for β (given the data), reflecting our modified beliefs after conducting the experiment. The f which relates posterior and prior p.d.f.'s is our likelihood function $\mathcal{L}(\text{data}|\beta)$.

Actually, we do not equate $f(\hat{\beta}|\beta)$ appearing in the above equation with the full likelihood function $\mathcal{L}(\beta_s, \beta_{t\bar{t}}, \beta_{nt})$ defined in Chapter 7. Instead we first reduce $\mathcal{L}$ to a function of the single variable β_s by maximizing $\mathcal{L}$ for each value of β_s with respect to $\beta_{t\bar{t}}$ and β_{nt}. The resulting "reduced likelihood function" $\mathcal{L}(\beta_s)$ represents the probability, for each value of fit parameter β_s (itself signifying a possible Poisson mean $n_s = \beta_s \mu_s$ about which N_s would be distributed), of observing the data we do observe. It is this reduced likelihood function $\mathcal{L}(\beta_s)$ that is equivalent to $f(\hat{\beta}|\beta)$.

A conventional choice of prior p.d.f. for β, reflecting the claim of no advance knowledge of β except that it lie in the physical domain, is the so-called "flat" prior

$$\begin{aligned} P(\beta) &= k \text{ for } \beta \geq 0 \\ &= 0 \text{ for } \beta < 0 \end{aligned}$$

Then the equation for the posterior p.d.f. becomes

$$g(\beta|\text{data}) = \frac{\mathcal{L}(\text{data}|\beta)}{\int_0^{+\infty} \mathcal{L}(\text{data}|\beta)d\beta}, \quad \beta \geq 0$$

[1] Note that Bayesians and frequentists debate only the *interpretation* of this theorem, not its mathematical correctness.

Integrating both sides up to β_{95} such that the probability (given the data) that $\beta \leq \beta_{95}$ is 95%, which on the LHS means integrating the posterior p.d.f. for β up to β_{95} such that $\int_{-\infty}^{\beta_{95}} g(\beta|\text{data})d\beta = 0.95$, we obtain

$$0.95 = \frac{\int_0^{\beta_{95}} \mathcal{L}(\text{data}|\beta)d\beta}{\int_0^{+\infty} \mathcal{L}(\text{data}|\beta)d\beta}$$

Thus to calculate β_{95} in a Bayesian treatment, we simply integrate the product of (reduced likelihood function $\mathcal{L}(\text{data}|\beta)$ $\times$ flat prior) up to that value $\beta = \beta_{95}$ such that 95% of the area under the curve occurs for $\beta < \beta_{95}$ [2]. To perform the numerical integration we use software developed by H. S. Kim [150].

8.1.4 Comparing Bayesian and frequentist limits

In this analysis, we use the Bayesian method to calculate β_{95}. Then, we check whether β_{95} satisfies the frequentist definition of upper limit at 95% confidence level. To perform this frequentist check, we throw pseudoexperiments with input $\beta = \beta_{95}$, fit each pseudoexperiment to obtain $\hat{\beta}$, and smear each $\hat{\beta}$ according to systematic uncertainty $\sigma(\hat{\beta})$. (Chapter 9 will describe both the calculation of systematic uncertainty $\sigma(\hat{\beta})$ and the procedure used to smear. The frequentist check will be applied to the limit obtained from CDF Run 1 data in Chapter 10.) Plotting the values of smeared $\hat{\beta}$ for each pseudoexperiment yields distribution $f(\hat{\beta}|\beta_{95})$. Next we integrate this distribution out to β_{exp}, or the particular value of $\hat{\beta}$ obtained from the fit to real CDF data. If $\int_{-\infty}^{\beta_{exp}} f(\hat{\beta}|\beta_{95})d\hat{\beta} = 5\%$, we know $\beta_{exp} = \gamma(\beta_{95})$, so our Bayesian-calculated β_{95} satisfies the frequentist definition of upper limit exactly. If the integral is $<(>)$ 5%, β_{exp} $<(>)$ $\gamma(\beta_{95})$, so the Bayesian-calculated β_{95} is too high(low), corresponding to frequentist over(under)coverage. Therefore if the integral is $\leq$ 5%, our Bayesian-derived result is at least conservative by frequentist standards.

[2] So far this treatment neglects systematic uncertainty. We will discuss how to fold systematic uncertainty into the likelihood function in Chapter 9.

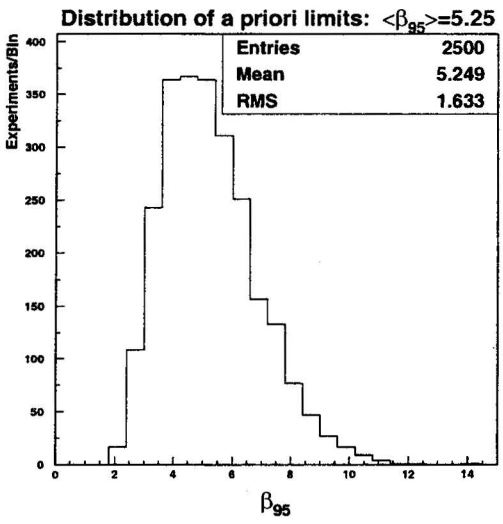


Figure 8.4: Distribution of *a priori* 95% C.L. upper limits β_{95} for 2500 pseudoexperiments. The mean value of this distribution, $\langle\beta_{95}\rangle$, is an estimate of the limit we can expect to set in CDF data.

8.2 An *A Priori* Upper Limit on $\sigma_{single\ top}$

Having defined an algorithm for obtaining β_{95}, we can now make an *a priori* estimate of the upper limit on we expect to be able to set in CDF Run 1 data. We throw many pseudoexperiments with the expected amounts of signal and background, calculate β_{95} for each, and interpret the mean value of this distribution $\langle\beta_{95}\rangle$ as an estimate of the *a priori* β_{95} we expect to achieve in CDF data. The expected limit on the single-top cross section follows immediately: $\sigma_{95} = \beta_{95} \times \sigma_{SM}$.

8.2.1 Expected limit from H_T fit

The distribution of β_{95} for 2500 pseudoexperiments is shown in Figure 8.4. Thus our *a priori* $\beta_{95} = 5.25$. Since the Standard-Model-predicted combined cross section for $Wg + W^*$ single top is $1.7 + 0.73 = 2.43$ pb, $\beta_{95} = 5.25$ corresponds to $\sigma_{95} = 5.25 \times 2.43$ pb $= 12.8$ pb.

8.2.2 Expected limit from counting experiment

It is interesting to compare the above expected limit, obtained using information both on the predicted amounts of S and B and on the shapes of the S and B H_T distributions, to the expected limit obtained from a pure counting experiment. In a counting experiment we do not use shape information; all we know is the number of observed events and the predicted amount of background.

8.2.2.1 1996 PDG method for Poisson process with background

To paraphrase the 1996 Particle Data Group treatment [151], suppose we observe N_{obs} events in a Poisson process with two components, signal and background. Let μ_s denote the unknown signal mean and μ_b the predicted mean for the sum of all backgrounds. For the time being, assume μ_b is known with negligible error. We do not know N_b, the *actual* number of background events in this particular experiment, but we do know that $N_b \leq N_{obs}$. We seek N_{95}, the upper limit on μ_s at 95% confidence level, which is defined to be

- "that value of μ_s such that any random repeat of the current experiment with $\mu_s = N_{95}$ and the same μ_b would observe *more* than N_{obs} events in total *and* would have $N_b \leq N_{obs}$, all with probability 95% [3]."

N_{95} is the solution to the equation

$$0.05 = \frac{e^{-(\mu_b+N_{95})} \sum_{n=0}^{N_{obs}} \frac{(\mu_b+N_{95})^n}{n!}}{e^{-\mu_b} \sum_{n=0}^{N_{obs}} \frac{\mu_b^n}{n!}}$$

[3] This definition appears to be a frequentist/Bayesian hybrid. The parallel with the generic frequentist definition of β_{95} is immediate; in Section 8.1.2 β_{95} was defined to be such that any random repeat of the experiment with $\beta = \beta_{95}$ yields a measurement $\hat{\beta}$ *greater* than data-observed value β_{exp} with probability 95%. Yet no analogue to the restriction $N_b \leq N_{obs}$ was found in that definition. Rather, renormalizing to unit area a subset of the p.d.f. for possible values of N_b recalls the Bayesian handling of a measurement near an unphysical boundary [148].

which can be understood as follows. Since the p.d.f. for N_b, $P_b(N_b) = \frac{e^{-\mu_b}\mu_b^{N_b}}{N_b!}$, is normalized to 1, we can write

$$\begin{aligned} 1 &= \sum_{N_b=0}^{\infty} P_b(N_b) \\ 1 &= \sum_{N_b=0}^{N_{obs}} P_b + \sum_{N_b=N_{obs}+1}^{\infty} P_b \end{aligned}$$

Obviously, since the second term on the RHS is nonzero, the probability that $N_b \leq N_{obs}$ is less than 1. But we desire that the new p.d.f. $P_b'(N_b)$ for $\{N_b$, subject to the constraint $N_b \leq N_{obs}\}$, should be normalized to 1, *i.e.* that $P_b'(N_b = 0) + \ldots + P_b'(N_b = N_{obs}) = 1$. We achieve this by means of normalization factor C:

$$P_b'(N_b) = C\,\frac{e^{-\mu_b}\mu_b^{N_b}}{N_b!}$$

The probability for observing N_s signal is the usual $P_s(N_s) = \frac{e^{-\mu_s}\mu_s^{N_s}}{N_s!}$. Then the probability of observing a total of N_{obs} events due to signal and background, subject to the constraint $N_b \leq N_{obs}$, is just

$$\begin{aligned} P_{tot}(N_{obs}) &= C \sum_{N_s=0}^{N_{obs}} \frac{e^{-\mu_s}\mu_s^{N_s}}{N_s!}\,\frac{e^{-\mu_b}\mu_b^{(N_{obs}-N_s)}}{(N_{obs}-N_s)!} \\ &= C\,\frac{e^{-(\mu_s+\mu_b)}(\mu_s+\mu_b)^{N_{obs}}}{N_{obs}!} \end{aligned}$$

The probability of observing N_{obs} or *fewer* events due to S and B is just $P_{tot}(n)$, summed from $n = 0$ to $n = N_{obs}$. The N_{95} we seek is that value of μ_s such that this probability is 0.05:

$$\begin{aligned} 0.05 &= \sum_{n=0}^{N_{obs}} P_{tot}(n) \\ &= C \sum_{n=0}^{N_{obs}} \frac{e^{-(\mu_b+N_{95})}(\mu_b+N_{95})^n}{n!} \end{aligned}$$

But what is C? Requiring $P_b'(N_b)$ to be normalized to 1 means $\sum_{N_b=0}^{N_{obs}} P_b'(N_b) = 1$, *i.e.*

$$C \sum_{n=0}^{N_{obs}} \frac{e^{-\mu_b}\mu_b^n}{n!} = 1$$

or

$$C = \frac{1}{\sum_{n=0}^{N_{obs}} \frac{e^{-\mu_b}\mu_b^n}{n!}}$$

Thus we recover the formula for N_{95}.

In practice, we solve for N_{95} by means of pseudoexperiments. This method has the added advantage of permitting us to incorporate the nonnegligible uncertainties on our background prediction and signal acceptance. We proceed as follows:

- Guess N_{95};
- Gaussian-fluctuate signal and background means within their respective uncertainties to obtain $N_{95\,fluc}$, $\mu_{b\,fluc}$;
- Generate $(N_s + N_b)$ using Poisson statistics about $(N_{95\,fluc} + \mu_{b\,fluc})$ and increment Counter #1 if $(N_s + N_b) \leq N_{obs}$;
- Generate N_b using Poisson statistics about $\mu_{b\,fluc}$ and increment Counter #2 if $N_b \leq N_{obs}$;
- Take the ratio of Counter #1/Counter #2 over many trials; if this ratio $>(<)$ 0.05, increase(decrease) N_{95}, and repeat.

To obtain an *a priori* estimate of the limit we expect to be able to set in a pure counting experiment, we take the number of observed events, N_{obs}, to be the sum $(\mu_s + \mu_b)$, rounded to the nearest integer. For our analysis cuts, $\mu_s = 4.29 \pm 0.69$ and μ_b is the sum of 8.38 ± 2.68 and 54 ± 12 (Table 6.9.) Thus we take "N_{obs}" to be 67 events. With the above algorithm for computing N_{95}, we obtain $N_{95} = 35.4$, which corresponds to $\beta_{95} = 8.25$. Thus we see the *a priori* utility of including shape information in setting a limit: the expected β_{95} from shape-fitting was 5.25.

8.3 Optimization Studies

Finally we check whether an improved *a priori* limit can be achieved by restricting our search to a subset of the $W + 1, 2, 3$-jet bins. We consider three cases in addition to the default ($W + 1, 2, 3$-jet) case:

- $W + 2$-jet bin only
- $W + 1, 2$-jet bin
- $W + 2, 3$-jet bin

For each case we recalculate the expected numbers of signal and background. These numbers are shown in Table 8.1. Note that if we used $S/\sqrt{B}$ as the figure of merit, we would choose $W + 2, 3$-jet as the best channel ($S/\sqrt{B} = 0.65$ *vs.* 0.54). However, when we conduct

CDF Preliminary

	Default	$W + 2j$ only	$W + 1, 2j$	$W + 2, 3j$
μ_{Wg}	3.00	1.45	2.25	2.20
μ_{W^-}	1.29	0.79	1.07	1.02
$\mu_{t\bar{t}}$	8.38	2.27	2.48	8.17
μ_{nt}	54.0	13.9	51.3	16.6
TOTAL # predicted	66.7	18.4	57.1	28.0
$S/\sqrt{B}$	0.543	0.557	0.453	0.647
$\langle\beta_{95}\rangle$	5.25	5.71	5.80	5.22

Table 8.1: Predicted numbers of events for alternate-jet-bin search channels

pseudoexperiments [4] and examine the distributions of β_{95} for each case, we discover that there is in fact no significant *a priori* advantage to changing our $W + 1, 2, 3$-jet search strategy (Figure 8.5).

[4] Using, of course, new templates $F_j(H_T)$ and normalizations μ_j appropriate to each new, more restrictive set of cuts.

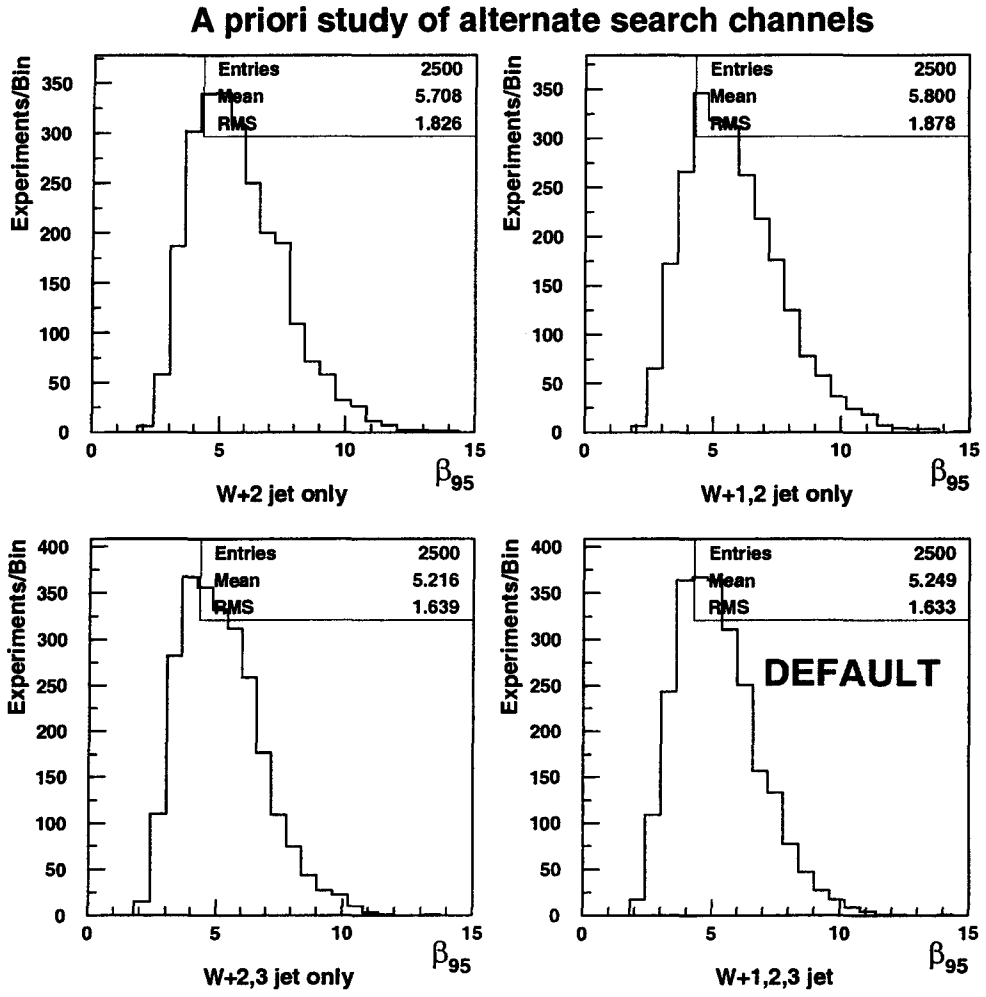


Figure 8.5: Distributions of *a priori* β_{95} for alternate search channels. Note that there is no significant *a priori* advantage to changing the default $W + 1, 2, 3$-jet search strategy.

Chapter 9

Incorporating Systematic Uncertainty

Systematic uncertainty in this analysis arises from sources ranging from hardware to theory—from uncertainty on the calorimeter response to jets to uncertainty on the measured value of M_{top} to uncertainty on the theoretical modeling of gluon radiation. For our result properly to include this uncertainty, we must fold the uncertainty into the reduced likelihood function $\mathcal{L}(\beta)$ used to calculate β_{95}. This process is known as *smearing* the likelihood function. We derive a function $\sigma(\beta)$ representing the net effect of all sources of systematic uncertainty on β at a given value of β. We replace the contents of each bin of the unsmeared likelihood function $\mathcal{L}(\beta)$ with a Gaussian of width $\sigma(\beta)$ so that one bin of the unsmeared likelihood contributes to multiple bins of the smeared likelihood (Figure 9.1). The contribution of $\mathcal{L}$ bin β to $\mathcal{L}_{smeared}$ bin β' is

$$\mathcal{L}(\beta)\frac{1}{\sqrt{2\pi\sigma^2(\beta)}}e^{-(\beta'-\beta)^2/2\sigma^2(\beta)}$$

The contribution of all such bins to the $\mathcal{L}_{smeared}$ bin β' is the integral over all contributing bins β:

$$\begin{aligned}\mathcal{L}_{smeared}(\beta') &= \int_0^\infty \mathcal{L}(\beta)\frac{1}{\sqrt{2\pi\sigma^2(\beta)}}e^{-(\beta'-\beta)^2/2\sigma^2(\beta)}d\beta \\ &= \int \mathcal{L}(\beta)G(\beta;\sigma(\beta))d\beta\end{aligned}$$

That is, we use $\sigma(\beta)$ as the width of a Gaussian to convolve with $\mathcal{L}(\beta)$. In this way we implement the loss of resolution on $\mathcal{L}(\beta)$ caused by the systematic uncertainty inherent at

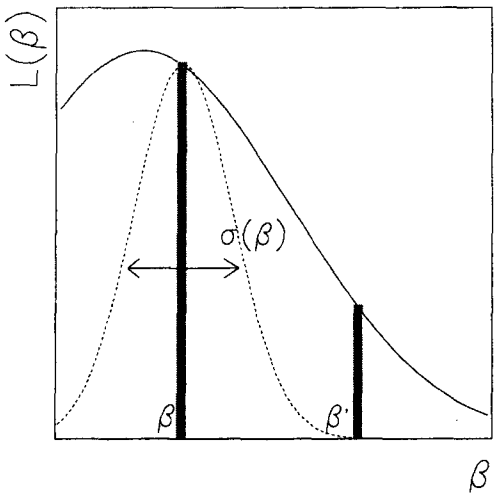


Figure 9.1: Cartoon showing how a representative bin β in the unsmeared likelihood is replaced by a Gaussian of width $\sigma(\beta)$ so that it contributes to multiple bins β' in the smeared likelihood.

each value of β. Qualitatively, the effect of convolving with the Gaussian is to reduce the sharp peak of the likelihood function, lengthen the tail, and increase the value of β_{95} at which 95% of the area under the curve occurs for $\beta < \beta_{95}$.

Most of the rest of this chapter is spent deriving the function $\sigma(\beta)$. We begin by discussing the ways in which a given systematic effect can produce uncertainty in β. Based on this discussion, we derive a general functional form of $\sigma(\beta)$. We then review the relevant systematic effects and discuss how each will be studied. Next we estimate the amount of uncertainty each effect contributes to $\sigma(\beta)$. We conclude the chapter with a revised estimate of the *a priori* limit we expect to set in CDF Run 1 data, this time taking systematic uncertainty into account. Since in this analysis the largest source of uncertainty is not systematic but statistical, including smearing does not increase the *a priori* expected limit much.

9.1 Form of $\sigma(\beta)$: Normalization *vs.* Shape Uncertainties

As an example of the ways in which a given systematic effect can lead to uncertainty on β, consider the specific case of uncertainty in jet energy scale. Suppose all our measurements of jet energies are off by a certain percentage. Since we apply cuts on jet energies, the predicted number of signal events μ_s might be affected. We call uncertainty on μ_s "normalization uncertainty" σ_{μ_s}. Since we also use jet energies in calculating H_T, the shape of the H_T distribution might be affected, giving rise to uncertainty in the number of signal n_s obtained from fitting the distribution. We call uncertainty on n_s "shape uncertainty" σ_{n_s} [1]. In other words, a given systematic effect gives rise to uncertainty on β through either or both the *normalization* or the *shape* of the predicted H_T distribution.

In "factorizing" systematic uncertainty thus into effects on normalization *vs.* effects on shape we follow the precedent of CDF's 1991 top search paper [152]. That paper justified this approach by appealing to the form of β itself. β_s is defined as the relative number of signal events required to fit the data, *i.e.*

$$\beta_s = \frac{n_s}{\mu_s}$$

From error propagation in which we assume no correlations between shape and normalization uncertainties, we can derive

$$\begin{aligned} {\sigma_{\beta_s}}^2 &= \left(\frac{\partial \beta_s}{\partial n_s}\right)^2 {\sigma_{n_s}}^2 + \left(\frac{\partial \beta_s}{\partial \mu_s}\right)^2 {\sigma_{\mu_s}}^2 \\ &= \left(\frac{1}{\mu_s}\right)^2 \left[{\sigma_{n_s}}^2 + \beta_s^2 {\sigma_{\mu_s}}^2\right] \end{aligned}$$

Transforming uncertainties on μ_s and n_s to units of β_s by dividing by μ_s, we define the normalization uncertainty on β_s to be $\beta_s \Delta_{NORM}$, where $\Delta_{NORM} = \frac{\sigma_{\mu_s}}{\mu_s}$, and the shape uncertainty on β_s to be $\Delta_{SHAPE} = \frac{\sigma_{n_s}}{\mu_s}$. With this terminology,

$$\sigma(\beta)^2 = {\Delta_{SHAPE}}^2 + \beta^2 {\Delta_{NORM}}^2$$

[1] The claim that n_s (and n_b) are determined purely by the shapes of the data distribution and the Monte Carlo templates is only true if we neglect the effect of the Gaussian constraints on background. In reality, n_b is constrained to μ_b by the Gaussian terms and n_s is in turn coupled to n_b via the Poisson term, so n_s depends partly on μ_b and not just on shape. Therefore to treat n_s as dependent only on shape (*i.e.* σ_{n_s} as dependent only on shape uncertainty) is an approximation.

This is the form of the width $\sigma(\beta)$ with which we will smear $\mathcal{L}(\beta)$. Note that while Δ_{NORM} is a relative quantity expressing the *fractional* uncertainty on μ, Δ_{SHAPE} is not a fractional but an *absolute* uncertainty on the magnitude of β. Paraphrasing Ref. [152], to see that this must be so, consider the case when the fit returns $\beta = 0$. If we treated Δ_{SHAPE} as a fractional uncertainty, we would obtain zero uncertainty on the number of signal in the fit—"a clearly erroneous result."

Having obtained an expression for $\sigma(\beta)$, we now address the task of computing the terms Δ_{NORM} and Δ_{SHAPE} appearing in it. The sources of uncertainty we consider are summarized in Table 9.1 below. In our choice of effects to study we follow the example of other CDF top analyses [136, 143], with some modifications specific to this analysis (*e.g.* choice of signal generator; background model.)

	Effect	Contributes to Δ_{NORM}	Contributes to Δ_{SHAPE}
1	Monte Carlo statistics	×	
2	Integrated luminosity	×	
3	Lepton-trigger efficiency	×	
4	b-tag efficiency	×	
5	Jet E_T scale	×	×
6	Modeling of initial-state radiation (ISR)	×	×
7	Modeling of final-state radiation (FSR)	×	×
8	Choice of parton distribution function (PDF)	×	×
9	Choice of signal generator	×	×
10	Top quark mass	×	×
11	Background model		×

Table 9.1: Sources of systematic uncertainty affecting this analysis. Whether each source contributes to Δ_{NORM} (normalization uncertainty on β due to σ_{μ_s}) or Δ_{SHAPE} (shape uncertainty on β due to σ_{n_s}), or both, is noted.

Our general approach will be to shift the ith systematic effect by one standard deviation—or some reasonable approximation thereof—and calculate the terms Δ_{NORMi}, Δ_{SHAPEi} resulting from this "1σ shift." After obtaining Δ_{NORMi}, Δ_{SHAPEi} for each effect we

combine the individual terms into a single Δ_{NORM}, Δ_{SHAPE} by adding in quadrature. The methods for shifting each effect by 1σ are summarized in Table 9.2; the specific techniques used to extract Δ_{NORMi}, Δ_{SHAPEi} from the shifted samples are discussed in the two following sections.

9.2 Normalization Uncertainties

We first consider the case of normalization uncertainty $\Delta_{NORM} \equiv \frac{\sigma_{\mu_s}}{\mu_s}$, the fractional uncertainty on the expected number of signal events μ_s. Since $\mu_s = \sigma \cdot \int \mathcal{L}dt \cdot A \cdot$ s.f., the systematic uncertainty on μ_s arises from uncertainties on integrated luminosity, data-to-Monte Carlo scale factors, acceptance, and theory cross section.

9.2.1 Methods for obtaining normalization uncertainties

We already discussed uncertainty on luminosity in Chapter 4 and on scale factors in Chapter 5. In this chapter we derive acceptance uncertainty by recalculating the acceptance A from the "1σ"-shifted samples described in Table 9.2 and comparing to the default A. We omit uncertainty on theory cross section because we do not want the smearing of our data result to depend on theoretical predictions. Note that since μ_s and A only differ by multiplicative constants, the desired fractional uncertainty on μ_s is the same as that on A. The specific methods used to extract the normalization uncertainty Δ_{NORMi} due to each systematic effect are summarized in Table 9.3. Two cases requiring lengthier explanation are discussed in subsections 9.2.1.1 and 9.2.1.2.

9.2.1.1 Normalization uncertainty due to Monte Carlo statistics $(\Delta_{NORM})_{MC\ stats}$

The default Wg and W^* acceptances A_{Wg} and A_{W^*} already have uncertainties due to Monte Carlo statistics. From Bernoulli statistics [2] we can derive a formula for the

[2] Acceptance calculations are analogous to Bernoulli trials in which the probability of success, p, is the ratio of number of successes r to number of trials N. The variance on number of successes is $(\delta r)^2 = Np(1-p)$. In an acceptance calculation $A = (\frac{n_{pass}}{N}) \times$ s.f.. Assuming the scale factors are errorless (and recalling that

	Effect	"1σ Shift" Implementation	Samples Affected	
			S	B
1	MC statistics	—		
2	Luminosity	—		
3	Lepton-trigger efficiency	—		
4	b-tag efficiency	—		
5	Jet E_T scale	Shift E_T scale for all jets in event by $\pm 1\sigma$ using ET_CORR_SYS [153]	×	×
6	ISR	Generate new MC with ISR turned off in PYTHIA	×	
7	FSR	Effectively "turn off FSR" by identifying no-ISRFSR events within a no-ISR sample (all final-state jets match a GENP parton within $\Delta R = 0.4$)	×	
8	PDF	Generate new MC using CTEQ3M PDF in PYTHIA rather than EHLQ default	×	
9	Signal generator	Generate new MC using HERWIG instead of PYTHIA	×	
10	Top mass	Generate new MC with M_{top} set to 170, 180 GeV/c^2 rather than 175 GeV/c^2 default	×	$t\bar{t}$ only
11	Background model	Generate new MC using PYTHIA Wc; mix with VECBOS $Wb\bar{b}$ in the approximate ratio 30%$Wb\bar{b}$:70%Wc (Table 6.6 shows that after all cuts except $M_{\ell\nu b}$, the non-top background is ~30% $Wb\bar{b}$ and $Wc\bar{c}$, ~70% other); use $Wb\bar{b}/Wc$ admixture to model non-top background instead of pure VECBOS $Wb\bar{b}$.		non-top only

Table 9.2: Method for shifting each systematic effect by "1σ." In cases where new Monte Carlo samples for signal or background needed to be generated, the samples were described in the appropriate Signal or Background chapter (Table 5.2 for signal; Tables 6.2 and 6.5 for background.)

uncertainty on an acceptance measured in a sample of N Monte Carlo events and corrected

there is no error on N), $\delta A = (\frac{\text{s.f.}}{N})\delta n_{pass}$. The error on n_{pass} is just the Bernoulli error on number of successes, *i.e.* $(\delta n_{pass})^2 = N(\frac{A}{\text{s.f.}})(1 - \frac{A}{\text{s.f.}})$. This leads to a formula for δA.

	Effect	Method for Estimating Δ_{NORMi}
1	MC statistics	2% statistical uncertainty on individual A_{Wg}, A_{W^*} and combined A_{Wg+W^*} (see Section 9.2.1.1)
2	Luminosity	4% uncertainty on $\int \mathcal{L}dt$ derived from [100] (Section 4.2.5)
3	Lepton-trigger efficiency	10% uncertainty on lepton-trigger & ID-efficiency scale factor sf_{lep} (Section 5.4.1.1)
4	b-tag efficiency	10% uncertainty on b-tag efficiency scale factor ($sf_{lep} \times sf_{1A}$) (Section 5.4.1.2)
5	Jet E_T scale	Computed $A_{\pm 1\sigma}$ using jet-E_T-scale-shifted samples. $(\Delta_{NORM})_{E_T\ scale}$ = the larger of the % differences between $\{A_{+1\sigma}, A_{default}\}$ and $\{A_{-1\sigma}, A_{default}\}$.
6	ISR	Computed A_{noISR} using no-ISR samples. $(\Delta_{NORM})_{noISR} = \frac{1}{2}$ the % difference between A_{noISR}, $A_{default}$.
7	FSR	Derived $A_{noISRFSR}$ using no-ISR samples with match criterion (see Section 9.2.1.2.) $(\Delta_{NORM})_{FSR} = \frac{1}{2}$ the % difference between A_{noISR} and $A_{noISRFSR}$.
8	PDF	Derived A_{CTEQ3M} using CTEQ3M samples. $(\Delta_{NORM})_{PDF}$ = % difference between A_{CTEQ3M}, $A_{default}$.
9	Signal generator	Derived A_{HERWIG} using HERWIG samples. $(\Delta_{NORM})_{generator}$ = % difference between A_{HERWIG}, $A_{default}$.
10	Top mass	Computed $A_{Mtop170}$, $A_{Mtop180}$ using M_{top}-shifted samples. $(\Delta_{NORM})_{Mtop}$ = the larger of the % differences between $\{A_{Mtop170}, A_{default}\}$ and $\{A_{Mtop180}, A_{default}\}$.
11	Background model	—

Table 9.3: Methods for computing the normalization uncertainties Δ_{NORMi}. Note that % differences in acceptance A_s are the same as those in number of signal μ_s.

by (presumed errorless) scale factors:

$$\delta A = \sqrt{\frac{A(\text{s.f.} - A)}{N}}$$

From this formula, and recalling that the sample sizes N used to calculate acceptances were listed in Table 5.1, we obtain the following statistical uncertainties on the default Wg and W^* acceptances:

$$\begin{aligned} A_{Wg} &= 0.01664 \pm 0.00031 \quad (2\% \text{ error}) \\ A_{W^*} &= 0.01672 \pm 0.00034 \quad (2\% \text{ error}) \end{aligned}$$

We now show by error propagation that these uncertainties lead to 2% uncertainty on the total acceptance $A_{Wg+W^*} = A_{Wg}\left(\frac{\sigma_{Wg}}{\sigma_{Wg}+\sigma_{W^*}}\right) + A_{W^*}\left(\frac{\sigma_{W^*}}{\sigma_{Wg}+\sigma_{W^*}}\right)$. We assume zero error on cross sections. Since we do not know the correlations between A_{Wg}, A_{W^*} we use the Schwarz inequality (with abbreviations $A \equiv A_{Wg+W^*}$ and $\sigma \equiv \sigma_{Wg+W^*}$):

$$\begin{aligned} \delta A &\leq \left|\frac{\partial A}{\partial A_{Wg}}\right| \delta A_{Wg} + \left|\frac{\partial A}{\partial A_{W^*}}\right| \delta A_{W^*} \\ &= \frac{\sigma_{Wg}}{\sigma}\delta A_{Wg} + \frac{\sigma_{W^*}}{\sigma}\delta A_{W^*} \end{aligned}$$

Thus to get an upper limit on the statistical error on A_{Wg+W^*} we take the cross-section-weighted sum of the errors on the individual acceptances. $A_{Wg+W^*} = 0.01666 \pm 0.00032$ (2% error), so $\Delta A_{MC\ stats} = 2\%$.

9.2.1.2 Normalization uncertainty due to FSR $(\Delta_{NORM})_{FSR}$

More detail must be supplied on the method used to compute the normalization uncertainty associated with modeling of final-state radiation (FSR). In case of inital-state radiation (ISR) modeling, we generated a sample with ISR turned off, compared the acceptance calculated from this sample with the default acceptance, and took half the difference as the measure of uncertainty due to ISR modeling. Unfortunately, we cannot proceed analogously in the case of FSR because turning off FSR in the Monte Carlo spoils the parton showering of jets. Therefore we must obtain FSR uncertainty in a more roundabout way. We start from the no-ISR sample, infer $A_{noISRFSR}$, and take half the difference between this and A_{noISR} as the uncertainty due to FSR. Inferring $A_{noISRFSR}$ is not entirely trivial. We now describe how this is done.

Starting from the no-ISR sample, we define "match" events to be events in which all jets match a final-state parton from the single-top reaction within a cone of 0.4. "Match" events lack extra jets from gluons and so are presumably no-ISRFSR. The desired $A_{noISRFSR}$ is then the ratio of match events passing all cuts to match events generated:

$$A_{noISRFSR} = \frac{N_{passing\ all\ cuts,\ match}}{N_{gen,\ match}}$$

The problem is that the denominator is unknown. When generating Monte Carlo events, we only save events passing the PARTFILT requirement of a 15-GeV lepton with $|\eta| < 3$; we do not know the number of match-events generated upstream of this filter. Nevertheless, we can deduce the acceptance denominator by making two assumptions.

- **Assumption #1:** Match events are no more or less likely to pass PARTFILT, so the fraction of the sample they occupy is the same before and after the filter. That is, we assume:

  $$\frac{N_{gen,\ match}}{N_{gen}} = \frac{N_{PARTFILT,\ match}}{N_{PARTFILT}}$$

  Unfortunately, the numerator on the RHS, *i.e.* the number of match events passing PARTFILT, is also unknown. This is because leptons that fail the lepton-ID cuts can fake non-match jets unless we explicitly remove them, but in analyzing events, we did not retain the information on each bad lepton necessary to do so. We get around this problem by making a second assumption:

- **Assumption #2:** Match events are no more or less likely to have a "golden" lepton [3], so the match-event fraction of the sample is the same before and after the lepton requirement. That is, we assume:

  $$\frac{N_{PARTFILT,\ match}}{N_{PARTFILT}} = \frac{N_{lepton,\ match}}{N_{lepton}}$$

All quantities on the RHS are known, so we can easily solve for $N_{gen,\ match}$:

$$N_{gen,\ match} = \frac{N_{lepton,\ match}}{N_{lepton}} \times N_{gen}$$

[3] *i.e.* a lepton passing all the lepton-ID cuts

and consequently for the desired $A_{noISRFSR}$:

$$A_{noISRFSR} = \frac{N_{passing\ all\ cuts,\ match}}{N_{lepton,\ match}} \times \frac{N_{lepton}}{N_{gen}}$$

9.2.2 Normalization uncertainty results

Focusing now on those systematic effects which contribute to acceptance uncertainty (E_T scale, gluon radiation, PDF, signal generator, top mass), we present the results of the above-described methods in Tables 9.4 and 9.5 [4]. Note that shifting a given systematic

	Effect	Acceptance ($\times 10^{-2}$)	ΔA_i ($\times 10^{-2}$)	ΔA_i (%)
0	Default	1.664 ± 0.031	—	—
5	Jet E_T scale $+1\sigma$	1.663 ± 0.031	$+0.013 \pm 0.044$	+0.8%
	Jet E_T scale -1σ	1.677 ± 0.031		
6	ISR	1.562 ± 0.041	-0.051 ± 0.026	−3.1%
7	FSR	1.629 ± 0.081	$+0.034 \pm 0.045$	+2.1%
8	PDF	1.569 ± 0.041	-0.095 ± 0.051	−5.7%
9	Signal generator	1.556 ± 0.043	-0.108 ± 0.053	−6.5%
10	Top mass 170 GeV/c^2	1.567 ± 0.044	-0.097 ± 0.054	−5.8%
	Top mass 180 GeV/c^2	1.676 ± 0.046		
	TOTAL ACCEPTANCE UNCERTAINTY		$\Delta A_{Wg} = 0.184 \pm 0.068$	11.1%

Table 9.4: Sources of systematic uncertainty on Wg-fusion signal acceptance. Statistical uncertainties are shown in order to emphasize that while the uncertainties on the individual acceptance shifts are fairly large, the quadrature sum or "total acceptance uncertainty" is known to within reasonable precision.

effect (say, turning off initial-state radiation) does not necessarily have the same impact on Wg and W^* signal acceptance (ΔA_{noISR} for $Wg = -3\%$; ΔA_{noISR} for $W^* = +2\%$.) This is not surprising, given the different kinematics of the two signal processes [5].

[4] To review the methods used to obtain errors appearing in these tables: Statistical error on ith **acceptance** is $\delta A_i = \sqrt{A_i(\text{s.f.} - A_i)/N_i}$. (Table 5.2 lists N_i for each sample.) Acceptance shift $\Delta A_i = A_i' - A_{default}$ (in some cases, multiplied by a factor of $\frac{1}{2}$); error on **acceptance shift** $\delta(\Delta A_i) = \sqrt{(\delta A_i')^2 + (\delta A_{default})^2}$ (neglecting correlations between A_i', $A_{default}$.) Total acceptance uncertainty $(\Delta A)^2 = \sum(\Delta A_i)^2$; error on **total acceptance uncertainty** $\delta(\Delta A) = \frac{1}{\Delta A}\left(\sum(\Delta A_i)^2(\delta A_i)^2 + (\delta A_{default})^2(\sum \Delta A_i)^2\right)^{\frac{1}{2}}$, again neglecting correlations between A_i', $A_{default}$.

[5] Specifically, when ISR is turned off, PYTHIA uses the $2 \to 2$ process ($qb \to q't$) to generate Wg. The loss of the second b jet in the final state contributes to a reduced b-tag efficiency for the event. (See Figures 5.3

	Effect	Acceptance ($\times 10^{-2}$)	ΔA_i ($\times 10^{-2}$)	ΔA_i (%)
0	Default	1.672 ± 0.034	—	—
5	Jet E_T scale $+1\sigma$	1.665 ± 0.034	$+0.030 \pm 0.048$	+1.8%
	Jet E_T scale -1σ	1.702 ± 0.034		
6	ISR	1.729 ± 0.049	$+0.029 \pm 0.030$	+1.7%
7	FSR	1.846 ± 0.102	$+0.059 \pm 0.057$	+3.4%
8	PDF	1.672 ± 0.048	$+0.0 \pm 0.059$	+0.0%
9	Signal generator	1.835 ± 0.050	$+0.163 \pm 0.060$	+9.7%
10	Top mass 170 GeV/c^2	1.675 ± 0.048	-0.062 ± 0.058	−3.7%
	Top mass 180 GeV/c^2	1.610 ± 0.047		
	TOTAL ACCEPTANCE UNCERTAINTY		$\Delta A_{W^*} = 0.189 \pm 0.069$	11.3%

Table 9.5: Sources of systematic uncertainty on W^* signal acceptance.

Note also that despite the fact that the acceptance-shifts reported in Tables 9.4 and 9.5 are not very statistically significant [6], the total quadrature sum of all the shifts (*i.e. total acceptance uncertainty*, the quantity of interest) is known to within reasonable precision. This is expected, since it can be demonstrated that the uncertainty on the quadrature sum of a list of uncorrelated numbers is on the order of the uncertainty on any one of the numbers in the list.

Combining acceptance uncertainty ΔA with uncertainties due to Monte Carlo statistics, integrated luminosity, and scale factors, we obtain the desired total normalization uncertainty. For the separate processes we obtain $11.1\% \oplus 2\% \oplus 4\% \oplus 10\% \oplus 10\% = 18.5\%$ uncertainty on μ_{Wg} and $11.3\% \oplus 2\% \oplus 4\% \oplus 10\% \oplus 10\% = 18.6\%$ uncertainty on μ_{W^*}. However, since in this analysis we treat $Wg + W^*$ as a single signal, we need the normalization uncertainty on the combined amount of signal $\mu_s = \mu_{Wg+W^*}$. We show the shifts in combined acceptance A_{Wg+W^*} in Table 9.6. (Statistical errors are omitted here, but the same

and 5.4 for the kinematics of the second b quark in a Wg-fusion event.) This effect outweighs the improved $M_{\ell\nu b}$ resolution and lower loss to the $W \leq 3$-jet cut that we would expect in a no-ISR sample. These effects *do* cause improved acceptance for W^*, which does not lose a final-state b jet when ISR is turned off. Hence, when ISR is turned off we see the Wg acceptance go down while the W^* acceptance goes up.

[6] Of course, we could reduce the statistical uncertainty by generating more shifted Monte Carlo. However, the improved precision on the overall acceptance uncertainty that would be achieved by halving the statistical uncertainties on the acceptance shifts was not judged worth quadrupling Monte Carlo samples.

argument regarding uncertainties on individual shifts *vs.* uncertainties on the quadrature sum of shifts hold as in the cases of the individual acceptances.) Note that since some of

	Effect	A_{Wg} ($\times 10^{-2}$)	A_{W^*} ($\times 10^{-2}$)	A_{Wg+W^*} ($\times 10^{-2}$)	ΔA_{Wg+W^*} (%)
0	Default	1.664	1.672	1.666	—
5	Jet E_T scale $+1\sigma$	1.663	1.665	1.664	1.1%
	Jet E_T scale -1σ	1.677	1.702	1.685	
6	ISR	1.562	1.729	1.612	1.6%
7	FSR	1.629	1.846	1.694	2.5%
8	PDF	1.569	1.672	1.600	4.0%
9	Signal generator	1.556	1.835	1.640	1.6%
10	Top mass 170 GeV/c^2	1.567	1.675	1.599	4.0%
	Top mass 180 GeV/c^2	1.676	1.610	1.656	
	TOTAL UNCERTAINTY ON COMBINED ACCEPTANCE				$\Delta A_{Wg+W^*} = 6.7\%$

Table 9.6: Systematic uncertainty on total signal acceptance A_{Wg+W^*}.

the acceptance shifts are anticorrelated for Wg and W^*, the uncertainty on the combined acceptance (~7%) is smaller than that on either of the individual acceptances (~11%). The uncertainties on A_{Wg+W^*}, along with all other contributions Δ_{NORMi} to uncertainty on μ_s, are summarized in Table 9.7. Adding in quadrature, we obtain a total fractional uncertainty on μ_s of $\Delta_{NORM} = 16\%$.

9.3 Shape Uncertainties

Now we turn to the problem of estimating $\Delta_{SHAPE} \equiv \frac{\sigma_{n_s}}{\mu_s}$, or the $\frac{1}{\mu}$-scaled absolute uncertainty on fitted number of signal events n_s due to uncertainty on the shape of the H_T distribution. The method for estimating the uncertainty Δ_{SHAPEi} due to the ith systematic effect is as follows. We again use the signal (and in some cases, background) samples in which the effect of interest is shifted by some approximation of one standard deviation as described in Table 9.2. We reapply cuts; remake smoothed H_T distributions for events passing cuts; throw pseudoexperiments from the new, shifted distributions; and fit the pseudoexperiments to the default, unshifted templates. In all cases we throw pseudoexperiments using the

	Effect	% normalization uncertainty (Δ_{NORMi})
1	MC statistics	2%
2	Luminosity	4%
3	Lepton-trigger efficiency	10%
4	b-tag efficiency	10%
5	Jet E_T scale	1.1%
6	ISR	1.6%
7	FSR	2.5%
8	PDF	4.0%
9	Signal generator	1.6%
10	Top mass	4.0%
	TOTAL	$\Delta_{NORM} = 16\%$

Table 9.7: Summary of contributions Δ_{NORMi} to fractional uncertainty on μ_s ("normalization uncertainty.")

default expected amount of signal and background as shown in Table 6.9. We plot the distribution of $\frac{1}{\mu_s}$-scaled [7] fitted n_s, *i.e.* β_s, for these "shifted" pseudoexperiments and note the mean of the distribution, $\langle\beta_s\rangle$. We compare it to $\langle\beta_s\rangle_{default}$, the mean of the analogous distribution from a set of pseudoexperiments thrown from the default templates. Figure 9.2 shows the distribution of β_s for default pseudoexperiments. We take the difference $\Delta\langle\beta_s\rangle_i = \langle\beta_s\rangle_i - \langle\beta_s\rangle_{default}$ between the means of the β distributions for shifted *vs.* default pseudoexperiments to be an estimate of the absolute shape uncertainty Δ_{SHAPEi} on β due to the *ith* systematic effect. Figure 9.3 shows schematically how $\Delta\langle\beta_s\rangle_i = \Delta_{SHAPEi}$ might be obtained for one systematic effect.

9.3.1 Calculating Δ_{SHAPE} for $\beta = 1$

For each of the seven sources of shape uncertainty listed in Table 9.1, we threw default and shifted pseudoexperiments using the expected amounts of signal and background ($\beta_s = \beta_{t\bar{t}} = \beta_{nt} = 1.0$). The techniques used to generate the samples for the shifted pseudoexperiments were described in Table 9.2. The techniques used to extract estimates

[7] Recall the default amount of signal is $\mu_s = 4.29$

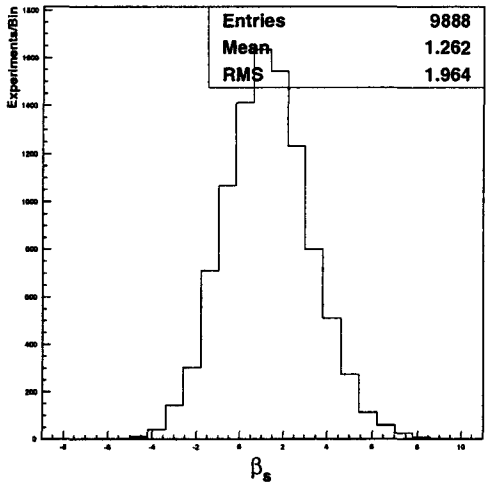


Figure 9.2: Distribution of β_s for pseudoexperiments thrown from the default templates and using input $\beta_s = \beta_{t\bar{t}} = \beta_{nt} = 1.0$ ("default pseudoexperiments.") The distribution of n_s for these same pseudoexperiments was shown in Figure 7.3. A discussion of the bias that causes the mean of this distribution, $\langle \beta_s \rangle_{default}$, to be offset from 1.0 may also be found in that chapter.

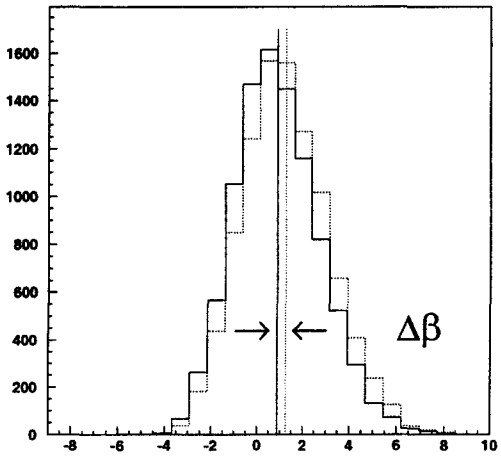


Figure 9.3: Cartoon showing how $\Delta\langle \beta_s \rangle_i$ is calculated.

of Δ_{SHAPEi} from differences $\Delta\langle \beta_s \rangle_i$ are summarized in Table 9.8 below. Table 9.9 lists $\langle \beta_s \rangle_i$ and Δ_{SHAPEi} for each source of shape uncertainty we considered. Adding in quadrature, we

	Effect	Method for Computing Δ_{SHAPEi}
5	Jet E_T scale	Made new templates for signal and bkgd. $(\Delta_{SHAPE})_{E_T\ scale}$ = symmetrized $\frac{1}{2}$-difference between $\langle\beta\rangle_{\pm 1\sigma}$.
6	ISR	Made new templates for signal; used default templates for bkgd. $(\Delta_{SHAPE})_{noISR}$ = symmetrized $\frac{1}{2}$-difference between $\langle\beta\rangle_{noISR}$ and $\langle\beta\rangle_{default}$.
7	FSR	Made new templates for signal from the no-ISRFSR subsample; used default templates for bkgd. $(\Delta_{SHAPE})_{noISRFSR}$ = symmetrized $\frac{1}{2}$-difference between $\langle\beta\rangle_{noISRFSR}$ and $\langle\beta\rangle_{noISR}$.
8	PDF	Made new templates for signal; used default templates for bkgd. $(\Delta_{SHAPE})_{PDF}$ = symmetrized difference between $\langle\beta\rangle_{CTEQ3M}$ and $\langle\beta\rangle_{default}$.
9	Signal generator	Made new templates for signal; used default templates for bkgd. $(\Delta_{SHAPE})_{generator}$ = symmetrized difference between $\langle\beta\rangle_{HERWIG}$ and $\langle\beta\rangle_{default}$.
10	Top mass	Made new templates for signal and $t\bar{t}$ bkgd; used default template for non-top bkgd. $(\Delta_{SHAPE})_{Mtop}$ = symmetrized $\frac{1}{2}$-difference between $\langle\beta\rangle_{Mtop170}$ and $\langle\beta\rangle_{Mtop180}$.
10	Background model	Made new template made for non-top bkgd; used default templates for signal and $t\bar{t}$ bkgd. $(\Delta_{SHAPE})_{bkgdmodel}$ = symmetrized difference between $\langle\beta\rangle_{bkgdmodel}$ and $\langle\beta\rangle_{default}$.

Table 9.8: Methods for computing the shape uncertainties Δ_{SHAPEi}.

obtain total shape uncertainty $\Delta_{SHAPE}|_{\beta=1} = 0.27$.

9.3.2 Calculating Δ_{SHAPE} for other values of β

Having calculated Δ_{SHAPE} at $\beta = 1$, we repeat the procedure for other values of β. That is, we measure absolute shifts Δ_{SHAPEi} in $\langle\beta_s\rangle_i$ for pseudoexperiments thrown with

i	Effect	$\langle\beta_s\rangle_i$	Δ_{SHAPEi}
0	Default	1.262	
5	Jet E_T scale $+1\sigma$	1.406	$\pm$0.21365
	Jet E_T scale -1σ	0.9787	
6	ISR	1.131	$\pm$0.0655
7	FSR	1.095	$\pm$0.018
8	PDF	1.246	$\pm$0.016
9	Signal generator	1.141	$\pm$0.121
10	Top mass 170 GeV/c^2	1.255	$\pm$0.043
	Top mass 180 GeV/c^2	1.169	
11	Background model	1.201	$\pm$0.061
	TOTAL		$\Delta_{SHAPE}\|_{\beta=1} = 0.27$

Table 9.9: Absolute uncertainties on β_s due to sources of error that affect the shape of the H_T distribution, evaluated at $\beta_s = 1$.

input $\beta_s = 0, 2, 4, \ldots$ *etc.* For each different input value of β_s, we adjust input values of $\beta_{t\bar{t}}$ and β_{nt} so that the total expected number of events is always equal to the predicted number 66.7 (Table 6.9). This accounts for the fact that for a given number of data events, the Poisson term in the likelihood ensures that a fit with a larger (smaller) value of β_s will have correspondingly smaller (larger) values of $\beta_{t\bar{t}}$, β_{nt} [8].

For each set of pseudoexperiments we make a table of $\{\Delta_{SHAPEi}\}$'s and add them in quadrature to obtain an overall Δ_{SHAPE} exactly as done in the previous section for $\beta = 1$. In this way, we assess the dependence of Δ_{SHAPE} on input β. Since each measurement of $\Delta_{SHAPE}|_\beta$ requires many tens of thousands of pseudoexperiments, only a few representative points in (β, Δ_{SHAPE})-space are sampled. The values of β and $\Delta_{SHAPE}|_\beta$ are summarized in Table 9.10.

These points are plotted in Figure 9.4 and fitted to a quadratic function. This function permits us to interpolate a value of Δ_{SHAPE} for any given value of β.

[8] Recall that the Poisson term $\frac{e^{-n}n^{N_{obs}}}{N_{obs}!}$, where $n = n_{Wg} + n_{W^*} + n_{t\bar{t}} + n_{nt}$, keeps n from deviating from N_{obs} with a Poisson penalty.

Input β	Input n_s	Input $n_{t\bar{t}}$	Input n_{nt}	$\Delta_{SHAPE}\|_\beta$
0	0.0	8.96	57.7	**0.27**
1	4.29	8.38	54.0	**0.27**
2	8.58	7.80	50.3	**0.37**
4	17.16	6.65	42.9	**0.61**
8	34.32	4.35	28.0	**0.95**
12	51.48	2.04	13.2	**0.80**
15	64.35	0.31	2.01	**0.18**

Table 9.10: Shape uncertainties evaluated at various values of β_{signal}.

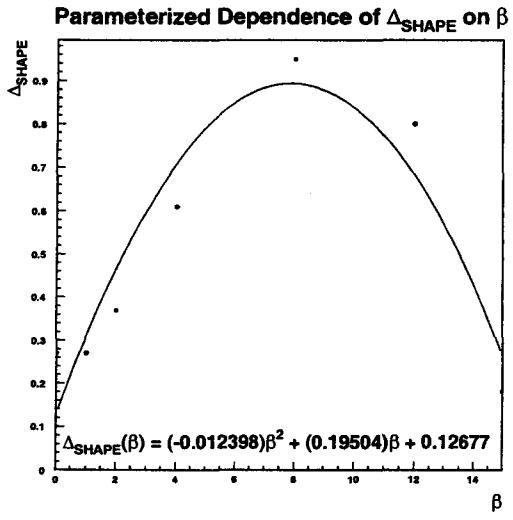


Figure 9.4: Δ_{SHAPE} can be extracted for β in the range $(0, 15)$ using this interpolating function $\Delta_{SHAPE}(\beta)$.

9.4 Smearing the Likelihood Function

Having obtained expressions for both Δ_{NORM} and $\Delta_{SHAPE}(\beta)$, we have now derived the function $\sigma(\beta)$ to be used as the width of the Gaussian for smearing the reduced likelihood function $\mathcal{L}(\beta)$. The smeared likelihood is

$$\mathcal{L}_{smeared}(\beta') = \int_0^\infty \mathcal{L}(\beta) G(\beta; \sigma(\beta)) d\beta$$

where $\sigma(\beta) = \sqrt{\Delta_{SHAPE}(\beta)^2 + \beta^2 \Delta_{NORM}{}^2}$.

We perform the smearing by Monte Carlo methods. Treating $\mathcal{L}(\beta)$ as a p.d.f. for β, we sample ten million values of β. For each value of β, we compute $\sigma(\beta)$ using fixed Δ_{NORM} and interpolated $\Delta_{SHAPE}(\beta)$. We then fluctuate the value β by a Gaussian of width $\sigma(\beta)$. The distribution of the thus-smeared β values, after re-smoothing, constitutes our smeared likelihood function $\mathcal{L}_{smeared}(\beta')$.

9.4.1 A priori limit, with smearing

Using this method for smearing the likelihood function, we can now calculate upper limits β_{95} that incorporate systematic uncertainty. Figure 9.5 shows the distributions of β_{95} calculated with and without smearing for 1000 pseudoexperiments.

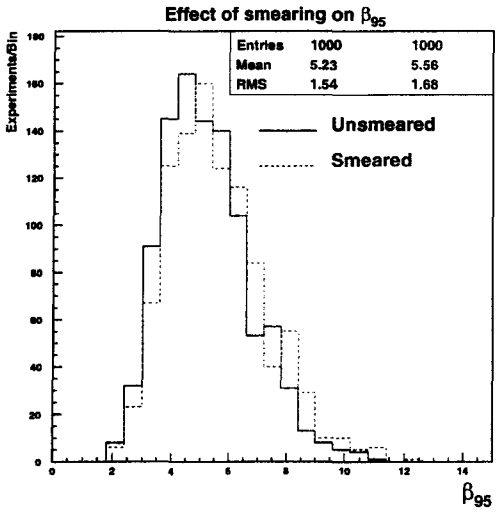


Figure 9.5: Smearing with systematic uncertainty causes the average value of β_{95}, or the limit we can expect to set in CDF data, to increase from 5.2 to 5.6.

That the effect of including systematic uncertainty only raises the *a priori* β_{95} from 5.2 to 5.6 reflects the fact that the dominant source of uncertainty in this analysis is not systematic, but statistical. Recall that in Figure 7.3 we saw that in default pseudoexperiments using the expected amounts of signal and background, the mean value of the fit parameter

n_s was 5.4 [9] but the mean value of the *error* on that fit parameter was 8.1. If the statistical error is on the order of 150%, we do not expect the additional systematic smearing by $\sigma(\beta)/\beta \simeq \sqrt{[\Delta_{SHAPE}(1.26)]^2 + (1.26)^2[\Delta_{NORM}]^2}/1.26 = \sqrt{(0.35)^2 + (1.26)^2(0.16)^2}/1.26 = 0.40/1.26 \simeq 30\%$ to have a large effect.

Thus the limit we expect to set in CDF data, with smearing, is 5.56×2.43 pb= 13.5 pb.

[9] Which corresponds to a mean value of $\beta_s = 5.421/4.29 = 1.26$

Chapter 10

Results from CDF Data

Thus far we have developed a set of techniques—for selecting events, for extracting an upper limit on the signal content of the H_T distribution, for incorporating systematic uncertainty into the calculation of that upper limit—using only Monte Carlo events. In this chapter we at last apply these techniques to real CDF data. We begin by describing the results of applying the selection cuts to data. Sixty-five events pass all cuts. We next compare the data numbers of events per jet-multiplicity bin and the H_T distribution to the predictions of Monte Carlo and find excellent agreement. Then we fit the data to extract a measure of the single-top content and obtain, predictably, a value with too large of a statistical uncertainty to be very useful. Finally we integrate the product of likelihood function $\times$ flat prior to obtain the desired upper limit on the single-top content of data at 95% confidence level. We check this Bayesian-calculated value against the frequentist definition and find it is conservative by frequentist standards.

10.1 Data Strip

Events from CDF Run 1A and Run 1B are organized into streams according to which Level 3 trigger each event passed. In this analysis we use events from high-p_T central electron and muon triggers, which correspond to Stream 2 in Run 1A and Stream A in Run 1B [154]. Inclusive electron and muon data sets are drawn from the larger data sample

using loose cuts. For both Run 1A and Run 1B our cuts are standard enough that we can use preexisting stripped samples made for previous analyses. For Run 1A, we use tapes ICE1A1 and IMU1A1 provided by CDF Top Group convenor R. Roser. These tapes contain approximately 130,000 electron events and 80,000 muon events. For Run 1B, we use the TOPFND strip of the so-called "Gordon-Glenzinski" inclusive lepton sample. This strip, described in [120], applies tight lepton-ID cuts (no cut on isolation) to reduce the Gordon-Glenzinski electron sample from roughly 750,000 to 130,000 events and the muon sample from 570,000 to 90,000 events. Table 10.1 lists the samples we used to select events. In summary, we effectively analyzed all data from CDF Run 1A + Run 1B, corresponding to a total integrated luminosity of 106.0 ± 4.1 pb^{-1}.

Run	Lepton	Dataset	Level 3 Triggers	Filenames	# Events
1A	Electron	ICE1A1	COMBINED_ELE2_CEM	ICE_CCxxxx.PAD	133,805
			(p_T cut 18 GeV/c)		
1A	Muon	IMU1A1	COMBINED_MUO2_CMU	WZ_CM_CCxxxx.STRX_3P	84,645
			(p_T cut 18 GeV/c)		
1B	Electron	EIA	ELEA_CEM_18	NxxxxxMM_EIA_TOPFND.PAD	128,761
			ELEA_CEM_50		
1B	Muon	MUA	MUOA_CMU_18	NxxxxxMM_MUA_TOPFND.PAD	90,908
			MUOA_CMX_18	(N=A,E; MM=2 letters)	
			MUOA_CMP_18		
			MUOA_CMU_AND_CMP_18		
			MUOA_CMU_ETA_GAP		

Table 10.1: Run 1 high-p_T inclusive central lepton samples used to select events. Level 3 trigger names are taken from Table 2 of [155].

10.1.1 Lepton-z fix to data

Before applying the analysis cuts we perform a minor fix to CDF data. Unlike Monte Carlo events, data events may contain multiple vertices due to multiple interactions per bunch crossing. Only one of these vertices corresponds to the high-p_T event of interest; any others, if present, almost certainly come from low-q^2 "minimum-bias" scatterings. The

high-p_T lepton certainly came from the high-p_T vertex. To make sure that calorimetry quantities are calculated with respect to the best guess at the primary vertex, we recluster jets using the z coordinate of the P.V. closest to the lepton if more than one candidate vertex is identified by CDF's vertex-finder software. CDF routine ZLECHG [156] implements this change.

10.1.2 Events passing cuts

Tables 10.2 and 10.3 list good-run events in the 1, 2, or 3-jet bins which pass our lepton, $\not\!\!E_T$, and b-tag criteria. Events which also pass Z and dilepton removal and the kinematic cut $M_{\ell\nu b} = 175 \pm 35$ GeV/c^2 are indicated with an asterisk. Sixty-five events in CDF Run 1 data pass all cuts.

CDF Run 1A

Run	Event	# Jets	Lepton	Pass All Cuts
41477	57938	1	e	*
41838	168601	1	e	
42686	8608	1	e	*
43670	146935	1	e	
40759	21592	1	μ (CMUP)	*
47615	20806	1	μ (CMUP)	*
40190	98182	2	e	*
41540	127085	2	e	
41627	87219	2	e	*
43066	66455	2	e	
45776	386857	2	e	*
45801	316268	2	e	
46935	266805	2	e	
47439	128290	2	e	*
46357	511399	2	μ (CMX)	
47689	80060	2	μ (CMUP)	
43096	47223	3	e	
46269	44897	3	e	*
46818	221912	3	μ (CMX)	*
TOTAL			19	10

Table 10.2: Good-run Run 1A events in the $1, 2, 3$-jet bins passing lepton, $\not\!\!E_T$, and b-tag requirements. Events which in addition pass removals and $M_{\ell\nu b}$ cut are flagged with an asterisk.

CDF Run 1B

Run	Event	# Jets	Lepton	Pass All Cuts
58341	40319	1	e	*
58400	180048	1	e	
58485	4226	1	e	*
60293	212047	1	e	*
61094	173847	1	e	
61169	37139	1	e	
61226	210745	1	e	
63191	132125	1	e	
63209	565078	1	e	*
63209	659741	1	e	
63581	13494	1	e	
63684	165025	1	e	*
64021	46988	1	e	
64311	43078	1	e	*
64880	623030	1	e	
64916	472453	1	e	*
65022	54018	1	e	
66367	882795	1	e	
66435	3103	1	e	
67327	103984	1	e	*
67561	296209	1	e	
67757	460709	1	e	*
67899	292190	1	e	
67989	25379	1	e	*
68110	123757	1	e	*
68185	122336	1	e	*
68333	44335	1	e	
68570	669666	1	e	*
68636	597980	1	e	
68637	252145	1	e	*
68638	8084	1	e	*
68685	411056	1	e	
69007	19019	1	e	*
69007	60885	1	e	
70543	235167	1	e	*
70965	207907	1	e	

CDF Run 1B, cont'd

Run	Event	# Jets	Lepton	Pass All Cuts
58400	230496	1	μ (CMX)	
58887	12132	1	μ (CMX)	
60004	87159	1	μ (CMX)	*
61188	287356	1	μ (CMU)	*
61525	167543	1	μ (CMUP)	
62912	4552	1	μ (CMP)	
63127	173868	1	μ (CMUP)	*
63370	109338	1	μ (CMUP)	*
63541	173943	1	μ (CMUP)	*
63946	194706	1	μ (CMU)	*
64185	420652	1	μ (CMUP)	
64700	98497	1	μ (CMUP)	*
65277	404855	1	μ (CMUP)	*
65298	736831	1	μ (CMU)	
65426	273116	1	μ (CMUP)	
65721	139367	1	μ (CMUP)	*
65751	179487	1	μ (CMUP)	
65834	199574	1	μ (CMU)	
66185	669100	1	μ (CMX)	
67561	450608	1	μ (CMX)	
67842	5949	1	μ (CMUP)	
68570	802975	1	μ (CMX)	
68685	755190	1	μ (CMUP)	*
69408	39712	1	μ (CMP)	*
69500	24560	1	μ (CMUP)	*
69876	48956	1	μ (CMUP)	
69956	562345	1	μ (CMUP)	
70559	489196	1	μ (CMUP)	
70577	158396	1	μ (CMUP)	

CDF Run 1B, cont'd

Run	Event	# Jets	Lepton	Pass All Cuts
57621	45230	2	e	
60766	299452	2	e	
61167	368226	2	e	*
63883	935	2	e	
64126	52063	2	e	*
64293	534789	2	e	
64916	499208	2	e	*
64997	46557	2	e	*
65022	34157	2	e	*
65298	907072	2	e	*
65384	266051	2	e	*
65470	4390	2	e	*
66103	563542	2	e	
66412	121506	2	e	*
67692	420568	2	e	
68044	53510	2	e	*
68374	364586	2	e	
68464	275644	2	e	
68986	185443	2	e	
69520	136405	2	e	
69683	21986	2	e	*
69709	173294	2	e	*
70965	296522	2	e	
59103	196685	2	μ (CMX)	
60705	93795	2	μ (CMUP)	
61377	114526	2	μ (CMX)	
63603	4029	2	μ (CMX)	*
63946	43019	2	μ (CMUP)	*
65741	654870	2	μ (CMX)	
65750	106257	2	μ (CMUP)	*
66103	743101	2	μ (CMUP)	
66518	203555	2	μ (CMUP)	
68231	157759	2	μ (CMUP)	
68423	3326	2	μ (CMUP)	
68593	37659	2	μ (CMUP)	*
68637	225974	2	μ (CMUP)	
68774	150313	2	μ (CMUP)	*
69498	36574	2	μ (CMUP)	*
69761	157205	2	μ (CMUP)	

CDF Run 1B, cont'd

Run	Event	# Jets	Lepton	Pass All Cuts
56911	114159	3	e	
61074	103772	3	e	
63417	142823	3	e	
64934	416715	3	e	*
66573	107219	3	e	
67824	281883	3	e	
67899	82457	3	e	*
69761	99801	3	e	
69808	639398	3	e	
70627	56836	3	e	*
56669	21631	3	μ (CMU)	
59124	31243	3	μ (CMUP)	*
61024	217119	3	μ (CMUP)	
61548	284898	3	μ (CMUP)	*
65025	152	3	μ (CMUP)	*
65277	209495	3	μ (CMUP)	*
65581	322592	3	μ (CMU)	
67879	407958	3	μ (CMX)	*
67971	55023	3	μ (CMX)	
TOTAL			123	55

Table 10.3: Good-run Run 1B events in the $1, 2, 3$-jet bins passing lepton, $\not\!\!E_T$, and b-tag requirements. Events which in addition pass removals and $M_{\ell\nu b}$ cut are flagged with an asterisk.

10.2 Comparing Data to Monte Carlo Prediction

The numbers of data events passing cuts in each jet-multiplicity bin are summarized in Table 10.4. Also shown are the predicted numbers of events from each process in each bin, as taken from Table 6.9.

Comparing the data H_T distribution to the sum of predicted H_T shapes from signal and background Monte Carlo, weighted by their respective predicted abundances, we find excellent agreement (Figure 10.1). Applying the Kolmogorov-Smirnov test to the two

Process	$W+1$ jet	$W+2$ jet	$W+3$ jet	$W+1,2,3$ jet
Wg	0.80	1.45	0.75	3.00 ± 0.56
W^*	0.28	0.79	0.23	1.29 ± 0.24
$t\bar{t}$	0.21	2.27	5.90	8.38 ± 2.68
non-top	37.4	13.9	2.7	54 ± 12
Total	38.7	18.4	9.6	66.7
Observed	**33**	**22**	**10**	**65**

Table 10.4: Numbers of events predicted and observed in CDF data.

distributions, we find the probability of compatibility between the data histogram and the summed Monte-Carlo histogram to be 50%.

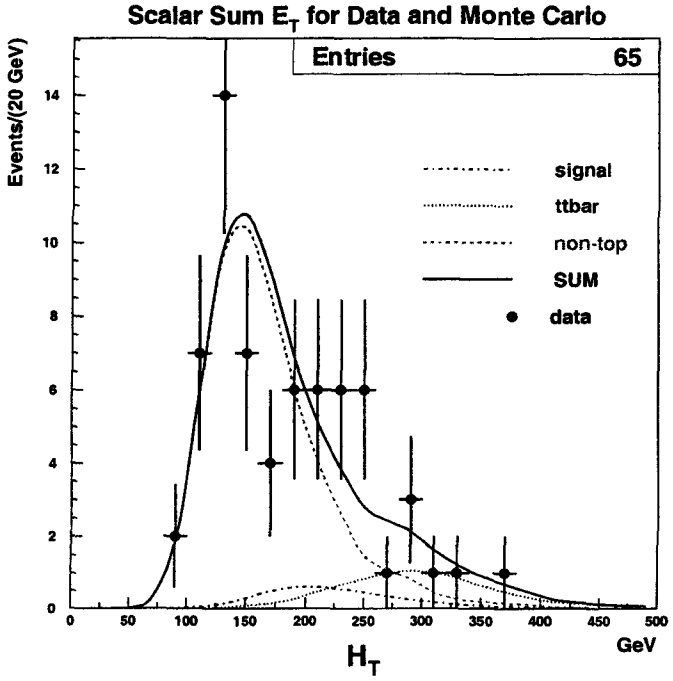


Figure 10.1: H_T distribution for data, with smoothed Monte Carlo predictions for signal and backgrounds. (In this plot, the smooth appearance of the Monte Carlo templates is enhanced with PAW; the actual smoothed templates used in the fit were shown in Figure 7.2. K-S comparison of data and Monte Carlo is of course performed using *unsmoothed* templates.)

10.3 Fitting the Data H_T Distribution

We apply the likelihood fit method of Chapter 7 to the H_T distribution from CDF data. The result we obtain, expressed both in absolute and relative number of events, is

$$n_s = 8.73 \pm 7.76 \quad \text{events, or}$$

$$\beta_s = 2.04 \pm 1.81 \quad \text{(statistical errors only)}$$

The complete fit results for signal and background are summarized in Table 10.5. The size

Process	# Predicted	# Fitted	Fitted β
$Wg + W^*$ signal	4.3 ± 0.7	8.7 ± 7.8	2.04 ± 1.81
$t\bar{t}$ background	8.4 ± 2.7	7.6 ± 2.5	0.90 ± 0.30
non-top background	54 ± 12	50.8 ± 7.6	0.94 ± 0.14

Table 10.5: Fit results: Comparing predicted and fitted amounts of signal and background in CDF data. Errors are statistical only.

of the statistical uncertainty on the fitted n_s confirms our expectation that we do not have sufficient data to measure the single-top cross section. Instead we must set an upper limit on this quantity.

10.4 Extracting β_{95} from Data

To obtain the upper limit on β_s in CDF data, we perform the Bayesian procedure for integrating the product of the reduced likelihood function and the flat prior as outlined in Chapter 8. To recap, we use fixed $\Delta_{NORM} = 16\%$ [1] and interpolated $\Delta_{SHAPE}(\beta)$ (Figure 9.4) in width $\sigma(\beta) = \sqrt{\Delta_{SHAPE}(\beta)^2 + \beta^2 \Delta_{NORM}{}^2}$ of the Gaussian to smear the reduced likelihood function $\mathcal{L}(\beta)$ for data. Integrating $\mathcal{L}(\beta)$, before and after smearing, up to β_{95} such that

[1] We emphasize again that Δ_{NORM} omits theory uncertainty on the single-top cross sections in order that our result not depend on this prediction. If σ_{theory} changes, so does our predicted μ_s. But since our fit parameter is ultimately n_s, the product $\beta_s \mu_s$ does not change. Neither does the product $\sigma_{95} = \beta_{95}\sigma_{theory}$. So our limit would be unaffected by a change in σ_{theory}. [This discussion is not precisely correct once smearing is taken into account, since the smearing width depends on the magnitude of β.]

95% of the area under the curve occurs for $\beta < \beta_{95}$, we obtain the following estimates of β_{95} for data:

$$\text{Before smearing}: \quad \beta_{95} = 5.509$$

$$\text{After smearing}: \quad \beta_{95} = 5.895$$

Figure 10.2 shows the shape of the reduced likelihood function for data with smeared and unsmeared values of β_{95} indicated.

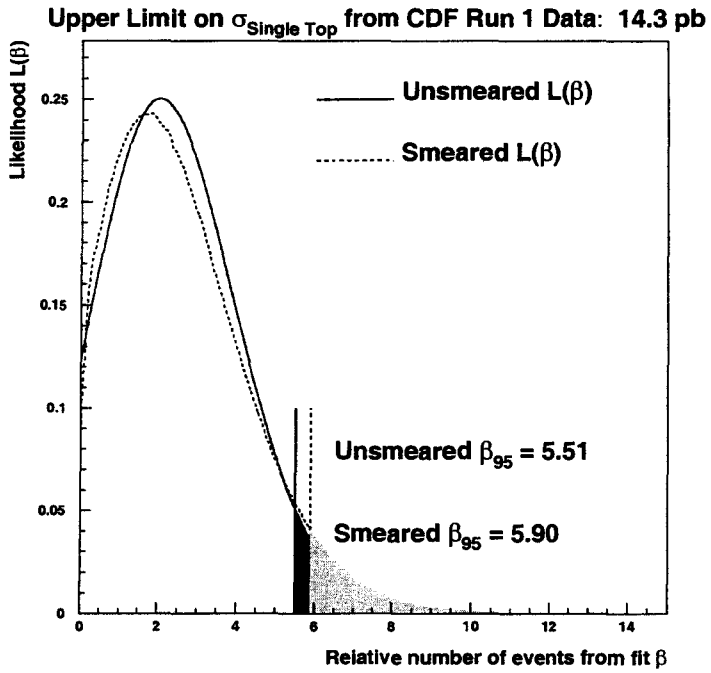


Figure 10.2: Likelihood function for data.

10.5 Frequentist Check of Upper Limit

Having obtained a value for β_{95} from data, we are now in a position to apply the frequentist check of this value as described in Section 8.1.4. We throw pseudoexperiments

with input $\beta = \beta_{95}$ [2], fit each pseudoexperiment to obtain $\hat{\beta}$, and smear each $\hat{\beta}$ according to a Gaussian of width $\sigma(\hat{\beta}) = \sqrt{\Delta_{SHAPE}(\hat{\beta})^2 + \hat{\beta}^2 \Delta_{NORM}{}^2}$. Plotting the values of smeared $\hat{\beta}$ for each pseudoexperiment yields distribution $f(\hat{\beta}|\beta_{95})$. Next we integrate this distribution out to $\beta_{exp} = 2.04$, *i.e.*. the particular value of $\hat{\beta}$ obtained from the fit to real data. If this integral is less than or equal to 5%, then the Bayesian-calculated value of β_{95} coincides with or exceeds the frequentist definition. Figure 10.3 shows the distribution of smeared $\hat{\beta}$ values for the Bayesian-calculated $\beta_{95} = 5.895$. We find the integral is equal to 3%, indicating that

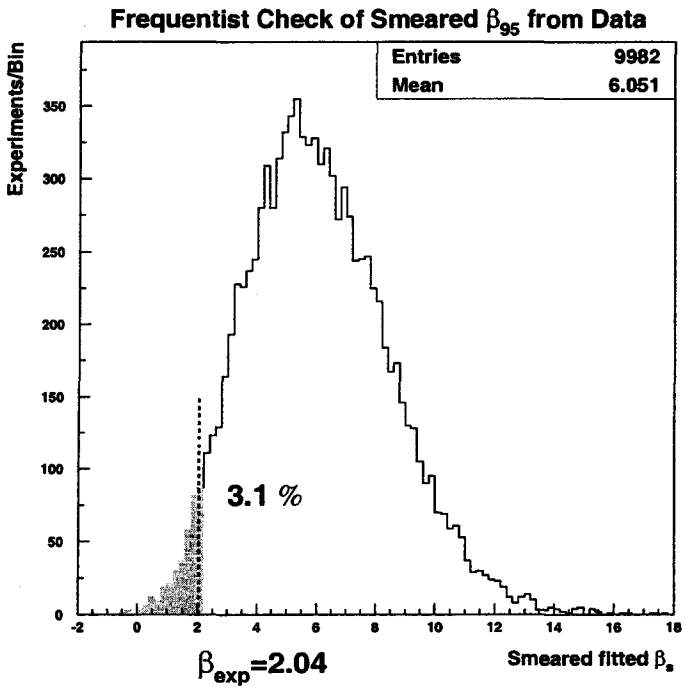


Figure 10.3: Frequentist check: Integrating the distribution of smeared $\hat{\beta}$ values $f(\hat{\beta}|\beta_{95})$ up to $\beta_{exp} = 2.04$. (The actual integration was performed using PAW++ on a much more finely-binned histogram than is shown here.)

the Bayesian-calculated value $\beta_{95} = 5.895$ corresponds to frequentist *overcoverage* at 95% C.L., or to frequentist *coverage* at 97% C.L.

[2]Pseudoexperiments were thrown with backgrounds uniformly scaled by a factor of $(66.7 - \beta_{95}\mu_s)/(\mu_{t\bar{t}} + \mu_{nt})$ to make the total number of expected events still equal to 66.7 (see analogous discussion in Section 9.3.2). That is, we threw the "frequentist check" pseudoexperiments using signal mean $\beta_{95} \times 4.29 = 25.3$, $t\bar{t}$ mean 5.56, and non-top mean 35.8.

Thus we obtain an upper limit at 95% confidence level on the single-top-production cross section of

$$\sigma_{95} = \beta_{95} \times \sigma_{SM} = 5.895 \times 2.43 \text{ pb} = 14.3 \text{ pb}.$$

Chapter 11

Conclusions

We search for single top production in 106.0 pb^{-1} of CDF Run 1 data by performing an unbinned maximum likelihood fit to the H_T distribution in the tagged $W + 1, 2, 3$-jet sample after a cut on reconstructed mass of the lepton, neutrino, and leading b jet of $140 < M_{\ell\nu b} < 210$ GeV/c^2. From acceptance studies normalized to theory predictions, we predict 4.29 signal to pass the analysis cuts. The fit prefers $n_{signal} = 8.7 \pm 7.8$, or in terms of parameter β expressing the deviation from the Standard-Model prediction, $\beta = 2.04 \pm 1.81$. However, with a background prediction of 62.4 events—for a $S/\sqrt{B} = 0.54$—we knew from the outset that the goal of this analysis would be not to measure but to put an upper limit on the single-top content of CDF Run 1 data. To do this we form the reduced likelihood function $\mathcal{L}(\beta)$, convolve with a Gaussian incorporating systematic uncertainties, and integrate the resulting smeared distribution up to β_{95}, the Bayesian definition of upper limit at 95% confidence level. We check that the result we achieve, $\beta_{95} = 5.895$, is conservative according to the frequentist definition of upper limit at 95% confidence level. This corresponds to an upper limit on the combined $Wg + W^*$ single-top cross section of 14.3 pb at 95% C.L.

To translate our measured value of and upper limit on β into a measured value of and upper limit on the Standard-Model parameter V_{tb}, we recall that the Standard-Model-calculated cross section contained an implicit factor of $|V_{tb}|^2 = 1$. That is, $\sigma_{SM} = |V_{tb,\ SM}|^2 X$. If we absorb into $V_{tb,\ fit}$ all deviation of the fitted cross section σ_{fit} from the

predicted value σ_{SM}, we can write $\sigma_{fit} = |V_{tb,\ fit}|^2 X$, or

$$\frac{\sigma_{fit}}{\sigma_{SM}} = \frac{|V_{tb,\ fit}|^2}{1} = \beta_{fit}$$

i.e. $|V_{tb,\ fit}| = \sqrt{\beta_{fit}} = 1.43$. The uncertainty on this measurement is one half the quadrature sum of the experimental and theoretical uncertainties on the cross section, *i.e.* $\delta|V_{tb}|/|V_{tb}| = \frac{1}{2}\sqrt{\delta\sigma^2_{stat.} + \delta\sigma^2_{sys.} + \delta\sigma_{theor.}}$, where

- statistical uncertainty $\delta\sigma_{stat.} =$
  $1.81/2.04 = 89\%$;
- systematic uncertainty $\delta\sigma_{sys.} =$
  $\sqrt{[\Delta_{SHAPE}(2.04)]^2 + (2.04)^2[\Delta_{NORM}]^2}/2.04 = \sqrt{(0.47)^2 + (2.04)^2(0.16)^2}/2.04 = 28\%$;
- theoretical uncertainty $\delta\sigma_{theor.} =$
  14%;

or 47%. That is, we measure $|V_{tb}| = 1.4 \pm 0.7$. The 95%-C.L. upper limit on $|V_{tb}|$ is $|V_{tb,\ 95}| = \sqrt{\beta_{95}} = 2.43$. These results—a limit on the combined $Wg + W^*$ cross section of six times the Standard-Model prediction, or the result $|V_{tb}| < 2.4$—do not strongly constrain any models of new physics in the top sector.

The analysis strategy laid out in this paper is tailored to the small statistics of a Run 1 search. With the 20× larger sample of CDF Run 2, the strategy will shift toward more-stringent cuts and separate searches for the Wg and W^* modes. A lower bound on the performance of a combined search in Run 2 data can be estimated from Table 8.1. Using the same cuts as this analysis but restricting ourselves to the $W + 2, 3$-jet bin, with 20× the data we would already achieve $S/\sqrt{B} = 2.9$, the benchmark of evidence for single-top production.

Appendices

Appendix A

The CDF Collaboration

T. Affolder,[21] H. Akimoto,[42] A. Akopian,[35] M. G. Albrow,[10] P. Amaral,[7] S. R. Amendolia,[31] D. Amidei,[24] J. Antos,[1] G. Apollinari,[35] T. Arisawa,[42] T. Asakawa,[40] W. Ashmanskas,[7] M. Atac,[10] P. Azzi-Bacchetta,[29] N. Bacchetta,[29] M. W. Bailey,[26] S. Bailey,[14] P. de Barbaro,[34] A. Barbaro-Galtieri,[21] V. E. Barnes,[33] B. A. Barnett,[17] M. Barone,[12] G. Bauer,[22] F. Bedeschi,[31] S. Belforte,[39] G. Bellettini,[31] J. Bellinger,[43] D. Benjamin,[9] J. Bensinger,[4] A. Beretvas,[10] J. P. Berge,[10] J. Berryhill,[7] S. Bertolucci,[12] B. Bevensee,[30] A. Bhatti,[35] C. Bigongiari,[31] M. Binkley,[10] D. Bisello,[29] R. E. Blair,[2] C. Blocker,[4] K. Bloom,[24] B. Blumenfeld,[17] B. S. Blusk,[34] A. Bocci,[31] A. Bodek,[34] W. Bokhari,[30] G. Bolla,[33] Y. Bonushkin,[5] D. Bortoletto,[33] J. Boudreau,[32] A. Brandl,[26] S. van den Brink,[17] C. Bromberg,[25] N. Bruner,[26] E. Buckley-Geer,[10] J. Budagov,[8] H. S. Budd,[34] K. Burkett,[14] G. Busetto,[29] A. Byon-Wagner,[10] K. L. Byrum,[2] M. Campbell,[24] A. Caner,[31] W. Carithers,[21] J. Carlson,[24] D. Carlsmith,[43] J. Cassada,[34] A. Castro,[29] D. Cauz,[39] A. Cerri,[31] P. S. Chang,[1] P. T. Chang,[1] J. Chapman,[24] C. Chen,[30] Y. C. Chen,[1] M. -T. Cheng,[1] M. Chertok,[37] G. Chiarelli,[31] I. Chirikov-Zorin,[8] G. Chlachidze,[8] F. Chlebana,[10] L. Christofek,[16] M. L. Chu,[1] S. Cihangir,[10] C. I. Ciobanu,[27] A. G. Clark,[13] M. Cobal,[31] E. Cocca,[31] A. Connolly,[21] J. Conway,[36] J. Cooper,[10] M. Cordelli,[12] J. Guimaraes da Costa,[24] D. Costanzo,[31] J. Cranshaw,[38] D. Cronin-Hennessy,[9] R. Cropp,[23] R. Culbertson,[7] D. Dagenhart,[41] F. DeJongh,[10] S. Dell'Agnello,[12] M. Dell'Orso,[31]

R. Demina,[10] L. Demortier,[35] M. Deninno,[3] P. F. Derwent,[10] T. Devlin,[36] J. R. Dittmann,[10] S. Donati,[31] J. Done,[37] T. Dorigo,[14] N. Eddy,[16] K. Einsweiler,[21] J. E. Elias,[10] E. Engels, Jr.,[32] W. Erdmann,[10] D. Errede,[16] S. Errede,[16] Q. Fan,[34] R. G. Feild,[44] C. Ferretti,[31] I. Fiori,[3] B. Flaugher,[10] G. W. Foster,[10] M. Franklin,[14] J. Freeman,[10] J. Friedman,[22] Y. Fukui,[20] S. Gadomski,[23] S. Galeotti,[31] M. Gallinaro,[35] T. Gao,[30] M. Garcia-Sciveres,[21] A. F. Garfinkel,[33] P. Gatti,[29] C. Gay,[44] S. Geer,[10] D. W. Gerdes,[24] P. Giannetti,[31] P. Giromini,[12] V. Glagolev,[8] M. Gold,[26] J. Goldstein,[10] A. Gordon,[14] A. T. Goshaw,[9] Y. Gotra,[32] K. Goulianos,[35] H. Grassmann,[39] C. Green,[33] L. Groer,[36] C. Grosso-Pilcher,[7] M. Guenther,[33] G. Guillian,[24] R. S. Guo,[1] C. Haber,[21] E. Hafen,[22] S. R. Hahn,[10] C. Hall,[14] T. Handa,[15] R. Handler,[43] W. Hao,[38] F. Happacher,[12] K. Hara,[40] A. D. Hardman,[33] R. M. Harris,[10] F. Hartmann,[18] K. Hatakeyama,[35] J. Hauser,[5] J. Heinrich,[30] A. Heiss,[18] B. Hinrichsen,[23] K. D. Hoffman,[33] C. Holck,[30] R. Hollebeek,[30] L. Holloway,[16] R. Hughes,[27] J. Huston,[25] J. Huth,[14] H. Ikeda,[40] M. Incagli,[31] J. Incandela,[10] G. Introzzi,[31] J. Iwai,[42] Y. Iwata,[15] E. James,[24] H. Jensen,[10] M. Jones,[30] U. Joshi,[10] H. Kambara,[13] T. Kamon,[37] T. Kaneko,[40] K. Karr,[41] H. Kasha,[44] Y. Kato,[28] T. A. Keaffaber,[33] K. Kelley,[22] M. Kelly,[24] R. D. Kennedy,[10] R. Kephart,[10] D. Khazins,[9] T. Kikuchi,[40] M. Kirk,[4] B. J. Kim,[19] H. S. Kim,[23] S. H. Kim,[40] Y. K. Kim,[21] L. Kirsch,[4] S. Klimenko,[11] D. Knoblauch,[18] P. Koehn,[27] A. Köngeter,[18] K. Kondo,[42] J. Konigsberg,[11] K. Kordas,[23] A. Korytov,[11] E. Kovacs,[2] J. Kroll,[30] M. Kruse,[34] S. E. Kuhlmann,[2] K. Kurino,[15] T. Kuwabara,[40] A. T. Laasanen,[33] N. Lai,[7] S. Lami,[35] S. Lammel,[10] J. I. Lamoureux,[4] M. Lancaster,[21] G. Latino,[31] T. LeCompte,[2] A. M. Lee IV,[9] S. Leone,[31] J. D. Lewis,[10] M. Lindgren,[5] T. M. Liss,[16] J. B. Liu,[34] Y. C. Liu,[1] N. Lockyer,[30] M. Loreti,[29] D. Lucchesi,[29] P. Lukens,[10] S. Lusin,[43] J. Lys,[21] R. Madrak,[14] K. Maeshima,[10] P. Maksimovic,[14] L. Malferrari,[3] M. Mangano,[31] M. Mariotti,[29] G. Martignon,[29] A. Martin,[44] J. A. J. Matthews,[26] P. Mazzanti,[3] K. S. McFarland,[34] P. McIntyre,[37] E. McKigney,[30] M. Menguzzato,[29] A. Menzione,[31] E. Meschi,[31] C. Mesropian,[35] C. Miao,[24] T. Miao,[10] R. Miller,[25] J. S. Miller,[24] H. Minato,[40] S. Miscetti,[12] M. Mishina,[20] N. Moggi,[31] E. Moore,[26] R. Moore,[24] Y. Morita,[20] A. Mukherjee,[10] T. Muller,[18]

A. Munar,[31] P. Murat,[31] S. Murgia,[25] M. Musy,[39] J. Nachtman,[5] S. Nahn,[44] H. Nakada,[40] T. Nakaya,[7] I. Nakano,[15] C. Nelson,[10] D. Neuberger,[18] C. Newman-Holmes,[10] C.-Y. P. Ngan,[22] P. Nicolaidi,[39] H. Niu,[4] L. Nodulman,[2] A. Nomerotski,[11] S. H. Oh,[9] T. Ohmoto,[15] T. Ohsugi,[15] R. Oishi,[40] T. Okusawa,[28] J. Olsen,[43] C. Pagliarone,[31] F. Palmonari,[31] R. Paoletti,[31] V. Papadimitriou,[38] S. P. Pappas,[44] A. Parri,[12] D. Partos,[4] J. Patrick,[10] G. Pauletta,[39] M. Paulini,[21] A. Perazzo,[31] L. Pescara,[29] T. J. Phillips,[9] G. Piacentino,[31] K. T. Pitts,[10] R. Plunkett,[10] A. Pompos,[33] L. Pondrom,[43] G. Pope,[32] F. Prokoshin,[8] J. Proudfoot,[2] F. Ptohos,[12] G. Punzi,[31] K. Ragan,[23] D. Reher,[21] A. Ribon,[29] F. Rimondi,[3] L. Ristori,[31] W. J. Robertson,[9] A. Robinson,[23] T. Rodrigo,[6] S. Rolli,[41] L. Rosenson,[22] R. Roser,[10] R. Rossin,[29] W. K. Sakumoto,[34] D. Saltzberg,[5] A. Sansoni,[12] L. Santi,[39] H. Sato,[40] P. Savard,[23] P. Schlabach,[10] E. E. Schmidt,[10] M. P. Schmidt,[44] M. Schmitt,[14] L. Scodellaro,[29] A. Scott,[5] A. Scribano,[31] S. Segler,[10] S. Seidel,[26] Y. Seiya,[40] A. Semenov,[8] F. Semeria,[3] T. Shah,[22] M. D. Shapiro,[21] P. F. Shepard,[32] T. Shibayama,[40] M. Shimojima,[40] M. Shochet,[7] J. Siegrist,[21] G. Signorelli,[31] A. Sill,[38] P. Sinervo,[23] P. Singh,[16] A. J. Slaughter,[44] K. Sliwa,[41] C. Smith,[17] F. D. Snider,[10] A. Solodsky,[35] J. Spalding,[10] T. Speer,[13] P. Sphicas,[22] F. Spinella,[31] M. Spiropulu,[14] L. Spiegel,[10] L. Stanco,[29] J. Steele,[43] A. Stefanini,[31] J. Strologas,[16] F. Strumia, [13] D. Stuart,[10] K. Sumorok,[22] T. Suzuki,[40] R. Takashima,[15] K. Takikawa,[40] M. Tanaka,[40] T. Takano,[28] B. Tannenbaum,[5] W. Taylor,[23] M. Tecchio,[24] P. K. Teng,[1] K. Terashi,[40] S. Tether,[22] D. Theriot,[10] R. Thurman-Keup,[2] P. Tipton,[34] S. Tkaczyk,[10] K. Tollefson,[34] A. Tollestrup,[10] H. Toyoda,[28] W. Trischuk,[23] J. F. de Troconiz,[14] S. Truitt,[24] J. Tseng,[22] N. Turini,[31] F. Ukegawa,[40] J. Valls,[36] S. Vejcik III,[10] G. Velev,[31] R. Vidal,[10] R. Vilar,[6] I. Vologouev,[21] D. Vucinic,[22] R. G. Wagner,[2] R. L. Wagner,[10] J. Wahl,[7] N. B. Wallace,[36] A. M. Walsh,[36] C. Wang,[9] C. H. Wang,[1] M. J. Wang,[1] T. Watanabe,[40] T. Watts,[36] R. Webb,[37] H. Wenzel,[18] W. C. Wester III,[10] A. B. Wicklund,[2] E. Wicklund,[10] H. H. Williams,[30] P. Wilson,[10] B. L. Winer,[27] D. Winn,[24] S. Wolbers,[10] D. Wolinski,[24] J. Wolinski,[25] S. Worm,[26] X. Wu,[13] J. Wyss,[31] A. Yagil,[10] W. Yao,[21] G. P. Yeh,[10] P. Yeh,[1] J. Yoh,[10] C. Yosef,[25] T. Yoshida,[28] I. Yu,[19] S. Yu,[30] A. Zanetti,[39] F. Zetti,[21] and S. Zucchelli[3]

(CDF Collaboration)

[1] *Institute of Physics, Academia Sinica, Taipei, Taiwan 11529, Republic of China*

[2] *Argonne National Laboratory, Argonne, Illinois 60439*

[3] *Istituto Nazionale di Fisica Nucleare, University of Bologna, I-40127 Bologna, Italy*

[4] *Brandeis University, Waltham, Massachusetts 02254*

[5] *University of California at Los Angeles, Los Angeles, California 90024*

[6] *Instituto de Fisica de Cantabria, University of Cantabria, 39005 Santander, Spain*

[7] *Enrico Fermi Institute, University of Chicago, Chicago, Illinois 60637*

[8] *Joint Institute for Nuclear Research, RU-141980 Dubna, Russia*

[9] *Duke University, Durham, North Carolina 27708*

[10] *Fermi National Accelerator Laboratory, Batavia, Illinois 60510*

[11] *University of Florida, Gainesville, Florida 32611*

[12] *Laboratori Nazionali di Frascati, Istituto Nazionale di Fisica Nucleare, I-00044 Frascati, Italy*

[13] *University of Geneva, CH-1211 Geneva 4, Switzerland*

[14] *Harvard University, Cambridge, Massachusetts 02138*

[15] *Hiroshima University, Higashi-Hiroshima 724, Japan*

[16] *University of Illinois, Urbana, Illinois 61801*

[17] *The Johns Hopkins University, Baltimore, Maryland 21218*

[18] *Institut für Experimentelle Kernphysik, Universität Karlsruhe, 76128 Karlsruhe, Germany*

[19] *Korean Hadron Collider Laboratory: Kyungpook National University, Taegu 702-701; Seoul National University, Seoul 151-742; and SungKyunKwan University, Suwon 440-746; Korea*

[20] *High Energy Accelerator Research Organization (KEK), Tsukuba, Ibaraki 305, Japan*

[21] *Ernest Orlando Lawrence Berkeley National Laboratory, Berkeley, California 94720*

[22] *Massachusetts Institute of Technology, Cambridge, Massachusetts 02139*

[23] *Institute of Particle Physics: McGill University, Montreal H3A 2T8; and University of Toronto, Toronto M5S 1A7; Canada*

[24] *University of Michigan, Ann Arbor, Michigan 48109*

[25] *Michigan State University, East Lansing, Michigan 48824*

[26] *University of New Mexico, Albuquerque, New Mexico 87131*

[27] *The Ohio State University, Columbus, Ohio 43210*

[28] *Osaka City University, Osaka 588, Japan*

[29] *Universita di Padova, Istituto Nazionale di Fisica Nucleare, Sezione di Padova, I-35131 Padova, Italy*

[30] *University of Pennsylvania, Philadelphia, Pennsylvania 19104*

[31] *Istituto Nazionale di Fisica Nucleare, University and Scuola Normale Superiore of Pisa, I-56100 Pisa, Italy*

[32] *University of Pittsburgh, Pittsburgh, Pennsylvania 15260*

[33] *Purdue University, West Lafayette, Indiana 47907*

[34] *University of Rochester, Rochester, New York 14627*

[35] *Rockefeller University, New York, New York 10021*

[36] *Rutgers University, Piscataway, New Jersey 08855*

[37] *Texas A&M University, College Station, Texas 77843*

[38] *Texas Tech University, Lubbock, Texas 79409*

[39] *Istituto Nazionale di Fisica Nucleare, University of Trieste/ Udine, Italy*

[40] *University of Tsukuba, Tsukuba, Ibaraki 305, Japan*

[41] *Tufts University, Medford, Massachusetts 02155*

[42] *Waseda University, Tokyo 169, Japan*

[43] *University of Wisconsin, Madison, Wisconsin 53706*

[44] *Yale University, New Haven, Connecticut 06520*

Bibliography

[1] F. Abe et al., The CDF Collaboration, *Observation of top quark production in $p\bar{p}$ collisions with the Collider Detector at Fermilab*, Phys. Rev. Lett. **74**, 2626 (1995); S. Abachi et al., The D0 Collaboration, *Observation of the top quark*, Phys. Rev. Lett. **74**, 2632 (1995).

[2] *Summary Tables of Particle Properties*, extracted from the Particle Listings of the *Review of Particle Physics*, D. Groom *et al.*, Eur. Phys. Jour. **C15**, 1 (2000), available at `http://pdg.lbl.gov/`.

[3] G. Kane, *Modern Elementary Particle Physics*, Addison-Wesley, New York (1993). Chapter 2 of this paper is deeply indebted to Gordy Kane's book. Unless otherwise footnoted, material in Section 2.2.1 is based on Kane ch. 3 and 4; Section 2.2.2 on Kane ch. 6; Section 2.2.3 on Kane ch. 7; Section 2.2.4 on Kane ch. 22; Section 2.4 on Kane ch. 7 and 8. Section 2.3 is based loosely on material in Kane ch. 1, p. 5; ch. 2, p. 33; and ch. 9, p. 117.

[4] J. Donoghue, E. Golowich, and B. Holstein, *Dynamics of the Standard Model*, Cambridge UP, New York (1992).

[5] B. Kayser, *Neutrino Mass*, in D. Groom *et al.*, Eur. Phys. Jour. **C15**, 1 (2000), available at `http://pdg.lbl.gov/`.

[6] D. Carlson, *Physics of Single-Top Quark Production at Hadron Colliders*, Ph.D. thesis, Michigan State University, 1995 (`hep-ph/9508278`).

[7] F. Gilman, K. Kleinknecht, and B. Renk, *The Cabibbo-Kobayashi-Maskawa Quark-Mixing Matrix*, in D. Groom *et al.*, Eur. Phys. Jour. **C15**, 1 (2000), available at `http://pdg.lbl.gov/`.

[8] M. Veltman, *Diagrammatica*, Cambridge UP, New York (1994).

[9] D. Griffiths, *Introduction to Elementary Particles*, Wiley, New York (1987).

[10] J. Erler and P. Langacker, *Electroweak Model and Constraints on New Physics*, in in D. Groom *et al.*, Eur. Phys. Jour. **C15**, 1 (2000), available at `http://pdg.lbl.gov/`.

[11] P. Schewe, J. Riordon, and B. Stein, *The Muon's Magnetic Moment is Misbehaving*, Physics News Update **524** No. 4, (2001).

[12] S. Willenbrock and D. Dicus, *Production of Heavy Quarks from W-Gluon Fusion*, Phys. Rev. D **34**, 155 (1986).

[13] C.–P. Yuan, *New Method to Detect a Heavy Top Quark at the Fermilab Tevatron*, Phys. Rev. D **41**, 42 (1990).

[14] R. Ellis and S. Parke, *Top-Quark Production by W-Gluon Fusion*, Phys. Rev. D **46**, 3785 (1992).

[15] S. Cortese and R. Petronzio, *The Single Top Production Channel at Tevatron Energies*, Phys. Lett. B **253**, 494 (1991).

[16] T. Stelzer and S. Willenbrock, *Single-Top-Quark Production via* $q\bar{q} \to t\bar{b}$, Phys. Lett. B **357**, 125 (1995).

[17] A. Heinson, A. Belyaev, and E. Boos, *Single Top Quarks at the Fermilab Tevatron*, Phys. Rev. D **56**, 3114 (1997).

[18] S. Moretti, *Single-Top-Quark Production in the* $tW^{\pm}$ *Channel and Higgs Boson Signals via* $H \to W^+W^-$ *at the CERN Large Hadron Collider*, Phys. Rev. D **56**, 7427 (1997).

[19] T. Tait and C.–P. Yuan, *The Phenomenology of Single Top Quark Production at the Fermilab Tevatron*, hep-ph/9710372 (1997).

[20] T. Stelzer, Z. Sullivan, and S. Willenbrock, *Single-Top-Quark Production at Hadron Colliders*, Phys. Rev. D **58**, 094021 (1998).

[21] S. Dawson and S. Willenbrock, *Heavy Fermion Production in the Effective W Approximation*, Nucl. Phys. B **284**, 449 (1987).

[22] R. Bonciani *et al.*, *NLL Resummation of the Heavy-Quark Hadroproduction Cross Section*, hep-ph/9801375 (1998).

[23] M. Smith and S. Willenbrock, *QCD and Yukawa Corrections to Single-Top-Quark Production via* $q\bar{q} \to t\bar{b}$, Phys. Rev. D **54**, 6696 (1996).

[24] A. Heinson, *Measuring the CKM Matrix Element* V_{tb} *at D0 and CDF*, invited talk at the *Second International Conference on B Physics and CP Violation* (Honolulu, March 24-27, 1997) and hep-ex/9707026.

[25] T. LeCompte and R. Roser, *Measurement of* $BF(t \to Wb)$ *and the CKM Matrix Element* $|V_{tb}|$ *in Top Decays*, CDF Note 3056, 1995 (unpublished).

[26] D. Amidei and R. Brock, *Report of the* tev_{2000} *Study Group*, FERMILAB-Pub-96/082 (1996).

[27] Ref. [24] cites the following reference for CDF limits on $|V_{tb}|$:
G. Chiarelli, for the CDF Collaboration, to appear in the proceedings of the XXXII Rencontres de Moriond, QCD and Hadronic Interactions, Les Arcs, Savoie, France, March 1997, Fermilab-Conf-97/143-E.

[28] I. Bigi *et al.*, *Production and Decay Properties of Ultra-Heavy Quarks*, Phys. Lett. B **181**, 157 (1986).

[29] S. Willenbrock, *Why Single Top Measures V_{tb} and Not $\Gamma(t \to Wb)$*, informal document, 2001 (unpublished).

[30] S. Dawson, *The Effective W Approximation*, Nucl. Phys. B **249**, 42 (1985).

[31] D. Carlson and C.–P. Yuan, *Studying the Top Quark via the W-Gluon Fusion Process*, Phys. Lett. B **306**, 386 (1993).

[32] G. Mahlon and S. Parke, *Improved Spin Basis for Angular Correlation Studies in Single Top Quark Production at the Tevatron*, Phys. Rev. D **55**, 7249 (1997). (hep-ph/9611367)

[33] M. Jeżabek, Nucl. Phys. B (Proc. Suppl.) **37**, 197 (1994).

[34] R. Peccei, S. Peris, and X. Zhang, *Non-Standard Couplings of the Top Quark and Precision Measurements of the Electroweak Theory*, Nucl. Phys. B **349**, 305 (1991).

[35] R. Chivukula and J. Womersley, *Dynamical Electroweak Symmetry Breaking*, in C. Caso *et al.*, The European Physical Journal **C3**, 1 (1998) and 1999 off-year partial update for the 2000 edition (URL: `http://pdg.lbl.gov/`)

[36] F. Larios and C.–P. Yuan, *Top quark Interactions and the Search For New Physics*, Phys. Rev. D **55**, 7218 (1997).

[37] T. Tait and C.–P. Yuan, *Single Top Production as a Window to Physics Beyond the Standard Model*, Phys. Rev. D **63**, 014018 (2001). (hep-ph/0007298).
Sections 3.3.1, 3.3.2, and 3.3.3 are based wholly on this paper and references therein, unless otherwise footnoted.

[38] A. Datta and X. Zhang, *Nonuniversal Correction to $Z \to b\bar{b}$ and Single Top Quark Production at Fermilab Tevatron*, Phys. Rev. D 55, R2530 (1997).

[39] Ref. [35] cites the following papers on top condensates:
V. Miransky, M. Tanabashi, and K. Yamawaki, Phys. Lett. **221B**, 177 (1989) and Mod.

Phys. Lett. **4**, 1043 (1989);

Y. Nambu, EFI-89-08 (1989);

W. Marciano, Phys. Rev. Lett. **62**, 2793 (1989); and

R. Chivukula, hep-ph/9803219 for a recent review.

[40] Ref. [35] cites the following paper on the too-large value of m_t produced by top-quark condensation alone:

W. Bardeen, C. Hill, and M. Lindner, Phys. Rev. D **41**, 1647 (1990).

[41] Ref. [35] cites the following paper on topcolor-assisted technicolor:

C. Hill, Phys. Lett. B **345**, 483 (1995) and Phys. Lett. **266B**, 419 (1991).

[42] Ref. [37] cites the following papers on the top-color model:

C. Hill, Phys. Lett. B **345**, 483 (1995);

B. Dobrescu and C. Hill, Phys. Rev. Lett. **81**, 2634 (1998);

R. Chivukula *et al.*, Phys. Rev. D **59**, 075003 (1999);

M. Popovic and E. Simmons, Phys. Rev. D **62**, 035002 (2000);

H.–C. Cheng, B. Dobrescu, and C. Hill, hep-ph/9912343;

H. Collins, A. Grant, and H. Georgi, Phys. Rev. D **61**, 055002 (2000);

H. Georgi and A. Grant, hep-ph/0006050;

A. Aranda and C. Carone, hep-ph/0007020.

[43] Ref. [37] cites the following papers on the top-flavor model. The last reference in the list is cited specifically in conjunction with (1) additional fermions and (2) an extra $SU(2)_h$ gauge symmetry involved in generating the top mass through a seesaw mechanism.

R. Chivukula, E. Simmons, and J. Terning, Phys. Lett. B **436**, 284 (1995);

E. Malkawi, T. Tait, and C.–P. Yuan, Phys. Lett. B **385**, 304 (1996);

D. Muller and S. Nandi, Phys. Lett. B **383**, 345 (1996);

J. Lee, K. Lee, and J. Kim, Phys. Lett. B **424**, 133 (1998);

K. Lynch *et al.*, hep-ph/0007286;

E. Malkawi and C.–P. Yuan, Phys. Rev. D **61**, 015007 (2000);
H.–J. He, T. Tait, and C.–P. Yuan, Phys. Rev. D **62**, 011702 (2000).

[44] With regard to scalar particles that exist as bound states of top and bottom quarks, Ref. [37] cites [42], [43] and the following:
W. Bardeen, C. Hill, and M. Linder, Phys. Rev. D **41**, 1647 (1990).
With regard to composite top-pions that are produced through $c\bar{b}$ fusion, Ref. [37] cites
H.–J. He and C.–P. Yuan, Phys. Rev. Lett. **83**, 28 (1999).

[45] For low-energy $b \to s\gamma$ data that constrains κ^R_{Wtb}, Ref. [37] cites
T. Skwarnicki, talk at ICHEP98, Vancouver, Cananda (1998);
M. Alam *et al.*, CLEO collaboration, Phys. Rev. Lett. **74**, 2885 (1998).

[46] Ref. [37] cites the following paper on constraints on the strength $|\kappa_{Ztc}|$ of the Ztc operator:
T. Han, R. Peccei, and X. Zhang, Nucl. Phys. B **454**, 527 (1995).

[47] Ref. [23] cites the following references on the $\mathcal{O}(\alpha_s^2)$ correction to the Drell-Yan initial state:
R. Hamberg, W. van Neerven, and T. Matsuura, Nucl. Phys. B **359**, 343 (1991);
W. van Neerven and E. Zijlstra, Nucl. Phys. B **382**, 11 (1992).

[48] All information in Sections 3.4.1.1, 3.4.1.2, and 3.4.1.3 is taken from Ref. [23], unless otherwise footnoted.

[49] L. Demortier and the Top Averaging Group, *Combining the Top Quark Mass Results for Run 1 from CDF and D0*, CDF Note 5024 and FERMILAB-TM-2084, 1999 (unpublished).

[50] T. Stelzer, Z. Sullivan, and S. Willenbrock, *Single-Top-Quark Production via W-Gluon Fusion at Next-to-Leading Order*, Phys. Rev. D **56**, 5919 (1997).
All information in Sections 3.4.2, 3.4.2.1, 3.4.2.2, and 3.4.2.3 is taken from this paper unless otherwise footnoted.

[51] Ref. [20] cites the following paper on gluon distribution function uncertainty:
J. Huston *et al.*, hep-ph/9801444.

[52] F. Abe *et al.*, *The CDF Detector: An Overview*, Nucl. Inst. and Methods, **A271**, 387 (1988).

[53] J. Thompson, *Introduction to Colliding Beams at Fermilab*, FERMILAB-TM-1909, 1994 (unpublished). Section 4.1 is based wholly on this document and references therein, unless otherwise footnoted.

[54] Fermilab Click N' Tour the Accelerators, http://www.fnal.gov/pub/tour.html, 1999.

[55] J. D. Jackson, *Classical Electrodynamics*, John Wiley and Sons, New York (1975).

[56] F. Abe *et al.*, *Evidence for Top Quark Production in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV*, Phys. Rev. D **50**, 2966 (1994).

[57] The Particle Data Group, *Review of Particle Properties*, Phys. Rev. D **45**, III.15 (1992).

[58] D. Amidei *et al.*, *The Silicon Vertex Detector of the Collider Detector at Fermilab*, FERMILAB-Pub-94/024-E, 1994 (unpublished).
Section 4.2.1.1 is based heavily on this document and on [59], unless otherwise footnoted.

[59] P. Azzi *et al.*, *SVX', the New CDF Silicon Vertex Detector*, FERMILAB-Conf-94/205-E, 1994 (unpublished).

[60] R. K. Bock and A. Vasilescu, *The Particle Detector BriefBook*, `home.cern.ch/~rkb/titleD.html`, *Semiconductor Detectors* page (1998).

[61] This information was gleaned from [56]. No published documentation or internal notes describing the new VTX, which was installed in 1992, exist.

[62] Obtained from DF (Run 1 event display) pictures of the VTX.

[63] F. Snider *et al.*, *The CDF Vertex Time Projection Chamber System*, Nucl. Inst. and Methods, **A268** 75 (1988).

This article describes the old VTPC, which the VTX replaced. The global form and function of the two devices are similar, but the VTX features many more chambers (28 vs. 8). Details of chamber construction differ significantly between the two.

[64] R. Wagner, private communication.

Aspects of VTX chamber construction which differ from the description in [63] were provided by R. Wagner. In some cases, interpolation between the decriptions provided by [63] and [64] were necessary; any resulting mistakes are solely the responsibility of this author.

[65] D. Amidei, private communication.

Thanks to D. Amidei for pointing out that the drift velocity in the VTX is not constant, but a measured polynomial function of time.

[66] F. Abe *et al.*, *The CDF II Detector Technical Design Report*, FERMILAB-Pub-96/390-E, chapter 3, 1996 (unpublished).

[67] Jackson, pp. 581-582.

[68] F. Bedeschi *et al.*, *Design and Construction of the CDF Central Tracking Chamber*, Nucl. Inst. and Methods, **A268** 50 (1988).

Section 4.2.1.3 is based heavily on this document, unless otherwise footnoted.

[69] F. Abe *et al.*, *Measurement of the W-Boson Mass in* 1.8-*TeV* $p\bar{p}$ *collisions*, Phys. Rev. D **43** 2070 (1991).

[70] R. K. Bock and A. Vasilescu, *The Particle Detector BriefBook*, `home.cern.ch/~rkb/titleD.html`, assorted calorimetry-related pages, (1998).

All information in the general calorimetry discussion in Section 4.2.2 was taken from these pages, unless otherwise footnoted.

[71] M. Campbell, private communication.

[72] The Particle Data Group, *Review of Particle Properties*, Phys. Rev. D **45**, III.15 (1992).

[73] L. Balka *et al.*, *The CDF Central Electromagnetic Calorimeter*, Nucl. Inst. and Methods, **A267** 272 (1988).
Information in section 4.2.2.1 is taken from this article and from [76], unless otherwise footnoted.

[74] F. Abe *et al.*, *The CDF II Detector Technical Design Report*, FERMILAB-Pub-96/390-E, chapter 8, 1996 (unpublished).

[75] R. Wagner *et al.*, *Cosmic Ray Test of the CDF Central Calorimeters*, Nucl. Inst. and Methods, **A267** 330 (1988).

[76] S. Bertolucci *et al.*, *The CDF Central and Endwall Hadron Calorimeter*, Nucl. Inst. and Methods, **A267** 301 (1988).

[77] Y. Fukui *et al.*, *CDF End Plug Electromagnetic Calorimeter Using Conductive Plastic Proportional Tubes*, Nucl. Inst. and Methods, **A267** 280 (1988).
Information in section 4.2.2.2 is taken from this article and from [78].

[78] W. Carithers *et al.*, *Performance and Calibration Studies of the CDF Endplug Hadronic Calorimeter*, CDF Note 368, 1985 (unpublished).

[79] G. Brandenburg *et al.*, *An Electromagnetic Calorimeter for the Small Angle Regions of the Collider Detector at Fermilab*, Nucl. Inst. and Methods, **A267** 257 (1988).
Information in section 4.2.2.3 is taken from this article and from [80].

[80] S. Cihangir *et al.*, *The CDF Forward/Backward Hadron Calorimeter*, Nucl. Inst. and Methods, **A267** 249 (1988).

[81] W. S. C. Williams, *Nuclear and Particle Physics*, Clarendon Press, New York (1991), p. 247.

[82] The Particle Data Group, *Review of Particle Properties*, Phys. Rev. D **45**, III.14 (1992).

[83] F. Abe *et al.*, *The CDF II Detector Technical Design Report*, FERMILAB-Pub-96/390-E, chapter 10, 1996 (unpublished).

[84] J. D. Lewis *et al.*, *The 1992 CDF Muon System Upgrade*, CDF Note 2858, 1995 (unpublished).

Sections 4.2.3.2 and 4.2.3.3 are based on this article.

[85] G. Ascoli *et al.*, *CDF Central Muon Detector*, Nucl. Inst. and Methods, **A268** 33 (1988).

Section 4.2.3.1 is based on this article.

[86] R. K. Bock and A. Vasilescu, *The Particle Detector BriefBook*, `home.cern.ch/~rkb/titleD.html`, *Gaseous Detectors, Operational Modes* page (1998).

[87] F. Abe *et al.*, *Measurement of the antiproton-proton total cross section at* $\sqrt{s} = 546$ *and* 1800 *GeV*, Phys. Rev. D **50** 5550 (1994).

[88] S. Belforte *et al.*, σ_{BBC} *Derivation and Results*, CDF Note 2361, 1994 (unpublished).

[89] M. Campbell, private communication.

[90] D. Amidei *et al.*, *A Two Level FASTBUS Based Trigger System for CDF*, Nucl. Inst. and Methods, **A269** 51 (1988).

Section 4.2.4 uses much information from this article.

[91] S. Catani *et al.*, *The QCD and Standard Model Working Group: Summary Report*, for the workshop "Physics at TeV Colliders", Les Houches, France 8-18 June 1999 (unpublished; available as CDF Note 5302).

[92] E. Barsotti *et al.*, *FASTBUS Data Acquisition for CDF*, Nucl. Inst. and Methods, **A269** 82 (1988).

[93] G. W. Foster *et al.*, *A Fast Hardware Track-Finder for the CDF Central Tracking Chamber*, Nucl. Inst. and Methods, **A269** 93 (1988).

[94] D. Cronin-Hennessy *et al.*, *Luminosity Monitoring and Measurement at CDF*, CDF Note 4956, 1999 (unpublished).

[95] J. D. Lewis and K. T. Pitts, *A Summary of CFT Operation in Run 1*, CDF Note 3999, 1997 (unpublished).

[96] D. Quarrie, *YBOS Users Guide*, CDF Note 156, 1987 (unpublished).

[97] U. Joshi *et al.*, *The Upgraded Level 3 Trigger for the CDF Experiment at the Fermilab Tevatron Collider*, CDF Note 1288, 1990 (unpublished).

[98] V. D. Barger and R. J. N. Phillips, *Collider Physics*, Addison-Wesley, New York (1987), p. 24.

[99] R. K. Bock and A. Vasilescu, *The Particle Detector BriefBook*, `home.cern.ch/~rkb/titleD.html`, *Luminosity* page (1998).

[100] D. Cronin-Hennessy and A. Beretvas, *Luminosity at CDF*, CDF Note 4721, 1998 (unpublished).

This is the primary reference for section 4.2.5 (Luminosity Monitoring). Unless otherwise footnoted, all information found in that section may be assumed to have been taken from here.

[101] S. Belforte *et al.*, σ_{BBC} *Updates*, CDF Note 2535, 1994 (unpublished).

[102] F. Abe *et al.*, *Measurement of* $\sigma B(W \to e\nu)$ *and* $\sigma B(Z^0 \to e^+e^-)$ *in* $p\bar{p}$ *Collisions at* $\sqrt{s} = 1.8$ *TeV*, Phys. Rev. Lett. **76** 3070 (1996).

[103] T. Sjöstrand, *PYTHIA 5.7 and JETSET 7.4 Physics and Manual*, CERN-TH.7112/93, 1993 (revised August 1994), chapter 1. This manual is the long writeup of reference [106].

[104] V. Barger and R. Phillips, *Collider Physics*, Addison-Wesley, New York (1987), chapter 11.

[105] PYTHIA manual, chapter 4.

[106] T. Sjöstrand, Comput. Phys. Commun. **82**, 74 (1994).

[107] PYTHIA manual, chaper 7.

[108] J. Lewis and P. Avery, *CLEOMC: The CDF Interface to the CLEO Monte Carlo (QQ)*, CDF Note 2724, 1997 (unpublished).

[109] S. Jadach *et al.*, CERN-TH-5856 (1990); CERN-TH-6195 (1991); CERN-TH-6793 (1992).

[110] M. Shapiro *et al.*, *A User's Guide to QFL*, CDF Note 1810, 1992 (unpublished).

[111] G. Marchesini and B.R. Webber, Nucl. Phys. B **310**, 461 (1998); G. Marchesini *et al.*, Comput. Phys. Commun. **67**, 465 (1992).

[112] M. Kruse *et al.*, *High p_T Lepton ID Efficiencies for Run 1B*, CDF Note 2883, 1994 (unpublished).

[113] F. Abe *et al.*, *Measurement of $\sigma B(W \to e\nu)$ and $\sigma B(Z^0 \to e^+e^-)$ in $p\bar{p}$ Collisions at $\sqrt{s} = 1800$ GeV*, Phys. Rev. D **44** 29 (1991).
Much of Section 5.3.1.1, including descriptions of electron-identification cut variables, were taken from this article.

[114] F. Bedeschi *et al.*, *A Primary Vertex Finding Package*, CDF Note 1789, 1992 (unpublished).

[115] B. Winer *et al.*, `top$fnd:eleval.cdf`, 1993 (CDF code).

[116] J. Proudfoot *et al.*, `c$ele:elcshr.cdf`, 1991 (CDF code).

[117] J. Guimarães da Costa, Ph.D. thesis, University of Michigan, 2000 (unpublished).

[118] D. Gerdes, *Study of Conversion Removal for the Lepton + Jets Sample*, CDF Note 2903, 1995 (unpublished).

[119] G. Apollinari *et al.*, *Measurement of the $t\bar{t}$ Production Cross Section Using SECVTX Tags*, CDF Note 3855, 1997 (unpublished).

[120] CDF Top Group, *Sample selection for Run 1B top search*, CDF Note 2966, 1996 (unpublished).

[121] B. Winer *et al.*, `c$top:topfnd.cdf`, 1994 (CDF code).

[122] T. Phillips *et al.*, `c$met:metser.cdf`, 1991 (CDF code).

[123] C. Campagnari *et al.*, `c$top:get_corr_met.cdf`, 1994 (CDF code).

[124] F. Abe *et al.*, *Topology of Three-Jet Events in $p\bar{p}$ Collisions at $\sqrt{s} = 1800$ GeV*, Phys. Rev. D **45** 1448 (1992).

[125] T. Affolder *et al.*, *Measurement of the Top Quark Mass Using the Collider Detector at Fermilab*, Phys. Rev. D **63**, 032003 (2001).

[126] W. Yao *et al.*, `c$top:secvtx.cdf`, 1994 (CDF code).

[127] F. Abe *et al.*, *Observation of Top Quark Production in $p\bar{p}$ Collisions with the Collider Detector at Fermilab*, Phys. Rev. Lett. **74** 2626 (1995).

[128] P. Giromini, F. Ptohos, and W. Yao, *Search for New Particles $X \to b\bar{b}$ Produced in Association with W Bosons at $\sqrt{s} = 1.8$ TeV*, CDF Note 4112, 1997 (unpublished).

[129] Particle Data Group, *Particle Physics Booklet*, 1994. (From the *Review of Particle Properties*, Phys. Rev. D **50**, 1173 (1994).)

[130] Cross Section and Branching Ratios Working Group, *$t\bar{t}$ Acceptance Calculation for the Run 1B SVX b-Tag Analysis*, CDF Note 3403, 1995 (unpublished).

[131] T. Liss and R. Roser, *$t\bar{t}$ Production Cross Section for 110 pb^{-1}*, CDF Note 3481, Version 4.0, June 22, 1997 (unpublished).

[132] P. Derwent *et al.*; *Godparents Report on the New $t\bar{t}$ Cross Section Measurement in the Lepton + Jets Channel Using SECVTX Tags*, CDF Note 4939, Version 2, April 20, 1999 (unpublished).

[133] R. Hughes and T. Liss, *$t\bar{t}$ Cross Section for the SVX b-tag Analysis*, CDF Note 2992, Version 2.0, 1995 (unpublished).

[134] T. Liss, *A Fast Muon Trigger Simulation*, CDF Note 2998, 1995 (unpublished).

[135] T. Affolder *et al.*, *Measurement of the $t\bar{t}$ Production Cross Section in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV*, CDF Note 5375 and FERMILAB-PUB-01/007-E, submitted to Phys. Rev. D, 2001.

[136] F. Abe *et al.*, *Measurement of the $t\bar{t}$ Production Cross Section in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV*, CDF Note 4030, 1997 (unpublished).
This note cites $\sigma_{t\bar{t}} = 7.6^{+1.8}_{-1.5}$ pb, i.e. is not current. However, the $t\bar{t}$ acceptance uncertainty is unlikely to have changed much in the current measurement from the 13.5% quoted here—surely not to have *increased* so much that in quadrature with the 4% luminosity uncertainty it would exceed 20%, the figure we quote for $\delta(A_{t\bar{t}} \cdot \int \mathcal{L}dt)$.

[137] F.A. Berends, W.T. Giele, H. Kuijf and B. Tausk, Nucl. Phys. B **357**, 32 (1991).

[138] G. Unal, *$Wb\bar{b}$ and $Wc\bar{c}$ Backgrounds in Top SVX Channel*, CDF Note 3389, 1995 (unpublished).

[139] G. Unal, *Background Summary for the Top SVX Analysis*, CDF Note 3513, 1995 (unpublished).

[140] C. Campagnari, *Background in the W+Jets Sample*, CDF Note 2476, 1994 (unpublished).

[141] Y. Cen *et al.*, *Backgrounds for the SVX b-Tag Top Search: Run 1A+1B*, CDF Note 2965, 1995 (unpublished).

[142] G. Apollinari *et al.*, *Method II Measurement of the Top Production Cross Section Using W+3, 4 Jet Events with SECVTX, JPB, and SLT Tags*, CDF Note 4303, 1997 (unpublished).

[143] L. Demortier *et al.*, *Upper Limits on the Cross Sections for Single Top Quark Production Via the s-Channel W* and t-Channel W-Gluon Fusion Processes*, CDF Note 5104, 2001 (unpublished).

[144] B. Roe, "Discussion with Woodroofe", 1999 (unpublished).
We owe our derivation of the likelihood function to Roe and Woodroofe.

[145] F. James, *MINUIT: Function Minimization and Error Analysis Reference Manual, version 94.1*, CERN Program Library Long Writeup D506, Geneva (1994).

[146] The Particle Data Group, *Review of Particle Physics*, Phys. Rev. D **54**, p. 162 (1996).
Our discussion of the frequentist definition of upper limit borrows much from this treatment and from that of Roe, below.

[147] B. Roe, *Probability and Statistics in Experimental Physics, 2nd ed.*, Springer-Verlag, New York (2001).

[148] The Particle Data Group, *Review of Particle Physics*, Phys. Rev. D **54**, p. 165 (1996).
This discussion of the Bayesian method of handling a measurement near an unphysical boundary provides a concise introduction to the Bayesian method of constructing confidence intervals. We borrow liberally from this treatment.

[149] R. Cousins, *Why Isn't Every Physicist a Bayesian?*, Am. J. Phys. **63**:398, (1995); G. Feldman and R. Cousins, *Unified Approach to the Classical Statistical Analysis of Small Signals*, Phys. Rev. D **57**, 3873 (1998).
Cousins clarifies the identification between $P(\text{data}|\beta)$ and the likelihood function.

[150] H. Kim, likelihood-integrating software SUM95PCT (1997).

[151] The Particle Data Group, *Review of Particle Physics*, Phys. Rev. D **54**, p. 166 (1996). This prescription for a Poisson process with background references O. Helene, Nucl. Instr. and Meth. **212**, 319 (1983), a thoroughly Bayesian article. Nevertheless, it is difficult to see how Bayes's Theorem is used in the formula for N_{95}.

[152] F. Abe *et al.*, *Top-quark search in the electron+jets channel in proton-antiproton collisions at* $\sqrt{s} = 1.8$ *TeV*, Phys. Rev. D **43**, 664 (1991).

[153] L. Galtieri *et al.*, `top$fnd:et_corr_sys.cdf`, 1997 (CDF code); L. Galtieri and J. Lys, *How well do we understand jets in Run 1? Study of the jet energy scale for raw jet* $E_T \geq 8$ *GeV*, CDF Note 3983, 1997 (unpublished).

ET_CORR_SYS computes the change in raw and corrected jet E_T due to 1σ variation of several sources of systematic error in jet energies, namely: (1) calorimeter stability; (2) underlying event due to multivertices; (3) relative energy scale; (4) absolute energy scale; (5) soft gluon radiation; (6) out-of-cone losses.

[154] E. Buckley-Geer, *How to find out what data is available*, unpublished text file `how_to_find_data.txt`, (1993); E. Buckley-Geer and S. Lammel, *How to find out what Run 1B data is available*, CDF Note 2657, 1994 (unpublished).

Among other things, these documents describe the breakdown of data into streams for Run 1A and 1B, respectively.

[155] T. Handa et al., *Search for technicolor particles in* $W + 2$ *jet with b-tag channel*, CDF Note 4372, 1997 (unpublished).

Table 2 of this note lists the inclusive 1A and 1B lepton samples the authors used to select events. The names are close enough to those of the samples used in this analysis that we believe they refer to the same samples–hence we list the same Level 3 triggers listed in that table.

[156] M. Binkley, `top$fnd:zlechg.cdf`, date unknown (CDF code).