Search for the Flavor Changing Neutral Current Decay $t \to qZ$ in $p\overline{p}$ Collisions at $\sqrt{s} = 1.96$ TeV

A dissertation presented

by

Ingyin Zaw

to

The Department of Physics

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Physics

Harvard University

Cambridge, Massachusetts

September 2007

Thesis advisor

Melissa E. B. Franklin

Author

Ingyin Zaw

Search for the Flavor Changing Neutral Current Decay
$t \to qZ$ in $p\overline{p}$ Collisions at $\sqrt{s} = 1.96$ TeV

Abstract

This dissertation presents a blind search for the flavor changing neutral current decay of the top quark $t \to qZ$ in $p\overline{p}$ collisions at $\sqrt{s} = 1.96 TeV$ using a data sample corresponding to an integrated luminosity of 1.12 fb^{-1} collected by the Collider Detector at Fermilab (CDF). This decay is extremely rare in the standard model, and a signal at the Tevatron would be an indication of new physics. Dividing candidate events with a Z boson and four or more jets into a sample of those which have a heavy flavor jet identified by a secondary vertex algorithm and those which do not, we observe data yields consistent with background expectations. We set a 95% C.L. upper limit on the branching fraction $\mathcal{B}$ $(t \to qZ)$ of 11.3%, consistent with an expected upper limit of 7.8% $\pm$ 3.3%.

Contents

List of Figures

List of Tables

For Yosi

Roads go ever ever on,
Over rock and under tree,
By caves where never sun has shown,
By streams that never find the sea;
Over snow by winter sown,
And through the merry flowers of June,
Over grass and over stone,
And under mountains in the moon.
Roads go ever ever on
Under cloud and under star,
Yet feet that wandering have gone
Turn at last to home afar.
Eyes that fire and sword have seen
And horror in the halls of stone
Look at last on meadows green
And trees and hills they long have known.
J. R. R. Tolkien

Acknowledgments

This thesis would not have happened without the help and support of many people. My advisor, Melissa Franklin, deserves much thanks for her advice and support over the years. I truly appreciate her unfailing physics intuition and her genuine concern for my academic and personal happiness. I have also learned a great deal from and enjoyed working with the other members of the Harvard CDF group. I owe a debt of gratitude to Sebastian Grinstein for my understanding of ALPGEN and heavy flavor fractions. I worked on Si dE/dx with Andy Foland and Josh Rosaler. I thank João Guimarães da Costa for his advice and many last minute recommendation letters. I am very grateful to Ayana Holloway for her wealth of knowledge and for being a great friend. I will always miss the "sunshine cookies" we shared. It was also fun to work with Sal Rappoccio, Daniel Sherman, and John Paul Chou. I would also like to thank all the people at the Laboratory for Particle Physics, in particular, George Brandenburg and Robyn Simpson. In the physics department, Sheila Ferguson deserves much gratitude for all her support through the years. My friends Emily Dunkel, Kathy Aidala, Ari Turner, and Ilya Finkler have broadened my horizons beyond high energy physics and given me a reason to walk over to Jefferson.

I am indebted to many people at Fermilab. I would like to acknowledge my collaborators on this analysis; Ulrich Husemann, who is always very knowledgeble and helpful, Jennifer Gimmell, who is fun with, Charles Plager, who knows his limits, David Saltzberg and Paul Tipton. Working on the CMX detector has been one of the most enjoyable parts of my time as a graduate student, mainly in thanks to Phil Schlabach and Dale Stentz. In addition to being great to work with, they have also become very good friends. Dee and Steve Hahn are truly wonderful. They have housed me, fed me, and welcomed me to all their parties. I am grateful to Rick Tesarek for his advice, support, and friendship. Müge Karagöz *Ünel*, Marcel Stanitzki, Eli Ben Haim, John Strologas, and Harry Cheung are all good friends who have enriched my time at Fermilab. I am grateful to Norman Gelfand for the fun lunches and for explaining how the accelerator works.

Last but not least, I would like to thank my friends and family who have given me much love and support over the years. Sonia, my best friend, has shared much laughter and many tears with me and has supported me on more occasions than I can count. I thank Kartini for motivating me to go to the gym and our fun Friday evenings. I appreciate Gitika for all the stories I can tell about the crazy things we've done, voluntarily and involuntarily. Elio, Bob, Bob, Denis, and Jack, thanks for being great. I thank my parents, grandmother, and Daw Ohn for who I am. Their wisdom, support, and love are always with me no matter how far away I am geographically. I thank my brother for being there when I need him and my nephew, Christopher, who never fails to make me smile. I owe more than I can describe to Yosi, my husband for his unconditional love, unwavering support, and endless patience. I cannot imagine the last six years or, indeed, my future, without him.

Chapter 1

Introduction

You see things; and you say, "Why?"
But I dream things that never were; and I say, "Why not?"
George Bernard Shaw (from "Back to Methuselah")

In this thesis, we present a search for the flavor changing neutral current (FCNC) decay of the top quark, $t \to qZ$, in data collected by the Collider Detector at Fermilab (CDF) at the Fermilab Tevatron. Since this decay is extremely rare in the standard model (SM) of particle physics, below our current detection capabilities, or even those of the experiments at Large Hadron Collider (LHC) which will collide particles at seven times that Tevatron energy. However, many extensions to the standard model predict much higher rates for this decay.

Ultimately, the reason to search for this decay is that, empirically, top is the least well studied quark. It is the "youngest" experimentally, having been discovered only in 1995 [1]. It is also the rarest quark since its large mass renders it energetically expensive to produce. Consequently, the experimental bounds on the $t \to qZ$ decay leave room for exploration. The Tevatron is the only place in the world where top

quarks are produced in a controlled environment and our current dataset contains ten times as many quarks as were produced in the previous run of the Tevatron. We would be remiss if we did not check if nature has a surprise for us that has not been dreamt of in our many theories.

1.1 The Standard Model of Particle Physics

The fundamental constituents of matter and their interactions (except for gravity) are described up by the standard model of particle physics. The smallest building blocks of matter fall into two groups: quarks which are found only in hadrons, bound states of quark-antiquark pairs (mesons) or three (anti-)quarks (baryons), and leptons, which exist individually. There are six quarks and six leptons which come in three groups known as *families*. Each family consists of a doublet of quarks, one of charge 2/3 and one of charge -1/3, and a doublet of leptons, one of unity charge, and its corresponding neutral neutrino. Each successive family is a heavier version of the previous one. All ordinary matter is made up of fermions from the first family, the up quark, the down quark, the electron, and the electron neutrino.[1]

There are four known fundamental forces. They are, in order from strongest to weakest, the strong force, the electromagnetic force, the weak force, and gravity. The first three are described in the SM and are mediated by integer spin particles known as gauge bosons.[2] The electromagnetic force is mediated by the photon, the weak force

[1] A proton is made up of two up quarks and one down quark; a neutron is made up of one up quark and two down quarks.

[2] A mediator for gravity, the graviton, has been hypothesized but has not been discovered by experiment. Attempts at combining gravity with the rest of the standard model in general have so far been unrewarded by experimental evidence.

by the Z and the W bosons, and the strong force by the gluon. All the force carriers are electrically neutral except for the W boson which can carry an electric charge of ± 1. Furthermore, in the SM, the electromagnetic force and the weak force unify (electroweak symmetry) into one electroweak force at high energies. The particles in the standard model are summarized in Figure 1.1.

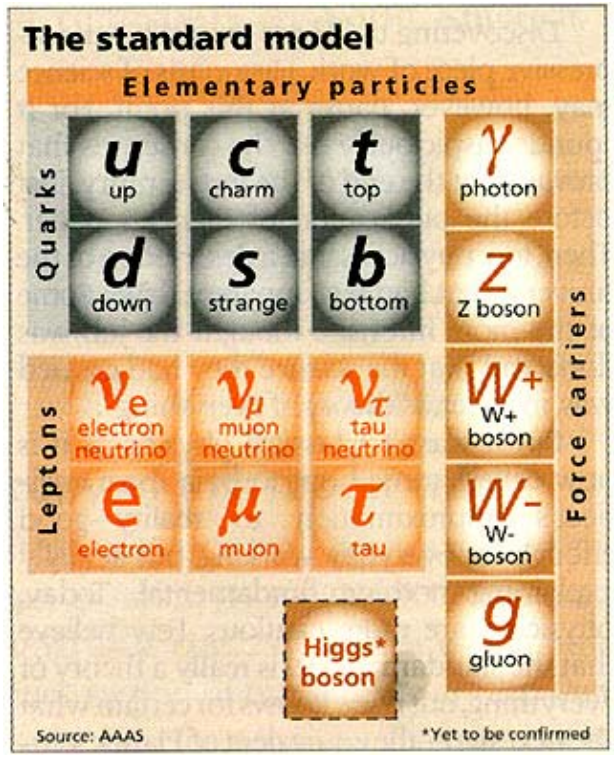


Figure 1.1: A summary of the standard model of particle physics.

Although the SM has been extremely successful and has withstood experimental tests since it was first proposed, it remains incomplete. The Higgs boson, the particle responsible for electroweak symmetry breaking (EWSB) and for giving particles mass, has yet to be discovered. To name a few other shortcomings: the SM has finely tuned parameters, it cannot explain why the weak energy scale and the Planck energy scale

are so different, and, on the astronomical end, the SM does not have any candidate particles for Dark Matter which is four to five times more abundant than ordinary matter and it cannot explain why the universe is made of matter and not anti-matter. Therefore, it is clear that physics beyond the SM must exist. Unfortunately, although many new physics theories have been proposed, none have been confirmed by experiment. A sensitive way to probe physics beyond the SM is to study the properties of the top quark, the most recently discovered particle in the SM.

1.2 The Top Quark

The top (t) quark was discovered at the Tevatron in 1995 [1]. For now, the Tevatron is still the only place in the world where t quarks are produced in a controlled environment. Although there were only a handful of events when it was first discovered, the current run of the Tevatron has produced a large number of tops, roughly 7500 $t\bar{t}$ pairs[3] in the dataset used in this analysis, providing us with a rich sample and a unique opportunity for exploring the properties of this "young" quark.

Other than the fact that it's the most recently discovered and, therefore, the least well studied, the top quark warrants further study for several fundamental reasons. The top quark is very massive, roughly 35 times more massive than the next most massive quark, the b, as shown in Figure 1.2. Its mass, 174.2 ± 3.3 GeV/c^2 [2], is of the same order as the EWSB scale, perhaps suggesting that it may have a role to play in the EWSB process. The top is also so massive that it decays before it can hadronize, form bound states, providing a rare opportunity to study a "bare" quark.

[3]Assuming the theoretical production cross-section

All this raise hopes that new physics might be discovered in the top quark sample. While many properties of the top, e.g. mass, charge, production cross-section, can be studied, of particular interest to this analysis is the way in which top decays. More specifically, we are interested in the extremely rare decay $t \to qZ$.

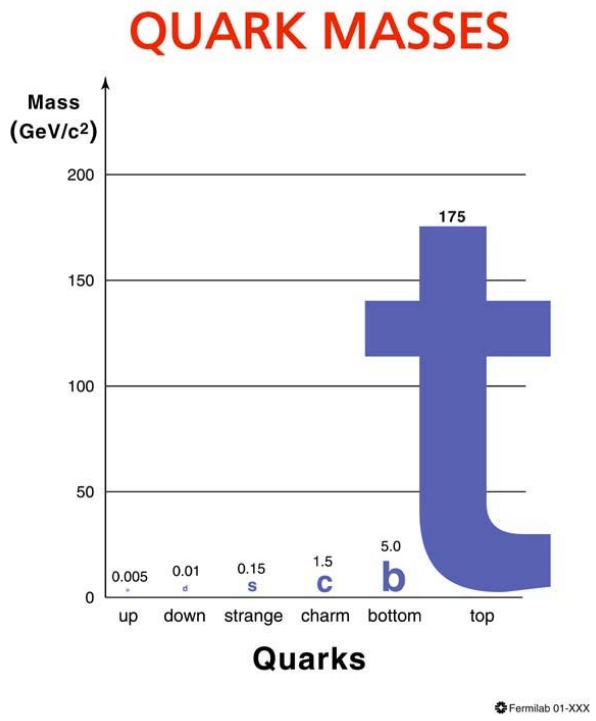


Figure 1.2: The masses of the quarks in the standard model.

1.2.1 Decays of the Top Quark

The vast majority of top quarks decay to a W boson and a b quark [3]. This is called a "charged current" since the top and bottom have different charges. We are interested in incidences when the top quark decays to a Z boson and another quark of the same charge (c or u). This is called a flavor changing neutral current (FCNC)

decay; flavor changing because it is decaying to a different species of quark, neutral current because the two quarks involved are of the same electric charge.

FCNC decays are absent at tree level (first order), shown in Figure 1.3(a), in the standard model. They are only allowed at higher order, e.g. in loop diagrams called penguin diagrams as shown in Figure 1.3(b), and even these higher order interactions are further suppressed. FCNCs via penguins loops have been discovered in lighter quarks, e.g. the CLEO collaboration discovered $b \to s\gamma$ in 1993 [4], but have never been seen for the top quark. Top FCNC decays are expected to be especially rare in the SM. The branching fraction for the top quark decay $t \to qZ$ is predicted to be $\mathcal{O}(10^{-14})$ [5], far below the experimental sensitivity of the Tevatron or even the Large Hadron Collider. However, as summarized by F. Larios and collaborators [6], there exist new physics models that predict much higher branching fractions, up to $\mathcal{O}(10^{-2})$. While these extensions to the SM are compelling, we conduct a model independent search, without relying on the properties of any particular new physics theory. Any detection of a top FCNC decay at the Tevatron would be an indication of physics beyond the standard model.

Previous searches for the FCNC $t \to qZ$ have been performed in CDF Run I and by the LEP experiments. The Run I analysis yielded an upper limit on the branching fraction $\mathcal{B}(t \to qZ)$ of 33% at 95% C.L. [7]. The current best 95% C.L. upper limit on $\mathcal{B}(t \to qZ)$ was obtained by the L3 experiment and amounts to 13.7% [8].

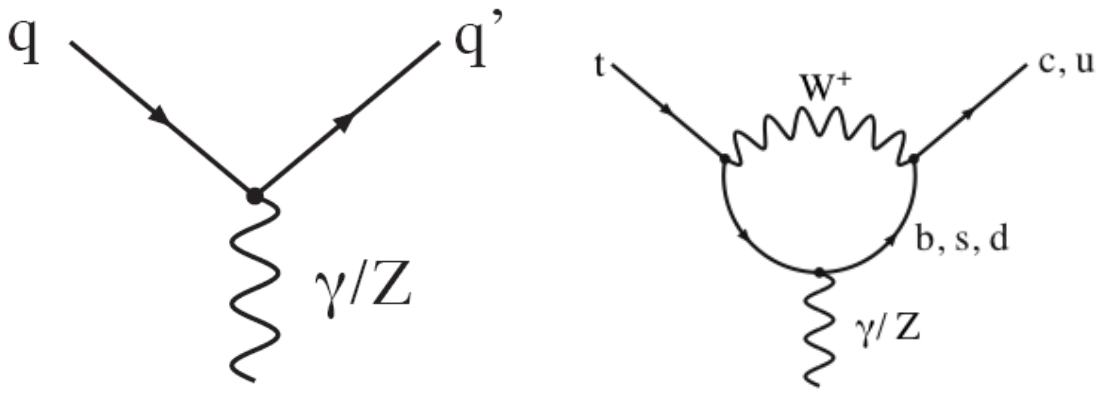


Figure 1.3: Feynman diagrams for the $t \to qZ$ FCNC decay. (a) Shows a tree level (first order) diagram which is absent in the standard model. (b) Shows an example of a higher order loop diagram, a.k.a. penguin diagram, which is allowed.

1.3 Signature of the $t \to qZ$ FCNC Decay

To search for the top FCNC decay $t \to qZ$, we target events with $t\bar{t}$ pairs where one top decays via the FCNC mode to a Z boson and a quark (c or u) and the other decays via the normal SM mode to a W boson and a b quark. We don't specifically require both tops to decay to a Z and q because the $t \to qZ$ decay is extremely rare and both tops decaying this way is suppressed compared to when only one top does. We do, however, allow for this decay to contribute to our signal acceptance. We are interested in the channel where the Z decays to a pair of oppositely charged leptons, e^+e^- or $\mu^+\mu^-$, and the W decays hadronically to a quark and an anti-quark of a different flavor, $q\overline{q}'$. Although the branching fraction of $Z \to e^+e^-/\mu^+\mu^-$ is only 3.33% per channel [2], we choose it because it is a very clean channel to identify that there was a Z present in the decay; a Z decaying to two quarks (i.e. jets) is very difficult to distinguish from the hadronic decay of the W, and a Z decaying to two neutrinos can only be detected by the presence of missing energy and cannot be

reconstructed with the CDF detector. We choose the $W \rightarrow q\overline{q}'$ decay mode because it has a large branching fraction of 68% [2] and contributes to a final state with a large jet multiplicity. The Feynman diagram for our signal of interest is shown in Figure 1.4. The final signature is a reconstructed Z and four or more jets, one of which is a b-jet that can be identified using a loose secondary vertex (SecVtx) tag. Our experimental signature does not include any neutrinos in the final state, and we are, therefore, able to fully reconstruct the event. We will make use of this fact by constructing a χ^2 from the masses of the reconstructed W, SM top, and FCNC top in the event.

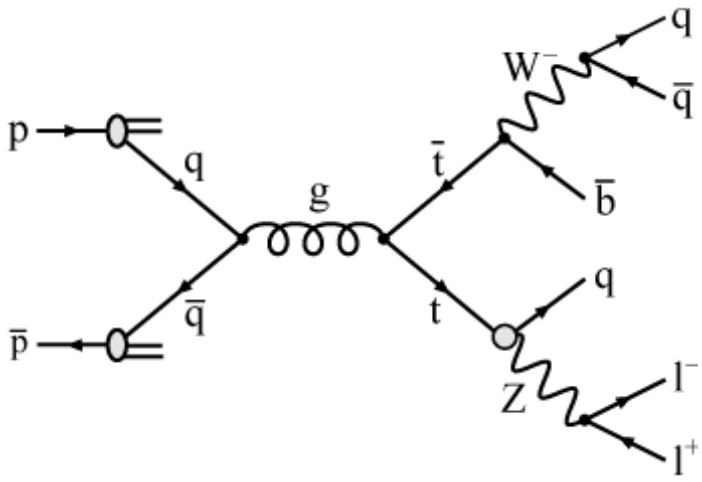


Figure 1.4: The Feynman diagram for our FCNC signal of interest; a $t\bar{t}$ event where one top decays via an FCNC to a Z and a q (c or u), the other top undergoes a SM decay to a W and a b, the Z decays to two leptons and and the W decays to two quarks.

Several more common SM processes can mimic the FCNC signal of interest. The most abundant of these background processes is Z bosons produced in association with jets (Z+jets). There are much smaller contributions from $t\bar{t}$ pairs where both

tops decay to a W and b. Although there are no Z bosons in these events, the two leptons when both Ws decay leptonically or the lepton and a jet when one W decays leptonically and the other decays hadronically can have an invariant mass within our Z mass window. A similar contribution comes from the electroweak production of a pair of gauge bosons where at least one is a Z, namely, ZZ and WZ. Finally, negligible contributions result from W+jets and WW production.

1.4 Analysis Method

We perform a blind search for the FCNC decay $t \to qZ$. The blinded region is initially defined as events with a reconstructed Z in the mass range of 76–106 GeV/c^2 and four or more jets. We later decrease the blinded region to Z + $\geq$4 jets events with mass χ^2 (constructed from reconstructed W, SM top, and FCNC top masses) less than 9.0.

We take the signal acceptances and efficiencies from a Monte Carlo (MC) simulation with appropriate scale factors and trigger efficiencies applied. We then normalize this acceptance of the expected number of signal events to the measured lepton+jets top cross section. The dominant standard model Z+jets production background is estimated using a rigorous combination of data and Monte Carlo techniques. SM top and diboson backgrounds are estimated using MC simulation, and W+jets production is estimated using both data and MC simulation.

We split the data sample into two subsamples that are analyzed separately, a "tagged" sample of events in which one or more of the four leading jets contain a secondary vertex identified by the loose SecVtx tagging algorithm, which indicates

that they contain heavy flavor (b or c hadrons), and an "anti-tagged" sample of events where none of the four leading jets have an identified secondary vertex. The event selection for the tagged and the anti-tagged samples are optimized for the best combined expected limit. We use both Feldman-Cousins (FC) and Bayesian frameworks for calculating expected limits; both frameworks take systematic uncertainties into account in the limit calculation. We find that the limits obtained in either framework track each other well. As the limit calculation in the FC framework is very CPU-intensive, we optimize our selection criteria for the best expected limit using the Bayesian framework and use the FC framework to obtain the limit for our signal event yields. After the final selection criteria have been chosen, we derive a limit on the branching fraction of the decay $t \to qZ$ from the number of events observed in the signal region and the number of expected background events.

Definition of Samples

We have several samples satisfying different selection criteria and we refer to them by different names. This section is an attempt to list the different samples, their selection criteria, and the different names by which I will refer to them throughout this thesis.

- Pre-tag sample:
  - Satisfies the base selection criteria, a reconstructed Z boson and four or more jets (Z + $\geq$4 jets)
  - This is our initial blinded region. Later, the blinded region is reduced only to events in the pre-tag sample which have mass $\chi^2 < 9.0$.

  - Events with a Z and three or fewer jets ($Z + \leq 3$ jets) are our initial control region. Later, $Z + \geq 4$ jets events with mass $\chi^2 > 9.0$ are added to the control region.

- Tagged sample:
  - Events where one or more of the four leading jets contain a loose SecVtx tag, and satisfy the optimized selection criteria.
  - This is our tagged signal region

- Anti-tagged sample:
  - Events where none of the four leading jets contain a loose SecVtx tag, and satisfy the optimized selection criteria.
  - This is our anti-tagged signal region.

Note that the tagged and anti-tagged sample are subsamples of the pre-tag sample and the tagged and anti-tagged signal regions are contained within the blinded region.

1.5 Overview of the Thesis

The rest of this thesis will describe the details of our search and the limit we set. Chapter 2 describes the theoretical predictions for this decay by both the standard model and new physics models, and the previous experiments which have searched for this vertex in both decay and production. Chapter 3 describes the experimental apparatus used for this search, namely the Fermilab Tevatron and the CDF II detector. In Chapter 4, the details of the particle identification and reconstruction algorithms

are described. In Chapter 5, the event selection criteria can be found. In Chapter 6, the measurement of acceptances and efficiencies for the FCNC signal (from Monte Carlo simulations) is explained. Chapter 7 contains the descriptions and estimations of the background processes which can masquerade as signal. Chapter 8 describes the limit calculation and optimization of event selection criteria. Chapter 9 has the inventory of the systematic uncertainties. The results are revealed in Chapter 10 and the conclusions can be found in Chapter 11.

Three appendices are also provided. The first contains a detailed description of the ALPGEN generator and the ALPGEN +PYTHIA samples used to study our main Z+jets background. The second contains cross-check and control plots for $Z + 4$jets events, our pre-tag sample. We have made several choices for our analysis, such as the generated top mass. The third appendix describes the other choices we could have made and how they affect our limit.

Chapter 2

Theoretical Predictions and Previous Limits

We have found a strange footprint on the shores of the unknown. We have devised profound theories, one after another, to account for its origins. At last, we have succeeded in reconstructing the creature that made the footprint. And lo! It is our own.
Sir Arthur Eddington

This chapter contains the theoretical predictions for the $t \to qZ$ decay, both in the standard model and in new physics models, and the limits that have been set by previous experiments. We will demonstrate the absence of flavor changing neutral currents at tree level in the SM. We discuss their presence in one loop diagrams and why even these higher order interactions are suppressed. The experimental evidence for this suppression in lighter quarks will be presented.

Despite these confirmations, we have reasons to suspect that the top quark may be special. A representative sample of new physics models which predict enhancements

for the $t \to qZ$ decay will be described. The limits set by previous experiments, in particular the CDF Run I limit and the limit set by the L3 experiment at the Large Electron Positron (LEP) collider, will be discussed.

2.1 Flavor Changing Neutral Currents in the Standard Model

The standard model [9] describes the constituents of matter, three generations of quarks and leptons, and their interactions the strong, weak, and electromagnetic forces. The governing symmetry of the SM is the $SU(3)_{color} \times SU(2)_{weak} \times U(1)_{hypercharge}$ gauge symmetry. The $SU(3)_{color}$ symmetry determines the interactions of the strong force, which does not allow quarks to change their flavor and, therefore, is irrelevant to flavor changing neutral current interactions. The $SU(2)_{weak} \times U(1)_{hypercharge}$ controls the electroweak interactions which permit flavor changing decays of quarks. The following sections will describe the parts of electroweak theory salient to the understanding of standard model FCNCs. The discussion will be limited to the quark sector, and further restricted to the parts of the electroweak Lagrangian relevant to flavor changing interactions.

2.1.1 Electroweak Theory

The electroweak quark eigenstate, ψ, is composed of a left-handed doublet and two right-handed singlets of the gauge group $SU(2)_{weak}$, which can be written as

$$\psi = \begin{pmatrix} \psi_{\rm L} \\ \psi_{\rm R} \end{pmatrix}, \quad \psi_{\rm L} = \begin{pmatrix} u \\ d \end{pmatrix}_{\rm L}, \quad \psi_{\rm R} = u_{\rm R},\, d_{\rm R}.$$

Since there are three generations of quarks, the u represents all up type quarks, u, c, and t, and the d represents all down type quarks, d, s, and t, i.e., the above notations are shorthand for

$$\overline{\psi}_{\rm L} = \left(\begin{pmatrix} \overline{u} \\ \overline{d} \end{pmatrix}_{\rm L}, \begin{pmatrix} \overline{c} \\ \overline{s} \end{pmatrix}_{\rm L}, \begin{pmatrix} \overline{t} \\ \overline{b} \end{pmatrix}_{\rm L} \right) \quad u_{\rm R} = \begin{pmatrix} u \\ c \\ t \end{pmatrix}_{\rm R}, \quad d_{\rm R} = \begin{pmatrix} d \\ s \\ b \end{pmatrix}_{\rm R}.$$

When these quarks interact, obeying the $SU(2)_{weak} \times U(1)_{hypercharge}$ symmetry, the Lagrangian for the kinetic energy is given by:

$$\mathcal{L}_{\rm Q} = \overline{\psi}_{\rm L}\, (i \not{D}_{\rm L})\, \psi_{\rm L} + \overline{\psi}_{\rm R}\, (i \not{D}_{\rm R})\, \psi_{\rm R}, \tag{2.1}$$

with the following definitions of the covariant derivatives:

$$D^{\mu}_{\rm L} = \partial^{\mu} + ig\, T^i W^i_{\mu} + ig' \frac{Y}{2} B_{\mu}$$
$$D^{\mu}_{\rm R} = \partial^{\mu} + ig' \frac{Y}{2} B_{\mu}.$$

The g and g' are the coupling constants of $SU(2)_{weak}$ and $U(1)_{hypercharge}$, respectively. The W^i_{μ}, $i = 1, 2, 3$, and B_{μ} are the $SU(2)_{\rm weak} \times U(1)_{\rm hypercharge}$ gauge fields. The physical electroweak gauge bosons ($W^{\pm}$, Z, and photon) have fields ($W^{\pm}_{\mu}$, Z_{μ}, and A_{μ}) which are a linear combinations of the W^i_{μ} and B_{μ}, namely,

$$W_\mu^\pm = \frac{1}{\sqrt{2}}\left(W_\mu^1 \mp W_\mu^2\right), \tag{2.2}$$

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_{\rm W} & \sin\theta_{\rm W} \\ -\sin\theta_{\rm W} & \cos\theta_{\rm W} \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix}, \tag{2.3}$$

where the mixing between the B_μ and the W_μ^i fields is characterized by the weak mixing angle $\theta_{\rm W}$.

2.1.2 Tree Level Flavor Changing Interactions

The Lagrangian for the interactions between quarks and gauge bosons can be written with three terms, one containing the electromagnetic current j_μ^{em}, one containing the weak charged current, J_μ^{CC}, and one containing the weak neutral current, J_μ^{NC}, as shown below.

$$\mathcal{L}_{\rm IA} = -e j_\mu^{\rm em} A^\mu - \frac{g}{2\sqrt{2}}\left(J_\mu^{\rm CC} W^{\mu+} + J_\mu^{\rm CC\dagger} W^{\mu-}\right) - \frac{g}{2\cos\theta_{\rm W}} J_\mu^{\rm NC} Z^\mu, \tag{2.4}$$

The currents are given by:

$$j_\mu^{\rm em} = \frac{2}{3}\,\overline{u}\,\gamma_\mu u - \frac{1}{3}\,\overline{d}\,\gamma_\mu d, \tag{2.5}$$

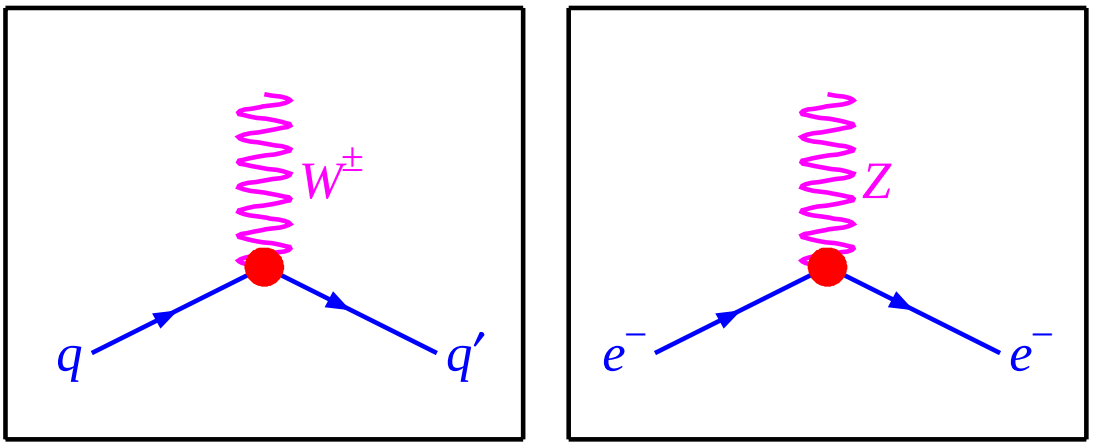
$$J_\mu^{\rm CC} = \overline{u}_{\rm L}\gamma_\mu\, d_{\rm L}, \tag{2.6}$$

$$\begin{aligned} J_\mu^{\rm NC} = {} & \overline{u}_{\rm L}\gamma_\mu \left(1 - \frac{4}{3}\sin^2\theta_{\rm W}\right) u_{\rm L} - \overline{u}_{\rm R}\gamma_\mu \frac{4}{3}\sin^2\theta_{\rm W} u_{\rm R} \\ & - \overline{d}_{\rm L}\gamma_\mu \left(1 - \frac{2}{3}\sin^2\theta_{\rm W}\right) d_{\rm L} + \overline{d}_{\rm R}\gamma_\mu \frac{4}{3}\sin^2\theta_{\rm W} d_{\rm R}. \end{aligned} \tag{2.7}$$

The details of theses currents are not important for our discussion. The crucial things to note are:

- The charged current is between down-type quarks and (anti-)up-type quarks.
  - i.e. it is a flavor changing interaction.
  - An example of this charged current, $t \to Wb$ decay, is shown in Figure 2.1(a).
- The terms in the neutral current (and also the electromagnetic current) only contain interactions of up-type quarks with anti-up-type quarks and down-type quarks with anti-down-type quarks.
  - i.e. it does not change flavors of quarks
  - An example of this neutral current, $\bar{s}s$ shown in Figure 2.1(b).

These currents describe leading order interactions, and their Feynman diagrams, shown in Figure 2.1, are called tree level diagrams.



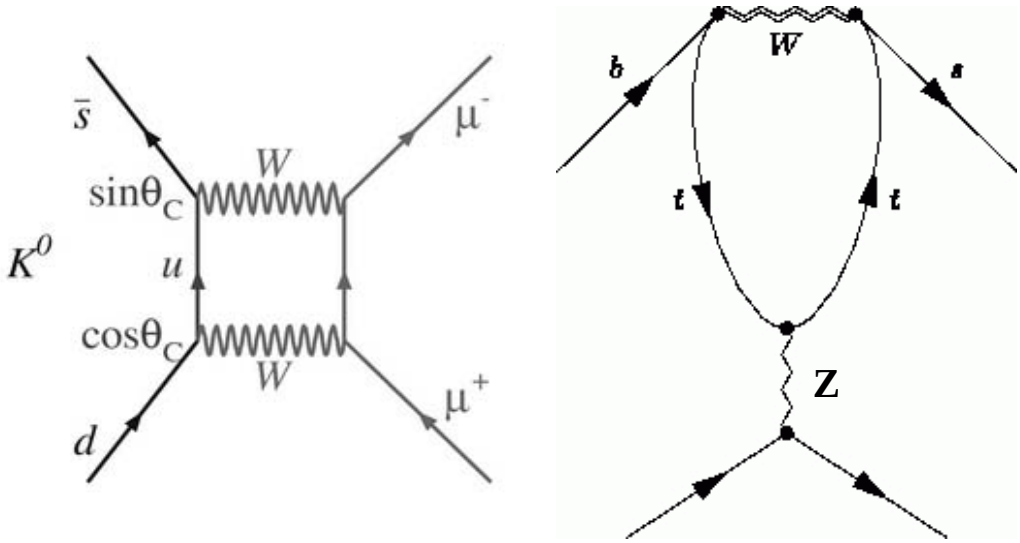
(a) (b)

Figure 2.1: Feynman diagrams of weak currents. (a) Shows an example of a charged weak current. (b) Shows an example of a neutral weak current.

2.1.3 One Loop FCNC Diagrams

We have seen that FCNCs are absent at tree level in the SM. However, electroweak theory allows for FCNCs at higher order. FCNCs can proceed via two types of one loop diagrams, box diagrams, such as the one shown in Figure 2.2(a), and penguin diagrams, such as the one shown in Figure 2.2(b).

The flavor changing charged current in the previous section is not enough to permit these one loop FCNC processes. An important ingredient is that the mass eigenstates, of the physical quarks, are rotated from the electroweak eigenstates described previously. This is a consequence of electroweak symmetry breaking via the Higgs mechanism. A complete description of the Higgs mechanism can be found in many textbooks (e.g. [10, 11, 12]. When the electroweak Lagrangian is rewritten in the mass basis, the neutral current keeps its form but the charged current transforms



(a) (b)

Figure 2.2: One loop FCNC Feynman diagrams. (a) Shows an example of a box diagram (b) Shows an example of a penguin diagram.

as shown below:

$$J_\mu^{\rm CC} = \overline{u}\left[\frac{1}{2}\gamma_\mu\left(1-\gamma_5\right)\right] d = \overline{u}_{\rm L}\,\gamma_\mu\, d_{\rm L} = \overline{u}'_{\rm L}\mathbf{U}_{\rm L}^{u\dagger}\,\gamma_\mu\,\mathbf{U}_{\rm L}^{d} d'_{\rm L} = \overline{u}'_{\rm L}\,\gamma_\mu\, d''_{\rm L} = \overline{u}'_{\rm L}\,\gamma_\mu\,\mathbf{V}_{\rm CKM}\, d'_{\rm L}, \tag{2.8}$$

where

$$d''_{\rm L} = \mathbf{U}_{\rm L}^{u\dagger}\mathbf{U}_{\rm L}^{d} d_{\rm L} = \mathbf{V}_{\rm CKM}\, d_{\rm L}.$$

The unprimed quarks are massless, the primed quarks are massive and $\mathbf{U}_{\rm L}$ are unitary matrices which transform the quarks from one basis to another. $\mathbf{V}_{\rm CKM}$ is the Cabibbo-Kobayashi-Maskawa (CKM) matrix [13], which parametrizes the couplings between different up-type and down-type quarks is given by:

$$\mathbf{V}_{\mathrm{CKM}} = \begin{pmatrix} V_{\mathrm{ud}} & V_{\mathrm{us}} & V_{\mathrm{ub}} \\ V_{\mathrm{cd}} & V_{\mathrm{cs}} & V_{\mathrm{cb}} \\ V_{\mathrm{td}} & V_{\mathrm{ts}} & V_{\mathrm{tb}} \end{pmatrix} \approx \begin{pmatrix} 1 & \lambda & \lambda^3 \\ -\lambda & 1 & \lambda^2 \\ \lambda^3 & -\lambda^2 & 1 \end{pmatrix}. \tag{2.9}$$

The parameter λ is the sine of the Cabbibo mixing angle [14], and its value is experimentally determined to be $\lambda \approx 0.22$. It emerges that while charged currents prefer interactions between quarks of different flavor in the same family, i.e. the diagonal entries are close to 1.0, they allow interactions between the up-type quarks of one generation and the (anti-)down-type quarks of another, rendering the diagrams shown in Figure 2.2 possible.

Further Suppression

Nevertheless, even these higher order interactions are further suppressed. We will focus on the penguin diagram since this is the one of interest in the $t \to qZ$ decay. The penguin loop diagram suppression is the result of the unitarity constraint for the CKM matrix, i.e. $\mathbf{V}_{\mathrm{CKM}}\mathbf{V}^{\dagger}_{\mathrm{CKM}} = \mathbf{V}^{\dagger}_{\mathrm{CKM}}\mathbf{V}_{\mathrm{CKM}} = \mathbf{1}$. This leads to relations between the rows and columns of the matrix, e.g. between the complex conjugate of the second column and the third column:

$$V^*_{\mathrm{us}}V_{\mathrm{ub}} + V^*_{\mathrm{cs}}V_{\mathrm{cb}} + V^*_{\mathrm{ts}}V_{\mathrm{tb}} = 0 \tag{2.10}$$

As an example, the matrix element of the FCNC process shown in Figure 2.2(b), $b \to sZ$, is obtained by the sum of the contributions from all three loops, one for each

up-type quark which can participate:

$$\mathcal{M} \propto F(x_{\rm u})\, V_{\rm us}^* V_{\rm ub} + F(x_{\rm c})\, V_{\rm cs}^* V_{\rm cb} + F(x_{\rm t})\, V_{\rm ts}^* V_{\rm tb}, \tag{2.11}$$

where $F(x_i)$ denotes a linear combination of Inami-Lin functions [15]. The important feature of the Inami-Lin functions to the FCNC decays is that they depend on a single parameter $x_i = {m_i}^2/M_W^2$, where m_i are the masses of the particles present in the loop, here u, c, and t. A comparison of the matrix element in Equation (2.11) to the unitarity relation of Equation (2.10) shows that $\mathcal{M}$ would vanish if the masses of the u, c, and t quarks and thus x_i were identical. This suppression was first theorized by Glashow, Iliopoulos, and Mainani, and is known as the GIM mechanism [16].

An important property of the GIM mechanism is that the suppression for the FCNC branching fractions in the top sector are more pronounced than in the bottom sector, because the particles in the loop are down-type quarks whose masses are more similar than those of the up-type quarks, as in the $b \to sZ$ example.

2.1.4 Flavor Changing Neutral Currents in Light Quarks

The flavor changing neutral current sector in light quarks has been a rich field of study, benefiting from the large datasets available at collider (e.g. CLEO, BaBar, Belle) and fixed target (e.g. KTeV, FOCUS) experiments. It has been confirmed that the branching fractions for FCNC decays of s, c, and b quarks are small and in accordance with SM expectations, either by discoveries or by setting stringent upper limits. A full list of FCNC decays in lighter quarks can be found in *The Review of Particle Physics* [2]. A few pertinent examples will be discussed below.

Strange FCNCs

The smallness of FCNCs was first discovered in the strange system. The $K_L^0 \to \mu^+\mu^-$ branching fraction, $6.87 \pm 0.11 \times 10^{-9}$ [17], was found to occur at a rate far below what was expected from the diagram shown in Figure 2.2(a). In fact, this led to the conception of the GIM mechanism and the prediction of the existence of the charm quark whose box diagram would cancel the contribution of the box diagram in Figure 2.2(a) to explain the rareness of this decay. The decay $K_L^0 \to e^+e^-$ is even smaller at $9^{+6}_{-4} \times 10^{-12}$ [2]. The FCNC decays of other kaons are also highly suppressed, e.g. $\mathcal{B}\left(K^+ \to \pi^+\nu\overline{\nu}\right) = 1.5^{+1.3}_{-1.9} \times 10^{-10}$ [2], as expected by the SM.

Charm FCNCs

As explained in Section 2.1.3, the FCNC suppression from the GIM mechanism is more severe for up-type quarks than for down-type quarks. Accordingly, FCNCs have never been observed for charm decays. Stringent limits have been set by many searches, e.g. $\mathcal{B}\left(D^0 \to \mu^+\mu^-\right) < 1.3 \times 10^{-6}$ at 90% C.L. [2], $\mathcal{B}\left(D^+ \to \pi^+e^+e^-\right) < 7.4 \times 10^{-6}$ at 90% C.L. [2], and $\mathcal{B}\left(D_s^+ \to K^+\mu^+\mu^-\right) < 3.6 \times 10^{-5}$ [2]. FCNC decays of c mesons via box diagrams are expected to have branching ratios $\sim 10^{-10}-10^{-9}$ [18].

Bottom FCNCs

Penguin decays were first discovered in the decay of b quarks. This is not surprising since b penguin decays are dominated by the contribution from top which is much heavier than u and c, and the GIM mechanism cancellation is less exact than for penguins decays of c and t quarks. In 1993, CLEO observed the evidence for the

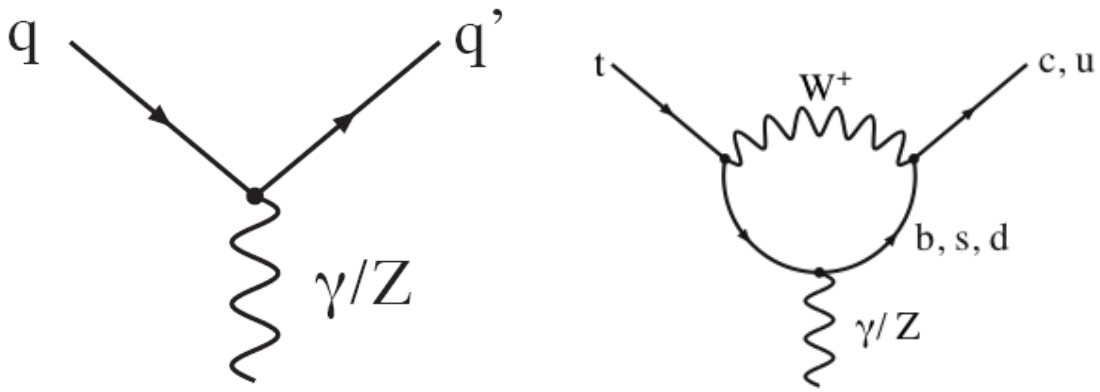
electromagnetic penguin decay $b \to s\gamma$ in the exclusive decay $B \to K^*\gamma$ with a branching fraction of $4.5 \pm 1.5 \pm 0.9 \times 10^{-5}$ [4]. This rate was within the expected range from the SM calculations. The electroweak penguin decay $b \to sZ$ was observed in $B \to K\ell^+\ell^-$ decays at Belle with a branching fraction of $0.75^{+0.25}_{-0.21} \pm 0.09 \times 10^-6$ [19] and in $B \to K^*\ell^+\ell^-$ decays at BaBar with a branching fraction of $0.88^{+0.33}_{-0.29} \times 10^{-6}$ [20]. The rates for these decays are also consistent with SM predictions.

2.2 Top Flavor Changing Neutral Current Decay $t \to qZ$

In the last section, we demonstrated that FCNCs are expected to be extremely rare in the SM and experimental limits and discoveries in lighter quarks have been consistent with these predictions. Nevertheless, we speculate that the top quark is an exception to the rule. Theoretically, its large mass may be an indication that it plays a role in electroweak symmetry breaking. Many new physics models predict much higher rates for the $t \to qZ$ decay compared to the SM expectations. Experimentally, the current bounds on the $t \to qZ$ decay leave room for exploration. Although the SM predicts that top decays to Wb more than 99.8% of the time [3], the experimental value for $R = \frac{t \to Wb}{t \to Wq}$ is much less stringent, with a 95% C.L. lower limit, $R > 0.61$ [21], leaving a large branching fraction window for other decays of top.

2.2.1 Standard Model Prediction for $t \to qZ$

A reminder: the $t \to qZ$ decay cannot proceed at tree level, shown in Figure 2.3(a), in the SM but is allowed in penguin diagrams, as shown in Figure 2.3(b). As described in Section 2.1.3, the penguin decays are suppressed by the GIM mechanism with very precise cancellations because the down-type quarks in the loop have very similar masses. Consequently, the SM prediction for the branching fraction of the $t \to cZ$ decay is $\mathcal{O}(10^{-14})$ [5]. The $t \to uZ$ decay is further Cabbibo supressed by a factor of $|V_{ub}/V_{cb}|^2 \simeq 0.0079$, i.e. $\mathcal{B}\ (t \to uZ) \sim 8 \times 10^{-17}$ [5].



(a) (b)

Figure 2.3: Feynman diagrams for the $t \to qZ$ decay. (a) Shows the tree level decay which is absent in the standard model. (b) Shows the standard model allowed penguin diagram.

2.2.2 New Physics Enhancements for $t \to qZ$

Many new physics models predict enhancements for the $t \to qZ$ decay orders of magnitude above the predicted rate in the standard model. These enhancement

can be in the form of extra particles in one loop diagrams or for some exotic models, even tree level terms. These models, summarized in Larios et al. [6] and J. A. Aguilar-Saavedra [5] (unless otherwise noted), span a large range in predictions of the branching fractions for the $t \to qZ$ decay, from $\mathcal{B}(t \to cZ) \sim 10^{-11} - 10^{-2}$. Table 2.1 lists the new physics models and their $\mathcal{B}(t \to cZ)$ predictions. The branching ratios for $t \to uZ$ are expected to be of the same order or less [5]. Each of these models is briefly described below. Note that while these models give us more hope than the SM for finding $t \to cZ$ decays, our search is model independent and does not rely on the properties of any particular model.

New Physics Model	$\mathcal{B}(t \to cZ)$
Two Higgs Doublet Models (THDM-II)	$\sim 10^{-11} - 10^{-8}$
Two Higgs Doublet Models (THDM-III)	$\sim 10^{-8} - 10^{-6}$
Supersymmetry (MSSM, R-parity conserved)	$\sim 10^{-6}$
Supersymmetry (MSSM, R-parity violate)	$\sim 10^{-4}$
Topcolor-assisted Technicolor	$\sim 10^{-5}$
LR SUSY, constrained	$\sim 10^{-6}$
LR SUSY, unconstrained	$\sim 10^{-5}$
Extra Quarks, Q=2/3 singlet	$\sim 10^{-4}$
Extra Quarks, fourth generation b'	$\sim 10^{-6}$
Effective Lagrangian Approach	$< 1.6 \times 10^{-2}$
Dynamical Electroweak Symmetry Breaking	0.053 ± 0.023

Table 2.1: Summary of $\mathcal{B}(t \to cZ)$ in new physics models. The enhancements result from extra particles in one loop interactions or even tree level interactions in some models.

Two Higgs Doublet Models

Two Higgs Doublet Models (THDM) are extensions of the standard model where the electroweak symmetry is broken by two $SU(2) \times U(1)$ scalar Higgs doublets,

instead of one. In these models, five physical Higgs bosons emerge from electroweak symmetry breaking; three neutral Higgs bosons, h and H which are CP-even and A which is CP-odd, and two charged Higgs bosons, $H^{\pm}$. There are three different types of THDM models. In models I and II (THDM-I, THDM-II) flavor is conserved in neutral current interactions. The $t \to cZ$ decay in these models proceed via one loop interactions mediated by the $H^{\pm}$. The branching fractions of the $t \to cZ$ in THDM-II (the larger of the two) is enhanced by three to six orders of magnitude compared to the SM, $\mathcal{B}\ (t \to cZ) \sim 10^{-11} - 10^{-8}$. In THDM-III, FCNCs involving the h and A Higgs bosons are allowed at tree level. Therefore, the $t \to cZ$ decay can proceed via one loop interactions which involve h and A as well as $H^{\pm}$. Consequently, the $t \to cZ$ rate is enhanced by up to another two orders of magnitude in THDM-III compared to THDM-I and THDM-II, $\mathcal{B}\ (t \to cZ) \sim 10^{-8} - 10^{-6}$.

Supersymmetric Models

Supersymmetry (SUSY) is a class of models which extend the standard model with a new symmetry which relates fermions and bosons. In SUSY, every SM fermion is paired with a bosonic partner (e.g. quark – squark, electron – selectron), and every SM boson is paired with a fermionic partner (e.g. gluon – gluino). The simplest SUSY extension to the standard model is known as the minimal supersymmetric standard model (MSSM). In MSSM, the possible misalignments between the rotation matrices which diagonalize the quark and squark sectors can lead to flavor violations which enhance FCNC rates. In MSSM models where R-parity[1] is conserved, the $t \to qZ$

[1] R = (-1)$^{3B+L+2S}$, where B = Baryon number, L = lepton number, and S = the spin of each field.

decay is enhanced by eight orders of magnitude compared to the SM, $\mathcal{B}\ (t \to cZ)$ $\sim 10^{-6}$. If R-parity is allowed to be broken, there is an additional enhancement and $\mathcal{B}\ (t \to cZ) \sim 10^{-4}$.

Topcolor-assisted Technicolor

In Technicolor theories, electroweak symmetry breaking is accomplished by a new, strongly coupled gauge interaction at TeV energy scales (instead of the Higgs mechanism). Embedding Technicolor in a larger gauge symmetry results and quark and lepton mass matrices. The Technicolor models developed to be consistent with LEP data and explain the large top quark mass are called Topcolor-assisted Technicolor (TC2) models. In these models, the $t \to cZ$ decay is enhanced by the virtual contributions of top-pions and the top-Higgs boson, resulting in $\mathcal{B}\ (t \to cZ) \sim 10^{-5}$.

Left-Right Symmetric Models

Left-Right (LR) symmetric models are built on the symmetries of the $SU(2)_L \times SU(2)_R \times U(1)_{B-L}$ gauge group, and were developed to study low energy parity violating weak interactions. In these models, the CKM matrix is not required to be unitary, due to the presence of vector-like heavy fermions, and FCNCs may exist at tree level. In SUSY versions of LR symmetric models, FCNCs can arise from only the non-unitary CKM matrix in the quark sector (constrained) or from the quark CKM matrix and mixings between the second and third generations in the squark mass matrix (unconstrained). The constrained LR SUSY models predict $\mathcal{B}\ (t \to cZ)$ $\sim 10^{-6}$ and the unconstrained LR SUSY models predict $\mathcal{B}\ (t \to cZ) \sim 10^{-5}$.

Models with Extra Quarks

In CKM matrix is no longer unitary in models where there are extra quarks. As, in LR symmetric models, this means that the tcZ coupling may occur at tree level. For models in which extra quarks which are $SU(2)_L$ singlets with electric charge Q=2/3, experimental data bounds allow large $t \to cZ$ branching ratios, $\mathcal{B}\ (t \to cZ)$ $\sim 1.1 \times 10^{-4}$. In models with $SU(2)_L$ Q = -1/3 singlets, however, the level of non-unitarity of the CKM matrix is stringently constrained by experiments and, therefore, $t \to cZ$ branching ratios are much smaller. If a fourth generation down-type quark, b', existed, the $t \to cZ$ decay rate will be enhanced to $\mathcal{B}\ (t \to cZ) \sim 10^{-6}$.

Effective Lagrangian Approach

The effective Lagrangian approach is where new physics effects from non-standard particles are parameterized as coupling constants of effective operators built from SM fields, i.e. a series of SM-gauge invariant operators are added to the SM Lagrangian. Each higher dimensional operator is added with a coefficient suppressed by an extra inverse power of Λ, the lowest energy scale for new physics. The largest contribution to top FCNCs result from dimension-6 operators. When constrained by existing experimental data, these models predict $\mathcal{B}\ (t \to cZ) < 1.6 \times 10^{-2}$.

Dynamical Electroweak Symmetry Breaking

A model with a particularly high $t \to cZ$ branching fraction is one in which electroweak symmetry breaking is achieved via a dynamical mechanism rather than the Higgs mechanism. In this particular model [22], electroweak symmetry is broken by

the addition of a gauge-invariant vertex of electroweak vector bosons' interaction. This interaction has a Λ cut off of a few TeV . This model predicts that $\mathcal{B}\ (t \to cZ)$ $= 0.053 \pm 0.023$.

2.2.3 Previous Limits on Top Flavor Changing Neutral Currents

Since the top quark is newly discovered and its production is energetically expensive due to its large mass, only a handful of top quarks have been available for study, and top FCNCs have not been as rigorously explored as FCNCs in lighter quarks. Currently, the experimental limits are consistent with the all the theoretical models described in the previous section (including the standard model).

Several experiments have searched for the top FCNC vertices $t - V - q$, where $V = Z, \gamma, g$ and $q = c, u$. Since top quarks have been produced only at the Tevatron, only CDF has searched for top FCNC decays. Other experiments have searched for single top production via an FCNC interaction. I will limit my discussion to two pertinent upper limits on the $t \to qZ$ branching fraction; the CDF Run I limit (33% [7]), set from an analysis similar to our own, and the L3 limit (13.7% [8]), which was the world's best limit prior to our analysis.

CDF Run I Top FCNC Search

The CDF Run I search [7] for the $t \to qZ$ decay in $p\overline{p}$ collisions is similar to ours. They searched for $t\bar{t}$ events where one decays to Zq and the other to Wb with the Z subsequently decaying to e^+e^- or $\mu^+\mu^-$ and the W to two quarks, i.e. $t\bar{t}$

$\rightarrow$ Zc Wb $\rightarrow \ell^+\ell^-$ q q' $\overline{q}$" b. Thus, the signature was also a Z boson reconstructed from an oppositely charged electron or muon pair and four or more jets. The dataset used was collected with high p_T electron and muon triggers and is equivalent to 110 pb^{-1} of integrated luminosity at the center-of-mass energy $\sqrt{s} = 1.8$ TeV.

They select events requiring:

- A reconstructed Z candidate
  - from e^+e^- or $\mu^+\mu^-$ with
  - 75 GeV/c^2 < $M_{e^+e^-/\mu^+\mu^-}$ < 105 GeV/c^2
- Four or more jets
  - Uncorrected $E_T > 20$ GeV
  - $|\eta| < 2.4$

They did not require a b-tag or further kinematic constraints. The signal acceptances and efficiencies were calculated from Monte Carlo simulations with events generated by the ISAJET generator and propagated through a parametric simulation of the CDF I detector. A sample where both tops decay via the FCNC mode $t \rightarrow qZ$ was also generated to account for additional acceptance gained from it.

The signal acceptances and efficiencies were normalized to the number of $t\overline{t}$ events observed in data, where both tops decay via the SM mode $t \rightarrow Wb$ and one W decays leptonically while the other decays hadronically. These events were identified by a high p_T ($p_T > 20$ GeV/c) electron or muon, $\geq$20 GeV missing transverse energy from the undetected neutrino, and three or more jets with uncorrected $E_T > 15$ GeV,

at least one which must have a secondary vertex tag indicating that it contains a b hadron. There were 34 $t\bar{t}$ candidate events (estimated background of 9 ± 1.5 events) in the 110 pb^{-1} of data.

The sources of background were Z bosons produced in association with jets (Z+jets), SM decays of $t\bar{t}$ events (SM top), and electroweak production of pairs of gauge bosons (dibosons), namely ZZ and WZ. The Z+jets background was simulated using VECBOS. The SM top and diboson backgrounds were simulated using ISAJET. The total background was estimated to be 1.2 events (0.5 events from Z+jets, 0.6 events from SM top, and 0.1 events from dibosons).

A single $Z \rightarrow \mu^+\mu^-$ candidate event passed all event selection requirements. This event yield is consistent with the background estimation. Furthermore, this event had kinematics consistent with Z+jets background. The 95% C.L. upper limit on the $t \rightarrow qZ$ decay was set at $\mathcal{B}\,(t \rightarrow qZ) < 33\%$.

CDF also searched for the decay $t \rightarrow q\gamma$ in the 110 pb^{-1} Run I dataset [7]. They searched for $t\bar{t}$ events where one decays to $q\gamma$ (FCNC decay) and the other to Wb (SM decay). They search for both the leptonic and hadronic decays of the W. For the leptonic decay, they select events with an electron or muon with $p_T > 20$ GeV, at least 20 GeV of missing E_T, one photon with $E_T > 20$ GeV, and two or more jets with $E_T > 15$ GeV. There must be a photon-jet combination with invariant mass between 140 GeV/c^2 and 210 GeV/c^2. For the hadronic decay, events are required to have a photon with $E_T > 50$ GeV , and four or more jets one of which must be identified as containing a b by a secondary vertex tag. The photon and one of the untagged jets must have an invariant mass within the range 140-210 GeV/c^2. The remaining

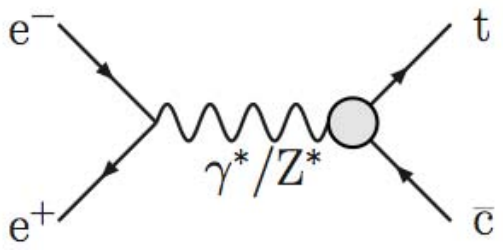
jets are required to have $\sum E_T > 140$ GeV . The acceptances and efficiencies for the signal were calculated from ISAJET Mont Carlo simulations.

The signal was normalized to the observed number of $t\overline{t}$ events where both tops decay to Wb, with one of the Ws decaying leptonically and the other hadronically, as it was for the $t \to qZ$ analysis. The main source of background was $W + \gamma + \geq 2$jets production. This background was derived from data and found to be less than 1 event for both channels combined. One event with a muon, photon, 3 jets, and missing E_T passed all selection requirements, consistent with background predictions. The 95% C.L. upper limit on the branching fraction was set at $\mathcal{B}\ (t \to q\gamma) < 3.2\%$.

L3 Top FCNC Search

L3, like other experiments, at the Large Electron Positron collider searched for single top production through FCNC interactions, $e^+e^- \to \gamma$ or $Z \to t\overline{c}$, as shown in the Feynman diagram in Figure 2.4. From the lack of this FCNC top production, a limit can be set on the coupling strength and, subsequently, on the branching fraction for the top FCNC decay. Note that in this search, it is impossible to distinguish between whether the intermediate state was a photon or a Z. Therefore the limit set on $t \to cZ$ depends on the limit set on $t \to c\gamma$ and vice versa. The upper limit on the branching fraction for $t \to qZ$ is set assuming that the branching fraction for $t \to q\gamma$ is zero. The L3 experiment at LEP set the world's best limit on the $t \to qZ$ branching fraction prior to our analysis.

The L3 search for single top production [8] was conducted using a dataset corresponding to an integrated luminosity of 634 pb^{-1} with a range of center-of-mass



(a)

Figure 2.4: Feynman diagram for the Single Top Production via FCNC

energies, $\sqrt{s} = 189$ to 209 GeV. They search for the production of a top quark and a charm quark where the top quark is produced almost at res and decays to immediately to Wb, leaving a W boson, a b quark, and a c quark in the final state. The energy of the c quark, E_c, has a fixed value for a particular center-of-mass energy, given by:

$$E_c = \frac{\sqrt{s}}{2}(1 - \frac{m_t^2}{s}). \tag{2.12}$$

and the energy of the b quark, E_b, has an almost constant value independent of the center-of-mass energy, assuming that the quark is at rest in the center-of-mass frame, given by:

$$E_b \simeq \frac{m_t}{2}(1 - \frac{m_W^2}{m_t^2}) \tag{2.13}$$

They search for both leptonic ($W \to \ell\nu$) and hadronic ($W \to q\overline{q}'$) decays of the W. The signal acceptances and efficiencies were calculated from PYTHIA Monte Carlo simulations. In searching for both channels, the events are selected in a three tier

process; a preselection, a channel-specific event selection, and finally, a requirement based on the output of a neural network. Both analyses use b-tag information from an algorithm mostly based on lifetime information in the neural net.

The event signature for the leptonic decay channel consists of an energetic lepton, large missing momentum, and two jets with a large difference in energy. The events are selected requiring the following criteria:

- Preselection criteria:
  - ≥ 3 tracks.
  - ≥ 15 calorimetric clusters.
  - $0.25 \times \sqrt{s} <$ visible energy $< 0.9 \times \sqrt{s}$.
  - Presence of a well defined lepton (electron, muon or tau); if there is more than one, the most energetic one is kept.
  - Clusters not associated with the lepton are combined to form two jets
  - Jet axes must be more than 15° away from the beam axis
  - Transverse momentum vector must be more than 26° away from the beam axis.
- Signature driven selection criteria:
  - Lepton energy > 10 GeV .
  - Missing momentum > 25 GeV .
  - Lepton+missing momentum invariant mass > 20 GeV .
  - Energy of leading jet > 60 GeV .

  - Energy of the softest jet less than cut off dependent on center-of-mass energy of of the collisions; e.g. 17 GeV at $\sqrt{s} = 188.6$ GeV (lowest) and 34 GeV at $\sqrt{s} = 206.6$ GeV (highest).
  - Width of softest jet < 0.4; width is defined as scalar sum of the transverse momenta of the jet clusters normalized to the jet energy.
  - Invariant mass of the two jets not within W mass window; window has a range of values, e.g. 14 GeV at $\sqrt{s} = 188.6$ GeV (lowest) and 3 GeV at $\sqrt{s} = 206.6$ GeV (highest).

- Neural network output:
  - 10 input nodes of variables related to missing momentum vector, b-tag value, and invariant masses of the two-jet and lepton+missing momentum systems.
  - 2 output nodes: $\mathcal{O}_{tc}$ and $\mathcal{O}_{back}$.
  - discriminating variable $\mathcal{O} = \mathcal{O}_{tc} \times (1 - \mathcal{O}_{back})$

The sources of background for this channel are $e^+e^- \rightarrow W^+ W^-$ modeled by KORALW Monte Carlo simulations, $e^+e^- \rightarrow q\overline{q}(\gamma)$, modeled by PYTHIA and KK2f Monte Carlo simulations. There were 346 events in data when 357.0 $\pm$ 1.8 events were expected from background processes.

The event signature for the hadronic channel is four jets, two of which has an almost fixed energy (one of them a b-jet) and the other two resulting from a decay of the W. The sources of background for this channel are also $e^+e^- \rightarrow W^+ W^-$ and $e^+e^- \rightarrow q\overline{q}(\gamma)$.

The events are selected requiring the following criteria:

- Preselection criteria:
  - Visible energy $> 0.7 \times \sqrt{s}$.
  - Effective center of mass energy (after removing photon radiation in the initial state) $> 0.85 \times \sqrt{s}$.
  - > 20 reconstructed tracks.
  - > 2 jets.
- Signature driven selection criteria:
  - Events are forced into a four-jet topology.
  - The c-jet is identified as the the jet with the closest energy to the one defined in Equation 2.12 assuming $m_t = 174.3$ GeV .
  - The two jets with an invariant mass closest to the W mass are identified as W daughters.
  - The remaining jet is assumed to be the b-jet
- Neural net output:
  - 24 input nodes of variables related to jet characteristics, b-tag discriminant of the b-jet candidate, event shapes
  - 3 output nodes: two for the two dominant backgrounds, one to select signal.

- Value of the output to select $e^+e^- \to q\overline{q}(\gamma)$ required to be less than 0.1
- Value of the output to select $e^+e^- \to W^+ W^-$ required to be less than 0.3
- Output to select signal is used as the discriminant.

There were 321 events in data for the hadronic channel when 287.9 $\pm$ 1.6 events were expected from background processes. There is no significant deviation from the standard model background predictions in data in either channel. Combining the leptonic and hadronic samples, limits were set on the single top production ($e^+e^- \to t\overline{c}$) cross section at different center-of-mass energies for $m_t = 170$ GeV/c^2, 175 GeV/c^2, and 180 GeV/c^2. A 10% deteoration of the limit is found for the $e^+e^- \to t\overline{u}$ process. The limits on the FCNC couplings are set and finally, the limits on the FCNC branching fraction are set. Figure 2.5 shows the exclusion region for the top FCNC decays in the $t \to qZ$ vs. $t \to q\gamma$ plane for top masses of 170 GeV/c^2, 175 GeV/c^2, and 180 GeV/c^2. Assuming that there is no $t \to q\gamma$, the 95% C.L. upper limit for the $t \to qZ$ branching fraction is $\mathcal{B}\ (t \to qZ) < 13.7\%$.

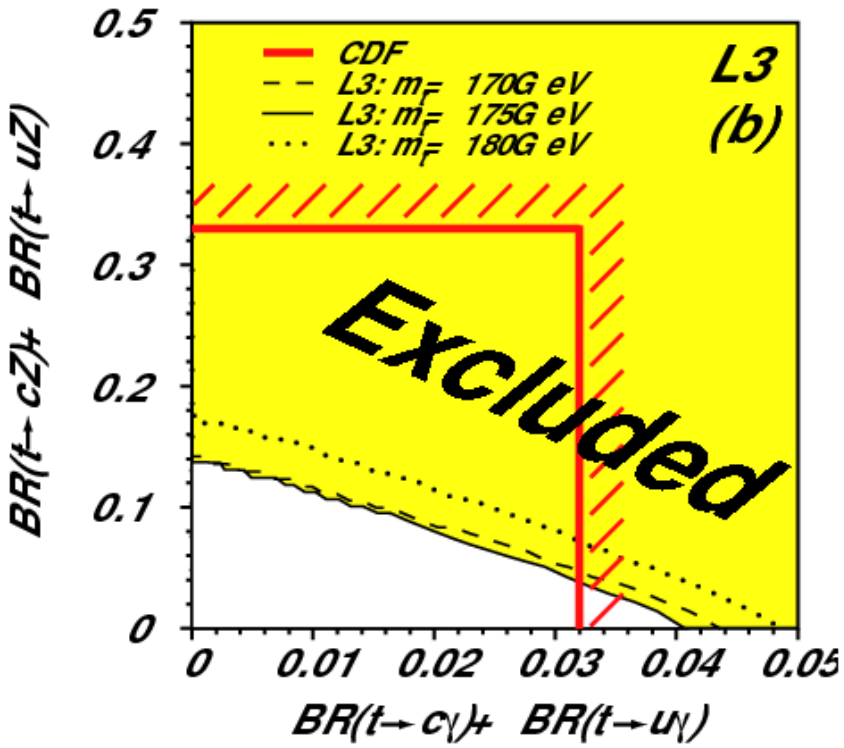


Figure 2.5: The L3 limit for the top FCNC decays $t \to qZ$ and $t \to q\gamma$. The analysis searched for single top production via the FCNC process. The dataset was from e^+e^- collisions, corresponds to 634 pb^{-1} of integrated luminosity. $\mathcal{B}\ (t \to qZ) < 13.7$ at 95 % C.L. for a top mass of 175 GeV/c^2 and assuming that the $t \to q\gamma$ decay is absent. The square box indicates the CDF Run I result. [8]

Chapter 3

Experimental Apparatus

A scientist in his laboratory is not only a technician:
he is also a child placed before natural phenomena which impress him
like a fairy tale.
Marie Curie

The data for this analysis is from proton-antiproton collisions produced by the Fermilab accelerator complex and collected by the Collider Detector at Fermilab (CDF) experiment.

3.1 The Fermilab Accelerator Complex

The Fermilab accelerator complex [23] has a series of acceleration systems. The last in the chain is the Tevatron, which accelerate protons and antiprotons and collide them at a center of mass energy of $\sqrt(s) = 1.96$ TeV. The equal energy beams are circulated in the same beampipe traveling in a helical pattern in opposite directions, clockwise for protons and anti-clockwise for antiprotons, and are brought to collide at

two points in the ring, named B0 and D0, where two multi-purpose detectors, the Collider Detector at Fermilab (CDF) and D0, are located to study the physics processes at the highest energies currently available. Figure 3.1 shows Fermilab's accelerator chain consisting of the proton source (Cockcroft-Walton pre-accelerator, the Linear Accelerator, the Booster), the antiproton source (target, debuncher, accumulator, recycler), the Main Injector, and the Tevatron Colliding Ring.

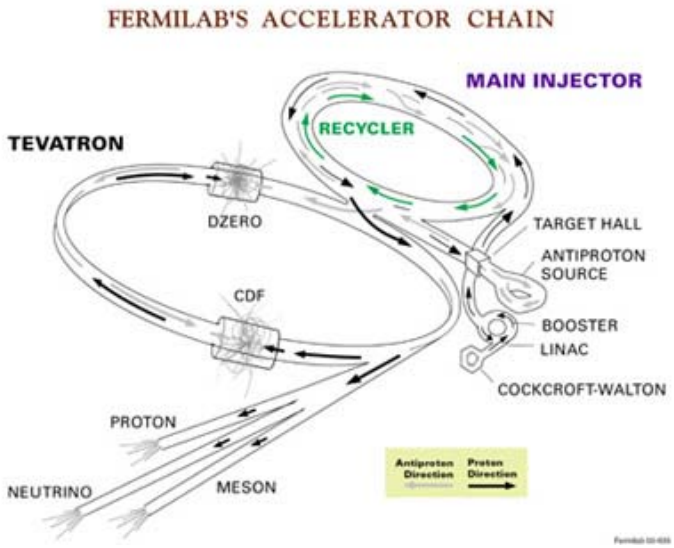


Figure 3.1: Fermilab's chain of accelerators

3.1.1 Proton Source

The acceleration process begins at the proton source which extract protons and accelerate them to 8 GeV to be injected into the Main Injector. The pre-accelerator, the linear accelerator, and the booster make up the proton source.

Cockcroft-Walton Pre-accelerator

The Cockcroft-Walton pre-accelerator consists of a cesium source in a dome, electrically charged to a potential of -750 kV, which ionizes hydrogen gas to negatively charged hydrogen (H^-) ions. The H^- ions then travel through a column to a grounded wall, accelerating to an energy of 750 keV. This process is repeated every 66 milliseconds and the hydrogen ions are transferred to the linear accelerator (Linac).

Linear accelerator (Linac)

The Linac takes the 750 keV H^- ions from the pre-accelerator and accelerates them to an energy of 400 MeV using a series of radio frequency (RF) cavities. The Linac has two sections; the low energy drift tube linac (DTL) and the high energy side coupled cavity linac (SCL). The Linac also accelerates beam every 66 milliseconds. The beam is focused by quadrupole magnets, located within the drift tubes in the DTL and between the accelerating modules in the SCL. After the Linac, the focused 400 MeV H^- ions are transferred to the booster to continue being accelerated or to the Neutron Therapy Center to strike a target and produce neutrons for cancer treatment.

Booster

The Booster is a synchrotron (circular) accelerator which takes the 400 MeV H^- ions, strips off the electrons by passing the ions through a carbon foil, and accelerates the remaining protons to an energy of 8 GeV. The booster is composed of a series of 75 magnets arranged in a 75-meter radius circle, interspersed with 18 RF cavities. The

Booster can also accelerate beam every 66 milliseconds, with phase offsets to accept the H^- ions from the Linac. After being accelerated in the booster, the protons are transferred to the Main Injector.

3.1.2 Main Injector

The Main Injector is a synchrotron accelerator with a radius of approximately half a kilometer and has 18 accelerating RF cavities. It takes the 8 GeV protons from the Booster and accelerates them to 120 GeV or 150 GeV depending on their next destination. If the protons are intended for the antiproton source, they are accelerated to 120 GeV and if they are meant for injection into the Tevatron, they are accelerated to 150 GeV. The Main Injector can also accept 8 GeV antiprotons from the antiproton source and accelerate them to 150 GeV for injection into the Tevatron. In addition, the Main Injector can send protons to fixed target experiments. The Main Injector accelerates particles every 2.2 seconds.

3.1.3 Antiproton Source

The antiproton source consists of a target and three systems, the debuncher, the accumulator, and the recycler. The process of producing antiprotons and collecting enough to be injected into the Main Injector is called stacking. This is the limiting process in the acceleration chain since on average only one or two antiprotons are captured and stored for every 10^5 protons striking the target. Although much effort has been invested to improve the process, it still requires hours to build up a suitable stack for a collider store. Initially, the recycler was designed to accept antiprotons

from the Tevatron at the end of a store and recycle them for the next store to reduce the time between stores. However, this has not been achieved and, currently, the recycler aides the accumulator in storing the antiprotons produced at the target.

Target

The 120 GeV protons from the Main Injector, focused into a small beam spot by quadrupole magnets, strike a nickel target, producing a shower of secondary particles. These secondary particles are focused by a lithium lens [24]. Following the lens is a dipole pulsed magnet which selects negatively charged particles with approximately 8 GeV of energy, the average energy of the antiprotons produced. The rest of the particles are absorbed by a beam dump.

Debuncher

The Debuncher is a rounded triangular synchrotron, with a mean radius of 90 meters, designed to receive antiprotons from the target and reduce their spread in momentum through RF bunch rotation [25] and adiabatic debunching [26]. This eases the transfer of antiprotons to the Accumulator which has a limited momentum aperture. The Debuncher also reduces the transverse energy of the antiproton beam in a process known as stochastic cooling **??**. In stochastic cooling, the beam passes by a scanner which detects the position of particles within it. The information is sent across the ring (instead of around) to a "kicker" on the other side. The kicker then processes the information and applies a correction (kick) to the beam to reduce it's momentum. The cooled 8 GeV antiprotons are injected into the Accumulator with a small momentum spread.

Accumulator

The Accumulator is another rounded triangular synchrotron, occupying the same tunnel as the Debuncher. Its purpose is to accumulate and store antiprotons for injection into the Main Injector and then to the Tevatron for a colliding store. The Accumulator receives unbunched antiprotons from the Debuncher and reduces their longitudinal and transverse momentum using RF and stochastic cooling processes. During a shot, the 8 GeV antiprotons are transferred to the Main Injector, accelerated to 150 GeV and injected into the Tevatron.

Recycler

The Recycler shares the same tunnel, located 47 inches above the Main Injector. Initially, the hope was for the Recycler to receive the antiprotons from the Tevatron left over at the end of a colliding store. Unfortunately, this has not been accomplished. However, it can still store antiprotons produced at the target, aiding the Accumulator. Storing antiprotons in the Accumulator gets progressively harder as a large number are stored because instabilities develop and the stack can be lost. Due to this, some of the antiprotons are stored in the Recycler. To do so, antiprotons are extracted from the Accumulator and injected into the Main Injector as if for acceleration for collisions. However, the Main Injector does not accelerate them, transferring them to the Recycler instead. The antiprotons are then cooled through stochastic and electron cooling and stored in the Recycler. For a shot for collisions, some protons can be taken from the Accumulator and some from the Recycler, in a "Mixed shot".

3.1.4 Tevatron

The Tevatron, the largest and last in the chain of accelerators at Fermilab, is a synchrotron with a radius of 1 km and has eight accelerating cavities (RF freq = 53.104 MHz), dipole, quadrupole, and octupole focusing magnets. The magnets in the Tevatron are made of superconducting niobium/titanium alloy and are cooled by liquid Helium to $\sim$4 K. The superconducting magnets are necessary because they allow the large currents needed for magnetic fields required to accelerate particles to TeV energy scales without the resistance which will heat and destroy magnets.

The Tevatron is not perfectly circular. It is divided into six sections, labeled A through F, and each section has five subsection, a large, straight "0" section and sections "1" through "4" which are curved. The straight "0" sections have special uses, notably, B0 houses the CDF detector and D0 is the location of the D0 detector. The Tevatron receives protons and antiprotons from the Main Injector, accelerates them from 150 GeV to 980 TeV in about 85 seconds. After acceleration, the Tevatron acts as a storage ring for collisions to occur. The proton and antiproton beams are squeezed down to narrow beam-lines and are forced to collide at B0 and D0. This is called a "store" and the collisions last until the luminosity drops below a threshold when it is no longer useful for colliding physics experiments, which takes roughly 20 hours.

3.1.5 Beam

The proton and antiproton beams circulate in opposite directions, clockwise for protons and anti- clockwise for antiprotons, in the same beam pipe and collide at B0

and D0 where the CDF and D0 detectors, respectively, are located. The beams cross every 392 ns. In order for them to collide only where the experiments are located and not everywhere along the ring, the beams are made to travel in a helical pattern winding around each other. Each beam is made up of 36 bunches spread out into three sections with 12 bunches each, separated by 2.6 μs abort gaps where there is no beam. When the beam needs to be aborted, it is redirected during an abort gap into a beam dump. Sometimes protons and antiprotons leak into the abort gaps which is dangerous when beam needs to be aborted because the particles and spray into and destroy detectors and magnets during the redirecting process.

On average, there are 2.7×10^{11} protons and 3.0×10^{10} anti-protons per bunch. The radius of the beam in the interaction region is $\sim$26 μm and the interaction region extends about 60 cm in the z-direction, along the beam line.

3.1.6 Luminosity

The amount of data collected by experiments is measured in terms of luminosity, which determines how many collisions occur. The total integrated luminosity has units of 1/Area and is the time integral of instantaneous luminosity defined as [27]:

$$\mathcal{L} = \frac{fBN_pN_{\bar{p}}}{2\pi(\sigma_p^2 + \sigma_{\bar{p}}^2)} \cdot F(\frac{\sigma_l}{\beta^*}) \tag{3.1}$$

where f is the revolution frequency in Hertz, B is the number of bunches, $N_{p(\bar{p})}$ is the number of protons (antiprotons) per bunch, $\sigma_{p(\bar{p})}$ is the RMS beam size at the interaction point, and $F(\frac{\sigma_l}{\beta^*})$ is a form factor which depends on the ratio of the bunch length in z, σ_l, and β^*, which is measured to be $\sim$38 cm at CDF. Figure 3.2 shows

the total luminosity delivered to and collected by CDF as a function of store number, covering the period from 2002 to 2007. Figure 3.3 shows the instantaneous luminosity at the beginning of each store (which is when it is the highest) as a function of store number, covering the period from 2002 to 2007.

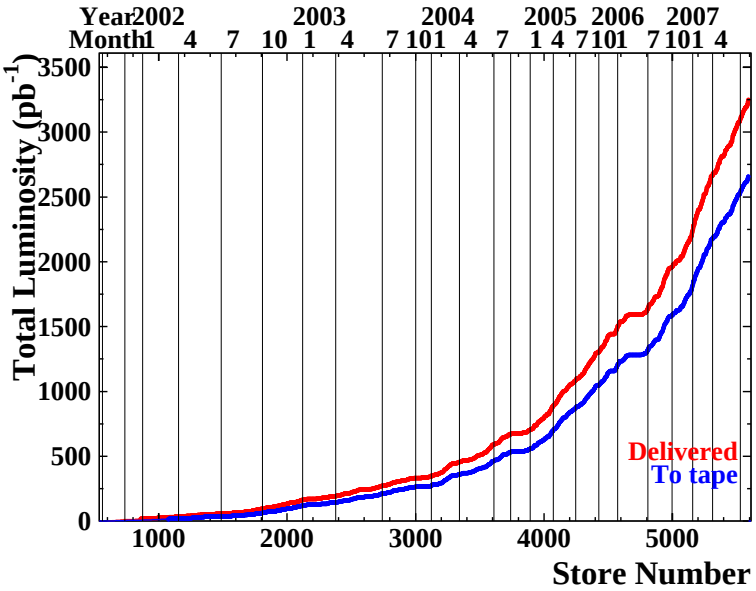


Figure 3.2: Total luminosity delivered by the Tevatron and recorded at CDF as a function of store number. This covers the period from 2002 to 2007 as indicated above the plot. The data used in this analysis was collected between March 2002 and September 2006, roughly 1.12 fb^{-1} [28].

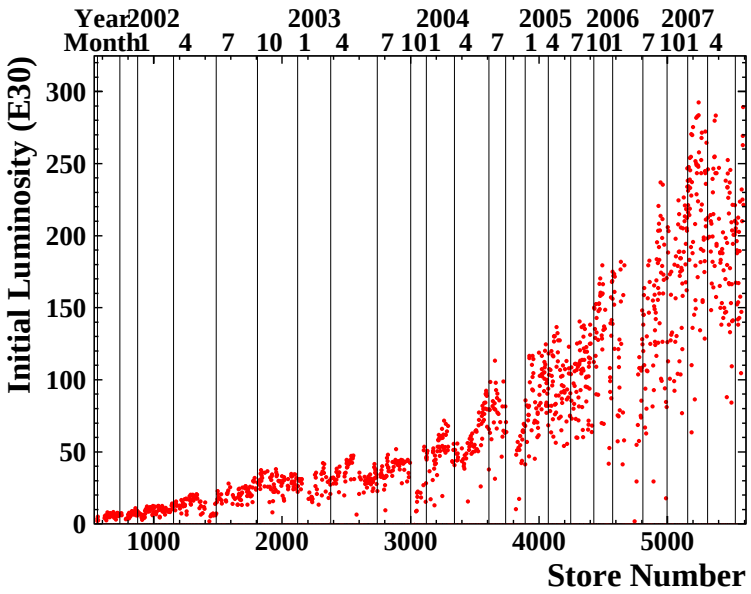


Figure 3.3: Instantaneous luminosity at the beginning of each store (peak) as a function of store number. This covers the period from 2002 to 2007 as indicated above the plot.The data used in this analysis was collected between March 2002 and September 2006, roughly 1.12 fb^{-1} [28].

3.2 The CDF Run II Detector

The Collider Detector at Fermilab (CDF) is a multi-purpose detector built at the B0 section of the Tevatron to study the proton-antiproton collisions. An elevation view of the detector is shown in Figure 3.4. The protons enter the detector from the west and anti-protons enter from the east. Collisions happen roughly at the center of the detector in a cylindrical region with a radial cross-section of $\sim$30 μm and a length of $\sim$60 cm. Even though the energy of the protons (anti-protons) are 980 GeV, the energy of the individual quarks (antiquarks) or gluons which interacted, and the boost along the beam axis, are not precisely known. Consequently, many quantities, such as momentum and energy, are measured in the transverse plane (p_T and E_T).

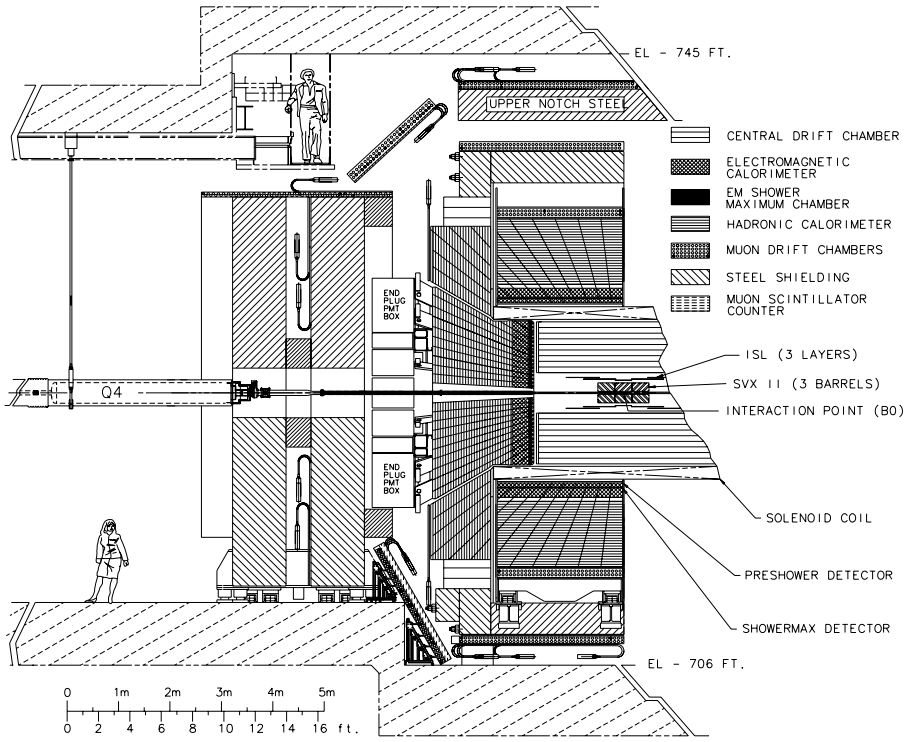


Figure 3.4: An elevation view of the CDF detector. The human figures are presented for size even though a woman physicist would not wear a short skirt in the collision hall. She would also know that she should wear a hard hat or be reprimanded by Dee Hahn.

At the center of the CDF detector are the tracking systems, enclosed within a 1.4 Tesla solenoidal magnetic field parallel to the beam axis. The curvature of the tracks in the magnetic field allow for measurement of particle momentum. The tracking at CDF is performed by two systems, eight layers of silicon microstrip detectors and a ninety six layer open cell drift chamber. The silicon detectors provide accurate impact parameter resolution and allows for the detection of displaced vertices while the COT provides seed tracks and accurate momentum resolution and tracks the passage of charged particles. Outside the solenoid are the electromagnetic and hadronic calorimeters which measure the energy of electrons, photons, pions, and jets. Beyond

the calorimeters, are the muon systems. Most of the particles which have reached the muon chambers are muons.

At hadron colliders, the collision rate is high and most of the events are low p_T physics which is not the object of study at CDF. Therefore, an important part of the data acquisition is a trigger system which differentiates between events which are "interesting" and those which will be discarded. CDF has a three-level, buffered trigger system, Level 1 which is a hardware trigger and synchronous with the 396 ns bunch crossing, Level 2 which is a programmable hardware trigger, and Level 3 which is a software trigger. Data is written to tape at a rate of 50-57 Hz.

A detailed description of the detector components can be found in the CDF Technical Design Report [29]. The systems used in this analysis are described briefly in the following sections.

3.2.1 The CDF Coordinate System

The origin of the CDF coordinate system is taken to be at the geometrical center of the detector. The positive z-axis is defined to be along the beam line in the direction that the protons are traveling (west to east). The positive y-axis is perpendicular to the beam line, pointing upward. The positive x-axis is perpendicular to the z- and y-axes defined to satisfy a right-handed coordinate system.

Since the detector is approximately cylindrical, cylindrical coordinates are used to describe the detector and the particles traversing it. The azimuthal angle, ϕ, is zero at the x-axis and increases in the counter clockwise direction when facing the positive z direction (east). The coordinate R, is the radial distance perpendicular to the beam

line. The polar angle, θ, is zero along the z-axis and is 90° at the y-axis. This angle, however, is unsuitable for physics analyses because $d\theta$ is not Lorentz invariant and the number of particles per unit angle ($dN/d\theta$) will not be the same for particles traveling at different velocity. Instead, we define *rapidity*,

$$\mathrm{Y} = \frac{1}{2} ln \frac{E + p_z}{E - p_z} \tag{3.2}$$

which preserves the density of particles per unit rapidity (dN/dY), i.e. dY is Lorentz invariant. In the limit where the momentum of the particle is much bigger than its mass, the rapidity is approximately equal to the *pseudo-rapidity*, defined as

$$\eta = -ln(tan(\frac{\theta}{2})) \tag{3.3}$$

3.2.2 The CDF Silicon Systems

The silicon detector lies at the center of CDF and is composed of three concentric systems, Layer 00 (L00), a single sided silicon layer mounted directly on the beam pipe, the Silicon Vertex Detector (SVX II), a five-layer double sided silicon detector varying in radius from 2.5 cm to 10.6 cm, and the Intermediate Silicon Layers (ISL), two layers of double sided silicon at radii of 20 cm and 28 cm. All three detectors have a twelve-fold symmetry. Together, they detect particles with pseudo-rapidity less than 2.0 and provide excellent position resolution imperative for heavy flavor identification by detecting secondary vertices. The silicon detector is operating with ∼92.5% of its ladders and ∼86% of them collect "good" data, i.e. with error rates below the acceptable threshold.

How Silicon Detectors Work

Silicon detectors are reversed-biased p-n junctions (diodes). Figure 3.5 shows the detection in a silicon microstrip system as a charged particle passes through. As the ionizing particle passes, it produces electron-hole pairs. The number of pairs is proportional to the amount of energy lost by the particle. Due to the electric field in the wafer produced by the voltage applied, the electrons drift to the anode and holes to the cathode. The charge is collected at the microstrips. Since charge is often deposited on two or more adjacent strips, a cluster is formed from these strips and a weighted average is used to find the centroid to determine the position of the hit.

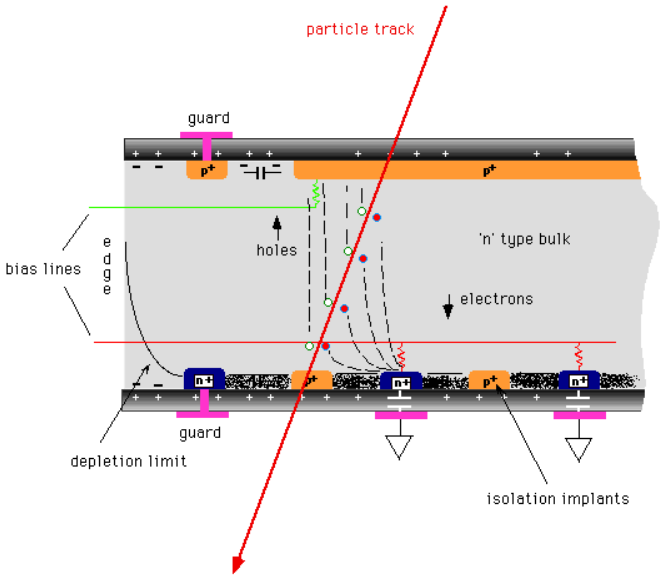


Figure 3.5: Detection in a silicon detector. Silicon detectors are reversed-biased p-n junctions. A passing ionizing particle produces electron-hole pairs which drift to the anode and cathode respectively.

Silicon detectors are used for several reasons. They have excellent spatial resolution, 5–15 μm. Only 3.6 eV of energy is needed to produce an electron-hole pair in

silicon, compared to 30 eV necessary in a gas. The response time is very fast, ∼5 ns. In addition, the bias voltages needed for microstrip detectors is less than 100 V. The challenge for silicon detectors is to have low noise and very fast electronics. Furthermore, due to the density of silicon, the detector must be kept thin to reduce multiple Coulomb scatterings. However, the amplitude of the signal in silicon is proportional to the thickness (no charge amplification in silicon detectors) so a compromise must be made. The thickness for optimal detection is ∼300 μm.

Layer 00 (L00)

Layer 00 (L00) [30] is a single-sided silicon detector mounted directly on the beam pipe (r = 1.2 cm). The detector consists of two overlapping hexagons at radii of 1.35 cm and 1.62 cm with sensors which are 7.84 cm in length. The implantation pitch of the silicon strips is 25 μm and the read-out pitch is 50 μm (alternate strips are read-out). The hit resolution is ∼6 μm. L00 has 13824 channels.

Due to its proximity to the beam, L00 enhances the impact parameter resolution of the Silicon Vertex Detector, especially for low momentum tracks. Improved impact parameter resolution is essential for identifying b quarks jets from the presence of secondary decay vertices. Figure 3.6 shows the impact parameter resolution as a function of track p_T with and without L00. There is a large improvement for tracks with p_T between 0.5 and 2.0 GeV.

Silicon Vertex Detector (SVX II)

The Silicon Vertex Detector (SVX II) [31] is the main detector of the silicon system. It consists of three cylindrical barrels aligned along the z-direction, separated by

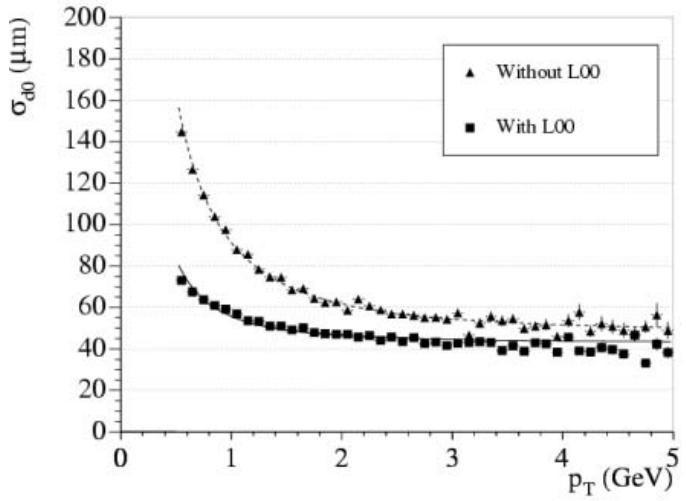


Figure 3.6: Impact parameter resolution with respect to p_T, with and without Layer 00.

gaps of 2.0 cm, as shown in Figure 3.7. Each barrel, mounted between two beryllium bulkheads, is 30 cm long, is divided into twelve azimuthal wedges of 30°, and has five double-sided silicon microstrip layers, ranging in radius from 2.54 cm to 10.64 cm. Each layer has 12 layers 29.0 cm in length. The barrel and bulkhead structures of SVX II are shown in Figure 3.7. All five layers have $r-\phi$ (axial) strips running lengthwise on the $p-n$ junction side of the layer. The first, second, and fourth layers, known as the "90° stereo" layers, have strips running crosswise on the ohmic contact side (n-side). The third and fifth layers, known as "small angle stereo" layers, have strips that run lengthwise at a 1.2° angle on the ohmic contact side. The strip pitch for the SVX II detector is 60 – 65 μm for the $r-\phi$ and small angle stereo strips, and 125 – 141 μm for the 90° strips. The SVX II has a total of 405,504 channels and is read out in approximately 10 μs. The hit resolution for SVX II is 12 μm (axial). The SVX II has impact parameter resolutions of $\sigma_\phi < 30$ μm and $\sigma_z < 60$ μm for central

high momentum tracks.

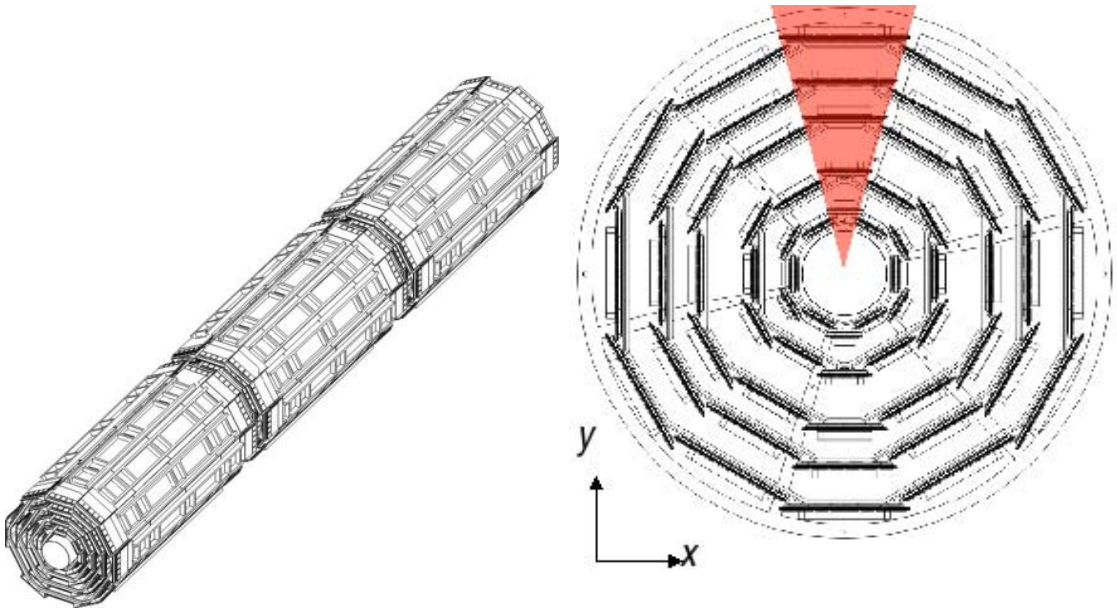


Figure 3.7: The SVX. The left figure shows the barrel structure and the right figure shows the bulkhead structure. The highlighted area is a wedge.

Intermediate Silicon Layers (ISL)

The Intermediate Silicon Layers (ISL) [32] constitute the outer-most system of the CDF silicon detector and provide accurate linking information to the Central Outer Tracker (COT) and, together with the SVX, allows for 3D tracking in the plug region, $1.0 \leq |\eta| \leq 2.0$, where there is only partial tracking information from the COT. In the central region, there is one double-sided silicon layer at a radius of 22 cm. In the plug region, there are two layers of silicon at radii of 20 cm and 28 cm. The structure of the ISL is shown in Figure 3.8. The ISL layers have small angle stereo design on the ohmic side and have strip pitches of 55 μm on the $r - \phi$ side and 73 μm on the stereo side. The ISL has a total of 268,800 channels. The hit resolution in the ISL is

16 μm (axial).

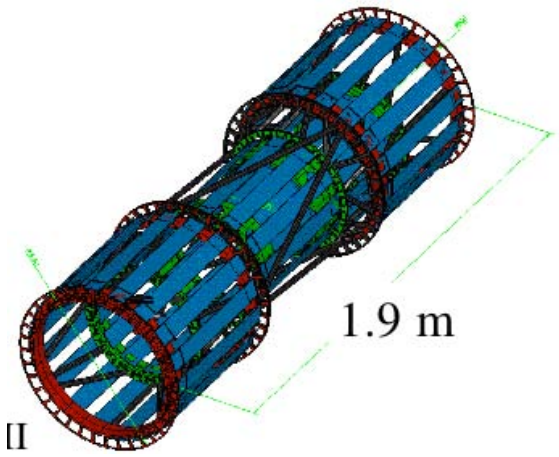


Figure 3.8: The ISL barrel structure.

3.2.3 The Central Outer Tracker (COT)

Tracking at large radii is performed by the Central Outer Tracker, COT, a multi-wire, open cell drift chamber [33]. The COT extends in radius from r = 44 cm to r = 132 cm, has an active length of 310 cm, and has full coverage in the $|\eta| < 1.0$ region and partial coverage up to $|\eta| < 2.0$. It has eight "superlayers", each comprised of twelve sense wires, alternating with shaper wires every 3.8 mm, providing a total of 96 measurements for each track. There are a total of 2,520 drift cells and 30,240 readout channels. The sense and shaper wires, gold plated tungsten with with a diameter of 40 μm, are sandwiched between gold-on-mylar (6.4 μm thick mylar, with ~350 A gold per side) cathode planes (at ground) separated by ~2 cm. Four of the superlayers provide $r - \phi$ measurements and four provide 2° stereo measurements. The stereo and axial superlayers alternate, starting with a stereo layer. The cells are tilted at a

35° angle to allow for ExB effects. Figure 3.9 shows the COT configuration as viewed from the east side endplate.

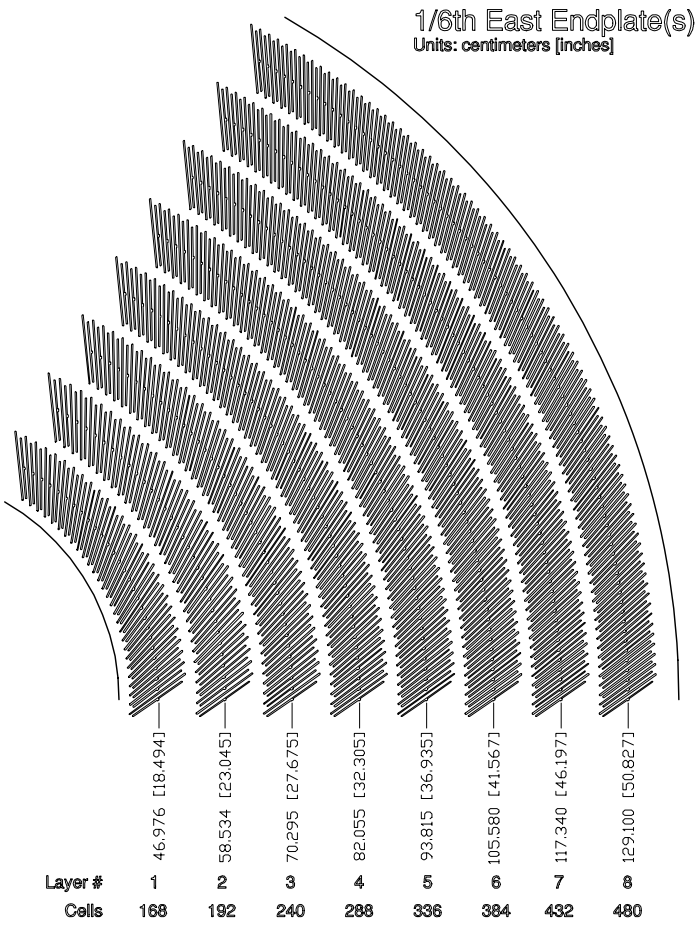


Figure 3.9: COT configuration as viewed from the east side endplate.

The COT is filled with an Argon-Ethane-Isopropyl Alcohol gas mixture with the ratio 49.5:49.5:1.0. It has a drift speed of ∼50 μm/ns and a maximum drift time of ∼100 ns. This is designed to reduce pile-up from the 396 ns beam crossings. The COT spatial hit resolution is $\sim$ 180 μm. The COT has a p_T resolution of $\sigma_{p_T}/p_T \sim$ 0.15% x p_T.

How Drift Chambers Work

Drift chambers consist of a cathode plane, and an anode (sense) wire immersed in a gas. As a charge particle enters the gas in the E field set up between the anode and the cathode, it ionizes the gas producing electron-ion pairs. The ionization energy needed is ~30 eV per pair. The electrons then drift toward the anode wire at constant velocity, until they reach the wire. Within a few radii of the sense wire, the E field becomes very strong and an avalanche multiplication of charges by electron-atom collisions occurs. This avalanche causes an amplification of $\sim 10^4$ and allows the signal to be detected. The drift time, the time difference between a start time t_0 (e.g. collision time) and the arrival time of electrons at the sense wire gives the distance from where the particle passed to the sense wire. Since a drift time is measured, there is a left-right ambiguity in the passage of the particle and layers of drift chamber cells are staggered from each other to determine directionality. Magnetic fields applied to measure momentum through curvature interact with the electric fields in the drift chamber and introduce a drift angle, the Lorentz angle. The solution is to tilt drift cells at the same angle to compensate.

The choice of gas for drift chambers is important. In general, several gases are mixed, one drift gas to be ionized by the particle and one or more quenchers for absorbing excess UV photons which could cause an avalanche in the drift region. The desired properties for drift gases are drift velocity independent of E field, low working voltage, high gain, and small diffusion coefficient for the electrons traveling through it. Argon is widely used in drift chambers for this purpose. For quenchers, it is desirable to have heavy organic molecules with many degrees of freedom. At CDF,

ethane is used as a quencher with a small admixture of alcohol.

3.2.4 Calorimetry

The CDF calorimeters lie outside the tracking systems and the solenoid. They cover a complete 2π in azimuth and extend up to 3.6 in pseudorapidity ($|\eta| < 3.6$). The pseudorapidity range is covered by two separate detectors, the central calorimeter which covers up to $|\eta| < 1.1$ and the plug calorimeter which covers the range $1.1 < |\eta| < 3.6$. Each calorimeter has an electromagnetic calorimeter and a hadronic calorimeter. All calorimeters at CDF are sampling calorimeters which have alternating layers of active scintillator and absorber material. The absorber for the electromagnetic calorimeters is lead and the absorber for hadronic calorimeters is iron. Additionally, a shower maximum detector is located six radiation lengths within each electromagnetic calorimeter, corresponding to the approximate depth of the shower maximum of electromagnetic showers. The shower shape helps in distinguishing electrons from pions. The calorimeters have projective tower geometry with angular segments in η and ϕ which point back to the origin of the CDF co-ordinate system. Figure 3.10 shows the segmentation of the CDF calorimeters.

How Calorimeters Work

Calorimeters sample the energy of particles. This is accomplished by measuring the energy as the particle has a series of interactions with the detector material and loosing energy until it is finally absorbed. The cascade of interactions is known as a shower. Segmenting the calorimeter into cells allows for the determination of the

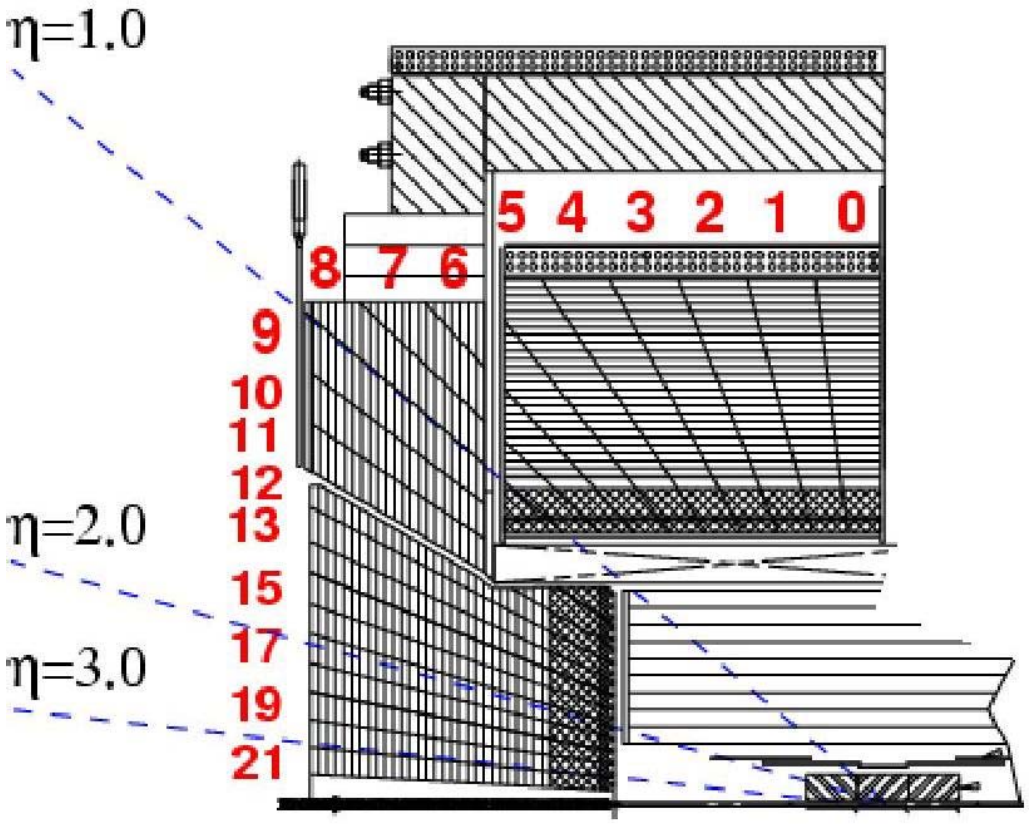


Figure 3.10: The segmentation of the CDF calorimeters.

original direction of the particle and for determining the shape of the shower which aids in particle identification. Calorimeters usually have an electromagnetic section and a hadronic section. In general, electromagnetic particles, photons and electrons, have showers entirely contained in the electromagnetic calorimeter while hadronic particles, particles consisting of quarks such as pions and protons, have showers which develop within the electromagnetic calorimeter but deposit most of their energy in the hadronic calorimeter.

Calorimeters need to both cause the showers and sample them. In homogeneous

shower counters, both functions are performed by one type of material. In heterogeneous shower counters (sampling calorimeters), such as those used at CDF, there are two types of material, a passive absorber material (e.g. lead, iron, uranium) sandwiched with an active signal readout material (e.g. liquid or solid scintillators, or proportional counters). In such calorimeters, only the portion of energy deposited in the active material is measured and the total energy extrapolated from the sampled energy.

The energy resolution in a calorimeter is dependent on both the detector performance and the statistical nature of the shower development process. Therefore, the resolution has two parts, a constant part which is the instrumental and calibration limit, and a variable part (resulting from the fluctuations in shower development) which is dependent upon energy in the shower and improves as the energy increases. The general formula is:

$$\frac{\sigma}{E} = \frac{\alpha}{\sqrt{E}} \oplus \beta \oplus \frac{\gamma}{E} \tag{3.4}$$

where σ is the standard deviation of the energy measurement, E is the energy of the particle, and α, β, and γ are constants which depend on the detector, e.g. the material and thickness of the active and absorber components.

Central Calorimeters

The central and wall calorimeters [34] cover the pseudorapidity range $|\eta| < 1.1$ and make up the components central electromagnetic calorimeter (CEM), central hadronic calorimeter (CHA), and wall hadronic calorimeter (WHA). All the central calorimeters are segmented into towers with granularity $\delta\phi = 15^\circ$ and $\delta\eta = 0.1$. While

the CEM covers the full $|\eta| < 1.1$ range, the hadronic coverage is split between the CHA which covers $|\eta| < 0.7$ and the WHA which covers $0.7 < |\eta| < 1.1$. The CEM and CHA for a barrel which extends in radius from 1.3 m to 2.2 m.

The CEM extends in radius from 173 cm to 208 cm and is composed of 31 layers of 5 mm thick polystyrene scintillator (active material) sandwiched with 0.32 cm thick lead sheets (absorber), which corresponds to 18 radiation lengths. The energy resolution for an electron in the CEM is

$$\frac{\sigma}{E} = \frac{13.5\%}{\sqrt{E_T}} \oplus 2\%$$

where the energy, E, is measured in GeV.

There is a set of proportional chambers between the solenoid and the CEM used to distinguish between photons which are likely to interact with the solenoid coil and pions which have a smaller chance of interacting. This central preradiator detector (CPR) [35] is made up of sense wires separated by 2.2 cm in Argon-Ethane gas. For each 15° CEM wedge, there are 2 CPR chambers in η.

The central shower max detector, CES, is made up of proportional chambers with strip and wire readout and is located at a radius of 184 cm (six radiation lengths within the CEM). The CES can distinguish between electrons and pions.

The CHA modules are in the same wedges as the CEM modules. They are composed of layers of 10 mm of PMMA naphthalene scintillator sandwiched with 1 in of iron, corresponding to 4.5 radiation lengths. The energy resolution of the CHA for a

single pion is

$$\frac{\sigma}{E} = \frac{75\%}{\sqrt{E_T}} \oplus 3\%$$

where the energy, E, is also measured in GeV. The WHA modules are the same as CHA modules and are mounted on the solenoid flux return and cover the region of 30° to 45° in azimuth.

Plug Calorimeters

The plug calorimeters [36] cover $1.1 < |\eta| < 3.6$. Like its central counterpart, the plug calorimeter has electromagnetic and hadronic calorimeters. Figure 3.11 shows a section of the plug calorimeter.

The plug electromagnetic calorimeter, PEM, is composed of 23 layers of 4 mm of scintillator sandwiched with 4.5 mm of lead, comprising a total of 21 radiation lengths. The energy resolution of PEM for an electron is

$$\frac{\sigma}{E} = \frac{16\%}{\sqrt{E_T}} \oplus 1\%$$

where energy is measured in GeV. The first layer is made of 10 mm of scintillator and used as a preshower detector, PPR which is similar to the CPR. At a depth of six radiation lengths within the PEM, there is a shower maximum detector, PES. It is composed of two layers of 5 mm wide scintillator strips which overlap at a 45° angle. There are 16 sectors in ϕ, each covering a 45° angle and two sectors in η.

The plug hadronic calorimeter, PHA, is composed of 23 layers of 6 mm scintillator sandwiched with 2 in iron, comprising 7 radiation lengths. The energy resolution of

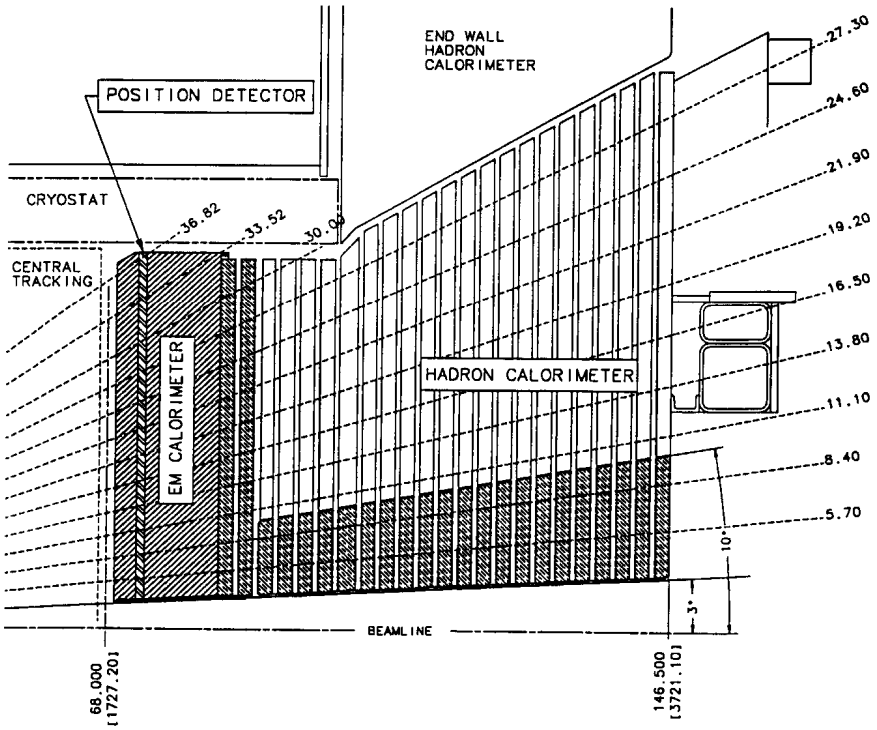


Figure 3.11: The schematic of the CDF plug calorimeters.

PHA for a pion is

$$\frac{\sigma}{E} = \frac{80\%}{\sqrt{E_T}} \oplus 5\%$$

where energy is also measured in GeV.

3.2.5 Muon Systems

Beyond the calorimeters and steel shielding lie the CDF Muon systems, made up of single-wire drift chambers, filled with gas which is an admixture of argon, ethane, and isopropyl alcohol in the proportion 49.5:49.5:1.0. The muon detection at CDF

is performed by four different systems which cover almost 2π in ϕ and up to $|\eta| < 2.0$. The central muon detector (CMU) and the central muon upgrade cover up to $|\eta| < 0.6$, the central muon extension (CMX) covers the region $0.6 < |\eta| < 1.0$, and the barrel muon detector covers the region $1.0 < |\eta| < 2.0$. For this analysis only the muons detected by the CMU and the CMP together (CMUP muons) and by the CMX are used due to the fact that triggering on BMU muons was available for only a fraction of the data and also because tracking is limited beyond $|\eta|$ of 1.0. The η and ϕ coverage of the CMU, CMP, and CMX is shown in Figure 3.12.The CMU, CMP, and CMX systems are described in more detail below.

Central Muon Detector (CMU)

The central muon detector (CMU) [37] is the inner-most muon system and is embedded in the central calorimeter wedges at their outer radius, at 350 cm, as shown in Figure 3.13. The CMU detects muons with $p_T > 1.5$ GeV/c and $|\eta| < 0.6$. There is a gap of 18 cm at $\eta = 0$ to make room for the central calorimeter support structure and high voltage fanout at this location. The CMU is cylindrical and divided into 12.6° wedges in ϕ. Each wedge is composed of three sections of 16 rectangular drift chambers, four wide and four high.

Each drift chamber has a sense wire, parallel to the beam (z-direction), which is 50 μm in diameter. The wires from adjacent chambers are ganged together and the pairs are offset from each other by 2 mm to resolve left-right ambiguity. The z position of hits is determined from the charge division along the wire and the ϕ position is determined by the drift time of the hit. Hits are connected together to

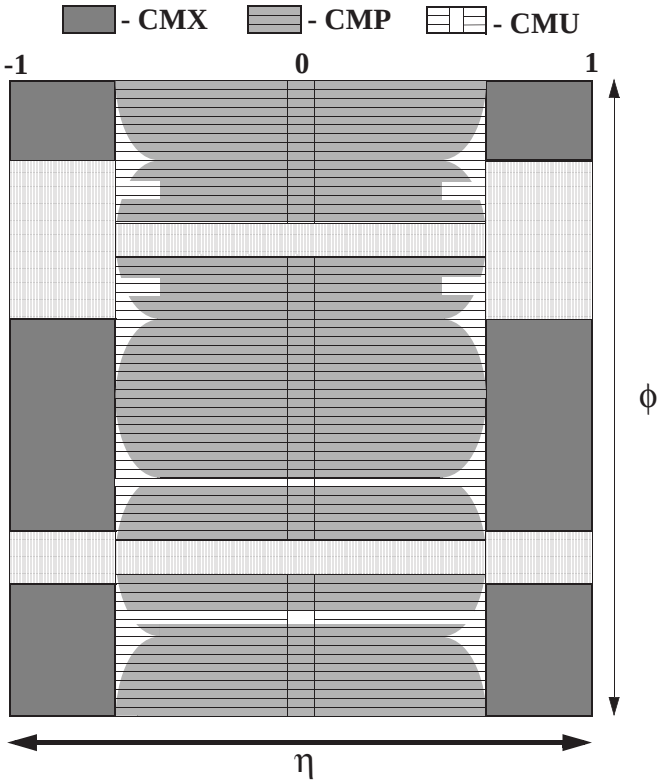


Figure 3.12: The η and ϕ coverage of the muon systems.

form a muon stub which can be matched to a COT track. The stub spatial resolution of CMU is 1.2 mm in z and 0.25 mm in ϕ. The n and n+2 chambers are radially aligned with x=y=0 and the comparison of the drift time from these wires gives a crude momentum measurement which is used in the level 1 muon trigger.

The central calorimeter provides shielding for the CMU of about 5.5 interaction lengths for pions. However, there is still a significant fraction of pions ($\sim$ 0.5%) which can penetrate the calorimeter and produce stubs in the CMU, known as punch throughs. To reduce these punch throughs, only muons detected by both the CMU

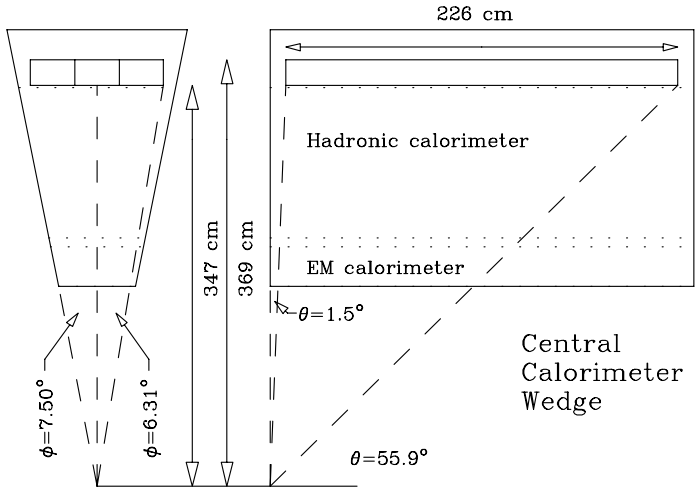


Figure 3.13: Location of the CMU detector in the central calorimeter wedge. On the left is the azimuthal cross section and on the right is the polar cross section.

and CMP are used.

Central Muon Upgrade (CMP)

The central muon upgrade (CMP) is located beyond an additional 60 cm of steel beyond the CMU. The extra shielding provides an additional 3.5 radiation lengths and reduces punch throughs by a factor of 20 as well as reducing low p_T muons from in-flight decays of pions and kaons. It detects muons with $p_T > 3.0$ GeV/c and $|\eta| < 0.6$. The CMP is made up of four walls of rectangular chambers, four layers thick and staggered by one half cell per layer. Each CMP drift chamber has a sense wire in the center and is 640 cm long, 15 cm wide and 2 cm high. The CMP provides r-ϕ information but no z information. The η coverage of the CMP is ϕ dependent because the chambers have fixed length. The CMU has a single layer of rectangular scintillator tiles beyond the outer most layer of the CMP, known as the central scintillator upgrade (CSP).

Central Muon Extension (CMX)

The central muon extension (CMX) is arguably "the most beautiful of the muon detectors although thinking about the conical geometry has driven some to drink" [38]. The CMX flanks CDF on the east and west, and has a conical structure with an opening angle of 41.4° and center of radius 10 m away from the interaction region along the beam direction. The CMX covers the region $0.6 < |\eta| < 1.0$ and nearly 2π in azimuth. It has three distinct sections, the CMX arches, two on each side which covers 240° in azimuth, the miniskirts which cover the bottom of the detector, 230° to 310° in azimuth, extending beneath the nominal detector floor, and keystone on the west side which covers 75° to 105° in azimuth. There is no keystone region on the east side because its place is occupied by the cryogenics input to the solenoid. The CMX has a hit resolution of 280 μm.

The CMX chambers are the same as the CMP chambers except that they are 183 cm in length (instead of 640 cm). They are arranged in groups of 48 chambers making up 15° wedges. Each wedge has eight layers of six chambers each. The eight layers are grouped into four pairs aligned radially and offset from each other by half a cell. The conical structure of the arches and keystone allow for the CMX z coordinate to be independent of ϕ but this is not true for the miniskirt section which does not have curvature along its width due to constraints for fitting in the available space.

The CMX has a corresponding scintillator system, the central scintillator extension (CSX). The arches and keystone stacks of chambers are sandwiched between rectangular scintillator tiles. The miniskirts only have one layer of scintillator counters on the side closer to the beam. The scintillator system provides timing information in

conjunction with the chambers and is used for rejecting accidental muons not coming from the interaction region. Information from only one layer of the scintillators on the arches is necessary to time-in passing muons.

3.2.6 Trigger Systems

At hadron colliders, the collision cross sections are many orders of magnitude higher than the rate at which data can be stored. At CDF, the beams cross once every 396 ns beam crossing. This means that the collision rate it 2.5 MHz while the rate of data written to tape is limited to 50-75 Hz. The limit in this rate is due to both the volume of data, ∼200 kB/event, and the time it takes to read out an entire event, ∼2 ms. In addition most of the collisions produce events which are not the high p_T physics which is of interest, which have cross sections many orders of magnitude smaller. For example, at $\sqrt{s} = 1.96$ TeV, the total hadronic cross section (elastic, inelastic, and diffractive) is ∼75 mb while the $t\bar{t}$ production cross section is ∼7.6 pb. Therefore, it is imperative to have a *trigger* which makes decisions on which events should be stored for physics analyses.

CDF has a three level nearly deadtimeless trigger where the output rate of one level is matched to the input rate of the next. The data rate is reduced at each level based on information available from the event. Level 1 and Level 2 use custom hardware to make decisions based on information from some subsystems while the Level 3 software trigger uses the full event readout for deciding whether to keep or discard the event. The data flow in the CDF trigger system is shown in Figure 3.14. The requirements that must be fulfilled to pass Level 1, Level 2, and Level 3 triggers

are known as the *trigger path*. CDF has about one hundred trigger paths through which an event may qualify to be written to storage. A physicist chooses the trigger path(s) which will have the events which correspond to the experimental signature of the physics he/she wishes to study.

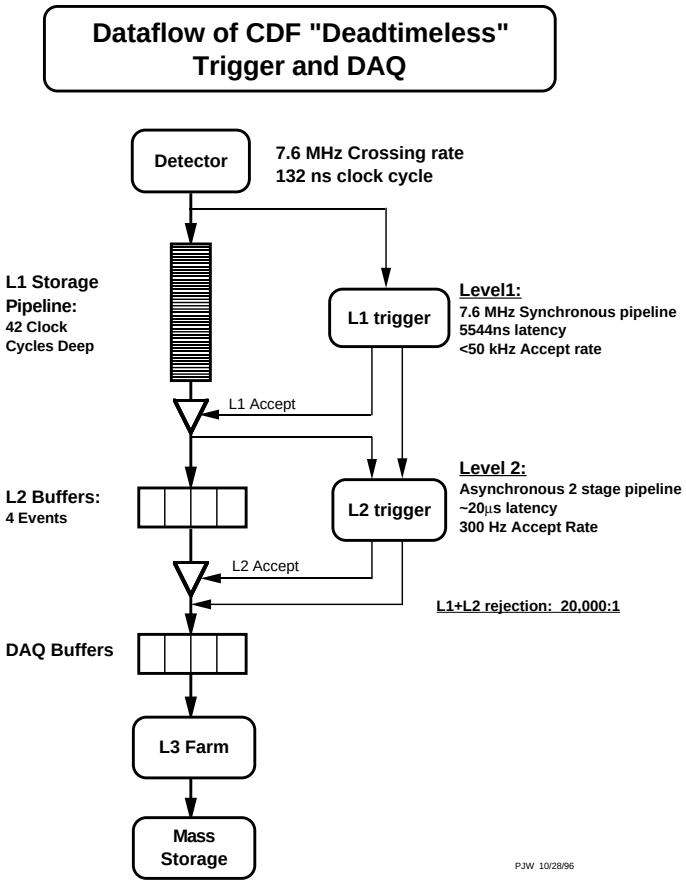


Figure 3.14: The data flow in the CDF trigger system.

Level 1

The Level 1 trigger rejects the majority of events. It is a synchronous trigger which use custom hardware to accept an event at every 396 ns bunch crossing and

makes decisions every 5.5 μs. This is accomplished by a pipeline with buffers which have 14 slots, each corresponding to one bunch crossing. At every bunch crossing, the event moves up one slot in the pipeline and after the fourteenth slot, a decision is made whether to accept or discard it. The 2.5 MHz collision rate is reduced to 50 kHz by Level 1.

The Level 1 trigger makes decisions based on partial information available from the COT, calorimeters, and muon systems. It has three parallel processing streams, L1CAL, L1MUON, and L1TRACK, which feed the Global Level 1 decision unit. The tracks are sent to the calorimeter and muon streams as well since electrons and muons need tracks pointed at the calorimeter or muon detector. Each stream has trigger *primitives*, coarse physics objects such as tracks, electrons, and muons on which the trigger decision is based. The decision can come from one of 64 predetermined binary combinations of the various streams.

Level 2

Events which pass the Level 1 trigger are passed on to the asynchronous Level 2 trigger which has two stages each taking 10 - 20 μs. There are four buffers accepting events and a buffer remains full until a Level 2 decision has been made for the event. If all four buffers are full when the next event passes the Level 1 trigger, the event is lost. The Level 2 trigger forms better primitives using more and improved data and has programmable processors which combines the primitive inputs to make more sophisticated decisions. The trigger has evolved during Run II due to its flexibility in programming.

Track information is improved by the Silicon Vertex Tracker (SVT). The XFT tracks are extrapolated into the silicon detector and information from silicon hits on the path formed by the extrapolated XFT tracks are added. The SVT requires four out of the five SVX layers to have hits. This provides better p_T and ϕ_0 information compared with Level 1. More importantly, adding silicon information allows for the determination of the impact parameter, d_0, of the track, allowing the SVT to trigger on tracks with large impact parameter and collecting events which have displaced secondary vertices. The SVT has an average processing time of 19 μs and has an impact parameter resolution of $\sim$35 μm (47 μm if convoluted with the beam profile of $\sim$30 μm) for tracks with $p_T > 2$ GeV/c.

The calorimeter information is also improved at Level 2. Individual trigger towers are clustered using seed and shoulder thresholds to form jets. Information from the CES is used to better distinguish between electrons and photons. In addition, matching track to the CES which has a finer spatial resolution than the towers reduces combinatorics for electron triggers. A diagram summarizing the Level 1 and Level 2 triggers is shown in Figure 3.15.

Level 3

The Level 3 trigger [39] accepts events which pass the Level 2 trigger. The data from the DAQ is polled after a Level 2 accept and passed on to the *event builder* which formats the data from the same beam crossing into banks. The data is then sent to an array of $\sim$300 dual-processor Linux PCs known as the Level 3 farm. The events are distributed among the nodes and each node is equipped with the current

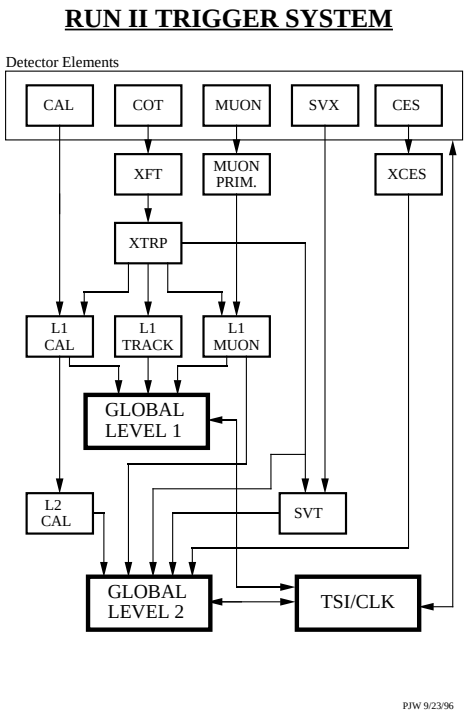


Figure 3.15: Summary of the CDF Level 1 and Level 2 trigger data flow.

reconstruction algorithms and detector calibration. The nodes perform an almost offline quality full reconstruction of the event to decide whether to accept the event for permanent storage or not. The events that Level 3 accepts are sorted according to the trigger which accepted them and written to tape. The stored events are characterized by their Level 1 trigger sequence number and the run number. A run is a continuous period of data taking with fairly constant detector and accelerator conditions. If the conditions change, e.g. a particular detector is removed from data taking, the current run is ended and a new one is started with the different conditions noted. The run-by-run data taking conditions are reflected in the Monte Carlo simulations to better represent data.

3.2.7 Luminosity Measurement

The luminosity measurement at CDF is performed by two Cerenkov Luminosity Counter (CLC) modules [40]. These modules are placed in 3° holes in the CDF end plug calorimeters in the forward and backward regions of the detector at small angles, covering the pseudorapidity region $3.7 < |\eta| < 4.7$. They replaced the scintillator systems which were traditionally used.

Each CLC module is composed of 48 thin, long, conical Cerenkov counters, arranged in three concentric layers centered around the beam, each consisting of 16 counters pointed toward the center of the interaction region. The counters in the outer two layers are 180 cm long and the ones in the innermost layer are 110 cm long. The counters have cross sections ranging between two and six centimeters in diameter. The counters are filled with isobutane and nominally operate at atmospheric pressure. The momentum threshold for emission of Cerenkov light in these counters is 9.3 MeV/c for electrons and 2.6 GeV/c for pions. The light is emitted at a Cerenkov angle, θ_c, of 3.1°.

The instantaneous luminosity in the CLC is given by the equation

$$\mathcal{L} = \frac{\mu_\alpha \cdot f_{BC}}{\sigma_\alpha(\mu)} \tag{3.5}$$

where μ_α is the number of interactions that can be detected by the CLC, f_{BC} is the frequency of beam crossings, and $\sigma_\alpha(\mu)$ is the cross section for these interactions. This cross section is the product of the total inelastic cross section, σ_{in}, and the detection efficiency, ϵ_α, i.e. $\sigma_\alpha = \sigma_{in} \cdot \epsilon_{alpha}$. The efficiency is, itself, the product of two

components, $\epsilon_\alpha = \varepsilon_\alpha \cdot \delta_\alpha$, where ε_α is the efficiency to detect a single $p\bar{p}$ interaction and δ_α accounts for the non-linearity of the CLC cross section from multiple interactions. The integrated luminosity is obtained by integrating the instantaneous luminosity with time. The systematic uncertainties for the luminosity measurement result from the uncertainties in the inelastic $p\bar{p}$ cross section, 3%, the CLC single interaction detection efficiency, 2%, and the non-linearity due to multiple interactions, 2%, giving a total systematic uncertainty of $\sim$6% [41].

Chapter 4

Data Reconstruction and Event Simulation

Any sufficiently advanced technology is indistinguishable from magic.
Arthur C. Clarke (from "Profiles of the Future")

The events which passed the Level 3 trigger and were stored for further analysis are reprocessed using *offline* reconstruction software. While the *online* systems are designed for speed, offline systems perform a much more careful analysis of the data and have access to better calibrations and alignment information. Approximately one third of the data used in this analysis was processed with version 5.3.1 of the offline software and the rest were processed with version 6.1.1. In addition, CDF also uses Monte Carlo event generation and simulation to help interpret the results from data. The event generation and simulation schemes will be described in this chapter.

Offline reconstruction software takes advantage of the fact that different particles interact differently with the various components of the CDF detector, leaving unique

signatures for their identification. Charged particles, such as electrons and muons, leave tracks in the silicon systems and the COT while neutral particles, such as photons and neutral pions, pass through the tracking systems undetected. Most particles deposit a majority of their energy in the calorimeters, the ones interacting mostly via the electroweak interaction, electrons and photons, depositing most of their energy in the EM calorimeter, and those interacting mostly via the strong interactions, hadrons, depositing most of their energy in the hadronic calorimeter. Since quarks cannot exist individually, when a bare quark is produced, it "hadronizes", pulling quark-antiquark pairs out of the vacuum and forming hadrons, leaving a jet, a collection of collimated tracks and deposition of energy in the calorimeters. Muons, due to their large mass, penetrate the calorimeters, leaving only minimum ionization energy in them. They are identified by the "stubs", short tracks, they leave in the muon chambers. The identification of particles in the CDF detector is summarized in Figure 4.1. The reconstruction algorithms for the physics objects used in this analysis, electrons, muons, jets, and heavy flavor identification via secondary vertex reconstruction, are described in further detail in the following sections.

4.1 Track Reconstruction

Tracks map the passage of charged particles through the CDF detector and are reconstructed using information from the COT and silicon detectors. Tracks are very important for this analysis because they form the basis for all other physics objects used, electrons, muons, and also for finding secondary vertices within jets.

Due to the presence of the magnetic field, particles curve in the CDF detector and

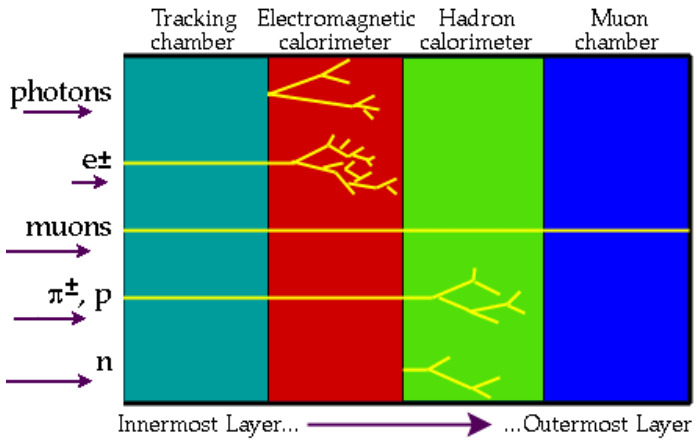


Figure 4.1: Summary of the signature left by the different particles in the CDF detector.

leave helical tracks and trace out circles in the $r-\phi$ plane. Tracks are reconstructed from *hits*. Tracking is performed in two overlapping steps; *pattern recognition*, assigning hits into tracks and *track fitting* using the assigned hits to determine the five spatial helix parameters, namely:

- *curvature*: the curvature of the circle traced by the track in the $r-\phi$ plane which is proportional to the transverse momentum, p_T, of the track,
- $cot(\theta)$: the cotangent of the polar angle θ which indicates the motion of the particle in the z-direction, along the beam axis,
- *impact parameter* (d_0): the distance of closest approach of the track to the interaction point in the $r-\phi$ plane,
- z_0: the distance of closest approach of the track to the interaction point in the z-direction, and
- ϕ_0: the azimuthal direction of the track at the point of closest approach.

Pattern recognition is performed starting at large radii where the density of hits is lower. Then new hits are added within a range of the preliminary track formed. After all the hits are attached to the track, the track parameters are refit for maximal precision and accuracy.

4.1.1 Pattern Recognition in the COT

Track finding in the COT begins with a search for hits in three consecutive layers of within a superlayer. Using a simple drift model, both possible drift direction of each hit is considered and the hits are fit to a straight line pointing back to the nominal origin of the CDF detector. Extrapolating from this seed segment, hits are searched for in other layers of the superlayer. For each layer, a hit is added to the segment or the layer is counted as a "missed" layer. The segments are required to have more than five hits to be accepted. There is a minimum p_T threshold of approximately 355 MeV/c due to the fact that the seed segment direction is required to be less than about 50° from the radial [42].

Once the segments have been found, two different algorithms are used to find track candidates, in order to maximize the efficiency for pattern recognition. *Segment linking* [42] takes segments in different axial layers with good position and angle matches and forms axial tracks. First, the segments in the outer two axial layers are linked. Then track parameters are refit. Finally, the segments in the inner axial layers are linked to the track or unlinked segments from outer superlayers. *Histogram linking* [43], begins with a single axial segment and the curvature of the track to which the segment belongs, known as a "telescope". Then a histogram is constructed for

each layer which could contain a hit for the track. The histogram is filled with the likelihood that the track could pass through the hit, as a function of the shift from the value of the nominal curvature. When the histograms are summed, a peak results close to zero and the correct hits for the track are those within the peak. The tracks found by both algorithms are kept and stereo information is added to each track. This is accomplished by either linking stereo segments to the axial tracks or, if that fails, adding individual stereo hits. Then hit positions are corrected using a better drift model and the tracks are refit. Since the two algorithms produce duplicates tracks, the track collection is pruned and only the best tracks are kept. Histogram linking has a bias for high p_T tracks and, therefore, only segment linked tracks are kept for tracks with $p_T < 400$ MeV/c. For other tracks, if the track from only one algorithm has at least 20 axial and 20 stereo hits with at least 3 axial and 3 stereo superlayers with 6 hits each, that track is kept. If the tracks from both algorithms pass these requirements, the segment linked track is kept. If the tracks from neither algorithm passes these requirements, the one with the most COT hits is kept [42].

4.1.2 Outside-In Silicon Hit Attachment

Once COT tracks have been found, silicon hits are attached to these tracks to improve the resolution on the track parameters, especially the track impact parameter whose resolution improves by more than an order of magnitude. In the "Outside-In" tracking algorithm [44], the COT track is extrapolated into the silicon and hits within a window of where the COT intersected the silicon wafer are added to the track, beginning with the outer-most silicon layer. The hits on the $r - \phi$ layers are added

first and then the hits on the 90° z layers are added. A new track candidate with a new error matrix is created for each additional hit and this new track is extrapolated to the next silicon layer. When all the hits have been added, the track candidates are ranked first by the number of added silicon hits and second by the fit χ^2. Only the best candidate based on this ranking is kept as the final track.

4.1.3 Phoenix Tracking

The COT provides full tracking coverage out to $|\eta| \sim 1.0$ while silicon coverage extends to $|\eta| \sim 2.0$. Therefore an alternate tracking method seeded with calorimeter information, called the Phoenix algorithm [45], has been developed for forward electrons. In this algorithm, a calorimeter object which passes electron requirements is chosen to seed the outside-in silicon tracking, replacing the COT. A primary vertex, a position in the calorimeter, and an energy measurement in the calorimeter are used to create two seed tracks, one for each possible sign of the electron. Silicon hits are added to these seeds and if hits are only found for one seed, that seed is kept. If silicon hits are found for both seeds the one with the best χ^2 per D.O.F. is kept and the charge of that seed is assigned as the charge of the electron. The reconstruction of phoenix electrons is further described in Section 4.3.2.

4.2 Muon Reconstruction

Muons are used in this analysis to reconstruct Z bosons which decay to two muons. Good muon reconstruction is crucial for increasing the efficiency to find Z bosons. Muons are characterized by tracks in the COT, minimum ionizing energy deposition

in the calorimeters, and a "stub" in the muon chambers. In this analysis, only the muons reconstructed in the CMX or both the CMU and the CMP are used.

Muon reconstruction [46] begins with hits in the muon chambers, obtained by converting TDC times to hit positions using a drift model which reflects the chamber geometry. In the CMU, charge division is used to determine the z position of the hit. In the CMP and CMX, no z position information is available. Then a stub is reconstructed from the hits. To find stubs, the algorithm looks in each region of the muon detectors and counts the number of hits and the number of layers with hits in the region. If more than three hits are found on more than three layers, stub finding begins in this region. A seed of two hits in two layers are used to construct a straight line pointing toward the interaction region and other hits are added to this seed stub. After the list of hits for a stub are determined, a simple drift model is used to fit the stub. The stub finding in the muon chambers is similar to COT tracking except that the muon chambers are not in a magnetic field and, therefore, have no curvature.

Once a muon stub has been found, its position and direction are used to match it to a COT track to form a muon. This process is complicated by the fact that calorimeters lie between the COT and muon chambers, and the path of the muon inside the calorimeters is undetermined. The COT track, with a muon hypothesis, is extrapolated through the calorimeters and the intersection of the track with the muon chambers is derived. The difference between the position of the stub and the extrapolated position of the track in the $r - \phi$ plane, Δx, is used to determine the match between a stub and a track. If more than one track is matched to a stub, the one with the best match quality is kept. Up to two stubs in two different muon

subdetectors can be attached to the same track. If there are more than two stubs matching to a track, the best matched stub in each subdetector is used. Unmatched stubs are discarded. The matched track-stub matches are considered reconstructed muons.

4.3 Electron Reconstruction

Electrons are used in this analysis to reconstruct Z bosons which decay to two electrons. Good electron reconstruction is essential for reconstructing as many Z bosons as possible. Electrons are characterized by tracks in the COT and a large deposition of energy in the electromagnetic calorimeters. This analysis uses electrons detected in both the central electromagnetic calorimeters and the plug calorimeters.

4.3.1 Central Electrons

The reconstruction of central electrons begins in the central electromagnetic (CEM) calorimeters. Individual towers with a transverse energy deposition of more than 3 GeV are used as "seed" towers. Then towers adjacent to the seed tower in η with 100 MeV of electromagnetic or hadronic energy deposition are added to form a cluster. An electromagnetic cluster is allowed to have only one tower in ϕ and up to three towers in η. Clusters are not allowed to have towers across a region boundary or the center of the detector. Then only the electromagnetic energy deposition in the towers is used to calculate the centroid and total energy of the cluster. Clusters with a total energy less than 5 GeV are discarded. Furthermore, in order to verify that the cluster is from an electromagnetic object, we require that the ratio of energy

deposited in the hadronic calorimeters to the energy deposited in the electromagnetic calorimeters is less than 0.125. Clusters which satisfy these requirements are classified as *CdfEmObject* and can be either an electron or a photon.

Once a cluster has been found, it must be matched to a track and a CES cluster to qualify as an electron. COT tracks are extrapolated to the plane of the CES within the calorimeter wedge of the cluster and required to be within 25 cm in x and 38 cm in z to center of the seed tower of the cluster. The extrapolated position of the track is used to seed a CES strip and wire clustering algorithm. This algorithm compares the energy deposition in the strips and wires to a template from electron test beam data and adjusts the center of the cluster. If the CES cluster profile matches well with the template, the EM cluster, track, and CES cluster are defined as an electron candidate. If more than one track matches the EM cluster, the track with the highest p_T, and the corresponding CES cluster seeded by this track, are chosen. In order to determine whether the energy was deposited by a single electron, a tower lateral energy sharing test variable, L_{shr}, is calculated from the position of the track and the electron energy given by the following equation [47]:

$$\mathcal{L}_{shr} = \frac{\sum_{towers}(E_{meas.} - E_{pred.})}{\sqrt{0.196 E_{EM} + \sum \sigma(E_{pred.})^2}} \tag{4.1}$$

4.3.2 Phoenix Electrons

The reconstruction of phoenix electrons begins in the plug electromagnetic (PEM) calorimeters. Plug clusters are reconstructed in a similar fashion to the central clusters, beginning with a seed tower with more than 3 GeV of transverse electromagnetic

energy and adding shoulder towers with more than 100 MeV of electromagnetic or hadronic energy. Plug clusters span up to three towers in η and three towers in ϕ. The profile of the energy deposition in this 3x3 region is compared to the profile of energy deposition by electrons from electron test beam data. The electromagnetic energy of the towers is used to calculate the total energy and centroid of the cluster. Clusters with a total energy greater than 5 GeV and a hadronic to electromagnetic energy deposition ratio of less than 0.125 are classified as CdfEmObjects [48].

A COT track is not required to be matched to the plug EM object. The energy deposition in the strip and wire cluster of the PES is compared to electron test beam data and the PES cluster is added to the PEM cluster. The PEM and PES clusters define a plug electron. Unfortunately, plug electrons are inadequate for our analysis due to the fact that charge information is unavailable for plug electrons not matched to tracks and we require opposite sign electrons to reconstruct a Z. Therefore, we resort to phoenix electrons where a silicon track is required in addition to the plug electron requirements. As described in Section 4.1.3, plug electrons are used to seed outside-in tracking in the silicon detector a silicon track is added to the plug electron. Phoenix electrons are required to have three or more silicon hits on the track. This provides charge information for the electron which is necessary in reconstructing Z bosons in our analysis.

4.4 Isolation

Most energetic particles at CDF are produced within jets. Consequently, the total energy in the in an $\eta - \phi$ cone surrounding the particles normally exceeds the energy

of the particle. In this analysis, we are interested in the leptonic decay products of the Z boson which would not have this surrounding energy since both leptons and electroweak gauge bosons are color-neutral. Consequently, isolation can be used as a discriminator between the electrons, muons, and tracks (used as electrons or muons) from the decay of Z bosons from the other tracks which mimic them.

Two definitions of isolation are used in this analysis. For electrons and muons, a calorimeter based isolation is used, defined as the ratio of the difference in energy within an $\eta - \phi$ cone of 0.4 to the lepton and the energy of the lepton to the energy of the lepton, $\sum_{\Delta R=0}^{0.4}(E^{cal} - E_l)/E_l$. This ratio is required to be low (typically < 0.1) for an isolated lepton. For tracks, we use a track based isolation which is the ratio of the sum of the transverse momenta of tracks within an $\eta - \phi$ cone of 0.4 to the transverse momentum of the track, $\sum_{\Delta R(i)=0}^{0.4} p_T^i / p_T^{track}$. This ratio is required to be large (typically > 0.9) for an isolated track.

4.5 Jet Reconstruction

Quarks are found only in bound states of two or three quarks called hadrons. Therefore, when a single quark is produced in an interaction, it pulls quark-antiquark pairs out of the vacuum to form hadrons in a process called *hadronization*. This results in a collection of particles traveling in the same general direction and depositing their energy in the calorimeter, forming a *jet*. In our analysis, we have four quarks, two from the W boson decay, a c or u quark from the FCNC decay of one top quark and a b quark from the SM decay of the other top quark, leading to four jets. Finding jets which are large enough to contain as much as possible of the energy from the

hadronization of each quark while being small enough to find four distinct jets in the event is the challenge for the jet clustering.

4.5.1 Jet Clustering

Jets used in this analysis are clustered by the `JetClu` algorithm [49], which requires a seed tower and a group of nearby towers. It takes as inputs the energy threshold for the seed tower and a cone in $\eta - \phi$ space to define the extent of the region in which adjacent towers are included in the clustering. For this analysis, we use a seed threshold of $E_T > 1$ GeV and a cone size of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.4$ rad. Jet studies show that $\sim$70% of the energy of the jet is contained within a cone size of 0.4. At CDF, cone sizes of 0.7 and 1.0 are also available but we do not use them due to the large jet multiplicity required in our analysis.

The algorithm begins with the seed tower and uses its position as the initial center of the jet (η_0,ϕ_0). Then it includes the towers within the cone with the predefined radius and recalculates the centroid of the jet using the following definitions:

$$\eta_{jet} = \sum_{tower=1}^{N_{towers}} \frac{\eta_{tower} E_{tower}}{E_{jet}} \tag{4.2}$$

$$\phi_{jet} = \sum_{tower=1}^{N_{towers}} \frac{\phi_{tower} E_{tower}}{E_{jet}} \tag{4.3}$$

The algorithm uses the recalculated η and ϕ as the centroid position and iterates until all the towers are added and the centroid position is stable.

The number of jets in an event is not robust. The use of seed towers means that

small changes in geometry and energy deposited in two towers versus one can change the clustering. Jets connected by soft radiation can lead to merged jets. There are gaps in the calorimeters which can lead to mismeasured or lost jets. Consequently, the number of partons and the number of jets do not correspond on a one-to-one basis and must be studied in simulations. Although not required for reconstructing a jet, tracks which fall within a cone of 0.4 around the centroid of the jet and have a z_0 within 2 cm of the primary vertex z position are associated with the jet.

4.5.2 Jet Energy Correction

The measured energy in a jet is most often not the total energy of the parton. This is due to instrumental reasons, arbitrariness of the jet definition, and experimental conditions. Uncertainties in energy result from the sampling nature of the calorimeter. The response of the towers is not uniform and calibrations need to be applied. There are gaps in the calorimeters and the energy of particles which fall within these gaps is lost. Some particles do not deposit all their energy in the calorimeters and their energies are mismeasured. These will decrease the energy of the jet. Particles which are not from the jet but from the *underlying event* can add extra energy to the jet as well as extra particles from additional $p\bar{p}$ collisions during the same beam crossing. Due to these, the energy of the jet must be corrected offline.

The corrections applied to jets are derived in several ways [50]. Calibrations of the calorimeter towers are derived from comparing the mass of Z bosons reconstructed from electrons to the world average Z mass. Furthermore the energy of the jets is studied in events where the jet recoils against an object whose energy is well measured.

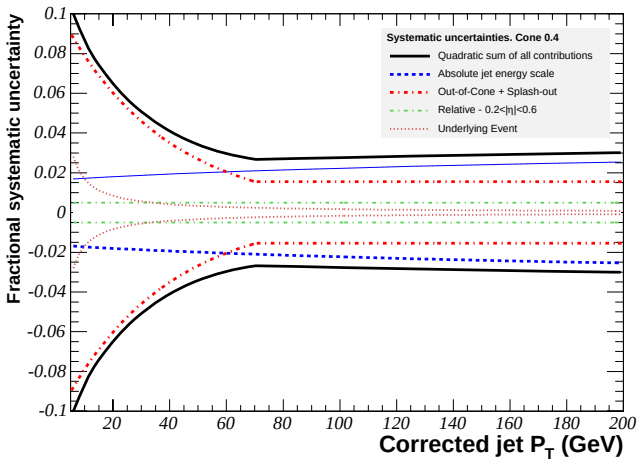


Figure 4.2: Jet energy scale systematic uncertainties for the different corrections as a function of corrected jet p_T.

This is done in photon-jet events, Z-jet events, and dijet events. In photon-jet events where a jet recoils against a photon, the energy of the photon is well measured in the central electromagnetic calorimeter. Therefore, it is possible to determine a correction factor for the energy of the jet to match that of the photon. Similarly, in events where a jet recoils against a Z boson, the energy of the jet can be compared to the p_T of the Z. Energy of the jets in back-to-back dijet events can also be used to determine the η dependence of the jet energy measurement. Figure 4.2 shows the jet energy correction systematic errors as a function of jet p_T.

In this analysis, jets are corrected to Level 5 which means that they are corrected for the following effects:

- Online/Offline Calibrations: Online calibrations correct the variation in calorimeter tower response to 3%. The rest is corrected at the offline level by using the latest calibration constants derived from particles with known energies in test

beams and well known reference particles such as the Z boson.

- Eta Dependence: These corrections are applied to make the detector response uniform in η.

- Multiple Interactions: Subtracts the contribution of extra particles from the extra interactions in the same beam crossing based on the number of z vertices in the event. The average correction is derived from minimum bias data.

- Absolute: Corrects for non-linearity in energy response and for particles lost in the uninstrumented regions of the detector and leakage.

Although corrections for the underlying event, energy outside the jet clustering cone, and energy that exits the calorimeter are available at CDF, we do not use them in this analysis. In addition, we apply the same corrections to light quark jets and b jets even though a b parton can decay to muons and neutrinos causing less of the b parton's energy to be visible in the calorimeter.

4.6 Heavy Flavor Identification

Heavy flavor quarks, bottom and charm, form hadrons that have a lifetime long enough that they travel a measurable distance before decaying. The distance they travel is given by the formula

$$\mathcal{D} = \beta ct = \beta\gamma c\tau \tag{4.4}$$

where β is the velocity of the particle divided by the speed of light (c), γ is the relativistic factor, t is the flight time of the particle in the lab reference frame, and τ is the proper lifetime of the particle. For b hadrons boosted from a top decay, this flight distance can be a few millimeters. When these hadrons decay, their products form a vertex, called a *secondary vertex*, which is displaced from the average primary vertex where the initial $p\bar{p}$ interaction took place as shown in Figure 4.3. Therefore b and c jets can be identified, *tagged*, by the presence of these secondary vertices. In our signal, there is a b jet from the SM top decay and a c jet from the FCNC top decay (assuming that $t \to cZ$ dominates over $t \to uZ$). Tagging these heavy flavor jets will help in differentiating between signal and background. In this analysis, we use the loose SecVtx algorithm to identify heavy flavor.

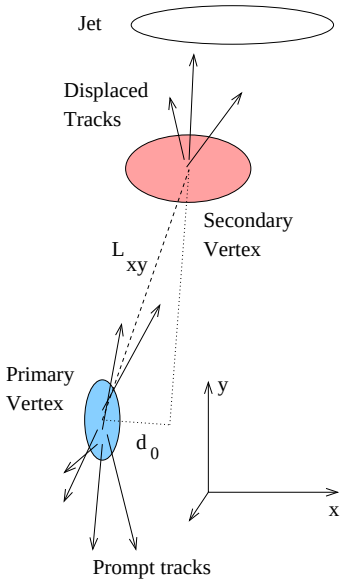


Figure 4.3: Diagram of a secondary vertex formed when a long-lived heavy flavor particle decays after travelling a macroscopic distance away from the primary interaction point.

4.6.1 Loose SecVtx Tagging Algorithm

The loose SecVtx b-tag algorithm [51] selects tracks with large impact parameter significance (indicating they are from a secondary vertex as opposed to the primary vertex), $(\frac{d_0}{\sigma_{d_0}}) > 2.0$, above a p_T threshold of 500 MeV/c and with silicon hits from two or more layers (improves position resolution). It makes two attempts to find secondary vertices. In the first attempt, Pass 1, it tries to form three track vertices. It begins by forming seeds with a pair of tracks. It takes the best seed and adds additional tracks which are consistent with being from the seed vertex. When all the tracks are added, the vertex is pruned by removing tracks contributing $\Delta\chi^2_{fit} > 50$ to the vertex. If a three-track vertex satifying vertex quality cuts remains after this procedure, it is retained. If not, it moves to the seed (in terms of quality) and repeats the procedure.

If Pass 1 fails to produce a secondary vertex, a second attempt, Pass 2, is made by finding two track vertices with tighter track selection criteria. Only tracks with $(\frac{d_0}{\sigma_{d_0}}) > 3.0$, $p_T > 1.0$ GeV/c, and with silicon hits from three or more layers attached to the track, are considered. Pass 2 forms two-track vertices from these tracks and searches for a vertex satifying the necessary quality cuts. A vertex must satisfy a two dimensional displacement (from the primary vertex) requirement, $L_{2D}/\sigma_{L_{2D}} > 6.0$ as well as other quality cuts to be accepted. This two dimensional displacement is defined as the dot product of the vertex position vector and the jet momentum direction, $L_{2D} = \vec{r}_{vtx} \cdot \vec{p}_jet$. A positive L_{2D} means that the vertex is consistent with being the decay of a particle travelling in the same direction as the jet while a negative L_{2D} appears to be from a particle travelling in the opposite direction.

Negative vertices can be used to study fake tags as will be further described in the following subsection.

The loose SecVtx algorithm was tuned to increase the tagging efficiency compared to the tight SecVtx algorithm [52] which had more stringent requirements for track quality. It is insufficient to measure the eifficiency of the tagger only in simulations because the tagging efficiency is different for data and simulations. The ratio of data tag efficiency to Monte Carlo tag efficiency for the loose SecVtx tagger is 0.95 $\pm$ 0.04 [53]. The efficiency for tagging jets depends on their transverse energy and pseudorapidity but the average per jet tagging efficiency for b jets is 48% $\pm$ 4% [53].

Mistags

While we want to maximize the efficiency to tag b-jets, we also want to have as low a rate of tags in light flavor jets, *mistags*, as possible. The loose SecVtx tagger traded a 20% increase in b-tag efficiency for a factor of $\sim$ 2.5 increase in the mistag rate, compared to the tight tagger. Mistags result from track resolution, decays of long lived particles such as K_0 and Λ, and from material interactions. Mistags from long lived light flavor particles are mitigated by requiring the tags to fall outside their mass window. Mistags from material interactions are reduced by requiring that the reconstructed vertex is within the beam pipe. Ideally, mistags should be estimated from measuring the rate of positive tags in a sample containing only light flavor jets. Unfortunately, such a sample cannot be constructed.

The rate of mistags is estimated in generic jet samples with heavy flavor contamination by looking at the negative tag rate. Mistags resulting from track resolution

are equally likely to form positive (mis)tags as well as negative (mis)tags while the majority of tags in heavy flavor are expected to be positive. Despite this, the mistag rate for heavy flavor jets is higher than the mistag rate for light flavor jets since heavy flavor jets contain more track candidates with large impact parameters and because tracks from two vertices can be reconstructed into a negative tag. The negative tag rate in generic samples is corrected for these two effects in order to accurately represent the mistag rate. The mistag probability for a jet depends on the jet's E_T, η, ϕ, number of tracks, as well as global event variables $\sum E_T$, number of z vertices and the primary vertex z [54]. The average per jet mistag rate for the loose SecVtx tagger is $1.3\% \pm 0.2\%$.

4.7 Monte Carlo Simulation

In order to understand the data, we must study simulations based on theory of the physics processes of interest. At CDF, the Monte Carlo simulation involves two steps; the generation of physics processes of interest and the simulation of how the particles propagate in the CDF detector.

4.7.1 Event Generation

The collision of elementary particles can be modeled throught *Monte Carlo* techniques, by selecting outcomes at random from probabilty distributions for the observables in the final states of the physics processes of interest. The probability distributions come from matrix element calulations or from approximations since matrix element calculations are difficult to perform. Two approximations are used. The first

is to use finite order expansions of differential cross sections for the *hard scattering* between partons. The second is employed in modeling QCD effects. Measured parton distribution functions are used to describe the incoming protons and antiprotons. Initial and final state ratiation are also simulated. The event generation outputs the initial positions and initial momenta of stable and metastable particles as well as all the intermediate particles and the lineage and descendents of each particle.

The simulations used in this analysis are generated by either PYTHIA [55] or a combination of ALPGEN [56], which performs higher order matrix element calculations, and PYTHIA, which performs parton showering and hadronization. ALPGEN +PYTHIA is used to simulate the Z+jets background and PYTHIA is used to simulate the signal and the other backgrounds. The specific Monte Carlo samples will be described in more detail in later chapters.

4.7.2 Detector Simulation

The particles from the event generation are passed to a description of the CDF detector based on GEANT [57], which simulates the interaction of these particles with the various detector elements. After detector simulations, the output bank is functionally identical to the data from the online data acquisition system. Data conditions are further mimicked by applying the beam and detector conditions, e.g. calibrations and alignments, on a run-by-run basis. Simulated events are reconstructed in the same way as the data except that a trigger is not applied to the simulations.

Chapter 5

Data Sample and Event Selection

The first step to getting the things you want out of life is this:
Decide what you want.
Ben Stein

In this chapter, we will describe the data sample used and how it was selected.

5.1 Data Sample

We use data collected between March 2002 and September 2006, which covers runs 138425 to 222426 and corresponds to an integrated luminosity of 1.12 fb^{-1}. We reject data from runs with known problems in detector response or calibrations. We use the `GoodRun v.16.0` list [58] and reject events from runs when the COT or silicon tracking, the electromagnetic calorimetry, or the muon chambers were not properly functioning. We veto CMX muons from runs prior to 150145 because the CMX detector was not fully commissioned during the first five months after silicon detector commissioning. This requirement reduces the luminosity of the CMX data set by

13.4 pb^{-1}. Additionally, we exclude runs 179057 to 182842 ("COT compromised" period) corresponding to the three-month period of data-taking when wire aging in the COT reduced the gain in some of its regions. During this time, some superlayers were operated at a lower electrostatic potential while a solution was developed [59]. Duplicate events contained in some runs due to data acquisition errors are removed.

The data were collected with the inclusive lepton triggers. The electron trigger path we use requires an electromagnetic calorimeter tower with energy ≥ 8 GeV (CHA energy must be less than 1/8 of the CEM energy in the tower if the energy of the tower is above 14 GeV) and an XFT track in the same wedge with $p_T \geq 8$ GeV/c at Level 1 (L1_CEM_8_PT8), a cluster of towers with $E \geq 16$ GeV and hadronic to electromagnetic energy ratio less than 1/8, matched to an 8 GeV track at Level 2 (L2_CEM_16_PT8), and a fully reconstructed electron with $E_T > 18$ GeV matched to a track with $p_T > 9$ GeV/c and hadronic to electromagnetic ratio less than 1/8 at Level 3 (ELECTRON_CENTRAL_18). [60] The corresponding data sets are `bhel0d`, `bhel0h`, and `bhel0i` (periods 1 and 2).

The CMUP trigger path we use requires matched hits in the hits in the CMU and CMP within 2.5° of a track with $p_T \geq 4$ GeV/c at Level 1 (L1_CMUP6_PT4), a match between a stub and an 8 GeV/c track at Level 2 (L2_CMUP6_PT8), and a fully reconstructed muon with $p_T > 18$ GeV/c and track matched in $r\phi$ within 10(20) cm of the CMU (CMP) stub at Level 3 (L3_Muon_CMP_18). The CMX trigger path we use requires a muon stub with scintillator coincidence and XFT track with $p_T \geq 8$ GeV/c at Level 1 (L1_CMX6_PT8_CSX), match between a stub and a track with $p_T \geq 10$ GeV/c at Level 2 (L2_CMX6_PT10), and a reconstructed muon with track $p_T >$

18 GeV/c matched to a reconstructed CMX stub within 10 cm [61]. The data selected by the CMUP and CMX triggers correspond to the datasets `bhmu0d`, `bhmu0h`, and `bhmu0i` (periods 1 and 2). The 0d, 0h, and 0i datasets are from different data taking periods between annual shutdowns when maintenance and repair are performed on the detector.

The events for this analysis are selected in two steps. We begin by selecting events with two leptons (e^+e^- or $\mu^+\mu^-$) which form Z boson and four or more jets. This is our base selection criteria and also our blind region. I will also refer to this as our "pre-tag" sample. We then optimize our selection criteria to obtain the best expected limit, placing further constraints on the mass χ^2, transverse mass, and the transverse energy of the four leading jets. Using the optimized selection criteria, we set two signal regions, separating the sample into events with a loose SecVtx b-tag (tagged events) and those without (anti-tagged events).

5.2 Base Selection Criteria

The base event selection is a loose set of criteria derived from the signature of the FCNC signal which contains a Z boson which decays to two leptons, a W boson which decays to two jets, and two extra jets from the $t \to qZ$ decay and the $t \to Wb$ decay. We, therefore require a Z and $\geq$4 jets with the following requirements:

- Z selection
  - Two, and only two, leptons with the same flavor and opposite charge
  - One of the leptons must be a tight central lepton, the second leg can be a

tight central lepton, an isolated track, or, in the case of e^+e^- pairs, a PHX electron[1]

- $\Delta(z_0)$ between the two leptons less than 5 cm
- 76 GeV/c^2 $< M_{\ell^+\ell^-} <$ 106 GeV/c^2
- Conversion electrons and cosmic ray muons are vetoed

- Jet selection
  - Jets are clustered with `JetClu` with a cone of 0.4 and corrected up to Level 5, as described in Section 4.5.2
  - $E_T > 15$ GeV
  - $|\eta| < 2.4$

These $Z+ \geq 4$ jetsevents fall into our blinded region and are vetoed until the end. We use events with a Z and three or fewer jets as our control region to perform cross-checks for our background prediction methods. For some studies, we will expand the control region (reduce the blinded region) to include events with a Z and four or more jets that fail a requirement on the mass χ^2 variable. The base selection criteria do not make any b-tag requirements and will sometimes be referred to as our "pre-tag" sample.

[1] Allowing the second leg to be an isolated track roughly doubles our acceptance

5.3 Lepton Selection

This analysis uses tight central electrons (TCE), tight phoenix electrons (PHX), tight central (CMUP or CMX) muons which follow the CDF Joint Physics criteria for electrons [62] and muons [63]. For the tight track leptons, we use the requirements developed for the lepton+track top cross section measurement [64]. We impose an additional correction for the electron+track Z candidates to account for possible energy loss due to bremsstrahlung, using the E_T of the track if it is larger than the p_T.

The efficiency for lepton selection cuts is determined from a pure sample of leptons from Z decays. One leg is required to pass all the tight selection criteria and the other leg is selected without requiring the criteria in order to obtain an unbiased sample. The lepton identification criteria are then applied on the second leg and the efficiency for each requirement is measured in both data and Pythia inclusive Z Monte Carlo samples. Due to discrepancies between data and Monte Carlo simulations, the lepton identification and reconstruction efficiencies in the two samples are different. We assign the ratio of efficiencies as a scale factor, $(SF)_{lepID} = \frac{\epsilon_{data}}{\epsilon_{MC}}$. The trigger efficiencies and lepton identification and reconstruction scale factors are different for the four data taking periods, 0d, 0h, 0i1, and 0i2. The trigger efficiencies and lepton scale factors for the different types of leptons used in this analysis for the different data taking periods are given in Table 5.1.

	0d	0h	0i 1	0i 2
CEM				
Trigger Efficiency	0.962±0.007	0.976±0.006	0.979±0.004	0.959±0.007
Electron ID Scale Factor	0.991±0.005	0.985±0.005	0.974±0.004	0.974±0.004
PHX				
Electron ID Scale Factor	0.929±0.006	0.943±0.006	0.937±0.006	0.937±0.006
CMUP				
Trigger Efficiency	0.902±0.004	0.919±0.004	0.918±0.005	0.913±0.006
Muon ID Scale Factor	0.985±0.004	0.989±0.004	0.975±0.005	0.975±0.006
Muon Reconstruction	0.951±0.004	0.939±0.004	0.941±0.004	0.955±0.005
CMX Arches				
Trigger Efficiency	0.967±0.004	0.955±0.004	0.954±0.005	0.947±0.006
Muon ID Scale Factor	1.014±0.004	1.000±0.005	1.004±0.006	1.000±0.008
Muon Reconstruction	0.996±0.002	0.993±0.002	0.989±0.003	0.991±0.003
CMX Miniskirt/Keystone				
Trigger Efficiency	–	0.772±0.014	0.744±0.019	0.755±0.023
Muon ID Scale Factor	–	0.979±0.011	0.990±0.013	1.001±0.015
Muon Reconstruction	–	0.933±0.009	0.939±0.011	0.902±0.016
Tracks				
Track ID Scale Factor		0.954 ± 0.011		

Table 5.1: The scale factors and trigger efficiencies used to scale the MC simulation estimates for electron [65], [66], muon [61], and track [64] reconstruction. Scale factors are run-dependent, and are listed below with the corresponding datasets. All datasets together represent the full $1.12\,\mathrm{fb}^{-1}$ sample. The uncertainties shown are statistical uncertainties only.

5.3.1 Electron Selection

Electron candidates reconstructed by the algorithm described in Chapter 4 are required to satisfy additional selection criteria. For tight central electrons, the track is required to be fiducial to the CES and near a CES cluster in z and in "local x", the $r\delta\phi$ coordinate. The CES strip energy deposition profile is fit to data from electron beam data and the fit χ^2 is required to be less than 10.0. Electrons from conversions are vetoed. The electrons are required to have energy greater than 20 GeV and the tower profile variable L_{shr} is required to be less than 0.2. To reduce jet backgrounds, the ratio of hadronic energy deposition of the towers to the electromagnetic energy deposition (E_{Had} / E_{Em}) is required to be small. To reduce charge pion backgrounds,

the ratio of the energy to momentum, E/p, is required to be less than 2.0. The electron track is required to have $p_T \geq 10.0$ GeV/c and satisfy number of COT superlayer and hits requirements. The z_0 of the electron track is required to be within the interaction region, within 60 cm of the center of the detector. The electron candidates are required to be isolated, with calorimeter based isolation less than 0.1. All these requirements are summarized in Table 5.2.

Electron Variable	**Base Cut**
Fiducial to CES	Yes
From photon conversion	No
E_T	≥ 20 GeV
Fractional isolation	≤ 0.10
E/p	≤ 2 (unless $p_T > 50$ GeV/c)
$E_{\rm Had}$ / $E_{\rm Em}$	$\leq 0.055 + 0.00045E$(GeV)
Track z_0	≤ 60 cm
Track p_T	≥ 10 GeV/c
Track $L_{\rm shr}$	≤ 0.2
COT axial superlayer hits	2
COT stereo superlayer hits	3
Hits per COT superlayer	5
CES Δz	$(-3$ cm; 3 cm)
CES Δx times charge	$(-3$ cm; < 1.5 cm)
CES strip χ^2	≤ 10

Table 5.2: The standard Joint Physics selection criteria for tight central electrons [62]. The variables are explained in the text.

Figures 5.1 to 5.3 show the distributions of the tight central electron identification quantities for electrons from data and Pythia Z Monte Carlo. A pure sample of electrons is obtained from Z decays. One leg of the Z is required to satisfy the tight central electron selection criteria. The other leg is selected by requiring only that it is fiducial to the CES and not from a conversion. Then cuts are applied to the second leg. For each variable plotted, the central electron is required to satisfy the

requirement for all variables except the one plotted. The two legs are required to have opposite charge and the reconstructed invariant mass of the dielectron pair is required to be within 76 GeV/c^2 to 106 GeV/c^2.

Tight Phoenix electrons are required to have a Phoenix track, with three or more silicon hits and $z0$ within 60 cm of the center of the detector, which points to the PES within a two dimensional pseudorapidity, $1.2 < |\eta| < 2.8$. The energy deposition shape in the PEM and PES are required to match those from electron test beam data and the centroids of the PEM and PES clusters are required to be within 3.0 cm of each other. The E_T is required to be above 20 GeV and the ratio of hadronic to electromagnetic energy deposition is required to be small. Phoenix electron candidates are also required to be isolated with calorimeter based isolation less than 0.1. These requirements are summarized in Table 5.3.

Electron Variable	**Base Cut**
Matched to a Phoenix track	Yes
E_T	≥ 20 GeV
PES two-dimensional η	$1.2 < \|\eta\| < 2.8$
E_{Had} / E_{Em}	≤ 0.05
PEM χ^2	≤ 10
PES 5×9 U and V	≥ 0.65
Fractional isolation	≤ 0.10
ΔR between the PES and PEM centroids	≤ 3.0 cm
Silicon hits	≥ 3
Track z_0	≤ 60 cm

Table 5.3: The standard Joint Physics selection criteria for tight phoenix electrons [62]. Variables are explained in the text.

Figures 5.4 and 5.5 show the distributions of the tight phoenix electron identification quantities for electrons from data and Pythia Z Monte Carlo. A pure sample

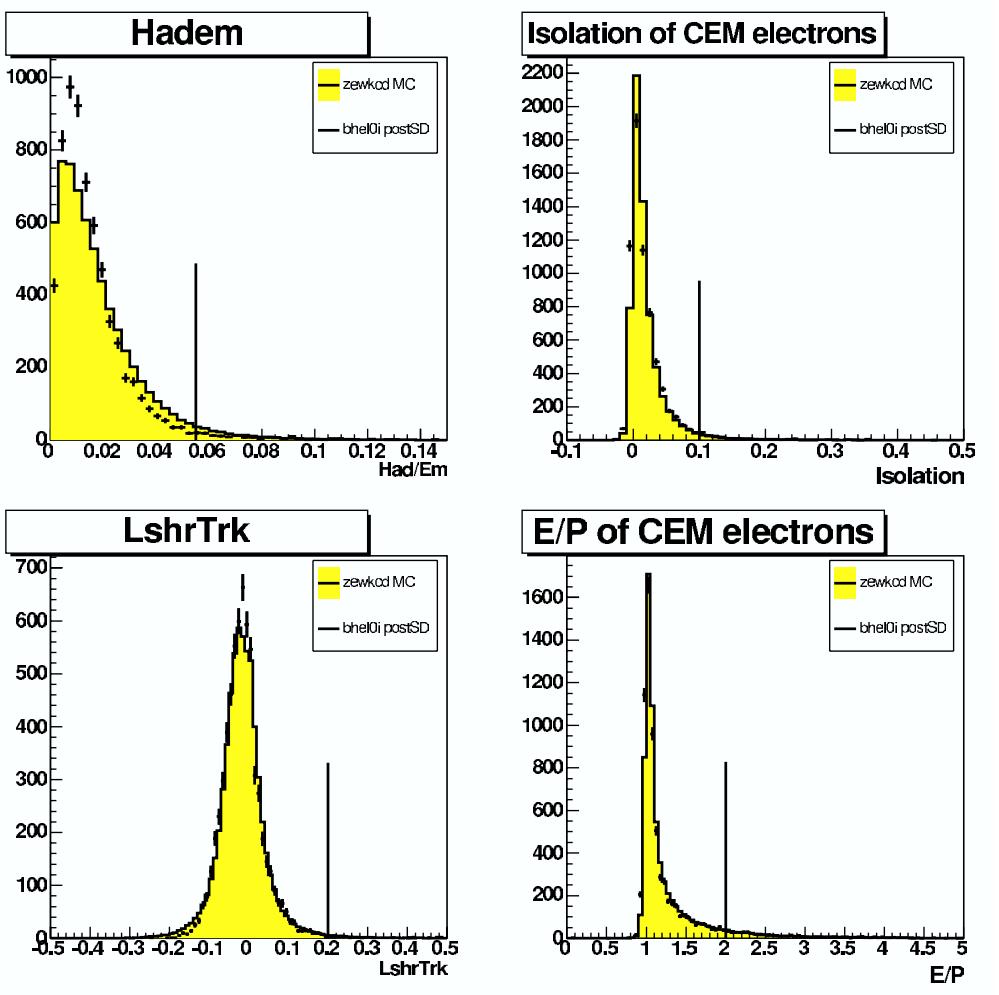


Figure 5.1: TCE ID Variables: Had/Em, Isolation, Lshr, and E/P [66]. The vertical lines show the selection requirement.

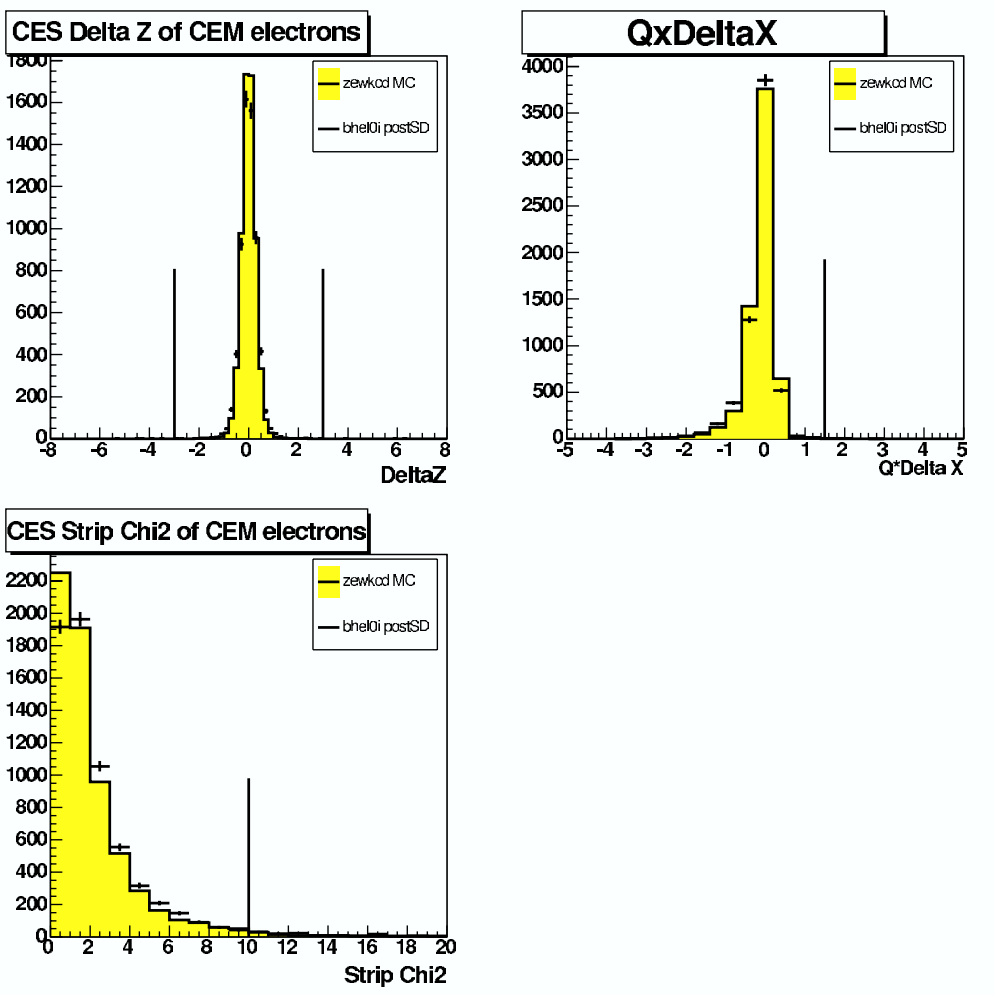


Figure 5.2: TCE ID Variables: Δz, Δx, and Strip χ^2 [66]. The vertical lines show the selection requirement.

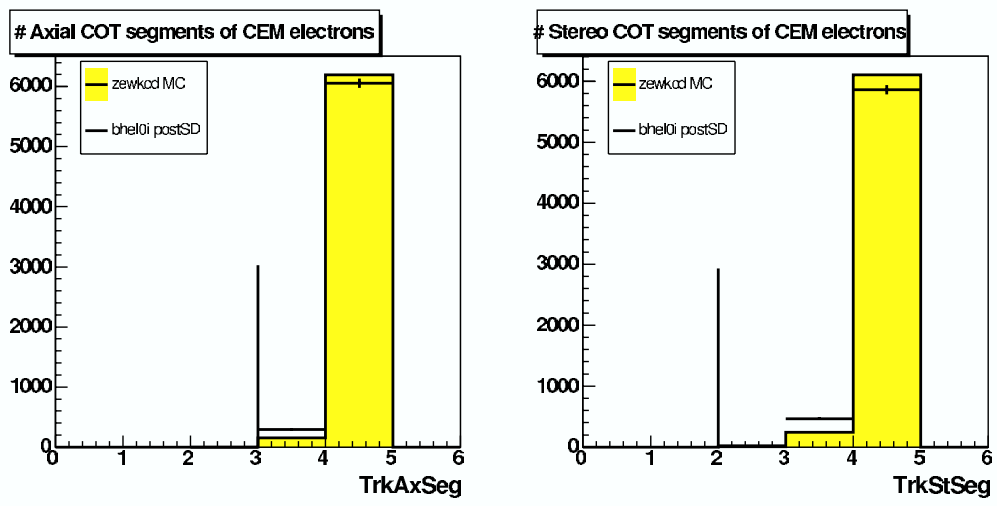


Figure 5.3: TCE ID Variables: Number of COT Axial and Stereo Segments [66]. The vertical lines show the selection requirement.

of electrons is obtained from Z decays. One leg of the Z is required to satisfy the tight central electron selection criteria. The other leg is selected by requiring only that it is fiducial to the PES. Then cuts are applied to the second leg. For each variable plotted, the phoenix electron is required to satisfy the requirement for all variables except the one plotted.The two legs are required to have opposite charge and the reconstructed invariant mass of the dielectron pair is required to be within 76 GeV/c^2 to 106 GeV/c^2.

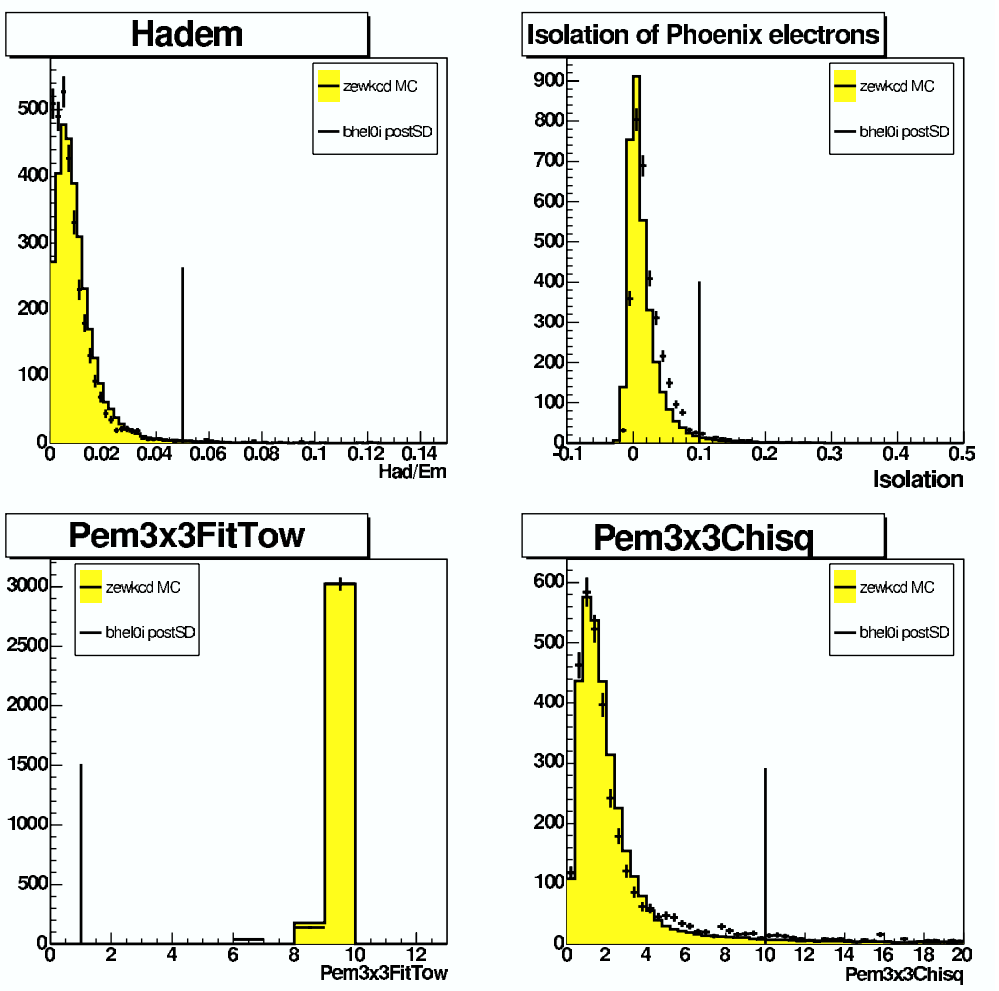


Figure 5.4: PHX ID Variables:Had/Em, Isolation, PEM variables [66]. The vertical lines show the selection requirement.

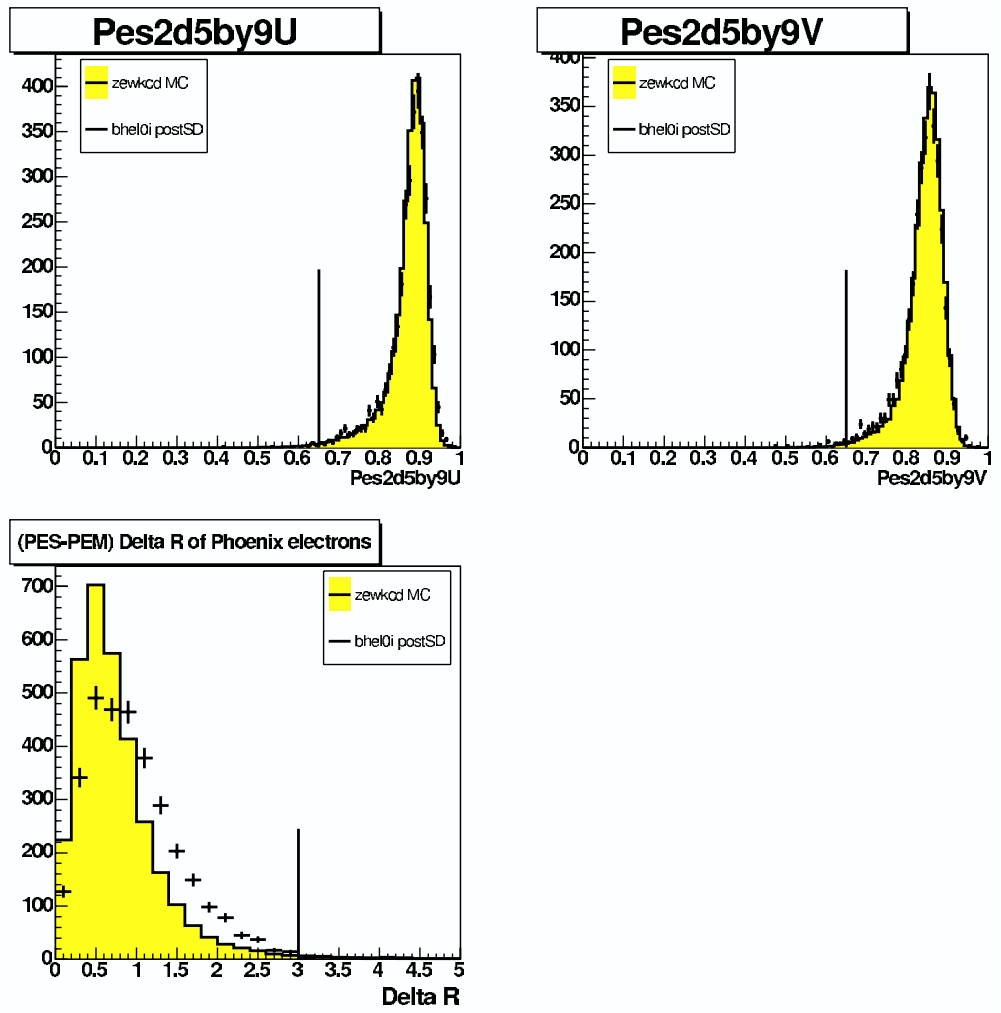


Figure 5.5: PHX ID Variables:PES variables [66]. The vertical lines show the selection requirement.

5.3.2 Muon Selection

Tight muon candidates constructed according to the algorithms described in Chapter 4 are also required to satisfy further quality cuts. Muons are minimum ionizing particles and the energy deposition in the calorimeters is required to be small with some additional allowance for higher p_T muons. Muon candidate tracks are required to have $p_T \geq 20$ GeV/c and must pass COT track quality cuts. The tracks must also point back to the interaction point, with $z0$ within 60 cm and impact parameter, d_0, 0.02 cm for tracks which have silicon hits and 0.2 cm for tracks which don't. The track is also required to be fiducial to the muon detectors (CMU and CMP, or CMX) and the extrapolated track position is required to be consistent with the stub position. Muon candidates are required to be isolated with fractional isolation less than 0.1. The requirements a muon must satisfy are summarized in Table 5.4.

Figures 5.6 to 5.8 show the distributions of the tight muon identification quantities for muons from data and Pythia Z Monte Carlo. A pure sample of muons is obtained from Z decays. One leg of the Z is required to satisfy the tight central muon selection criteria. The other leg is selected by requiring only that it is a reconstructed muon. Then cuts are applied to the second leg. For each variable plotted, the muon is required to satisfy the requirement for all variables except the one plotted. The two legs are required to have opposite charge and the reconstructed invariant mass of the dimuon pair is required to be within 76 GeV/c^2 to 106 GeV/c^2.

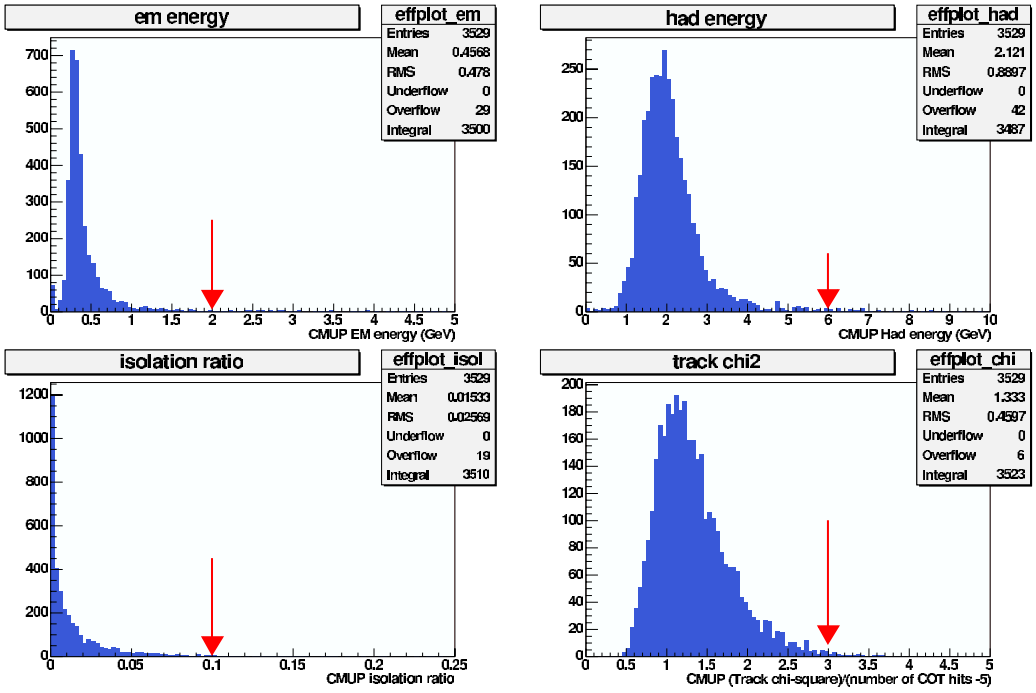


Figure 5.6: CMUP ID Variables:EM Energy, Had Energy, Isolation, Track χ^2 [67]. The arrows show the selection requirement.

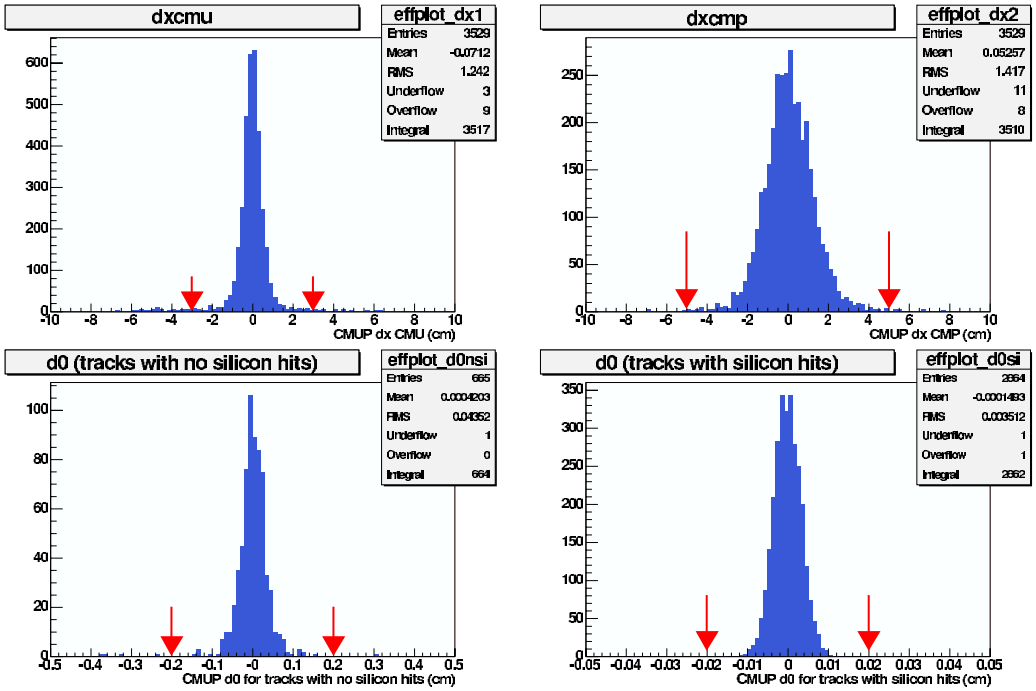


Figure 5.7: CMUP ID Variables:Δx and Impact Parameter [67]. The arrows show the selection requirement.

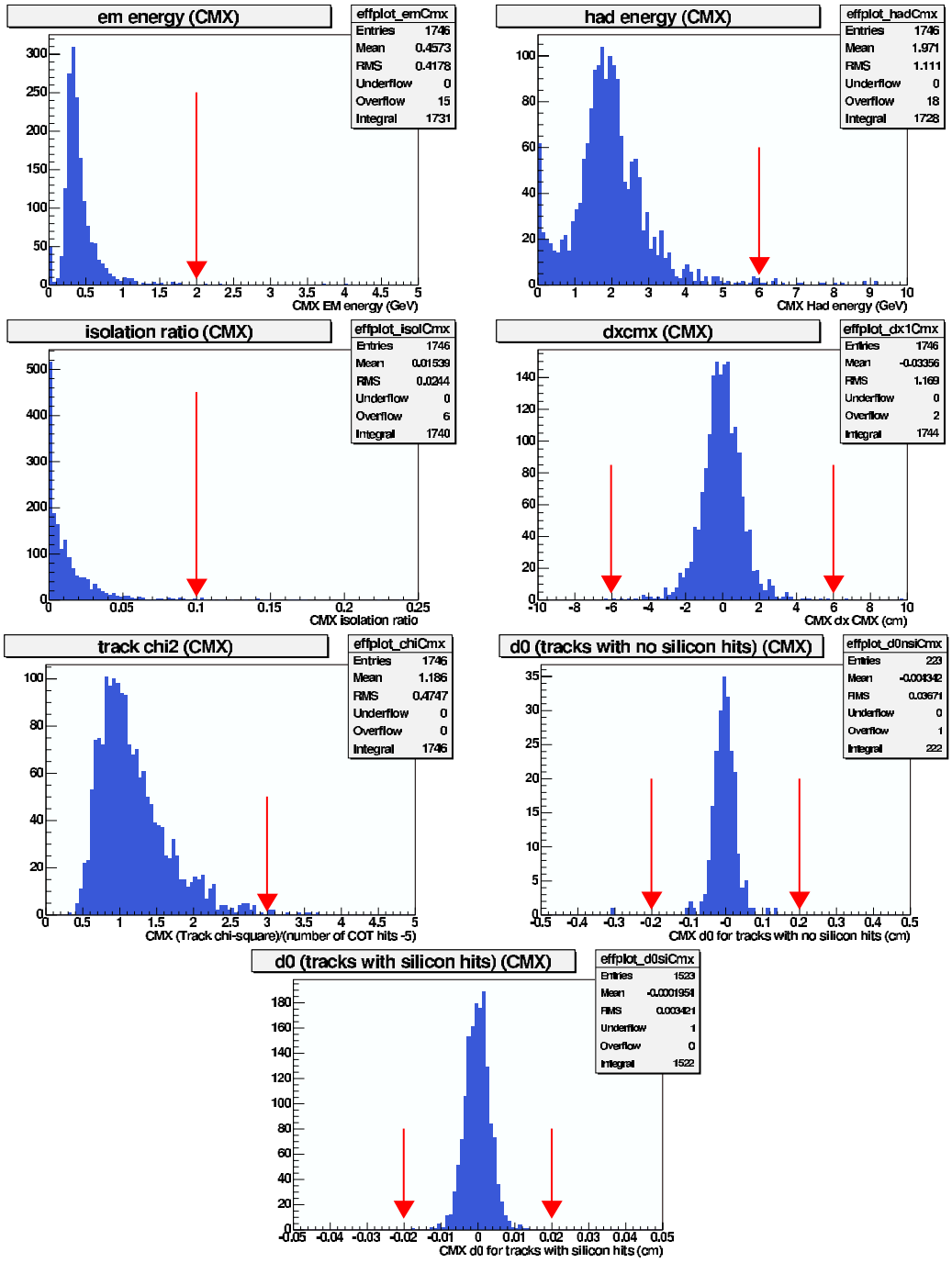


Figure 5.8: CMX ID Variables: EM Energy, Hadronic Energy, Isolation, Δx, Track χ^2, and impact parameter (d_0) without and with silicon hits [67]. The arrows show the selection requirement.

Muon Variable	**Base Cut**
p_T	≥ 20 GeV/c
E_T	< 2 GeV
Sliding addition for $E_{\rm Em}$	$0.115 \cdot (p - 100(\text{GeV}/c))$ for muons with $p > 100$ GeV/c
$E_{\rm Had}$	< 6 GeV
Sliding addition for $E_{\rm Had}$	$0.028 \cdot (p - 100(\text{GeV}/c))$ for muons with $p > 100$ GeV/c
Fractional isolation	≤ 0.10
COT axial superlayer hits	2
COT stereo superlayer hits	3
Hits per COT superlayer	5
Track z_0	≤ 60 cm
Impact parameter d_0	< 0.02 cm (w/ silicon) or < 0.2 cm (w/o silicon)
CMU $\lvert\Delta x\rvert$	< 7 cm, for CMUP
CMP $\lvert\Delta x\rvert$	< 5 cm, for CMUP
CMX $\lvert\Delta x\rvert$	< 6 cm, for CMX
Fiducial x_{CMP}	< 0 cm
Fiducial z_{CMP}	< 0 cm
Fiducial $\text{x}_{\text{CMU,CMUP,CMX}}$	< 0 cm
Fiducial $\text{z}_{\text{CMU,CMUP,CMX}}$	< -3 cm

Table 5.4: The standard Joint Physics selection criteria for CMUP and CMX muons [63]. Variables are explained in the text.

5.3.3 Track Selection

To find track lepton candidates, we begin with tracks in the `defTrack` collection produced by algorithms described in Chapter 4. We select tracks which have $p_T \geq$ 20 GeV/c. They are required to have 24 or more axial superlayer hits, 20 or more stereo super layer hits and 3 or more silicon hits (unless fewer than 3 silicon hits are expected). The impact parameter, d_0, of the track is required to be less than 0.025 cm for tracks with 3 or more silicon hits and 0.25 cm for tracks with fewer than 3 silicon hits. Tracks must be isolated with track based isolation greater than 0.9. These requirements are summarized in Table 5.5.

Track Variable	**Base Cut**
p_T	≥ 20 GeV/c
Track Isolation	> 0.9
COT axial superlayer hits	≥ 24
COT stereo superlayer hits	≥ 20
Silicon hits	≥ 3, unless < 3 expected
Impact parameter d_0	< 0.025 cm (≥ 3 silicon hits)
Impact parameter d_0	< 0.25 cm (< 3 silicon hits)
Track χ^2	Not used.

Table 5.5: The standard lepton+track top cross section measurement selection criteria for tight track leptons. We do not use the track χ^2 cut. For more information, see CDF Note 8696 [64].

For tracks which are paired with electrons an additional correction is applied to recover possible energy loss due to bremsstrahlung. If the electron radiates a photon, the track p_T will be reduced but the collinear photon will deposit energy into a nearby calorimetry tower and will contribute to the cluster E_T. To take advantage of this, we use the E_T if the E_T of the cluster associated with the track is bigger than the p_T of the track. We scale the track momentum four-vector by the ratio of E_T to p_T. We check

that energy loss from bremsstrahlung and its recovery is well modeled in Monte Carlo by comparing the effect of this correction on the reconstructed dielectron invariant mass in data and Pythia Z Monte Carlo. As seen in Figure 5.9, the reconstructed Z invariant mass in data and Monte Carlo match well before and after the correction.

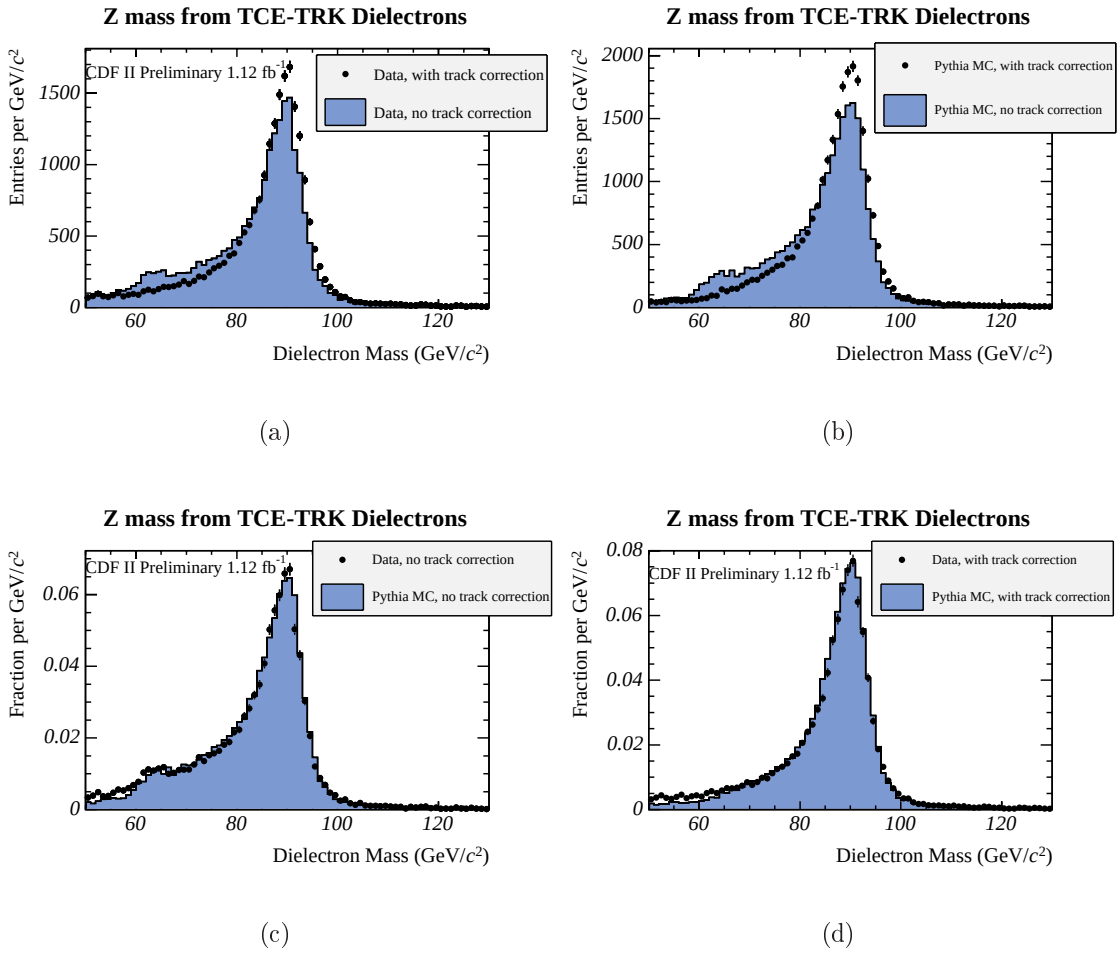


Figure 5.9: Track momentum correction for Z bosons reconstructed from tight central electrons and track leptons. In order to recover energy loss from bremsstrahlung, E_T of the cluster is used instead of the p_T of the track if $E_T > p_T$. The correction is well modeled in Pythia Z Monte Carlo simulation. (a) Comparison of Z invariant mass in bhelXd data before and after correction. (b) Comparison of Z invariant mass in the Pythia Monte Carlo simulation before and after correction. (c) Comparison of data and Pythia Monte Carlo simulation before correction. (d) Comparison of data and Pythia Monte Carlo simulation after correction.

5.4 Additional Selection Criteria

After the base selection, we apply further selection requirements on the event kinematics to improve background rejection. These criteria are optimized for the best expected limit. We impose additional constraints on the mass χ^2, event transverse mass, M_T, and a "sliding" transverse energy requirement on the four leading jets. The events which pass the optimized selection criteria are our signal candidate events. We further separate these events into two signal regions; those with a secondary vertex b-tag and those without. Since our signal is more likely to be tagged than our dominant Z+Jets background, these two signal regions will have different ratios of signal to background. We improve the expected limit by combining the information from both regions. The calculation of the expected limit and the optimization of the additional selection criteria are described in Chapter 8. The description of the additional criteria follows.

5.4.1 Mass χ^2

The signal we are interested in has no neutrinos, and, consequently, no (real) missing energy. Therefore, we can full reconstruct the event kinematics to separate signal from background. We have four masses to reconstruct in our events, namely, the Z, the W, the top from the $t \to qZ$ decay, and the top from the $t \to Wb$ decay. The Z is reconstructed from leptons and our event selection criteria already requires its invariant mass to fall within a window. Of the four jets in the event, two should form a W, which in turn should form a top with a third jet. The fourth jet should form a top with the reconstructed Z. From the reconstructed W and the two tops,

we can define a mass χ^2 as,

$$\chi^2 = \left(\frac{m_{W,\mathrm{rec}} - m_{W,\mathrm{PDG}}}{\sigma_{W,rec}}\right)^2 + \left(\frac{m_{t\to Wb,\mathrm{rec}} - m_{t,\mathrm{PDG}}}{\sigma_{t\to Wb}}\right)^2 + \left(\frac{m_{t\to qZ,\mathrm{rec}} - m_{t,\mathrm{PDG}}}{\sigma_{t\to qZ}}\right)^2, \tag{5.1}$$

where, for each permutation, the masses are obtained by the following prescription:

1. Correct the four leading jets four-vectors with level-5 jet energy scale corrections and top-specific corrections taken from the Top Mass Template (TMT) analysis [68].

2. Calculate the invariant mass of the first two jets to form the mass of the reconstructed W, $m_{W,\mathrm{rec}}$.

3. Vary the momentum four-vectors of both W daughters within their respective resolutions such that the W mass is fixed to its PDG value. The resolutions are taken from the TMT analysis.

4. Calculate the invariant mass of the W and the third jet to form the reconstructed top mass, $m_{t\to Wb,\mathrm{rec}}$.

5. Reconstruct a Z from two leptons.

6. Fix the Z mass to its PDG value by varying the two lepton four-vectors. We assume the leptons resolutions are constant percentages of their total momenta.

7. Calculate the invariant mass of the Z and the fourth jet, $m_{t\to qZ,\mathrm{rec}}$.

In the above calculation for the mass χ^2 distribution, we assume a top mass of 175 GeV/c^2. The widths used in the mass χ^2 are the RMS values of the HEPG-

matched reconstructed signal Monte Carlo, where $\sigma_{W,rec}$, $\sigma_{t\to Wb}$, and $\sigma_{t\to qZ}$ are 15 GeV, 24 GeV, and 21 GeV, respectively. The above prescription is repeated for all possible permutations of the four leading jets in the event, and the permutation with the lowest χ^2 is selected.

We do not use b-tagging information to reduce the combinatorics. We tried using b-tagging information and found that the separation between signal and background was worse than when we did not. This is most likely due to the fact that a significant fraction of the b-tags are in the charm jet from the $t \to cZ$ decay or a mistag in a light flavored jet, and using that jet as a b jet gives the wrong combination.

This mass χ^2 variable is the best discriminant between signal and background. We will also expand our control region by looking at events with high mass χ^2 in order to obtain a better estimate of the background contribution in our signal and blinded regions.

5.4.2 Transverse Mass

We expect top decays to be more central than events with an electroweak Z produced in association with jets. In order to take advantage of this, we construct the transverse mass of the full event, defined as

$$M_T = \sqrt{\left(\sum E_T\right)^2 - \left(\sum \vec{p}_T\right)^2}, \tag{5.2}$$

where the sums include the four leading jets and the reconstructed Z.

5.4.3 "Sliding" Jet Transverse Energy

The jets in our signal result from the decay of massive top quarks and are expected to be hard jets. We, therefore, expect these jets to have higher E_T than the softer jets from our backgrounds, some of which have to be the result of initial and final state radiation. We impose a "sliding" requirement on the transverse energy of the four leading jets, tighter energy requirements on the higher E_T jets in the event, to help suppress the backgrounds.

5.4.4 Heavy Flavor Jet Identification

For one of our signal regions, we require a heavy flavor jet identified by the loose SecVtx algorithm, to identify the b jet from the $t \to Wb$ decay or c jet from the $t \to qZ$ decay. (The other signal region requires that the events do not contain a b-tagged jet. Tagging effects shift events from one region to the other.) Tagging efficiency for heavy flavor jets as well as the rate for mistagging light flavor jets depend on track resolution and reconstruction efficiency. The Monte Carlo simulations we use are slightly optimistic in their modeling of these track parameters. Consequently, the Monte Carlo predicts a higher b-tag efficiency and a lower mistag rate than data. The data-based mistag prediction is described in Chapter 4. We apply the mistag parametrization to our Monte Carlo jets to correctly predict the mistag rate.

The b-tagging scale factor, the ratio of the tag rates in data and Monte Carlo, is measured in a heavy flavor enhanced sample. We select events where a jet contains a low p_T electron which recoils against a SecVtx tagged jet. This sample is 70% pure in heavy flavor. A complementary heavy flavor depleted sample is constructed

by selecting events where the low p_T electron in the jet is a conversion candidate. This allows for *in situ* background subtraction [69]. A similar measurement using jets containing muon fit $p_T^{rel} \equiv \overrightarrow{p}_T(\mu) \cdot \widehat{p}_T(jet)$ to measure the background [70]. The combined loose SecVtx correction factor is 0.95 ± 0.05 [53].

5.5 Lepton+Jets Selection

Our expected number of signal events is normalized to the measured lepton+jets SecVtx cross section, in order to extract a branching fraction. To correct for our FCNC signal falling within the lepton+jet acceptance, we apply the lepton+jets event selection criteria to the FCNC signal Monte Carlo sample and measure the acceptance. The lepton+jets analysis [**?**] measures the cross section for the standard model top decay $t\overline{t} \rightarrow W^+ \; b \; W^- \; \overline{b}$, where one W decays leptonically to a lepton and neutrino and the other decays hadronically to two jets, resulting in the lepton + three or more jets + missing E_T event signature. The lepton+jets event selection requires one tight lepton (TCE, CMUP, CMX), missing $E_T > 20$ GeV, three jets with $E_T > 20$ GeV corrected to jet energy correction level 5, and at least two loose SecVtx b-tags. These selection criteria are listed in Table 5.6.

Selection Criteria	**Base Cut**
Lepton type	Tight central leptons: TCE, CMUP, CMX
Number of tight leptons	exactly 1
Number of jets	≥ 3
Jet E_T, JCL 5	≥ 20 GeV
Missing E_T	≥ 30 GeV
Z veto	Yes
Dilepton Veto	Yes
$\|z\|$ jet vertex	≤ 60 cm
Δz lepton-jet vertex	≤ 5 cm
H_T	≥ 200 GeV
Number of loose SECVTX tags	≥ 2

Table 5.6: Selection criteria for the double loose SecVtx tag $t\bar{t}$ cross section analysis [71].

Chapter 6

Signal Expectation

Nature is wont to hide herself.
Heraclitus (from "On the Universe")

Our signal expectation is quantified through both a pretag acceptance and an efficiency to find one or more secondary vertex tags (for our base selection criteria). The pretag acceptance includes the geometric acceptance, the Z reconstruction efficiency, and the efficiency to find four or more jets. The acceptance and efficiency measured in the Monte Carlo are corrected to better represent expected data. The following sections describe the signal MC generation and the calculations of acceptance and efficiency for the signal.

6.1 Event Generation

The FCNC signal Monte Carlo samples were generated using the PYTHIA event generatior, version 6.126 [72]. We assume a top mass of 175 GeV/c^2. The simulation is "run dependent", taking into account the beam and detector conditions of each data

taking run, and covers the 1.12 fb^{-1} run range (runs 141544-222426). The underlying event is also simulated and overlaid onto our signal. The number of underlying events is based on the instantaneous luminosity in each run.

The FCNC decay $t \to qZ$ is not in the standard set of decays in PYTHIA. We have forced this decay in PYTHIA by hijacking the $t \to Wd$ decay, which has a negligible branching fraction, and redefining the decay products to a Z boson and a c or u quark. We have chosen to redefine a channel with a very small branching fraction so that the total decay width of the top quark will not be affected by the extra FCNC decay. The consequence of this redefinition is that PYTHIA losses the (V-A)x(V-A) interactions for the decay of the top and the Z, and decays both isotropically in their rest frames. We remedy this by reweighting the events according to the helicity of the decay vertex as described in the next section.

For the main FCNC signal sample, we force one top quark into our FCNC decay $t \to cZ$ and the other top quark to the SM decay $t \to Wb$ letting PYTHIA decide randomly whether the t or $\overline{t}$ decays to a Z and a c quark. This decay is the dominant contribution to the FCNC signal acceptance. We have generated 0.1 events per 0.5 nb^{-1}, yielding a total of 539,445 signal events. The decay of the Z is forced into e^+e^- or $\mu^+\mu^-$ pairs, and the decay of the W is forced into $W \to q\overline{q}'$.

In addition to the main FCNC signal MC sample, we have generated smaller samples, to study various effects for events with an FCNC decay. We generated a sample to study the additional acceptance gained from events in which both top quarks decay via the FCNC decay, or the "double FCNC decay". In the double FCNC sample, each of the two Z bosons is allowed decay to e^+e^-, $\mu^+\mu^-$ or $q\overline{q}$ pairs. No

$t\bar{t}$ **Decays**	**Sample Name**	**Sample Size**	**Description**
$ZcWb$	$Z(ll)W(q\overline{q}')$	539,445	Signal Monte Carlo Sample: $Z \to e^+e^-$ or $Z \to \mu^+\mu^-$, and $W \to q\overline{q}'$
$ZcWb$	$Z(ll)W(l\nu)$	111,181	$Z \to e^+e^-$ or $Z \to \mu^+\mu^-$ and $W \to e\nu$, $W \to \mu\nu$, or $W \to \tau\nu$
$ZcWb$	$Z(incl.)W(incl.)$	116,573	Inclusive Z and W decays
$ZcZc$	$Z(ll, q\overline{q})Z(ll, q\overline{q})$	116,573	Double FCNC decay: $Z \to e^+e^-$, $Z \to \mu^+\mu^-$ or $Z \to q\overline{q}$
$ZuWb$	$Z(ll)uW(q\overline{q}')$	116,573	$Z \to e^+e^-$ or $Z \to \mu^+\mu^-$ and $W \to q\overline{q}'$

Table 6.1: List of signal Monte Carlo samples generated. The abbreviation "incl" refers to the inclusive decay of the boson shown.

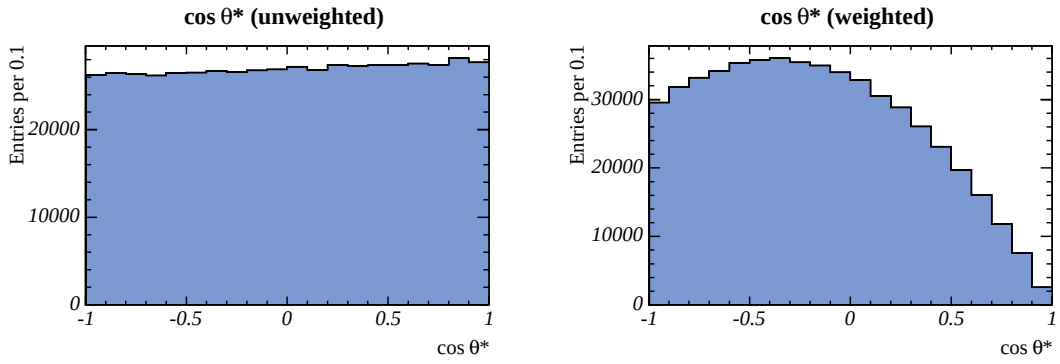
correlation was forced between the decays of the two Z bosons. The events where one Z decays to e^+e^- or $\mu^+\mu^-$ and the other to $q\overline{q}$ contribute to our acceptance.

Although our channel of interest is where the Z decays to e^+e^- or $\mu^+\mu^-$ and W decays to $q\overline{q}'$, we can gain additional acceptance from events where the Z and W decay to other particles. To measure this additional acceptance, we generated a sample in which the Z and the W forced to decay leptonically, and another sample in which we allow the fully inclusive decays of the Z and the W. The former sample was for the channel that contributed the most additional acceptance. The latter sample is also used to study how much of the FCNC signal will fall within the lepton+jets SecVtx cross section acceptance.

We also generated a sample in which the c quark is replaced by a u quark to study the reduction in the b-tagging efficiency. Each sample was generated with 0.1 events per $1\,\mathrm{nb}^{-1}$, for a total of approximately 110,000 events per sample, and spans the full $1.12\,\mathrm{fb}^{-1}$ run range. All signal MC samples are listed in more detail in Table 6.1.

6.1.1 Helicity of the Z Boson

Due to the fact that we have redefined the $t \to Wd$ decay channel and redefined the decay products of the FCNC decay, PYTHIA does not know the type of interaction involved at the decay vertex. Hence PYTHIA decays the particles isotropically in their rest frame, i.e. flat in $\cos\theta^*$, the cosine of the angle between the top boost direction and the lepton with the same charge sign as the top. Fig. 6.1a shows the $\cos\theta^*$ distribution of the FCNC top decays in the FCNC signal MC. Note that the distribution for $\cos\theta^*$ is only approximately flat (3% slope). We correct for this by fitting a straight line to the distribution and applying a correction factor of $1/(1 + 0.032\cos\theta^*)$.



(a) (b)

Figure 6.1: Distribution of $\cos\theta^*$ for the FCNC signal Monte Carlo sample: (a) before re-weighting, (b) after re-weighting according to a Z helicity of 65% longitudinal and 35% left-handed.

The helicity of the W from a SM top decay is determined by the fact that the longitudinal degree of freedom of the W is acquired from the Higgs field. Any theory beyond the SM must provide a similar mechanism to create massive gauge bosons. Therefore, the best guess for the longitudinal fraction of the Z helicity, f^0, is given

by the SM prediction [73]:

$$f^0 = \frac{{m_t}^2}{2m_Z^2 + {m_t}^2} \approx 0.65. \tag{6.1}$$

The Higgs mechanism does not predict the fraction of left-handed to right-handed helicity. However, for a Z decay, changing from left-handed to right-handed is equivalent to looking at the oppositely charged lepton. Since the CDF detector is approximately symmetric with respect to positively and negatively charged leptons, the acceptance is very similar when switching from a left-handed to a right-handed decay.

For all acceptance calculations, we re-weight the signal MC sample such that the Z helicity is 65% longitudinal and 35% left/right-handed. For the double FCNC decay helicity weight, we use the product of the weights for the two $t \to cZ$ decays. Our ignorance of the exact nature of the interaction is taken into account by a assigning a corresponding systematic uncertainty, as will be described in Section 9.1.4. The overall correction to the pretag acceptance due to the helicity reweighting is $\sim$4.3%.

6.2 Acceptance

The acceptance for the FCNC signal MC is determined by applying the base event selection (Z candidate and four jets with $E_T > 15$ GeV and $|\eta| < 2.4$) to the signal MC and then correcting for some known limitations of the MC simulation. For all FCNC signal samples, we reweight the whole sample for the Z helicity before acceptance calculations. The numerator of the acceptance is defined as the number of reconstructed events in the $Z+ \geq 4$ jets region after helicity re-weighting and applying

all scale factors and efficiencies. We define the denominator of the acceptance as the sum of the helicity weights for events with generated Z masses between 76 GeV/c^2 and 106 GeV/c^2.

6.2.1 Acceptance Corrections

As discussed in the last chapter, the lepton identification efficiencies are different in data and Monte Carlo simulations. Scale factors have been measured to account for these differences. These scale factors are not only different for the different detectors used to identify the leptons, but they also vary among the MC simulations corresponding to different data taking periods. We correct for their affect on the Z reconstruction efficiency. We apply these scale factors to each lepton. Tight leptons which have tracks that pass our track selection criteria they may fall within the acceptance as a tight lepton or a track. Each event receives a scale factor, $\mathcal{SF}$:

$$\mathcal{SF} = \mathcal{SF}_{L1} \cdot \mathcal{SF}_{L2} + (1 - \mathcal{SF}_{L1}) \cdot \mathcal{SF}_{L2} \cdot \mathcal{SF}_{T1} + (1 - \mathcal{SF}_{L2}) \cdot \mathcal{SF}_{L1} \cdot \mathcal{SF}_{T2}. \quad (6.2)$$

In the above equation, $\mathcal{SF}_{L1}$ and $\mathcal{SF}_{L2}$ are the scale factors for the leptons, and $\mathcal{SF}_{T1}$ and $\mathcal{SF}_{T2}$ are the track scale factors for the tracks matched these leptons. The added efficiency from recovering both tight leptons as tracks appears in the last two terms of the above equation. Note that for the case when neither lepton has been matched to a track which passes our track lepton selection criteria, both $\mathcal{SF}_{T1}$ and $\mathcal{SF}_{T2}$ are zero, and the above equation reduces to the more familiar event scale factor

MC Sample Name	**Acceptance**	**BR(Z)**	**BR(W/Z)**	**Acc×BR**
$Z(ll)W(q\overline{q}')$	15.95±0.05	6.73±0.01	67.60±0.27	0.73±0.01
$Z(ll)W(l\nu)$	2.50±0.05	6.73±0.01	32.40±0.27	0.05±0.02
$Z(ll,q\overline{q})Z(ll,q\overline{q})$	2.50±0.05	$(6.73+69.91)^2$		1.47±0.04

Table 6.2: The corrected acceptances for the FCNC signal Monte Carlo samples in %, shown with statistical uncertainties only.

for a Z formed from two tight leptons:

$$\mathcal{SF} = \mathcal{SF}_{L1} \cdot \mathcal{SF}_{L2}. \tag{6.3}$$

We apply the lepton trigger efficiencies in a similar manner. We assume that all tight leptons may fire the trigger, and that phoenix electrons and tracks do not. Per event, we assign a weight for the trigger efficiency, $\mathcal{E}$:

$$\mathcal{E} = \mathcal{E}_{L1} + \mathcal{E}_{L2} - \mathcal{E}_{L1} \cdot \mathcal{E}_{L2}. \tag{6.4}$$

In the above efficiency equation, $\mathcal{E}_{L1}$ and $\mathcal{E}_{L2}$ are the trigger efficiencies for leptons. The lepton ID scale factors and lepton trigger efficiencies are given in Table 5.1 in Chapter 5. The corrected acceptances for our signal Monte Carlo samples are given in Table 6.2.

6.3 Tagging Efficiency

Simply taking the fraction of b-tagged events in the MC simulation as the tagging efficiency is insufficient. As discussed in Chapter 5, the tagging efficiency in data is

lower than in the MC simulation for heavy flavor jets, and the mistag rate, i.e. the fraction of light flavor jets that are falsely tagged, is higher in data than in the MC simulation. We correct for these discrepancies in the data and MC tag rates on a jet-by-jet basis.

We identify heavy flavor jets in the MC simulation by matching b and c hadrons from the list of observed particles (OBSP) to reconstructed jets in the event. If a b or c hadron is within an (η, ϕ) cone of $\Delta R < 0.4$ of a jet, we consider that jet to be a heavy flavor jet, otherwise a jet is classified as a light flavor jet. Heavy flavor jets that are not b-tagged in the MC are considered untagged. We re-weight the number of tagged heavy flavor jets by the loose SecVtx b-tagging scale factor, 0.95 ± 0.05. We assume that the b-tagging scale factor is the same for b and c jets.

For light flavor jets we ignore the tagging information from the MC simulation altogether. Instead we get an estimate for the background from mistagged light flavor jets by applying the Gen6 mistag parameterization [74, 54]. The mistag parameterization contains the probability for a light jets to be tagged as a function of the jet properties (E_T, $|\eta|$, number of good tracks), and event properties (the number of tags, the number of Z vertices, and the v_z of the primary vertex). The values of the entries in the mistag matrix are determined from negative tags in generic dijet data. The mistag parameterization includes the $\alpha\beta$ correction [75, 76], which accounts for asymmetries in the positive and the negative tag rates due to decays of long-lived K^0_S and Λ particles, photon conversions, and material interactions.

To account for the scale factor and mistag corrections, we assign a weight to each MC event that represents the probability that at least one of the jets in the event is

tagged, either as a genuine heavy flavor jet or as a mistagged light flavor jet:

$$\begin{aligned} P_{\text{event,tag}} &= 1 - \prod_i \text{probability that jet } i \text{ is not tagged} \\ &= 1 - \prod_j (1 - P_{\text{mistag},j}) \cdot \prod_k (1 - \mathcal{SF}_k) \cdot \prod_l 1. \end{aligned} \tag{6.5}$$

The index i runs over all jets, j runs over all light flavor jets, k runs over all tagged and matched heavy flavor jets, and l runs over the remaining non-tagged but matched heavy flavor jets. P_{mistag}, j is the mistag probability, and $\mathcal{SF}_k$ is the tagging scale factor. The final event tagging efficiency of a given MC sample is the sum of all per-event weights divided by the total number of events.

For signal events satisfying our base selection criteria, the event tagging efficiencies and, consequently, the efficiencies for the signal events to be anti-tagged are given in Table 6.3, for the different signal Monte Carlo samples.

We want to set a limit for $t \to qZ$ while our main signal Monte Carlo samples are $t \to cZ$. Since the tagging rate for c and u are different, we check for the effect of having Zc versus Zu on the tagging efficiency for the signal. We check the loose SecVtx b-tag rates in the $Z(ll)uW(q\overline{q}')$ sample and find that the tagging efficiency for this sample is $92 \pm 1\%$ of the tagging efficiency for the$Z(ll)W(q\overline{q}')$ sample. The tagging efficiencies for the two samples are similar because the tag rate is dominated by the presence of the b jet. (Note: For this analysis, we require one or more tags. If we had required two tags or one and only one tag, the difference between havin a u quark instead of a c quark would be more pronounced. Since $t \to cZ$ is expected to dominate over $t \to uZ$, we assign the difference as a systematic uncertainty as

	Tag Efficiency	**Anti-tag Efficiency**
$Z(ll)W(q\overline{q}')$	54.6±0.2%	45.4±0.1%
$Z(ll)W(l\nu)$	51.9±1.0%	48.1±1.0%
$Z(ll)uW(q\overline{q}')$	38.2±0.9%	61.8±0.9%

Table 6.3: Loose SecVtx tagging efficiencies for signal Monte Carlo samples.

described in Section 9.1.7.

6.4 Branching Fraction Calculation

In the FCNC signal MC simulation, we generate $t\overline{t}$ events with one top decaying via the FCNC mode and the other via the standard model mode (or FCNC mode). All of the acceptances and efficiencies so far have been for finding these $t\overline{t}$ events. However, ultimately, we are interested in setting a limit on $\mathcal{B}(t \to qZ)$. We have to properly translate the acceptances for $t\overline{t} \to ZcWb$ and $t\overline{t} \to ZcZc$ to extract the limit for $t \to qZ$. In order to do this, we assume top can only decay to Wb or Zq, i.e. $\mathcal{B}(t \to Wb) + \mathcal{B}(t \to Zc) = 1$. The following table shows the possible decay modes for the $t\overline{t}$ pair.

	$\mathcal{B}(t \to Wb)$	$\mathcal{B}(t \to Zc)$
$\mathcal{B}(t \to Wb)$	$\mathcal{B}_W \cdot \mathcal{B}_W$	$\mathcal{B}_W \cdot \mathcal{B}_Z$
$\mathcal{B}(t \to Zc)$	$\mathcal{B}_W \cdot \mathcal{B}_Z$	$\mathcal{B}_Z \cdot \mathcal{B}_Z$

where

$$\begin{aligned} \mathcal{B}_Z &\equiv \mathcal{B}(t \to Zc), \\ \mathcal{B}_W &\equiv \mathcal{B}(t \to Wb) = 1 - \mathcal{B}_Z(\text{assumption}), \text{ and} \\ \mathcal{B}_W \cdot \mathcal{B}_W &\text{ is the SM } t\bar{t} \text{ background.} \end{aligned}$$

The mode where both top quarks decay to Wb is a background. We can derive the probabilites $\mathcal{P}$ for one or both of the top quarks in the $t\bar{t}$ pair to decay to Zq and derive the number of expected events, namely,

$$\begin{aligned} \mathcal{N}_{\text{signal}} &= [(\mathcal{P}(t\bar{t} \to WbZq) \cdot \mathcal{A}_{WZ}) + (\mathcal{P}(t\bar{t} \to ZqZq) \cdot \mathcal{A}_{ZZ})] \cdot \sigma_{t\bar{t}} \cdot \int \mathcal{L}dt \\ &= \left(2\mathcal{B}_Z(1-\mathcal{B}_Z) \cdot \mathcal{A}_{WZ} + \mathcal{B}_Z^2 \cdot \mathcal{A}_{ZZ}\right) \cdot \sigma_{t\bar{t}} \cdot \int \mathcal{L}dt \\ &= \mathcal{B}_Z \underbrace{\left(2(1-\mathcal{B}_Z) + K_{ZZ/WZ} \cdot \mathcal{B}_Z\right)}_{\text{Acceptance Correction Factor}} \cdot \overbrace{\mathcal{A}_{WZ}}^{\text{From MC}} \cdot \sigma_{t\bar{t}} \cdot \int \mathcal{L}dt, \end{aligned}$$

where

$$\begin{aligned} \mathcal{A}_{WZ} &\equiv \text{Acceptance } \cdot \text{ efficiency for } t\bar{t} \to WbZc, \text{ and} \\ \mathcal{A}_{ZZ} &\equiv \text{Acceptance } \cdot \text{ efficiency for } t\bar{t} \to ZcZc, \text{ and} \\ K_{ZZ/WZ} &\equiv \text{Ratio of } \mathcal{A}_{ZZ} \text{ to } \mathcal{A}_{WZ}. \end{aligned}$$

The "acceptance correction" is the additional acceptance from the double FCNC

decay to the main channel $t\bar{t} \to WbZq$.

The number of signal events depends on the $t\bar{t}$ production cross section. We use the top cross section measured in the lepton+jets samples requiring SecVtx tags. The lepton+jets SecVtx cross section top cross section measurement assumes that the top quark can only decay by the dominant SM decay mode, $t \to Wb$. Consequently, we have to adjust the lepton+jets top cross section to allow for possible FCNC decay, $t \to qZ$. In doing so, the effective $t\bar{t}$ production cross section enhanced, due to the fact that the acceptance for both tops decay to Wb is reduced by the factor, $(1-\mathcal{B}_{\rm z})^2$.

We also account for the FCNC signal falling within the lepton+jets SecVtx top cross section acceptance. We apply the top cross section event selection to the $t\bar{t}$ $\to WbZc$ signal MC sample with inclusive decays of the W and Z, $Z(incl.)W(incl.)$, and to the double FCNC MC sample, $t\bar{t} \to ZcZc$ $(Z(ll, q\overline{q})Z(ll, q\overline{q}))$, to measure the acceptances. The algebra for this is shown below.

The top cross section is calculated as [71]:

$$\sigma_{t\bar{t}\,\text{“Lepton + Jets”}} = \frac{\mathcal{N}_{LJ} - B_{LJ}}{\mathcal{A}_{LJ} \cdot \int \mathcal{L} dt},$$

where $\mathcal{A}_{LJ}$ is acceptance for the lepton+jets selection convoluted with the corresponding efficiency. We modify the acceptance to allow for the FCNC decay:

$$\begin{aligned} \mathcal{A}_{LJ} &= (1-\mathcal{B}_{\rm z})^2 \cdot \mathcal{A}_{LJ_{\rm ww}} + 2 \cdot \mathcal{B}_{\rm z}(1-\mathcal{B}_{\rm z}) \cdot \mathcal{A}_{LJ_{\rm wz}} + \mathcal{B}_{\rm z}^2 \cdot \mathcal{A}_{LJ_{\rm zz}} \\ &= \mathcal{A}_{LJ_{\rm ww}} \cdot \left[(1-\mathcal{B}_{\rm z})^2 + 2 \cdot \mathcal{B}_{\rm z}(1-\mathcal{B}_{\rm z}) \cdot \mathcal{R}_{\rm wz/ww} + \mathcal{B}_{\rm z}^2 \cdot \mathcal{R}_{\rm zz/ww}\right] \end{aligned}$$

where

$$\mathcal{R}_{\mathrm{wz/ww}} \equiv \frac{\mathcal{A}_{LJ\,\mathrm{wz}}}{\mathcal{A}_{LJ\,\mathrm{ww}}}$$
$$\mathcal{R}_{\mathrm{zz/ww}} \equiv \frac{\mathcal{A}_{LJ\,\mathrm{zz}}}{\mathcal{A}_{LJ\,\mathrm{ww}}}$$

This gives an enhancement to the top cross section as shown below.

$$\text{“Lepton + Jets”} = \frac{\mathcal{N}_{LJ} - B_{LJ}}{\mathcal{A}_{LJ} \cdot \int \mathcal{L} dt}$$
$$= \overbrace{\frac{1}{(1-\mathcal{B}_{\mathrm{z}})^2 + 2 \cdot \mathcal{B}_{\mathrm{z}}(1-\mathcal{B}_{\mathrm{z}}) \cdot \mathcal{R}_{\mathrm{wz/ww}} + \mathcal{B}_{\mathrm{z}}^2 \cdot \mathcal{R}_{\mathrm{zz/ww}}}}^{\text{Enhancement Factor}} \cdot \underbrace{\frac{\mathcal{N}_{LJ} - B_{LJ}}{\mathcal{A}_{LJ\,\mathrm{ww}} \cdot \int \mathcal{L} dt}}_{\text{Standard Cross Section}}$$

We want this enhancement factor to be as large as possible. If the FCNC decay channels do not contribute to the acceptance of the cross section measurement, i.e. $\mathcal{R}_{\mathrm{wz/ww}}$ and $\mathcal{R}_{\mathrm{zz/ww}}$ are zero, the enhancement factor would be at its largest, $1/(1\text{-}\mathcal{B}_{\mathrm{z}})^2$). If the FCNC decay channels are as likely to be accepted by the lepton+jets event selection as the SM decay, i.e. $\mathcal{R}_{\mathrm{wz/ww}}$ and $\mathcal{R}_{\mathrm{zz/ww}}$ are 1.0, the enhancement factor would be equal to 1, and we would get no enhancement from using the measured $t\bar{t}$ cross section.

Inserting this top cross section back into our calculation for the number of events

gives:

$$
\begin{aligned}
\mathcal{N}_{\text{signal}} &= \mathcal{B}_Z\left(2(1-\mathcal{B}_Z)+K_{ZZ/WZ}\cdot\mathcal{B}_Z\right)\cdot\mathcal{A}_{WZ}\cdot\sigma_{t\bar{t}}\cdot\int\mathcal{L}dt \\
&= \mathcal{B}_Z\left(2(1-\mathcal{B}_Z)+K_{ZZ/WZ}\cdot\mathcal{B}_Z\right)\cdot\mathcal{A}_{WZ}\cdot \\
&\quad \frac{1}{(1-\mathcal{B}_{\text{z}})^2+2\cdot\mathcal{B}_{\text{z}}(1-\mathcal{B}_{\text{z}})\cdot\mathcal{R}_{\text{wz/ww}}+\mathcal{B}_{\text{z}}^2\cdot\mathcal{R}_{\text{zz/ww}}}\cdot\frac{\mathcal{N}_{LJ}-B_{LJ}}{\mathcal{A}_{LJ_{\text{ww}}}\cdot\int\mathcal{L}dt}\cdot\int\mathcal{L}dt \\
&= \mathcal{B}_Z\cdot(\mathcal{N}_{LJ}-B_{LJ})\cdot\frac{\mathcal{A}_{WZ}}{\mathcal{A}_{LJ_{\text{ww}}}}\cdot\underbrace{\frac{\left(2\cdot(1-\mathcal{B}_Z)+K_{ZZ/WZ}\cdot\mathcal{B}_Z\right)}{(1-\mathcal{B}_{\text{z}})^2+2\cdot\mathcal{B}_{\text{z}}(1-\mathcal{B}_{\text{z}})\cdot\mathcal{R}_{\text{wz/ww}}+\mathcal{B}_{\text{z}}^2\cdot\mathcal{R}_{\text{zz/ww}}}}_{\text{Full Running Acceptance Correction}}\cdot
\end{aligned}
\tag{6.6}
$$

This equation shows that including the FCNC decay mode to the lepton+jets top cross section yields an acceptance correction expression that is based on the limit we set for the FCNC decay. We say that this acceptance correction "runs" with the $t \to qZ$ limit.

The lepton+jets acceptances $\mathcal{A}_{LJ_{\text{ww}}}$, $\mathcal{A}_{LJ_{\text{wz}}}$, and $\mathcal{A}_{LJ_{\text{zz}}}$ and the FCNC acceptances $\mathcal{A}_{WZ}$ and $\mathcal{A}_{ZZ}$ in the above calculation are measured in the Monte Carlo simulation samples. We chose to normalize to the loose double tag SecVtx cross section which minimizes the ratio of the FCNC to SM lepton+jets acceptance, and therefore maximizes the enhancement factor; $\mathcal{R}_{\text{wz/ww}} = 32\%$ and $\mathcal{R}_{\text{zz/ww}} = 1.96\%$. As shown in Figure 6.2, normalizing to the lepton+jets cross section and algebraically acquiring the enhancement factor increases the expected limit on top's FCNC branching fraction. For more details on the expected limit calculation, refer to Chapter 8. The effects of normalizing to the loose *single* tag SecVtx cross section instead of the double tag cross section are discussed in Appendix C.2.

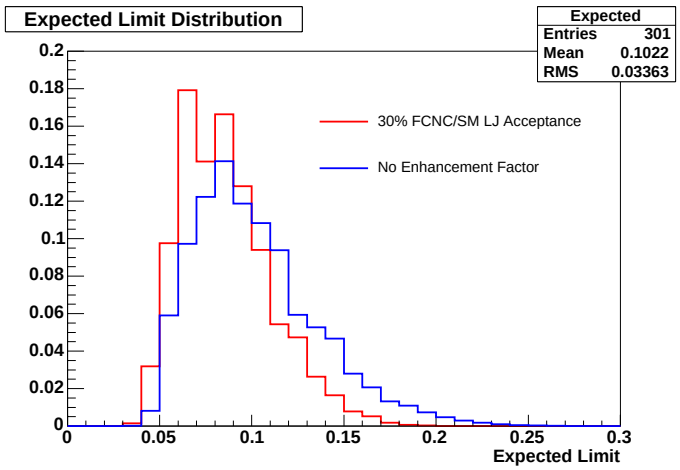


Figure 6.2: The expected limit distributions for the FCNC acceptance with and without the enhancement factor gained from normalizing to the lepton+jets cross section.

Additionally, normalizing to the measured top cross section removes our dependence on luminosity uncertainties. Using the ratio of acceptances, many of our systematic uncertainties also partially cancel; however, we assign a systematic uncertainty for the statistical and systematic uncertainties on the number of lepton+jets candidates and background estimate from the top cross section measurement.

Chapter 7

Background Processes

The man who is swimming against the stream knows the strength of it.
Woodrow Wilson

There are several standard model processes that have event signatures that satisfy our base selection criteria. This chapter describes these processes and the methods to estimate their contributions to our blinded region, pre-tag and with a loose SecVtx tag. These processes have cross sections much higher than our FCNC signal and need to be suppressed. The procedure for optimization of event selection to further reduce background compared to signal and, consequently, improve the expected limit, will be described in Chapter 8.

The dominant background contribution to this analysis comes from Z bosons produced in association with jets (Z+Jets). This background is also the most difficult to estimate because it is not well modeled by the Monte Carlo simulations available to us. Therefore, we use data and Monte Carlo to predict the contribution from this background process. The second process which contributes to the background is the

standard model decay of $t\bar{t}$ pairs, $t\bar{t} \to W^+ \, b \, W^- \, \bar{b}$. Although there are no real Z bosons in this process, the two leptons in the dilepton decay mode or the lepton and a jet faking a lepton in the lepton+jets decay mode can have an invariant mass within our Z mass window. Although this background contribution is small in the pre-tag sample, it has a much higher tag rate than other backgrounds because there are two real b-jets in the final state. It makes a significant contribution to the background event in our tagged sample, which is more sensitive to signal. The third background, similar in size to the standard model $t\bar{t}$ background, is from dibosons which has a real Z in the events, WZ and ZZ. Finally, very small contributions come from Ws produced in association with jet (W+Jets) and from WW production, which do not have real Z bosons.

7.1 Z + Jets

The vast majority of the background events in our blinded region are from Z bosons produced in association with jets. This background is also not well modeled by Monte Carlo, adding to the difficulty of accurately predicting it. We use a combination of the ALPGEN v2.10 MC generator [56]—using PYTHIA for parton showers—and data from control regions to estimate the Z+jets background. Much time and effort has gone into validating and tuning the ALPGEN + PYTHIA Monte Carlo as described further in Appendix A. As a result, the kinematic distributions match well between data and simulation. However, ALPGEN is a leading order event generator and underpredicts the inclusive Z cross section by ∼40% citeAcosta:2004uq. We take the overall normalization from inclusive Z bosons in data. It also under-

predicts the fraction of events with higher jet muliplicities, as shown in Figure 7.2, which is problematic, since we are interested in the number of Z+$\geq$4 jets events. We correct for this in two ways, by extrapolating the underprediction in the Z+3 jets events into the blinded region and by estimating the total Z+$\geq$4 jets background from the high tail of the mass χ^2 distribution.

We also need to measure the b-tag rate for the Z+jets background. The secondary vertex tags come from two samples; real heavy flavor jets which are tagged and light flavor tags which are mistagged. We measure the tag rate for Z+$\geq$4 jets events using both Monte Carlo simulations and data. We estimate mistags using the Gen6 mistag parameterization and scale heavy flavor tags by the b-tagging scale factor on a jet-by-jet basis. These numbers are combined to a per-event tagging probability as we did for signal. We check the predicted tagging fractions in data in events with a Z and three or fewer jets and in the high mass χ^2 region in Z+$\geq$4 jets events.

7.1.1 Monte Carlo Samples

We use ALPGEN v2.10 + Pythia to model the Z+jets background. ALPGEN models the hard scattering using leading order matrix element calculations and Pythia models the parton showers. We combine light flavor $Z + 0, 1, 2, 3, 4$ partons samples and heavy flavor, $Z + b\overline{b} + 0, 1, 2$ partons and $Z + c\overline{c} + 0, 1, 2$ partons, samples according to their generated cross sections. The Z bosons in these samples are forced to decay to e^+e^- or $\mu^+\mu^-$. The samples along with their cross sections and number of events are listed in Table 7.1. The different n-parton samples can populate the same Z+n-jet phase space, e.g. a Z+3p event can have an extra jet from the

parton shower and be a Z+4 jet event, overlapping with a Z+4p event which also has four jets. ALPGEN v2.10 contains a built-in mechanism to remove this overlap ("MLM matching"). It matches partons to generator level jets and removes overlapping events from the different n-p samples. The samples with the largest parton multiplicities, i.e. $Z + 4$ partons, $Zb\overline{b} + 2$ partons, and $Zc\overline{c} + 2$ partons, are generated using "inclusive" matching, i.e. each parton is required to match a generator level jet but extra jets unmatched to partons are allowed. For all other samples we used "exclusive" matching, i.e. each parton is required to match a jet and vice versa. The details of MLM matching can be found in Apendix A.

The light flavor samples contain heavy flavor from the parton shower (gluon splitting) and also have massless charm at the matrix element level, leading to an overlap between the light and heavy flavor samples. ALPGEN does not address this overlap, so we need to remove it by hand. We apply the jet-based overlap removal scheme (a.k.a. "Harvard overlap removal") developed for the SECVTX top cross section analysis [71], as summarized in Table 7.1.1. All light flavor events with massless (matrix element) charm are removed. For massive bottom and charm, the guiding principle is that soft, collinear $b\overline{b}$ and $c\overline{c}$ pairs are better modeled by the PYTHIA parton showers while ALPGEN is better at modelling hard, well separated heavy flavor pairs. This leads to a prescription that keeps events with $b\overline{b}$ or $c\overline{c}$ pairs in the light flavor sample only if they come from the parton shower (in PYTHIA: `STDHEP=2`) and are contained in the same reconstructed jet ($\Delta R < 0.4$). In the heavy flavor samples, all events with $b\overline{b}$ and $c\overline{c}$ pairs from the matrix element (`STDHEP=3`) that do not share the same jet are kept. This heavy flavor overlap removal is described in detail in Appendix A.

Sample	**Dataset Name**	**cross section (pb)**	**Number of Events**
$Z \to e^+e^- + 0\text{p}$	`ztopp0`	158	2,639,520
$Z \to e^+e^- + 1\text{p}$	`ztop1p`	21.6	536,159
$Z \to e^+e^- + 2\text{p}$	`ztop2p`	3.47	536,159
$Z \to e^+e^- + 3\text{p}$	`ztop3p`	0.550	528,491
$Z \to e^+e^- + 4\text{p}$	`ztop4p`	0.099	525,065
$Z \to \mu^+\mu^- + 0\text{p}$	`ztopp0`	158	2,665,104
$Z \to \mu^+\mu^- + 1\text{p}$	`ztop6p`	21.6	536,159
$Z \to \mu^+\mu^- + 2\text{p}$	`ztop7p`	3.47	530,843
$Z \to \mu^+\mu^- + 3\text{p}$	`ztop8p`	0.550	536,159
$Z \to \mu^+\mu^- + 4\text{p}$	`ztop9p`	0.099	536,159
$Z \to e^+e^- + b\overline{b} + 0\text{p}$	`ztopb0`	0.511	532,205
$Z \to e^+e^- + b\overline{b} + 1\text{p}$	`ztopb1`	0.134	525,955
$Z \to e^+e^- + b\overline{b} + 2\text{p}$	`ztopb2`	0.039	405,652
$Z \to \mu^+\mu^- + b\overline{b} + 0\text{p}$	`ztopb5`	0.511	530,793
$Z \to \mu^+\mu^- + b\overline{b} + 1\text{p}$	`ztopb6`	0.134	525,695
$Z \to \mu^+\mu^- + b\overline{b} + 2\text{p}$	`ztopb7`	0.039	536,159
$Z \to e^+e^- + c\overline{c} + 0\text{p}$	`ztopc0`	1.08	699,861
$Z \to e^+e^- + c\overline{c} + 1\text{p}$	`ztopc1`	0.331	710,734
$Z \to e^+e^- + c\overline{c} + 2\text{p}$	`ztopc2`	0.107	663,518
$Z \to \mu^+\mu^- + c\overline{c} + 0\text{p}$	`ztopc5`	1.08	710,734
$Z \to \mu^+\mu^- + c\overline{c} + 1\text{p}$	`ztopc6`	0.331	710,734
$Z \to \mu^+\mu^- + c\overline{c} + 2\text{p}$	`ztopc7`	0.107	705,108

Table 7.1: The dataset names, cross sections, and number of events for the Z+jets samples. The first block gives the Z+Light Flavor samples and the second and third block give the Z+Heavy Flavor samples.

Sample	**No HF**	**Matrix Element $b\bar{b}/c\bar{c}$**		**Parton Shower $b\bar{b}/c\bar{c}$**	
		$\Delta R < 0.4$	$\Delta R \geq 0.4$	$\Delta R < 0.4$	$\Delta R \geq 0.4$
LF	O	X	X	O	X
$c\bar{c}$	—	X	O	O	O
$b\bar{b}$	—	X	O	O	O

Table 7.2: Heavy flavor overlap removal scheme for ALPGEN +Pythia Z+jets Monte Carlo samples. In each sample, events which have criteria marked with an 'X' are removed, and events with criteria marked with an 'O' are kept.

7.1.2 Pre-tag prediction

The distribution of the number of jets in events with a reconstructed Z obtained from combining the above samples is shown in Figure 7.1. Although the sample weights were chosen such that the samples were normalized to a luminosity of 1.12fb^{-1}, we cannot use the prediction for number of events with $Z+ \geq 4$ jets directly since ALPGEN, a leading order MC generator, underestimates the inclusive Z cross section. The ALPGEN v2.10 generated inclusive Z cross section is 185.9 pb whereas the CDF measured cross section for $Z \to \ell^+\ell^-$ is $\sigma \times \mathcal{B}(p\overline{p} \to Z/\gamma^* \to \ell^+\ell^-) = 254.9 \pm 16.2$ pb [77]. In order to get an initial estimate of the number of Z+jets background events expected in the $Z+ \geq 4$ jets region, we normalize our sample to the number of inclusive Z candidates in data. We restrict the normalization to events with three or fewer jets in order not to unblind the analysis. In the full dataset of 1.12 fb^{-1} we find 104,245 events with a Z and three or fewer jets. ALPGEN predicts the ratio of events with four or more jets over events with three or fewer jets to approximately 0.001. This yields a first estimate of approximately 105 events with a Z and four or more jets in 1.12 fb^{-1} of data. We improve this estimate by checking the n-jet distribution predicted by the Monte Carlo against data.

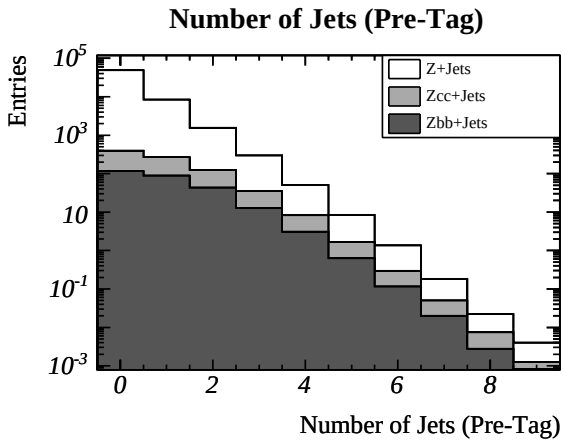
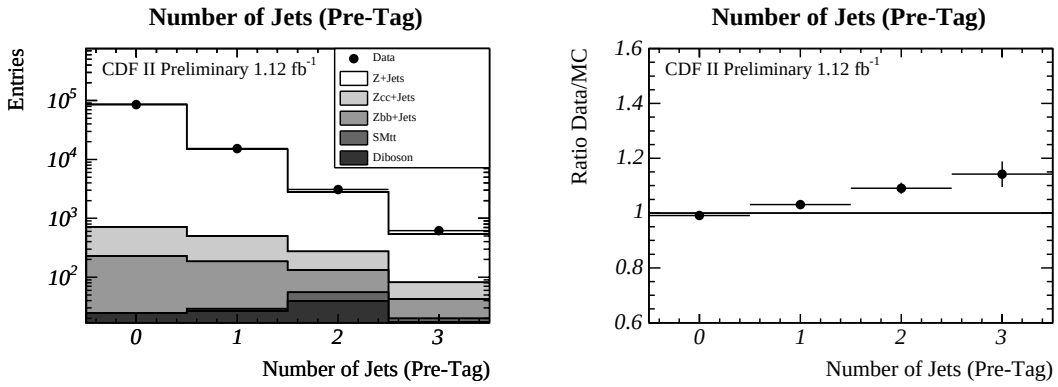


Figure 7.1: Distribution of the number of jets in events with a reconstructed Z from the combination ALPGEN Z+jets Monte Carlo samples.

In Fig. 7.2, a comparison of the predictions of the ALPGEN, SM $t\bar{t}$, and diboson MC samples are shown with data for events with a Z and three or fewer jets. For Z + 3 jets events, the MC simulation underestimates the number of jets by 15%; assuming the same deficit in the $Z+ \geq 4$jet bin, we estimate approximately 120 total background events in our signal region. The concern is that the rate of extra jet production is low, and this will potentially lead to an even larger underestimate in the number of Z+jets background events in the signal region. We have developed a more data-driven method to estimate the total $Z+\geq$4 jets background based on the mass χ^2 variable described in Chapter 5. This will be described in Section 7.5.

We verify the pre-tag kinematic distributions modelled by the ALPGEN + Pythia against kinematic distributions in data. We use Z+3 jets events because they are closest in jet multiplicity to our blinded region. The kinematic variables used in our optimization, E_T of jets and transverse mass, for these events are shown in Figure 7.3. Data and Monte Carlo simulations agree well.



(a) (b)

Figure 7.2: Data-MC comparison of the number of jets in events with a reconstructed Z. (a) Distribution of the number of jets before b-tagging. (b) Ratio of data over MC. The Z+jets MC samples are normalized to the 0–3-jet bins, and contributions from SM $t\bar{t}$ and diboson productions are added according to their predicted cross sections.

7.1.3 Tagging Rates

We estimate the tag rate in the Z+jets samples to predict their tagged contribution to the background. The secondary vertex tags come from both real heavy flavor (b and c) jets and from mistagged light flavor jets. We measure the tag rates in Monte Carlo according to the prescription described in Chapter 6 and check them against the control region (Z+1,2,3 jets events) in data.

To estimate the tag rate from Z bosons produced in association with heavy flavor jets, we begin by measuring the fraction of events with heavy flavor in the Z+jets Monte Carlo sample. The Feynman diagrams for the production of Z bosons in association with heavy flavor jets are shown in Figure 7.4. The heavy flavor events in the Z+jets Monte Carlo come from both the heavy flavor, $Z + b\bar{b} + 0, 1, 2$ partons and $Z + c\bar{c} + 0, 1, 2$ partons, samples and the light flavor, $Z + 0, 1, 2, 3, 4$ partons samples,

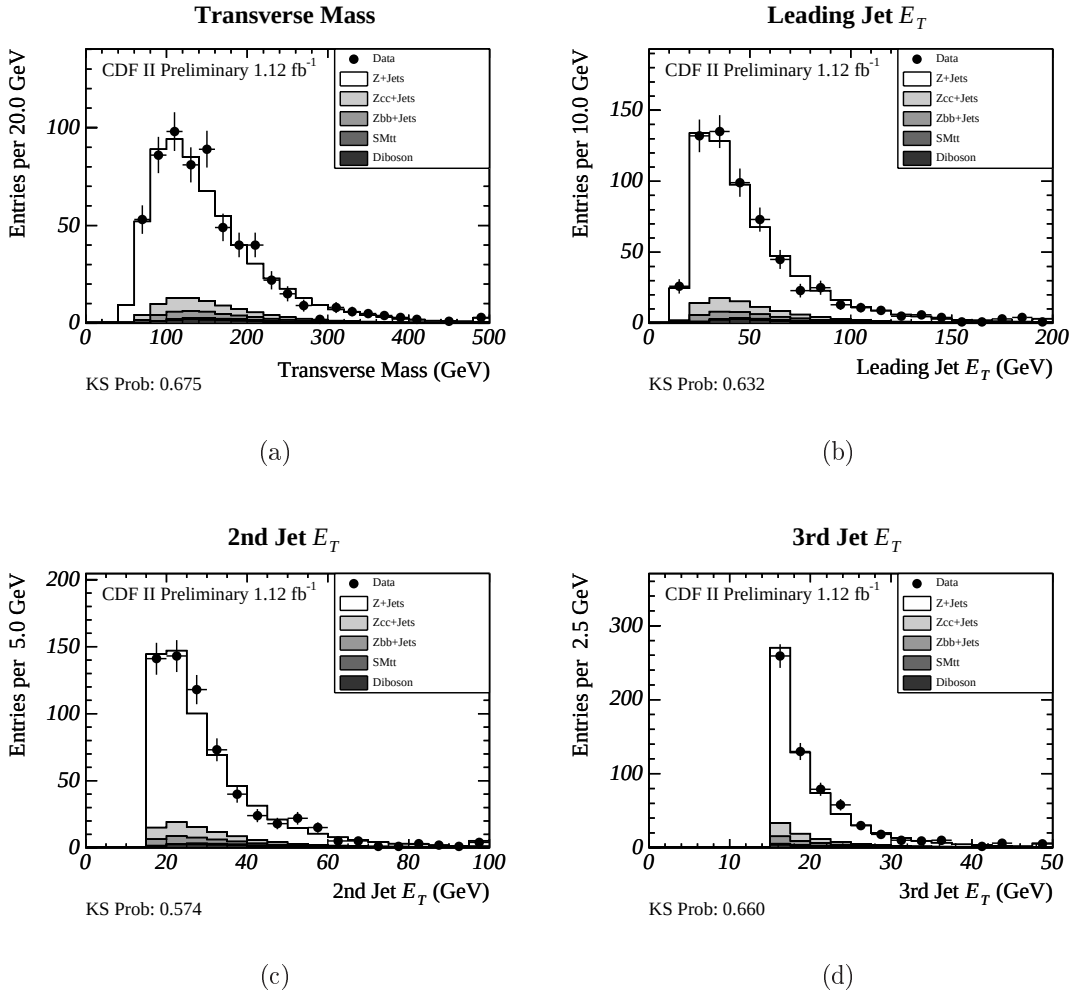
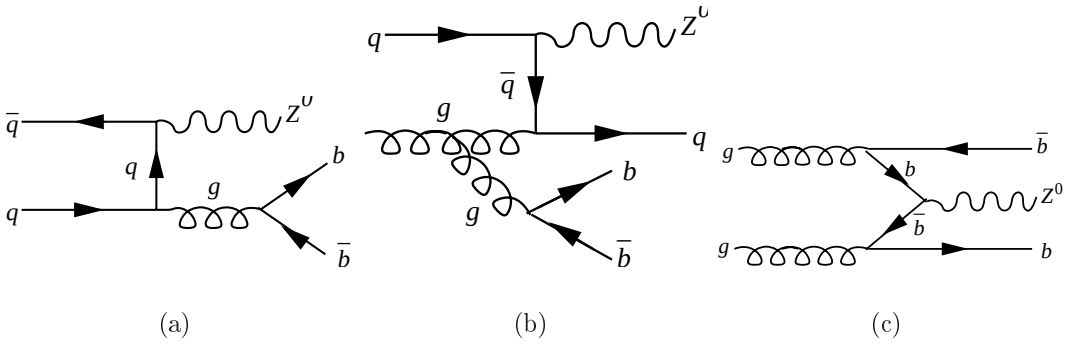


Figure 7.3: Data-MC comparison of kinematic variables in Z+3 jet events. The data are from high E_T (p_T) electron (muon) triggers. The Z+Jets, Zcc+Jets, and Zbb+Jets are from ALPGEN v2.10 + PYTHIA Monte Carlo samples. The SM top and dibosons are from PYTHIA Monte Carlo samples. (a) Transverse mass. (b)-(d) Jet transverse momenta. The expected backgrounds are normalized to the data event yields.



(a) (b) (c)

Figure 7.4: Feynman diagrams for Z+HF production.

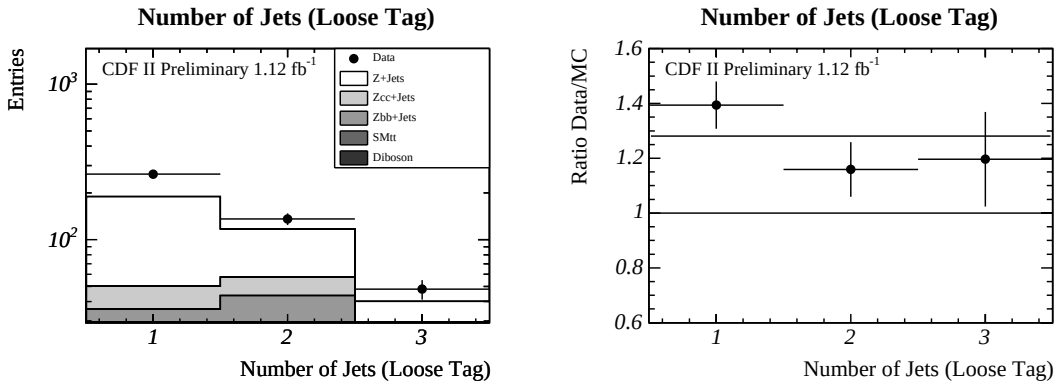
according to the prescription described in Section 7.1.1. We identify jets containing real heavy flavor by matching them to b and c hadrons in the observed particle (OBSP) list within a cone of $\Delta R < 0.4$. We measure the fraction of events which have one heavy flavor jet or two heavy flavor jets, since the events with two heavy flavor jets have two chances to get tagged. The heavy flavor fractions measured in Monte Carlo simulations are given in Table 7.3. We cross-checked these fractions against the heavy flavor fractions in W+Jets events [71] and found them to be ∼30% higher. This is consistent with theoretical expectations. Of the Feynman diagrams shown in Figure **??**, diagrams (a) and (b) are also responsible for heavy flavor production in W+Jets events while diagram (c) is only for Z+jets events, leading to extra heavy flavor in the Z+jets sample.

As described in Chapter 6, we take the tags from heavy flavor jets and scale them down by the loose SecVtx scale factor. To find the mistag rate from light flavor jets, we apply the Gen 6 mistag parametrization to light flavor jets in Monte Carlo (jets unmatched to b and c hadrons). The per-jet mistag rate depends on several event as

Sample	**1-jet (%)**	**2-jet (%)**	**3-jet (%)**	**4-jet (%)**
$Zb\bar{b}$, 1 b	0.83±0.01	1.54±0.03	2.37±0.05	3.30±0.10
$Zb\bar{b}$, 2 b	—	0.96±0.02	2.08±0.05	4.24±0.45
$Zc\bar{c}$, 1 c	1.90±0.03	3.65±0.07	5.48±0.14	7.66±0.81
$Zc\bar{c}$, 2 c	—	1.41±0.02	3.27±0.14	5.94±0.17

Table 7.3: Z+Heavy Flavor fractions in ALPGEN Monte Carlo samples. This table gives the fraction of Z+jets events (in %) that contain heavy flavor jets, for each physical process, sorted by the amount of heavy flavor and number of jets. Only statistical errors are given.

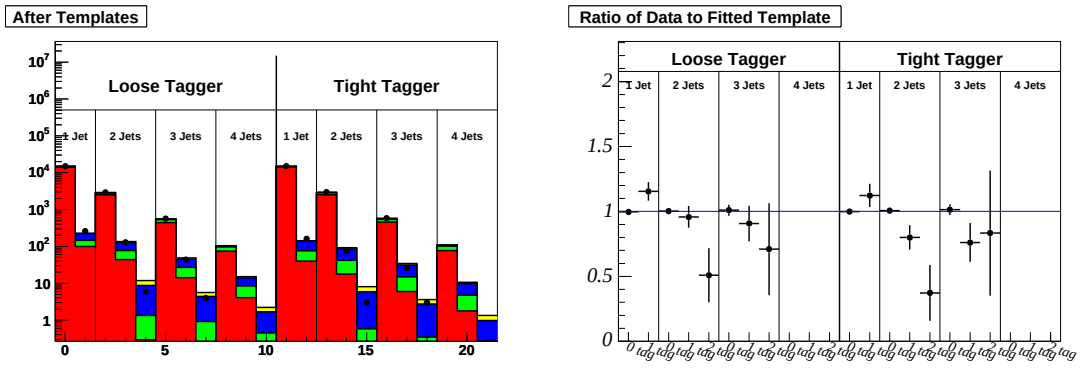
well as jet kinematic variables, such as the total transverse energy in the event, jet E_T, and number of tracks in the jet [54]. Since we have verified that the Monte Carlo models the event kinematics well, we can expect the mistag matrix parametrization applied to Monte Carlo light flavor jets to represent the mistag rate in data. The event tag probability is calculated according to the equation given in Chapter 6. We compare the event tag rates predicted by Monte Carlo against event tag rates in Z+1,2,3 jet events. We also include the contribution from SM $t\bar{t}$ and dibosons in this comparison. We normalized the tagged events to the pre-tag events in data, keeping the SM $t\bar{t}$ and diboson contributions fixed and scaling up the Z+jets contribution. Figure 7.5 shows data-MC comparison of the number of jets distribution for events with a Z and a loose SecVtx tag. It is evident that Monte Carlo systematically underpredicts the number of tags, compared to data. Naively fitting a straight line to data-MC ratios in Figure 7.5 (b) suggests that the tag rate in data is ∼28% higher than Monte Carlo in the Z+1,2,3 jets events. Extrapolating to the Z+≥4 jets region, where the tag rate predicted by Monte Carlo is 12%, we expect 15.4% tag rate. We can fit to the data in a more sophisticated manner using tagging templates.



(a) (b)

Figure 7.5: Data-MC comparison of the number of tagged events with a Z boson and 1–3 jets. The tagged events are normalized to the pre-tag number of events. (a) Distribution of the number of jets for at least one loose SECVTX b-tag. (b) Ratio of data to MC. Note that the MC distribution is normalized to the 0–3-jet bins of the Z+jets sample in the pre-tag data sample. A naive straight line fit indicates that the tag rate should be 28% higher than the Monte Carlo prediction.

We construct tagging templates which contain the tagging probability predictions from MC, separated by both the number of jets and the number of b-tags in each event. We do this for all background samples and for both the loose and tight tagger. Fitting to the templates allows subsets of these degrees of freedom to float. Knowing that the MC does not predict the jet multiplicities well, we let every jet bin float separately. The contribution of SM $t\bar{t}$ is fixed, while the light flavor Z+jets sample and the combination of the $Zc\bar{c}$+jets and the $Zb\bar{b}$+jets samples are allowed to float. The data together with the fitted template are shown in Fig. 7.6. We extract the tagging rate by comparing the fraction of MC events with at least one loose tag in the four jet bin after the fit. We extract a tagging rate of 14% from this method. Section 7.5 describes the more data-driven method of extracting tag rates directly



(a) (b)

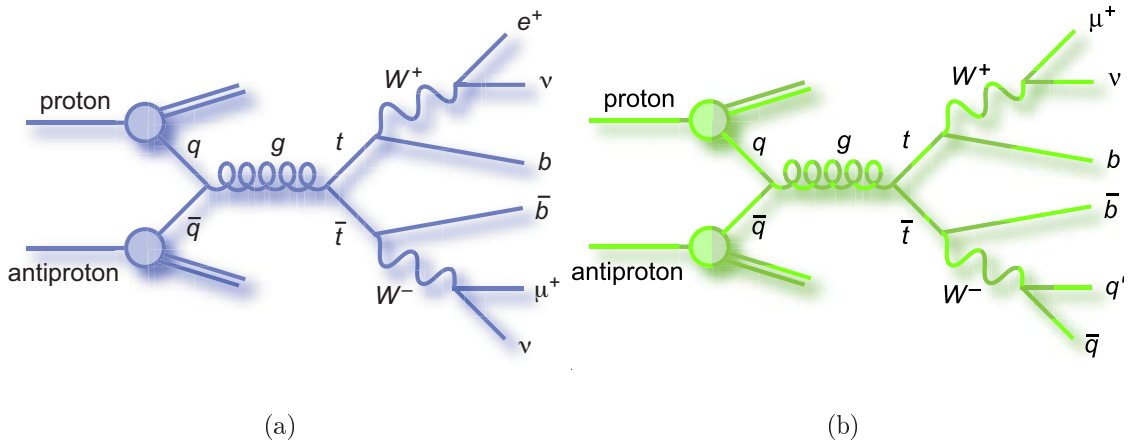
Figure 7.6: Extraction of tag fractions from a fit to tagging templates. (a) Tagging templates fitted to data. (b) Ratio of data to fitted template.

from data in $Z+\geq 4$ jets events with high mass χ^2 and how we combine the methods to get the tag rate prediction for the total background in our blinded region.

7.2 Standard Model $t\bar{t}$

A source of background for our analysis is the standard model decay of $t\bar{t}$ pairs, $t\bar{t} \rightarrow W^+ \; b \; W^- \; \bar{b}$. Although these events do not have real Z bosons, a Z candidate can be falsely reconstructed from the two leptons when both Ws decay leptonically (dilepton events) or, on rare occasions, from the lepton and a jet faking a lepton when one W decays leptonically and the other decays hadronically (lepton+jets events). The Feynman diagrams for these processes are shown in Figure 7.7. Since we allow the second lepton to be an isolated track, event dilepton events where one lepton is an electron and the other is a muon may satisfy our selection criteria.

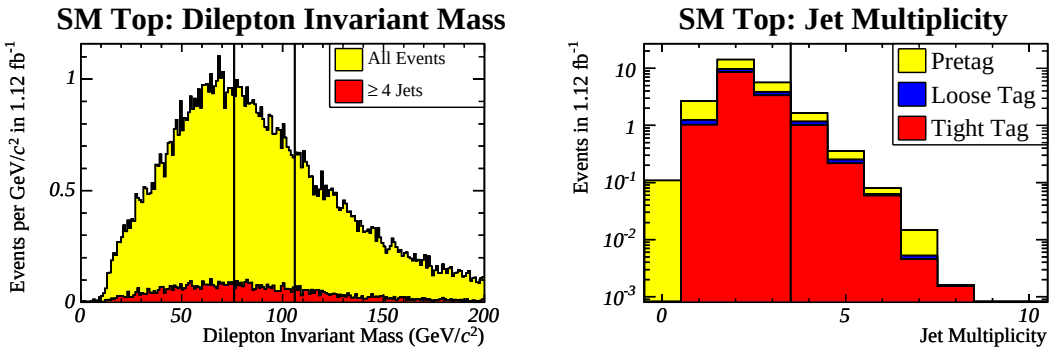
We estimate this background from Monte Carlo simulations, using 4.7 million $t\bar{t}$



(a) (b)

Figure 7.7: Feynman diagrams for standard model $t\bar{t}$ decay. (a) Dilepton decay process where both Ws decay leptonically. (b) Lepton+jets decay process where one W decays leptonically and the other decays hadronically.

events generated with the PYTHIA generator at a top quark mass of 175 GeV/c^2 (MC sample: `ttop75`). We normalized our signal to the measured lepton+jets SecVtx $t\bar{t}$ production cross section, 8.8 $\pm$ 1.1 pb [71]. We do the same when measuring our standard model $t\bar{t}$ background for an integrated luminosity of 1.12 fb^{-1}. The normalized dilepton reconstructed mass distribution from the PYTHIA $t\bar{t}$ Monte Carlo simulation is shown in Figure 7.8a. As expected, there is no peak at the Z mass, but there are events which have a reconstructed dilepton mass which falls within our Z mass window. The yellow histogram in Figure 7.8b shows the distribution of the number of jets for events inside the Z mass window. Most of these events have fewer than four jets. This is because in dilepton events, there are only two other hard jets from the decay of the $t\bar{t}$ pair and in lepton+jets events, one jet is misidentified as a lepton so there are only three jets left over. To pass the requirements of four or more jets, the events need to have extra jets from initial state and final



(a) (b)

Figure 7.8: (a) Reconstructed dilepton mass for all standard model $t\bar{t}$ PYTHIA MC events and events with four or more jets. The vertical lines show our Z mass selection window. (b) Jet multiplicity for standard model $t\bar{t}$ MC events before b-tagging and using the loose and tight SecVtx tagger. The events for the loose SecVtx tagger are the increase compared to the tight SecVtx tagger. The events to the right of the vertical line fall within our base selection criteria.

state radiation. Therefore, the $\geq$ 4 jet requirement rejects more than 90% of this background. Figure 7.8a shows the reconstructed dilepton mass also after requiring $\geq$4 jets. This leaves us with 2.4 ± 0.3 expected events in the $Z+ \geq 4$ jets region in 1.12 fb^{-1}, rendering SM $t\bar{t}$ a much smaller background than Z+jets (123.3 ± 28 events).

Although the contribution of SM $t\bar{t}$ events to the untagged background is small, there are two b-jets in these events. The probability for finding a b-tag is therefore much higher compared to Z+jets, and even FCNC signal. As described for the FCNC signal MC in Section 6.3, the b- and c-jet tags are scaled down by the scale factor, and light flavor tags are added according to mistag parameterization predictions to produce the expected tagging efficiency in data. In 1.12 fb^{-1} of data, we expect

1.7 ± 0.2 SM $t\bar{t}$ background events with a loose SECVTX b-tag (compared to an expected number of 17.6 ± 6 tagged Z+jets events for the loose SECVTX tagger). These results and the MC sample used are summarized in Table 7.5.

7.3 Dibosons

We expect background contributions on the same order as the SM $t\bar{t}$ production from the production of pairs of gauge bosons ("dibosons"). The Feynman diagrams for these processes are shown in Figure 7.9. ZZ and WZ events contain real Z bosons that can decay into leptons while the other boson is allowed to decay to a pair of jets, including heavy flavor jets. Within detector mass resolution, it is difficult to distinguish between a Z and a W from the dijet invariant mass. There can also be additional jets from initial and final state radiation which will make these events fall in our signal region of $Z + 4$ jets. WW diboson production does not contain a real Z in the final state and makes a negligible contribution to Z+$\geq$4 jets events.

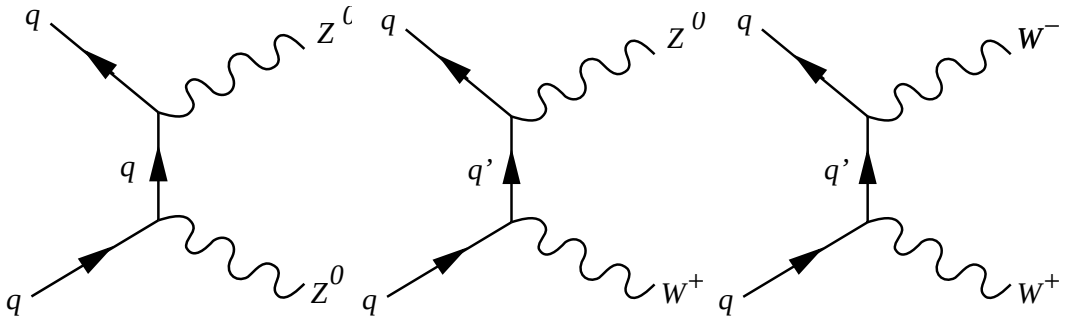


Figure 7.9: Feynman diagrams for ZZ, WZ, and WW production.

We used MC events simulated with PYTHIA, ∼2 M events per sample, and normalized to the theoretical cross sections for diboson production to estimate this background. The theoretical cross sections and their uncertainties are obtained from the MCFM event generator using MRS98 parton distribution functions [78]. To measure the tagged background, we scaled down the MC heavy flavor tags by the loose SecVtx scale factor and add light flavor tags according to the mistag parametrization, as we have done for the signal and other backgrounds. The Monte Carlo samples, theoretical cross-sections, and background contributions from diboson production are listed in Table 7.5.

7.3.1 $p + \overline{p} \rightarrow ZZ$

ZZ events satisfy our base selection criteria when on Z decays leptonically, the other decays hadronically, and there are two or more extra jets from initial and/or final state radiation. The theoretical cross section for ZZ production [78], 1.56 pb, was calculated for two on-shell Z bosons, with a width of 2.49 GeV , while the PYTHIA Monte Carlo sample we use requires only one Z on-shell and the other Z is generated to contain the Drell-Yan contribution, with γ^*/Z mass down to 0 GeV. The masses of the two generated Z bosons are shown in Figure 7.10. We calculate the effective cross-section for these events by finding the fraction of events with two on-shell Z bosons, approximately 1/3, and scaling up the theoretical cross-sections accordingly.

We calculate the fraction of events with two on-shell Z bosons in two different ways and they give similar results. The first is to simply take the fraction of events

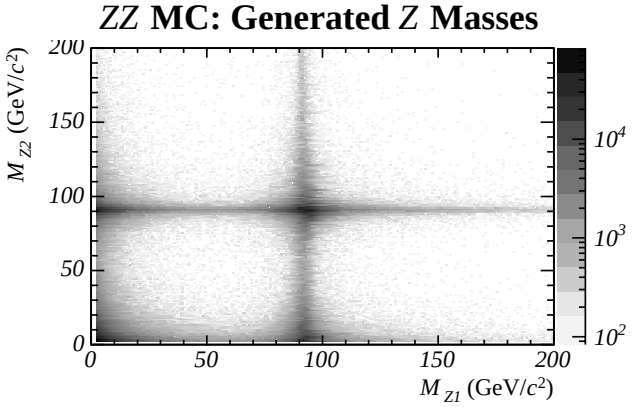


Figure 7.10: The masses of the two generated Z bosons in the PYTHIA ZZ Monte Carlo sample. One of the Z bosons is required to be on-shell (width of 2.49 GeV) and the other can have mass down to 0 GeV/c^2

where both Z bosons have a generated mass between 76 GeV/c^2 and 106 GeV/c^2. The second way is to fit a Breit-Wigner distribution to the two dimensional Z mass distribution, which gave a Z mass of 91.18 GeV/c^2 and a width of 2.54 GeV/c^2, and to normalize to it. Integrating this Breit-Wigner distribution between 76 GeV/c^2 and 106 GeV/c^2 gives the same result as integrating the Z mass distribution. These results are summarized in Table 7.4. We take the average of the two methods as our effecive ZZ production cross section and half the difference as the uncertainty, namely 4.50 $\pm$ 0.25 pb. The ZZ events contribute 2.29$\pm$0.17 $Z+$ $\geq$4 jets events pre-tag and 0.50$\pm$0.05 events after requiring a loose SecVtx tag. Although the ZZ pre-tag contribution is similar to WZ pre-tag contribution, ZZ events are roughly twice as likely to be tagged because Z decays to $b\bar{b}$ 15% of the time.

Area of interest	**Percentage of Events**	**Scaled Theoretical Cross Section(pb)**
Integral of 2D Histogram (76-106) GeV/c^2	32.9%	4.74
Normalization of Breit-Wigner $(-\infty,+\infty)$	36.7%	4.25
Integral of 2D Breit-Wigner (76-106) GeV/c^2	32.9%	4.75

Table 7.4: The fraction of PYTHIA ZZ events with two on-shell Z bosons and the effective cross sections using different methods.

7.3.2 $p + \overline{p} \rightarrow WZ$

WZ events satisfy our base event selection criteria when the Z decays leptonically, W decays hadronically, and there are two or more extra jets from initial and/or final state radiation. The theoretical cross section for WZ production is 3.96±0.06 pb [78]. WZ production contributes 1.98±0.14 events to our $Z+ \geq 4$ jets events before tagging and 0.22±0.02 events after requiring a loose SecVtx tag.

7.3.3 $p + \overline{p} \rightarrow WW$

Although WW has the biggest cross-section of the three diboson processes, 13.00 ± 0.05 pb [78], it does not contain a real Z and contributes a negligible amount to the background. There are no real Z bosons in these events. A Z candidate is misreconstructed from the two leptons in the case where both Ws decay leptonically or from a lepton and a jet when one W decays leptonically and the other decays hadronically. WW production contributes 0.07±0.02 events to the pre-tag background and 0.01±0.01 events to the tagged background.

Sample	Dataset Name	Cross Section (pb)	Generated Events	Events Pre-Tagged	Events w/ Loose SECVTX b-Tagged
SM $t\bar{t}$	ttop75	8.8±1.1	4,719,385	2.4±0.3	1.7±0.2
WZ	itopwz	3.96±0.06	2,340,145	2.0±0.1	0.2±0.1
ZZ	itopzz	4.50±0.25	2,323,812	2.3±0.2	0.5±0.1
WW	itopww	13.00±0.05	2,284,862	< 0.1	negligible

Table 7.5: The dataset names, cross sections, number of events in the MC sample, and expected number of events in 1.12 fb^{-1} for both the SM $t\bar{t}$ and the diboson samples. Also included in the uncertainty calculation is the 6% luminosity uncertainty.

7.4 W + Jets

The production of Ws in association with jets, W+jets, also contribute to our background. W+jets events where the W decays leptonically (22% branching fraciton for $W \to e\nu$ and $W \to \mu\nu$), a jet which fakes a lepton and the two have a reconstructed invariant mass within the Z mass window (76 GeV/c^2 to 106 GeV/c^2) contribute events to our blinded region. W+jets events are abundant at CDF; the ratio of the cross sections of $W+(n+1)$ jets to $Z+n$ jets is approximately 60% [79]. Fortunately, the probability for a jet to fake a lepton is small and further requirements of the invariant mass window, opposite sign leptons, and $\Delta(z_0)$ between the two leptons decrease this background far below Z+jets.

We estimate the W+jets events in e^+e^- and $\mu^+\mu^-$ final states using several different methods. For dimuon events, it is sufficient to estimate the number of background events from the number of same-sign muon pairs (or muon-track pairs) that fall into the Z mass window. We found only 15 same-sign Z candidates in a sample corresponding to 700 pb^{-1} of integrated luminosity. None of the events contained four or more jets; therefore, we conclude that the background from $W \to \mu\nu$ + jets events is negligible for our analysis. For the electron sample, using same-sign Z candidates

over-estimates the W+jets background since real $Z \to e^+e^-$ decays can form same-sign Z candidates. There is a significant contribution to the same-sign Z candidates from phoenix electrons whose charge sign was misidentified. Bremsstrahlung photons radiated from one or both electrons from the Z decay can be converted into e^+e^- pairs in the detector material ("trident electrons"), such that three electron tracks can be reconstructed per Z leg. One of the electrons has opposite charge to the original electrons. We use both the data and the MC simulation to predict the number of same sign $Z \to e^+e^-$ events we expect.

The main technique to estimate the W+jets background in $Z \to e^+e^-$ is to study same-sign electron-muon pairs with an invariant mass falling into the our Z mass window. The dominant source of same-sign $e\mu$ pairs is W+jets decays in which the W decays into $\mu\nu$ and one of the jets is misidentified as an electron. The charge sign of the muon can be trusted because the background in $Z \to \mu^+\mu^-$ is negligible. We expect the same number of same-sign and opposite-sign $e\mu$ pairs from these decays. To infer the number of same sign e^+e^- pairs from the number of same-sign $e\mu$ pairs, we have to take into account that the acceptances for muon identification are different than those for electron identification. We scale the number of events we find in the full $1.12\,\mathrm{fb}^{-1}$ dataset by the ratio $\mathcal{R}$ of the acceptances $\mathcal{A}$ for W decays into $e\nu$ and $\mu\nu$ final states [77]:

$$\mathcal{R} = \frac{\mathcal{A}(W \to e\nu)}{\mathcal{A}(W \to \mu\nu)} = 1.217 \pm 0.026. \tag{7.1}$$

The expected number of background events are compared to the number of opposite-sign $Z \to e^+e^-$ for three and fewer jets in Table 7.6. The fraction of such events is 6×10^{-4} such that they can be safely neglected. Note that our method to estimate the

W+jets background does not include track leptons. We believe that our conclusion that the W+jets background is negligible is not changed if track leptons are included.

Sample	**Without *b*-tag**	**Loose *b*-tag**	**OS $Z \to e^+e^-$**
CEM	2.4±1.7	0.02±0.16	19,325
PHX	21.9±5.2	0.04±0.19	22,107
Total	24.3±5.5	0.06±0.25	41,432

Table 7.6: Expected number of events from W+jets production in $1.12\,\mathrm{fb}^{-1}$ in events with three or fewer jets, before and after b-tagging. The number is estimated from the number of same-sign $e\mu$ pairs in the data and compared to the number of opposite-sign (OS) $Z \to e^+e^-$ candidates.

We have cross-checked the results of our W+jets background estimate by two alternative methods that rely more on MC simulations. For both methods we use an inclusive $Z \to e^+e^-$ MC sample generated with PYTHIA (`zewk6d`) to subtract the number of same-sign Z candidates expected from charge misidentification and tridents in real $Z \to e^+e^-$ events from the number of same-sign Z candidates in the data. In the first method we obtain the ratio of same-sign to opposite-sign Z candidates from the MC simulation,

$$N_{\mathrm{bkg}}^{\mathrm{data}} = N_{Z,\mathrm{SS}}^{\mathrm{data}} - \frac{N_{Z,\mathrm{SS}}^{\mathrm{MC}}}{N_{Z,\mathrm{OS}}^{\mathrm{MC}}} \cdot N_{Z,\mathrm{OS}}^{\mathrm{data}}. \tag{7.2}$$

For the second method we assume that the fraction of tridents in the same-sign Z candidates is correctly modeled in the MC simulation:

$$N_{\mathrm{bkg}}^{\mathrm{data}} = N_{Z,\mathrm{SS}}^{\mathrm{data}} - \frac{N_{Z,\mathrm{SS}}^{\mathrm{MC}}}{N_{\mathrm{trident,OS}}^{\mathrm{MC}}} \cdot N_{\mathrm{trident,OS}}^{\mathrm{data}}. \tag{7.3}$$

We find expected background rates of 0.01–0.04, compared to Z+jets, with these

methods, more than an order of magnitude larger than the result obtained with our main technique. Note however that these methods are conservative because both the effects of charge misidentification and the amount of material in the detector giving rise to tridents tend to be underestimated in the MC. The conclusion that the W+jets background is negligible remains valid.

7.5 Mass χ^2 Tail

The previous sections describes how each of the background processes contributing to our blinded region are estimated from Monte Carlo simulations and/or data. We have also developed a largely data-driven method for estimating the total background in Z+$\geq$4 jets events based on the mass χ^2 described in Section 5.4.1. Figure 7.11 shows the mass $\sqrt{\chi^2}$ distribution for the FCNC signal and the backgrounds. There is a good separation between signal and backgrounds. Furthermore, the different backgrounds have similar $\sqrt{\chi^2}$ distributions. Events with $\sqrt{\chi^2} > 3.0$ (to the right of the black vertical line in Figure 7.11) contain only 3% of the FCNC signal but nearly 30% of the background. We will, therefore, extend our control region to Z+$\geq$4 jets events with $\sqrt{\chi^2} > 3.0$ (and correspondingly decrease our blind region).

Before we unblind part of the Z+$\geq$4 jets events based on the mass χ^2 variable, we want to verify that it is well modelled in Monte Carlo. The control region available to us is the Z+3 jets events, where the full mass χ^2 cannot be constructed since it requires four jets. We can, however, break up the mass χ^2 into two parts. We assume that the the three jets are from (a) the W boson and the b quark or (b) the W boson and the c quark. We can reconstruct the W and either the SM top or the FCNC top,

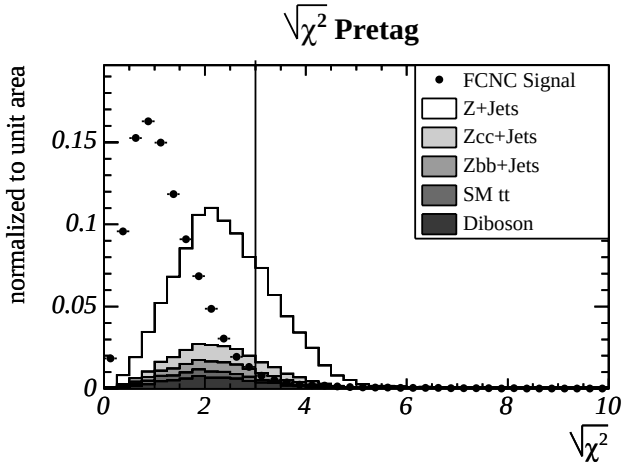


Figure 7.11: Mass χ^2 distribution of signal and background Monte Carlo events, with the vertical line showing the cut of $\sqrt{\chi^2} = 3$. The signal and background samples are normalized to unit area.

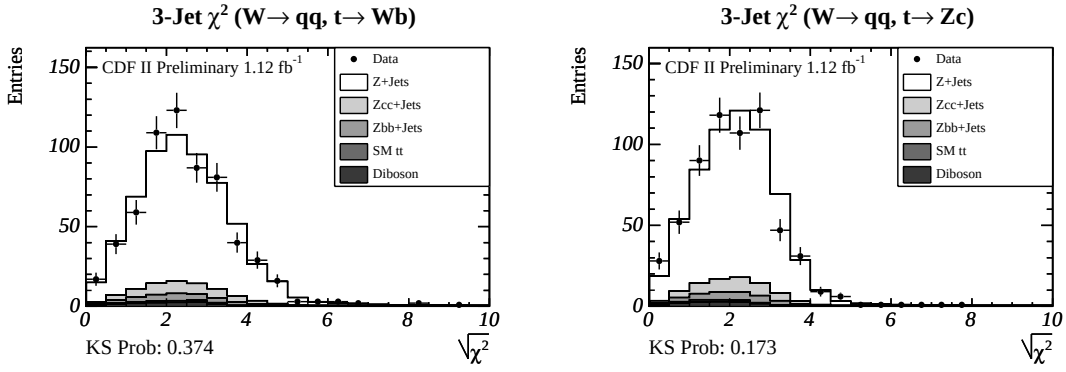
according to the following equations:

$$\chi^2_a = \left(\frac{m_{W,\mathrm{rec}} - m_{W,\mathrm{PDG}}}{\sigma_{W,\mathrm{rec}}}\right)^2 + \left(\frac{m_{t\to Wb,\mathrm{rec}} - m_{t,\mathrm{PDG}}}{\sigma_{t\to Wb}}\right)^2 + \tag{7.4}$$

$$\chi^2_b = \left(\frac{m_{W,\mathrm{rec}} - m_{W,\mathrm{PDG}}}{\sigma_{W,\mathrm{rec}}}\right)^2 + \left(\frac{m_{t\to qZ,\mathrm{rec}} - m_{t,\mathrm{PDG}}}{\sigma_{t\to qZ}}\right)^2 . \tag{7.5}$$

Figure 7.12 shows the data-MC comparison for these two $\sqrt{\chi^2}$ distributions in Z+3 jets events. Like the other kinematic distributions, the data $\sqrt{\chi^2}$ are well modelled in Monte Carlo simulations. This indicates that we can confidently unblind the Z+$\geq$4 jets events with $\sqrt{\chi^2} > 3.0$.

Figure 7.13 shows the Z+$\geq$4 jets events with $\sqrt{\chi^2} > 3.0$ fit to the background $\sqrt{\chi^2}$ shape from Monte Carlo simulations. Unfortunately, the bin at $\sqrt{\chi^2} = 3.0$ is empty. If we include this bin in the fit, we get a prediction for the total number of Z+$\geq$4 jets background events of 115 $\pm$ 21 events. If we remove this bin from the



(a) (b)

Figure 7.12: The full mass χ^2 requires four jets and therefore cannot be validated in Z+3 jets events; however, a W boson candidate mass and a top candidate mass can be reconstructed. These plots show a comparison of data (1.12 fb^{-1}) and Monte Carlo distributions in Z+3 jets events for (a) Mass χ^2 based on $W \to q\overline{q}'$ and $t \to Wb$. (b) Mass χ^2 based on $W \to q\overline{q}'$ and $t \to cZ$. In both plots, we assume that two of the three jets come from the W decay and the third jet is the b- or c-jet.

fit, we get a background estimate of 144 ± 26 events. We take the average of these two numbers as the background prediction from the mass χ^2 method, and summing the uncertainties in quadrature with half of the difference, we expect 130 ± 28 total background events in the $Z+ \geq 4$ jets region. This estimate is consistent with the prediction made in the ALPGEN MC simulation of 120 total background events.

Out of the 31 Z+≥4 jets events with $\sqrt{\chi^2} > 3.0$ in data, five events have one or more loose SecVtx tags. This indicates a tag rate of 16% ± 7%. The tag template method described in Section 7.1.3 gave a tag rate of 14%. We take the average of the mass χ^2 and tag template methods and assign a systematic error to cover the Monte Carlo method, taking a tag rate prediction of 15% ± 4% in our blinded region. Given a total Z+≥4 jets background prediction of 130 ± 28 events, we estimate the number

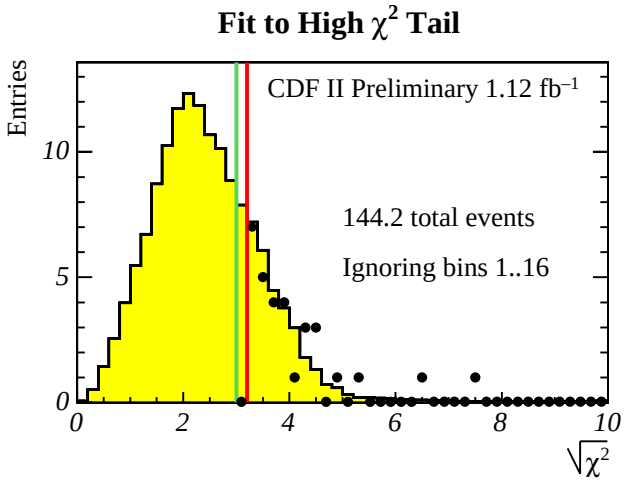


Figure 7.13: Mass χ^2 distribution in events with a Z and four or more jets, showing the fit to data (1.12 fb^{-1}) in the high χ^2 tail. The background shape is from Monte Carlo simulations. Vertical lines show cuts at $\sqrt{\chi^2} = 3$ and $\sqrt{\chi^2} = 3.2$.

of tagged events to be 20 ± 6 events.

7.6 Background Summary

In summary, the main background for the search for the FCNC decay $t \to qZ$ comes from Z boson production in association with jets. The backgrounds coming from SM $t\bar{t}$ production and diboson production are small, and contributions from W+jets events are negligible. We expect a total of 130 ± 28 pre-tagged and 20 ± 6 loose tagged background events in our 1.12 fb^{-1} dataset as summarized in Table 7.7.

Source	**Without *b*-tag**	**Loose SECVTX*b*-tag**
Z+Jets	123.3±28	17.6±6
Standard Model $t\bar{t}$	2.4±0.3	1.7±0.2
Diboson (WZ, ZZ)	4.3±0.2	0.7±0.1
WW, W+Jets	< 0.1	negligible
Total Backgrounds:	130±28	20±6

Table 7.7: Summary of all background contributions to the search for the FCNC decay $t \to qZ$. Given are the expected numbers of background events in $1.12\,\mathrm{fb}^{-1}$.

Chapter 8

Limits and Optimization

In a minute there is time
For decisions and revisions which a minute will reverse.
T. S. Eliot (from "The Love Song of J. Alfred Prufrock")

In the last chapter, we estimated the background contribution for our loose base event selection criteria; 130 $\pm$ 28 background events with a Z boson and four more jets, out of which we expect 20 $\pm$ 6 events to have one or more loose SecVtx tags. We have to further reduce this background in order to improve our chances of seeing a signal or to set a better upper limit on the $t \to qZ$ branching fraction.

The FCNC signal has different event kinematics than the dominant Z+jets background. We take advantage of this and further optimize our selection criteria using the mass χ^2, transverse mass, and the E_T of the four leading jets. We optimize our selection criteria to obtain the best expected limit, which takes systematic errors into account. We also separate the events into two samples; those with one or more loose SecVtx tags (tagged) and those with no loose SecVtx tags (anti-tagged). The limit calculations and the optimization procedure are described in this chapter.

8.1 Expected Limit Calculation

Once we look at the events in the currently blinded region, we will have a limit based on the data yield, background predictions, and signal acceptances. For the optimization, however, we have only the background predictions and signal acceptances to calculate the expected limit. The simplest assumption we could make is that we will see exactly the number of background events we have predicted. This assumption is inadequate because (1) there are large errors associated with the background prediction and (2) the general statistical nature of data makes for a sizable probability that the number of background events will fluctuate for an individual dataset even if the predictions are accurate on average. Consequently, the expected calculation has to allow for different possible data yields.

We construct an expected upper limit which is the weighted sum of the limits for different possible number of observed events. For a single signal region, the expected limit is calculated as:

$$\text{Expected Limit} = \sum_{n_{\text{obs}}} \underbrace{P(n_{\text{obs}}|n_{\text{back}})}_{\text{weight}} \cdot Lim(n_{\text{obs}}|A, n_{\text{back}}), \tag{8.1}$$

where n_{obs} represents the number of events observed, n_{back} is expected background, A is the signal acceptance convolved with efficiency, $P(n_{\text{obs}}|n_{\text{back}})$ is the Poisson probability that n_{back} background events fluctuated to n_{obs}, and Lim is any upper limit calculation. In words, this means that for a given number of observed events, the upper limit on the $t \to qZ$ branching fraction is calculated with the predicted background and signal acceptance, as described in the following subsections. That value,

then, receives a weight which is the Poisson probability that the number of background events as predicted would fluctuate to this number of observed events, assuming that there is no signal. Then the weighted upper limits for different possible number of observed events are summed.

In this analysis, we have two signal regions, tagged and anti-tagged, so we extend the expected limit calculation. Each individual limit is calculated for two sets of possible data yields, expected backgrounds, and signal acceptances. Then these limits are summed over a two dimensional grid of possible expected data yields for the two signal regions. The weight is the sum of the probability for getting the data yield for the first region, given the expected background for that region, and the probability for getting the data yield in the second region, given the expected background for that region. To incorporate systematic uncertainties for the background prediction in the weighted sum, we calculate the probability by Gaussian smearing the background prediction before the Poisson fluctuation, using a random number generator.

In practice, it is impossible to sum over all the possible number of observed events, since it is the sum over all positive integers. Therefore any finite sum, given by the equation below, is an *under-estimate* of the expected upper limit.[1]

$$\text{Expected Limit}_{\text{under-estimate}} = \sum^{n_{\text{used}}} P(\vec{n}_{\text{obs}}|\vec{n}_{\text{back}}) \cdot Lim(\vec{n}_{\text{obs}}|\vec{A}, \vec{n}_{\text{back}}). \tag{8.2}$$

This underestimate approaches the true expected limit as n_{used} increases. Given

[1] The possible values for the number of observed events are ordered from highest to lowest probability before summing.

this underestimate, the *best-estimate* is:

$$\text{Expected Limit}_{\text{best-estimate}} = \frac{\sum^{n_{\text{used}}} P(\vec{n}_{\text{obs}}|\vec{n}_{\text{back}}) \cdot Lim(\vec{n}_{\text{obs}}|\vec{A}, \vec{n}_{\text{back}})}{\text{weight}_{\text{sum}}}, \tag{8.3}$$

the under-estimate divided by the total weight, $\text{weight}_{\text{sum}}$.

We also construct an *over-estimate* where we assume that the rest of the limits not in the sum are the same as the current max limit in the sum, lim_{max}:

$$\begin{aligned} \text{Expected Limit}_{\text{over-estimate}} &= \frac{\sum^{n_{\text{used}}} P(\vec{n}_{\text{obs}}|\vec{n}_{\text{back}}) \cdot Lim(\vec{n}_{\text{obs}}|\vec{A}, \vec{n}_{\text{back}}) + (1 - \text{weight}_{\text{sum}}) \cdot lim_{\text{max}}}{\text{weight}_{\text{sum}}} \\ &= \text{Expected Limit}_{\text{best-estimate}} + \frac{(1 - \text{weight}_{\text{sum}})}{\text{weight}_{\text{sum}}} \cdot lim_{\text{max}} \end{aligned} \tag{8.4}$$

We added the possible values for the number of observed events until the under-estimate and the over-estimate are within 0.01% of each other. We then take the best-estimate as our expected limit.

8.1.1 Feldman Cousins Limit

To obtain our limit, we use the standard Feldman-Cousins [80] limit calculation because it has the following properties:

- It guarantees a physical answer.
- It guarantees coverage.[2]

[2]The frequentist definition of coverage does not include the idea of nuisance parameters. Our implementation of FC including systematic uncertainties does guarantee coverage with an extended definition, i.e. including the varying of nuisance parameters in the pseudo-experiments used to check the coverage.

- It decides if the result is a (single-sided) limit, or a measurement (two-sided limit).
- It has no metric dependence.

We calculate the Feldman-Cousins upper limit in two steps:

1. For each true value, μ, a range of expected observed events (confidence belt) is generated. This is done by considering both
   - the probability for observing a given number of events given the true value μ, $P(n_{\text{obs}}|\mu)$, and
   - the maximum probability to observe this number of events for all values of μ, $P(n_{\text{obs}}|\mu_{\text{best}})$

   In the discrete case, the ratio of these two values (likelihood ratio) is used to order the bins of observed values. The confidence belt is described as the bins are added, ordered from high to low in likelihood ratio, until the sum of $P(n_{\text{obs}}|\mu)$ reaches the desired confidence level (95% in this case).
2. The upper limit is found by locating the true value, μ, whose lower range almost contains the number of observed events. Likewise, a lower limit is found the locating the true value, μ, whose upper range almost contains the number of observed events.

We have extended the Feldman-Cousins implementation as compared to the original in two ways: (1) we have two signal regions and (2) we include systematic uncertainties. Extending to two signal regions similar to what we did when calculating

expected limits; each bin is now a set of two numbers which are the possible number of observed events for each signal region. $P(\vec{n}_{\rm obs}|\mu)$ is the total probability of observing n_i in each bin given the true value μ:

$$P(\vec{n}_{\rm obs}|\mu) = P(n_{\rm obs1}|\mu) \cdot P(n_{\rm obs2}|\mu)$$

To find the value of $P(\vec{n}_{\rm obs}|\mu_{\rm best})$, we perform a minimization using *Minuit*, since it cannot be calculated analytically.

We incorporate systematic uncertainties into the Feldman-Cousins method by following the prescription laid out by an earlier work of Cousins and Highland [81] as well as the published CDF *Measurement of* $\frac{B(t\to Wb)}{B(t\to Wq)}$ *at the Collider Detector at Fermilab* [82, 21]. Simply put, the only change is that $P(\vec{n}_{\rm obs}|\mu)$ is now calculated by generating pseudo-experiments: we implement Gaussian smearing of the acceptances, efficiencies, and backgrounds before Poisson fluctuating the numbers of observed events.

Unfortunately, this implementation of the Feldman-Cousins method is computationally expensive. While this is not a problem for calculating one limit, it is slow for calculating expected limits, and unacceptably slow for optimization where we calculate the expected limit for many different event selection criteria. For this reason, we have implemented an objective Bayesian limit calculation that profiles the systematic uncertainties. We use the Bayesian limit calculation for the optimization and then use the Feldman Cousins limit calculation to find the expected limit for the final event selection criteria and also for the final limit after unblinding the data.

8.1.2 Bayesian Limit

To calculate the Bayesian limit, we start with the standard Poisson likelihood:

$$\mathcal{L} = \prod_{\text{Signal Regions}} P(n_{\text{obs}i}|n_{\text{sig}_i} + n_{\text{back}i}) \cdot \prod_{\text{Systematics}} \mathcal{G}(\text{syst}_k|\mu_{\text{syst}_k}, \sigma_{\text{syst}_k}), \tag{8.5}$$

where $P(n_{\text{obs}i}|n_{\text{sig}_i}+n_{\text{back}i})$ is the Poisson probability and $\mathcal{G}(\text{syst}_k|\mu_{\text{syst}_k}, \sigma_{\text{syst}_k})$ is a Gaussian penalty term to allow for systematic shifts in acceptances and backgrounds.

First, we profile the likelihood. This is done by stepping through the likelihood on a grid of branching fraction values and maximizing the likelihood at each value while letting the nuisance parameters float. To calculate a limit, we multiply the profiled likelihood function by a flat prior (1 between the branching fraction values of 0 and 1 and 0 everywhere else) and find where we need to set the limits to have 95% of the area under posterior probability curve, i.e. the value of the upper limit of the branching fraction, UL, where:

$$\frac{\int_0^{UL} \mathcal{L} dBR \cdot 1}{\int_0^1 \mathcal{L} dBR \cdot 1} = 0.95. \tag{8.6}$$

8.2 Optimization of Event Selection

8.2.1 Scanning Expected Upper Limits

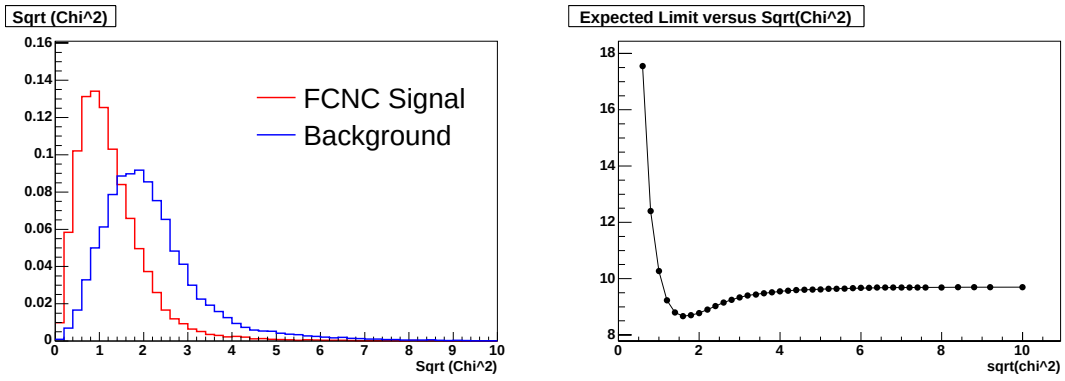
We optimize our event selection criteria by scanning over the variable of interest, i.e. we:

1. Tighten the event selection gradually by moving the cut value for the variable in question.

2. Recalculate all background and acceptance numbers with new event selection.

3. Calculate new expected limit based on new event selection.

Figure 8.1 shows an example of this scanning process. Here, we calculate the limits requiring that the events have a $\sqrt{\chi^2}$ below each given value (for other variables, we take events above the given value). Figure 8.1 (a) shows the $\sqrt{\chi^2}$ distributions for signal and background events. Figure 8.1 (b) shows the limit at different $\sqrt{\chi^2}$ selection points. We take the value where the expected limit is a minimum.



(a) (b)

Figure 8.1: (a) Shows the $\sqrt{\chi^2}$ distribution for signal and the background events (includes all background processes) from Monte Carlo simulations. (b) Shows the limit as a function of mass χ^2.

8.2.2 Comparison of Feldman Cousins and Bayesian Limits

To have confidence in using the Bayesian limit calculation in our optimization, we need to verify that it will find the same selection criteria as the Feldman Cousins limit. Figure 8.2 shows the expected limits calculated with the Feldman Cousins method and the Bayesian method for different values of mass χ^2. Although the two methods give different expected limits at each value, their behaviors track each other well as we scan the different variables, and we find the minimum expected limits at the same value. This gives us confidence that we will arrive at the same optimized selection criteria with the Bayesian method as we would with the Feldman cousins method. We optimize using the Bayesian method and calculate the expected limit for the optimized selection criteria using the Feldman Cousins method.

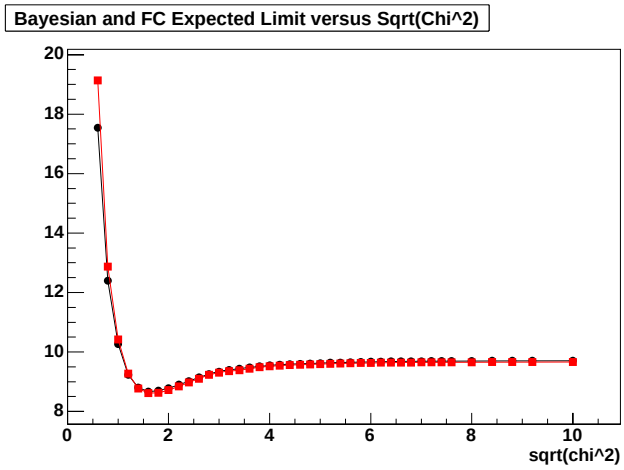


Figure 8.2: The behaviors of the Feldman Cousins limit (black circles) and the Bayesian limit (red squares) track each other well. In this example the optimization variable is mass χ^2.

8.2.3 Systematic Uncertainties for the Optimization

We take the main systematic uncertainties into account in our optimization. We determine these uncertainties for the pre-tag selection criteria and assume that the percent error is the same as we scan the different values of the kinematic variables. The procedure for determining the systematic uncertainties are described in Chapter 9. For the signal systematics, we assume that many cancel (at least partially) when we take the ratio of our FCNC acceptance to the $t\bar{t}$ cross section analysis acceptance. For background, we take the uncertainty from the total background estimate, described in Section 7.5. The systematic uncertainties considered in the optimization are:

- Signal systematic uncertainties
  - Uncertainties on the event counts and background estimates of the lepton + jets loose SecVtx $t\bar{t}$ cross section analysis, 10% citeCDF8767, see Section 9.2 for details
  - Helicity of the Z from FCNC decay, 1.2% [3], and
  - B-tagging, uncertainties on the loose SecVtx scale factor and mistag matrix parameterization, 5%. Note that this uncertainty moves events between the tagged and anti-tagged signal regions.
- Background systematic uncertainties
  - Pre-tag yield, 21.5% (130 $\pm$ 28 events), and

[3] We found that this systematic uncertainty was 3.5% after the optimized selection criteria were chosen.

– Event tag rate, 26.7% (15% ± 4%).

All the systematic uncertainties are carefully determined for the final selection criteria after optimization as described in Chapter 9.

8.2.4 Stability of Optimization

Since we incorporate only the major systematic uncertainties in the optimization, we test the robustness of the optimization to systematic uncertainties, in order to verify that we have found the true optimal selection criteria. The largest systematic uncertainty in this analysis is the uncertainty on the number of total background events for our pre-tag selection and the loose SecVtx tag rate in these events, determined from the high mass χ^2 region as described in Section 7.5. We test the effect of systematic errors on the optimal selection criteria by testing the optimization for 135 total background events and 150 total background events with a 14% or 20% tag rate. The effect of these variations on the expected limits for different values of χ^2 are shown in Figure 8.3.

The expected limits depend strongly on the total number of background events and the tag rates, as expected. However, the optimal selection point, i.e. where the minimum expected limit falls with respect to the kinematic variable, does not. This gives us confidence that we have found the optimal selection criteria.

8.2.5 Optimization variables and Multivariate Optimization

We optimize the selection criteria using the following variables:

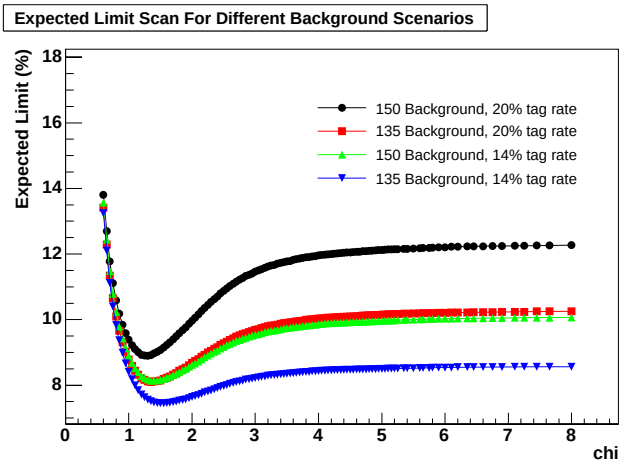


Figure 8.3: The expected limit scan shown for various different background scenarios. Even though the expected limit changes depending on the number of background events, the optimal selection point remains stable.

- the mass χ^2, $\chi^2 = \left(\frac{m_{W,\mathrm{rec}} - m_{W,\mathrm{PDG}}}{\sigma_{W,rec}}\right)^2 + \left(\frac{m_{t\to Wb,\mathrm{rec}} - m_{t,\mathrm{PDG}}}{\sigma_{t\to Wb}}\right)^2 + \left(\frac{m_{t\to qZ,\mathrm{rec}} - m_{t,\mathrm{PDG}}}{\sigma_{t\to qZ}}\right)^2$,
- the event transverse mass, $M_T = \sqrt{\left(\sum E_T\right)^2 - \left(\sum \vec{p}_T\right)^2}$, using the reconstructed Z and the leading four jets, and
- the transverse energy of the four leading jets

These variables are described in detail in Section 5.4. The distributions of these variables for our main FCNC signal ($t\bar{t} \to qZWb$, $Z \to \ell^+\ell^-$ and $W \to q\overline{q}\,'$) and the backgrounds are shown in Figure 8.4.

The kinematic variables used in the optimization are highly correlated and we want to take these correlations into account in our optimization procedure. We do this by scanning variables sequentially. For a two-dimensional example, variable A and B (e.g. χ^2 and transverse mass), this means that we:

1. Scan the expected upper limit using variable A and find the optimal value,

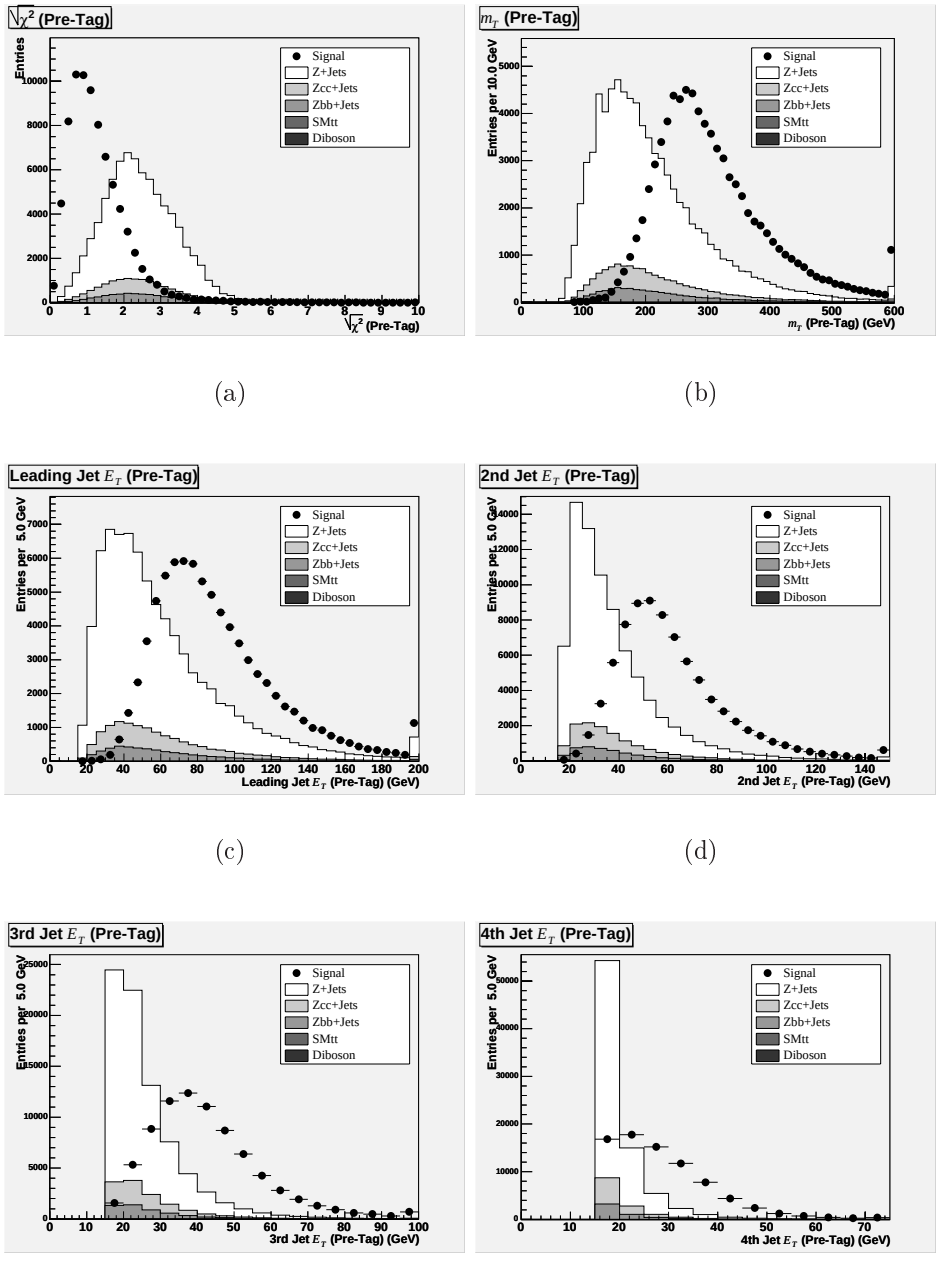


Figure 8.4: Signal and background distributions for kinematic variables used in the optimization determined from Monte Carlo simulations. The signal and background are normalized to equal area. (a) χ^2 (b) Transverse Mass (c) Transverse Energy of the Leading Jet (d) Transverse Energy of the Second Jet (e) Transverse Energy of the Third Jet (f) Transverse Energy of the Fourth Jet.

2. Apply the requirement on variable A to the selection criteria,
3. Scan the expected upper limit using variable B and find the optimal value,
4. Apply the requirement on variable B and drop the requirement on variable A,
5. Rescan the expected upper limit using variable A,
6. Stop if the optimal value in Step 5 is within a difference of δ to the optimal value in Step 1.

This is represented pictorially in Figure 8.5. We repeat this procedure on all the six variables we use in the optimization. Note that there are two important (opposing) principles governing this method; (1) The closer to the starting value to the optimal selection, the faster the optimization converges, but (2) we want to avoid local minima and to do so, we optimize each variable several times, starting with different values.

8.2.6 Optimized Selection Criteria

The optimized selection criteria are listed in Table 8.1. The criteria are the same for the tagged and anti-tagged samples except for mass χ^2 where the tagged events are required to have $\sqrt{\chi^2} < 1.6$ while the more numerous anti-tagged events are required to have $\sqrt{\chi^2} > 1.35$. With these criteria, 71% (56%) of the tagged (anti-tagged) pre-tag signal events remained compared to 16% (7%) of the tagged (anti-tagged) pre-tag background events. The expected limit for these optimized selection criteria is 7.4% $\pm$ 2.2% using the Bayesian method and 7.1% $\pm$ 3.0% using the Feldman Cousins method. The Bayesian expected limit distribution is shown in Figure 8.6. The optimized selection criteria define our tagged and anti-tagged signal regions.

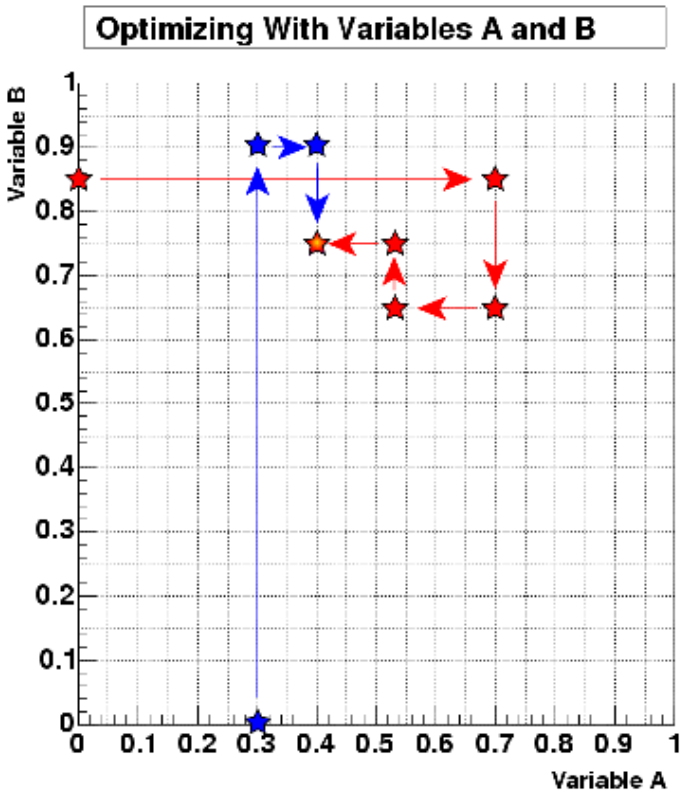


Figure 8.5: We scan variables sequentially as illustrated in this diagram. This takes into account the correlations between the different variables.

Kinematic Variable	**Optimized Cut**
Z Mass	$\in$ [76 GeV/c^2, 106 GeV/c^2]
Leading Jet E_T	$>$ 40 GeV
Second Jet E_T	$>$ 30 GeV
Third Jet E_T	$>$ 20 GeV
Fourth Jet E_T	$>$ 15 GeV
Transverse Mass	$>$ 200 GeV
$\sqrt{\chi^2}$	$<$ 1.6 in the loose SecVtx tagged sample, $<$ 1.35 in the anti-tagged sample

Table 8.1: The optimized event selection criteria.

Using the final event selection criteria for the two signal regions, we re-calculate all acceptances required for the limit calculation and the total background estimates.

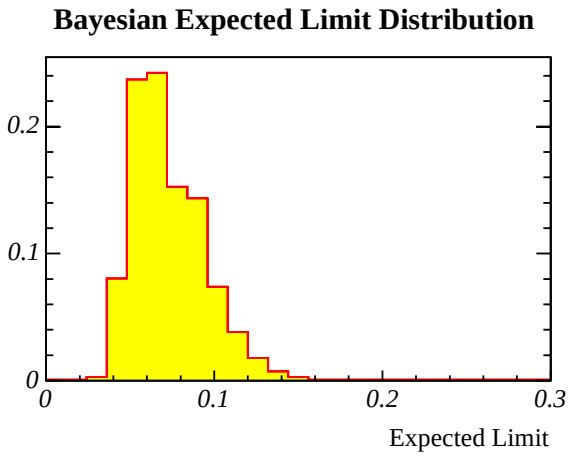


Figure 8.6: Bayesian expected limit distribution for the optimized selection criteria

The systematic uncertainties of the acceptances and backgrounds are calculated for the tagged and anti-tagged final event selections as described in the next chapter. The acceptances and background estimates for the base event selection, anti-tagged final selection, and tagged final selection are summarized in Table 8.2.

	Anti-Tagged	**Loose SecVtx Tagged**
$\mathcal{A}_{WZ}$ (%)	$0.1900 \pm 0.0097 \mp 0.0319$	$0.2924 \pm 0.0219 \pm 0.0199$
$\mathcal{A}_{ZZ}$ (%)	$0.4764 \pm 0.0243 \mp 0.0800$	$0.3961 \pm 0.0297 \pm 0.0269$
Total Background	$7.67 \pm 1.76 \mp 0.36$	$3.17 \pm 0.73 \pm 0.86$

Table 8.2: Final acceptances, in percent, for $\mathcal{A}_{WZ}$, $\mathcal{A}_{ZZ}$, and the total background. The first uncertainty given is the uncertainty correlated between the two signal regions, and the second uncertainty is the anti-correlated uncertainty.

Chapter 9

Systematic Uncertainties

Not to be absolutely certain is, I think, one of the essential things in rationality.

Bertrand Russell, (from "Am I An Atheist Or An Agnostic?")

We evaluate systematic uncertainties for our signal acceptance and the dominant Z+jets background. Unless noted otherwise, we calculate the systematic uncertainties by running the full event selection several times while varying parameters like scale factors (SF) within their uncertainties. The systematic uncertainty is then quoted as relative change in the parameter of interest (e.g. signal acceptance) with respect to the default set of parameters, for three sets of event selection; the base selection, and the two signal regions: the optimized anti-tagged selection, and the optimized tagged selection.

The two signal regions are complementary because of their b-tagging requirements. We will distinguish systematic uncertainties that are correlated between the regions (e.g. lepton scale factors), which shift the numbers of both tagged events and anti-tagged events up or down, and systematic uncertainties that are anti-correlated

between the regions (e.g. b-tagging scale factors), which shift events from one signal region to the other. Our limit calculation takes into account the correlations among the systematic uncertainties across the signal regions.

9.1 Signal Acceptance Systematic Uncertainties

Our final acceptance formula given in Equation (6.6) shows that the main figure of interest for the signal acceptance is the ratio $\mathcal{A}_{WZ}/\mathcal{A}_{LJ\mathrm{ww}}$, i.e. the ratio of acceptances for the FCNC signal, using our FCNC selection criteria, and the acceptance for SM $t\bar{t}$, using the selection criteria for the double loose tag SECVTX cross section analysis to which we normalize our search. We will quote our signal acceptance systematics as the relative change in $\mathcal{A}_{WZ}/\mathcal{A}_{LJ\mathrm{ww}}$. A summary of the results is given in Table 9.1. The following sections describe in detail how these uncertainties are estimated.

We have also determined the effect of systematic uncertainties on the "running" acceptance correction factor in Equation (6.6). The total systematic uncertainty of the correction factor is smaller than 1% for all three selections: base, anti-tagged, and tagged; subsequently, it is neglected in the estimations of signal systematic uncertainties described below.

9.1.1 Lepton Scale Factors and Trigger Efficiencies

We take into account the statistical uncertainties of the scale factors for lepton identification and reconstruction and the electron and muon trigger efficiencies as a systematic uncertainty for our analysis. To calculate these uncertainties, we have varied both the lepton SFs and the trigger efficiencies on a lepton-by-lepton basis. We

Systematic Uncertainty	**Base Selection (%)**	**Anti-Tagged (%)**	**Loose Tag (%)**
Lepton Scale Factor	0.5	0.5	0.5
Trigger Efficiency	0.2	0.2	0.2
Jet Energy Scale	3.1	2.6	1.9
ISR/FSR	1.3	2.6	6.5
Helicity Re-Weighting	3.5	3.4	3.2
Parton Distribution Functions	0.9	0.9	0.9
Total Correlated	5.0	5.1	7.5
B-Tagging Scale Factor	10.2	16.3	5.5
Mistag $\alpha\beta$ Correction	0.6	1.0	0.4
$\mathcal{B}(t \to cZ)$ versus $\mathcal{B}(t \to Zu)$	0.0	4.0	4.0
Total Anti-Correlated	10.2	16.8	6.8

Table 9.1: Summary of systematic shifts of the acceptance ratio $\mathcal{A}_{WZ}/\mathcal{A}_{LJ\mathrm{ww}}$. In the case of asymmetric uncertainties for the upwards and the downwards shift of a parameter, we chose the larger of the two. Note that the upper grouping contains those systematics that are correlated, and the lower grouping are those that are anti-correlated between the anti-tagged and the loosely tagged selection.

conservatively assume that all lepton scale factors are fully correlated and shift them simultaneously by $\pm 1\sigma$. Similarly, we assume 100% correlation between the trigger efficiencies, which we also shift by $\pm 1\sigma$. We found that the combined effect of lepton SF and trigger efficiencies on $\mathcal{A}_{WZ}/\mathcal{A}_{LJ\mathrm{ww}}$ is less than 1% for all three selections.

9.1.2 Jet Energy Scale

The choice of the jet energy scale (JES) influences corrections to the raw jet energy. Our FCNC event selection requires four or more jets and is therefore sensitive to the effect of gaining or losing jets at the lower E_T cut. Other selection criteria that are sensitive to jet energies comprise the top mass χ^2 and transverse mass. To estimate the effect of JES variations, we follow the recommendation of the CDF Jet Energy Resolution group [83]. We vary the jet energy scales by $\pm 1\sigma$ and calculate the difference in the acceptance ratio. The effect of jet energy scale variations on the

acceptance ratio amounts to approximately 3%.

9.1.3 Initial and Final State Radiation

The modeling of initial state radiation (ISR) and final state radiation (FSR) in the MC simulation is expected to have an effect on the number of reconstructed jets. As the FCNC event selection requires four or more reconstructed jets, it is influenced by the amount of ISR and FSR. We have generated MC samples to study the effect of the modeling of ISR and FSR on the FCNC signal acceptance. The generated samples include 50,000 signal events ($Z(ll)W(q\overline{q}')$) each for the scenarios of more ISR, less ISR, more FSR, and less FSR which are $\pm 1\sigma$ values. The relevant PYTHIA parameters and settings for these samples are those used for similar studiesof the systematics of $t\bar{t}$ production in Gen5. We modified the following PYTHIA parameters [72]:

- `PARP(1)`: nominal value of Λ_{QCD} (in GeV) for the running of α_S.
- `PARP(61)`: value of Λ_{QCD} (in GeV) for the running of α_S in space-like shower evolution.
- `PARP(72)`: value of Λ_{QCD} (in GeV) for the running of α_S in time-like shower evolution.
- `PARP(64)`: Q^2 scale times this factor is maximum parton virtuality in space-like showers.
- `PARP(71)`: Q^2 scale times this factor is maximum parton virtuality in time-like showers.

Setting	**PARP(1)**	**PARP(61)**	**PARP(72)**	**PARP(64)**	**PARP(71)**
Default	0.250	0.250	0.250	1.0	4.0
More ISR	0.146	0.292	0.146	0.5	4.0
Less ISR	0.146	0.072	0.146	2.0	4.0
More FSR	0.146	0.146	0.292	1.0	8.0
Less FSR	0.146	0.146	0.076	1.0	2.0

Table 9.2: PYTHIA settings for the MC samples to study systematic effects due to the amount of initial and final state radiation compared to the default settings used in CDF. The PYTHIA parameters are described in the text.

The settings for the four MC samples are summarized in Table 9.1.3.

We have evaluated the effect of varying the amount of ISR/FSR on the ratio $\mathcal{A}_{WZ}/\mathcal{A}_{LJ\mathrm{ww}}$. We obtain the numerator of the ratio by applying the FCNC event selection on the four ISR/FSR samples, and the denominator is taken from the ISR/FSR study done in the context of the lepton+jets double loose tag $t\bar{t}$ cross section analysis (see Table 16 of CDF Note 8767 [71]). The effect of ISR/FSR variations on $\mathcal{A}_{WZ}/\mathcal{A}_{LJ\mathrm{ww}}$ is 1.4% for the base selection, 2.6% for the anti-tagged selection, and 6.5% for the tagged selection. More details can be found in Table 9.1.3.

Sample	**Base Selection (%)**	**Anti-Tagged (%)**	**Loose Tag (%)**
More ISR	0.0	2.4	−1.6
Less FSR	0.4	−0.1	3.0
More FSR	−0.1	−0.9	2.9
Less FSR	1.3	−0.4	4.7
Total	1.3	2.6	6.5

Table 9.3: ISR/FSR systematics: Shown are shifts of the acceptance ratio $\mathcal{A}_{WZ}/\mathcal{A}_{LJ\mathrm{ww}}$ for different amounts of initial and final state radiation. The total uncertainty is obtained by adding the individual uncertainties in quadrature.

9.1.4 Z Helicity Reweighting

As described in Section 6.1.1, we re-weight the FCNC MC samples to force the helicities of Z bosons from $t \to qZ$ decays to 65% longitudinal and 35% left-handed. Due to the unknown nature of the FCNC interaction we study the systematic effects of other choices of the Z helicity. We re-weight the $Z(ll)W(q\overline{q})$ sample with different admixtures of longitudinal, left-handed, and right-handed components, and measure the effect on the acceptance ratio $\mathcal{A}_{WZ}/\mathcal{A}_{LJ\mathrm{ww}}$.

Table 9.1.4 shows the shift in the acceptance ratio $\mathcal{A}_{WZ}/\mathcal{A}_{LJ\mathrm{ww}}$ for a collection of re-weighted samples. As expected from the approximate symmetry of the CDF-II detector with respect to positively and negatively charged leptons, the shifts in the acceptance ratio are similar for the same amount of left-handed and right-handed helicity. The shifts are also independent of the details of the event selection, as the lepton requirements for the base selection and the anti-tagged and loosely tagged selections are identical. The biggest difference occurs between the extreme cases of a 100% longitudinal sample and a 100% left-handed or right-handed sample, with a relative acceptance change of 14%. Since these two values are absolute extremes, we consider half the difference as the 2σ uncertainty on the Z helicity and assign a systematic uncertainty of 3.5%.

9.1.5 Parton Distribution Functions

Parton distribution functions (PDF) depend only on $t\bar{t}$ production, not on the details of the decay; therefore, we take the systematic uncertainty on the choice of the PDFs of the proton and antiproton from the lepton+jets double loose tag

Helicity	**Base Selection (%)**	**Anti-Tagged (%)**	**Loose Tag (%)**
35% LH, 65% Long.		default	
Flat	−4.3	−4.2	−4.5
100% Longitudinal	5.0	4.7	4.5
100% Left-Handed	−9.2	−8.8	−8.3
100% Right-Handed	−8.6	−8.6	−9.5
35% RH, 65% Long.	0.2	0.1	−0.4
Total Uncertainty (%)	3.5	3.4	3.2

Table 9.4: Helicity systematics: Shown are shifts of the acceptance ratio $\mathcal{A}_{WZ}/\mathcal{A}_{LJ\mathrm{ww}}$ with respect to the default Z helicity of 35% left-handed and 65% longitudinal. The total systematic uncertainty is obtained from taking half of the largest difference as the 2σ systematic uncertainty.

SECVTX cross section analysis we are normalizing our acceptance to, see Table 16 of CDF Note 8767 [71]. This choice is conservative in that it neglects the correlation between the PDF uncertainties for the FCNC analysis and the lepton+jets cross section analysis. The PDF uncertainty used in the lepton+jets cross section analysis amounts to 0.9%.

9.1.6 *B*-Tagging

To estimate systematic uncertainties related to the SECVTX b-tagging algorithm, we vary the per-jet tagging probability (see Section 6.3). We independently shift the tagging probability, (b-tagging SF), and the mistag probability ($\alpha\beta$ correction) by $\pm 1\sigma$.

Our studies show that the b-tagging SF is the leading systematic uncertainty for the acceptance ratio $\mathcal{A}_{WZ}/\mathcal{A}_{LJ\mathrm{ww}}$. The b-tagging SF uncertainty also shows the largest difference among the event selections. The base selection uncertainty is 10%, the uncertainty of the anti-tagged selection amounts to 16%, and the uncertainty of the selection requiring one or more loose SECVTX tag is 6%. These numbers demon-

strate the effect of cancellations in the acceptance ratio: The lepton+jets selection requires two loose tags. The FCNC selection most similar to the lepton+jets selection is the loosely tagged selection. Consequently, the ratio of these acceptances has the smallest systematic uncertainty, smaller than the individual uncertainties. The anti-tagged acceptance is anti-correlated with the tagged acceptance, so that an increase (decrease) of the b-tagging SF results in a smaller (larger) acceptance. In this situation the systematic uncertainties are even slightly amplified. The anti-correlation is taken into account in the limit calculation. The systematic uncertainty due to the $\alpha\beta$ correction is approximately 1%.

9.1.7 Difference between $\mathcal{B}(t \to Zc)$ and $\mathcal{B}(t \to Zu)$

We quote our limit on top FCNC as a limit on the branching fraction $\mathcal{B}(t \to qZ)$ even though the MC samples we used for the acceptance calculation contain only the channel $t \to cZ$, and not $t \to Zu$. The difference in the two channels lies in the different SecVtx b-tagging probabilities for c and u quarks. With the help of a test sample that contains the decay $Z(ll)u\,W(q\overline{q})$, i.e. $t \to cZ$ is replaced by $t \to Zu$, we found that the probability for a $t \to Zu$ event to be b-tagged is $92 \pm 1\%$ of the b-tagging probability for $t \to cZ$. The event tagging rate is dominated by tagging of the b-quark from the SM $t \to Wb$ decay. The difference between the charm tagging rate and the mistag rate has a much smaller effect on the event tagging rate.

We assume that in models of physics beyond the standard model, the branching fraction of a top FCNC decay into the second generation c quark is generally larger than the branching fraction for a decay into a first generation u quark. This is

motivated by the size of the CKM matrix elements and the fact that the Higgs coupling is proportional to the mass of a particle. In a "worst case scenario," in which 50% of the $t \to qZ$ decays are into Zu final states, the b-tagging probability would be reduced to 96%. We add the 4% difference between the pure $t \to cZ$ case and "worst case" to the systematic uncertainties for both the anti-tagged and the loosely tagged selection.

9.1.8 Monte Carlo Statistics

All the signal MC samples used for this analysis contain a sufficiently large number of events, so that the statistical uncertainty of the MC samples statistics is always well below the leading systematic uncertainties we have studied.

9.2 Normalization to the Lepton+Jets SecVtx Top Cross Section Analysis

We normalize our measurement to the lepton+jets double loose SECVTX tagged $t\bar{t}$ cross section analysis. As shown in Eq. (6.6), the expected number of events depends on the number of signal and background events in that analysis: $\mathcal{N}_{\text{signal}} \propto \mathcal{N}_{LJ} - \mathcal{B}_{LJ}$. Consequently, the statistical uncertainty on the lepton+jets signal and the total uncertainty on the background need to be added to the total systematic uncertainty of our measurement. The systematic uncertainty on the signal has already been taken into account in the calculation of systematic uncertainties of the acceptance ratio $\mathcal{A}_{WZ}/\mathcal{A}_{LJ_{\text{ww}}}$. We take the numerical values needed for the calculation from Section 5

and Table 13 of CDF Note 8767 [71].

The result of the lepton+jets $t\bar{t}$ cross section measurement is $8.8 \pm 0.8\,(\text{stat.})$ pb. The 9.1% statistical uncertainty of the signal translates into an uncertainty of the number of signal events with three or more jets of $\mathcal{N}_{LJ} - \mathcal{B}_{LJ} = 148.7 \pm 13.5\,(\text{stat.})$. To calculate the total background uncertainty we add the background uncertainties in the n-jet bins in quadrature. The resulting number of expected background events with three or more jets is $\mathcal{B}_{LJ} = 16.3 \pm 5.6\,(\text{stat.} + \text{syst.})$. To obtain the total uncertainty on $\mathcal{N}_{LJ} - \mathcal{B}_{LJ}$ we add the uncertainties on signal and background in quadrature, resulting in $\mathcal{N}_{LJ} - \mathcal{B}_{LJ} = 148.7 \pm 14.7\,(\text{stat.} + \text{syst.})$ or a relative uncertainty of 9.8%.

9.3 Background Systematic Uncertainites

The dominant systematic uncertainty for background processes results from our estimation of the total background using mass χ^2, described in Section 7.5. Due to the small number of Z+$\geq$4 jets events with $\sqrt{\chi^2} > 3.0$, this method gave large errors on the background contribution in our blinded region. The additional background systematic uncertainties are a result of the variation in the shape of the χ^2 distribution. Since we use events with $\sqrt{\chi^2} > 3.0$ to predict the number of events with $\sqrt{\chi^2} < 1.6$, a change in the shape of the distribution will shift the fraction of events in each region, and, consequently, the predicted number of background events in the signal regions. We estimate this uncertainty by varying the same parameters of interest as we did for the signal systematics. For a given parameter, we quote the systematic uncertainty as the shift in the ratio of the number of events predicted for $\sqrt{\chi^2} < 1.6$ to the number of events predicted for $\sqrt{\chi^2} > 3.0$. The systematic uncertainties for the

background arising from the shape of the χ^2 distribution are summarized in Table 9.3. The following sections describe in detail how these uncertainties were estimated.

Systematic Uncertainty	**Anti-Tagged (%)**	**Loose Tag (%)**
Lepton Scale Factor	< 0.1	< 0.1
Trigger Efficiency	< 0.1	< 0.1
Jet Energy Scale	5.1	2.1
B-Tagging Scale Factor	< 0.1	0.3
Mistag $\alpha\beta$ Correction	0.2	0.4
ALPGEN MC Generator	10.0	5.9
Total Uncertainty	11.2	6.3

Table 9.5: Summary of systematic shifts of the ratio of events with $\sqrt{\chi^2} < 1.6$ to events with $\sqrt{\chi^2} > 3$. In the case of asymmetric uncertainties for the upwards and the downwards shift of a parameter, we chose the larger of the two.

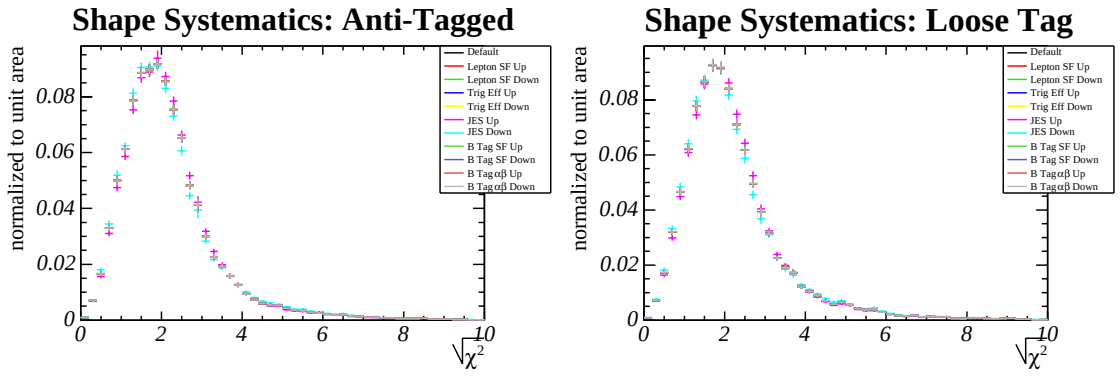
9.3.1 Total Event Yield and Tagging Rate

As discussed in Section 7.5, the main method to estimate the background is to normalize the Z+jets event yield from the MC simulation to the number of data events in the tail of the χ^2 distribution. From an expected 115 ± 21 events for $\sqrt{\chi^2} > 3$ and 144 ± 26 events for $\sqrt{\chi^2} > 3.2$ we estimated 130 ± 28 Z+jets background events. The uncertainty of this number (21.5%) is the dominant systematic uncertainty of the background estimate.

In Section 7.5 we derived a tagging rate for the Z+jets background of $15 \pm 4\%$. We obtained this number by averaging the expected tagging rates from counting the number of tagged events in the tail of the χ^2 distribution and from fitting tagging templates to the data. We expect 20 ± 6 tagged events, and 110 ∓ 6 anti-tagged events.

9.3.2 Top Mass χ^2 Shape Systematics

We have studied the systematic effects of lepton SFs, trigger efficiencies, jet energy scale, b-tagging SFs, mistag $\alpha\beta$ correction, and variations of the ALPGEN MC generator settings on the shape of the χ^2 distribution. The χ^2 distributions obtained after varying these variables are shown in Fig. 9.1. We quantify the uncertainty in the χ^2 distribution as the ratio of the number of events with $\sqrt{\chi^2} < 1.6$ to the number of events with $\sqrt{\chi^2} > 3$. The results are summarized in Table 9.3. The total systematic uncertainty is 5.1% for the anti-tagged sample, and 2.2% for the tagged sample. The uncertainty is dominated by the uncertainty of the JES, which visibly shifts the χ^2 distribution up and down. We discuss the uncertainty introduced by ALPGEN in greater detail below.



(a) (b)

Figure 9.1: Distribution of χ^2 for the Z+jets MC sample: (a) Anti-tagged event selection. (b) Loosely tagged event selection. The distributions are normalized to equal area.

The ALPGEN MC generator has several parameters which can be tuned, namely,

the p_T and E_T thresholds for MLM matching, the factorization and renormalization energy scale, the energy scale for each vertex, and the separation between the generated quarks from the matrix element calculation. The default values for these parameters are listed in Table 9.3.2. A more detailed description of the parameters can be found in Appendix A.

We have generated test samples to study the systematic effects of these settings on the shape of the χ^2 distribution. These affect the background prediction for our signal regions after optimizing the event selection. For each setting, we have generated 50,000 events each for $Z + 0, 1, 2, 3, 4$ partons, $Z + c\overline{c} + 0, 1, 2$ partons, and $Z + b\overline{b} + 0, 1, 2$ partons samples, where the Z decays into $\mu^+\mu^-$ pairs only. The list of settings tested are given in Table 9.3.2.

Parameter	**Setting**
MLM Matching (E_T, p_T)	$(15\,\text{GeV}, 15\,\text{GeV}/c)$ (default) $(10\,\text{GeV}, 10\,\text{GeV}/c)$ $(20\,\text{GeV}, 20\,\text{GeV}/c)$
Renormalization and Factorization Scale	$Q = 1.0 \times \sqrt{M_Z^2 + \sum p_T^2(p)}$ (default) $Q = 1.0 \times M_Z^2$ $Q = 2.0 \times \sqrt{M_Z^2 + \sum p_T^2(p)}$ $Q = 0.5 \times \sqrt{M_Z^2 + \sum p_T^2(p)}$ $Q = 2.0 \times M_Z^2$ $Q = 0.5 \times M_Z^2$
Vertex Energy Scale	$Q = 1.0 \times p_T$ (default) $Q = 1.0 \times m_T$ $Q = 2.0 \times p_T$ $Q = 0.5 \times p_T$ $Q = 2.0 \times m_T$ $Q = 0.5 \times m_T$

Table 9.6: Default settings for the ALPGEN MC generator and summary of systematic variations for samples with different ALPGEN settings.

In general, the different settings change the cross section of individual n-parton subsamples by as much as a factor of 20 for the samples with small cross sections; however, the inclusive cross sections remain stable to approximately 10%, and the shapes also do not change drastically. We found systematic shifts of the ratio of the number of events with $\sqrt{\chi^2} < 1.6$ to the number of events with $\sqrt{\chi^2} > 3$ that are comparable in size with the statistical uncertainty of the MC samples we used to evaluate the ALPGEN systematics, of the order of 10%.

Due to the small sample size and because of the fact that changes of the ALPGEN parameters may correspond to shifts of more or less than one standard deviation, we estimate this systematic uncertainty differently from the other sources of systematic error. The small sample size makes it difficult to isolate how much of the uncertainty results from systematic effects of changing the ALPGEN settings and how much is a result of statistical effects. This ambiguity is exasperated when taking a ratio of the number of events with $\sqrt{\chi^2} < 1.6$ to the number of events $\sqrt{\chi^2} > 3.0$, since both regions are smaller portions of the full distribution. We, therefore, take the ratio of the number of events with $\sqrt{\chi^2} < 1.6$ to the total number of Z+$\geq$4 jets events. We quote the systematic uncertainty due to the ALPGEN MC generator as the standard deviation of the shifts we obtained, instead of the biggest shift. This leads to a systematic uncertainty of 10.0% for the anti-tagged sample and 5.9% for the tagged sample.

9.4 Luminosity

As we are normalizing the FCNC acceptance to the acceptance calculated for the top cross section analysis, the common factor of the luminosity uncertainty cancels in the ratio. The background from Z+jets events is normalized to data, so that no luminosity uncertainties are required. The only pieces of our analysis that are affected by luminosity uncertainties are the backgrounds for which we predict an absolute rate, SM $t\bar{t}$ production and diboson production. We assume a 6% systematic uncertainty on these background estimates, which is already included in the estimates given in Table 7.5.

Chapter 10

Results

And the day came when the risk to remain tight in a bud
was more painful than the risk it took to blossom.
Anais Nin

We have carefully studied the background processes masquerading as the FCNC signal. We have optimized the event selection criteria to further suppress the backgrounds relative to the signal. We have determined the systematic uncertainties for the signal and backgrounds events which would fall within the final selection criteria. Armed with the final expectations for the signal, as listed in Table 8.2, we unblind the Z+$\geq$4 jets events with $\sqrt{\chi^2} < 3.0$.

We find 141 data events satisfying our pre-tag selection criteria (Z+$\geq$4 jets) while we expected 130 $\pm$ 28 events. Of these events, 17 events have one or more loose SecVtx tags; we expected 20 $\pm$ 6 tagged events. In our optimized signal region, we find 12 events satisfying the anti-tagged optimized selection criteria (listed in Table 8.1), when we expected 7.7 $\pm$ 1.8 events, and 4 events satisfying the tagged optimized selection criteria (listed in Table 8.1), when we expected 3.2 $\pm$ 1.1 events. The

properties of the 16 events in our signal regions (anti-tagged and tagged) are listed in Table 10. The properties of the events in the full pre-tag sample can be found in Appendix B.5. The data in the signal regions are consistent with background expectations.[1]

Using the Feldman-Cousins limit calculation, we set a 95% C.L. upper limit on the $t \to qZ$ branching fraction, $\mathcal{B}\ (t \to qZ) < 11.3\%$.[2] This is consistent with the expected limit of $7.8\% \pm 3.3\%$. Figure 10.1 shows the mass χ^2 distribution, all selection criteria applied except for mass χ^2, for the tagged and anti-tagged events. Data are compared against the background distribution and the signal distribution consistent with our 11.3% branching fraction limit. The events to the left of the vertical lines are within the signal regions.

We also check the kinematic distributions for events in our signal regions. We check the distributions we use in our optimization, namely mass χ^2, transverse mass, and the transverse energies of the four leading jets. We also cross check three other variables, the scalar sum of the transverse energies of the two leptons and all the jets in the event (G_T)[3], and the missing transverse energy. Although we did not use G_T and missing E_T in our analysis we check them for any anomalies that may be lurking in our event sample. Figures 10.2 and 10.3 show the kinematic variables for events satisfying our optimized anti-tagged selection criteria. Figures 10.4 and 10.5 show the kinematic variables for events satisfying our optimized tagged selection criteria. All

[1] The slight excess in data yield over background expectation is due to an excess in one bin of the transverse mass, as explained in Appendix C.4.

[2] This limit assumes a top mass of 175 GeV/c^2. We have also derived a limit assuming a top mass of 170 GeV/c^2, as described in Appendix C.1. Due to our secondary vertex tag requirement, we are slightly more sensitive to the $t \to cZ$ decay than the $t \to uZ$ decay. This difference is taken as a systematic uncertainty, as explained in Section 9.1.7.

[3] This variable is H_T without missing E_T

# run	event	lepton1		lepton2		mZ	njets	Et1	Et2	Et3	Et4	nloose	chi	m_T
#														
# anti-tagged ee														
155393	905522	e1	TCE	e0	TCE	102.76	5	87.70	55.05	30.34	20.20	0	1.1304	214.95
184868	1369737	e0	TCE	t0	TRK	86.35	4	94.66	78.39	26.43	19.75	0	1.2214	363.76
186321	13513	e1	PHX	e0	TCE	91.48	5	97.42	39.20	27.37	26.45	0	0.9520	213.66
191766	134752	t0	TRK	e0	TCE	88.14	4	59.09	46.05	25.56	17.73	0	0.7500	208.50
196354	3069410	t1	TRK	e0	TCE	88.66	4	71.46	52.35	23.76	15.93	0	0.8818	211.88
205640	9932309	e2	TCE	e0	TCE	89.51	4	68.02	34.11	30.90	15.14	0	1.1260	244.03
219606	782276	e1	TCE	e0	TCE	88.93	5	67.16	66.70	39.99	22.81	0	1.2833	301.35
221723	4413196	e0	TCE	t40	TRK	93.44	6	123.32	40.86	38.87	37.30	0	1.2267	276.39
#														
# anti-tagged mumu														
194382	5095623	t1	TRK	m0	CMUP	104.95	4	48.52	41.90	31.16	28.89	0	1.1193	202.80
194907	3171094	m1	CMUP	m0	CMUP	92.36	4	184.62	31.71	30.30	23.92	0	1.3249	436.76
199983	1771296	m0	CMUP	t0	TRK	89.33	5	60.10	55.39	30.06	21.17	0	1.0078	236.68
209535	406891	t1	TRK	m0	CMX	89.10	4	73.20	72.60	26.38	19.80	0	1.1559	223.68
#														
# tagged														
203509	331412	e1	TCE	e0	TCE	95.02	4	124.69	45.63	21.42	20.55	1	1.2307	407.68
204721	417698	t0	TRK	e0	TCE	92.38	4	52.43	33.35	31.11	29.16	1	1.5254	239.32
191545	898092	m1	CMUP	m0	CMUP	87.51	5	80.19	64.56	20.37	18.37	2	1.0601	202.52
196276	8472174	m0	CMUP	t0	TRK	101.89	4	82.93	61.59	53.10	45.43	1	1.2820	291.15

Table 10.1: List of the 16 candidate events after the final anti-tagged and tagged selections. The table gives the run and event number, the types of the two leptons that form the Z, the Z mass, the number of jets, the transverse momenta of the four leading jets, the number of loose tags, $\sqrt{\chi^2}$, and the transverse mass.

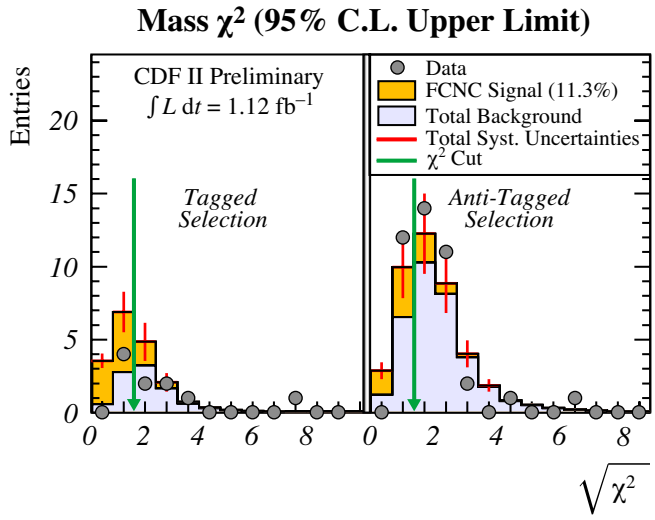


Figure 10.1: Mass χ^2 distributions for the optimized tagged and anti-tagged data samples. The data points are compared to the the background prediction and the expected FCNC yield at the observed 95% C.L. upper limit of $\mathcal{B}(t \to qZ) = 11.3\%$. The data is consistent with the background prediction.

plots are N-1 plots, i.e. the selection criteria were applied for all the variables except for the one plotted. The expected backgrounds are normalized to the data event yields by letting the Z+jets contribution float while keeping the SM top and diboson contributions fixed. All shapes are consistent with the background. The kinematic distributions of the full, unoptimized pre-tag sample are shown in Appendix B.

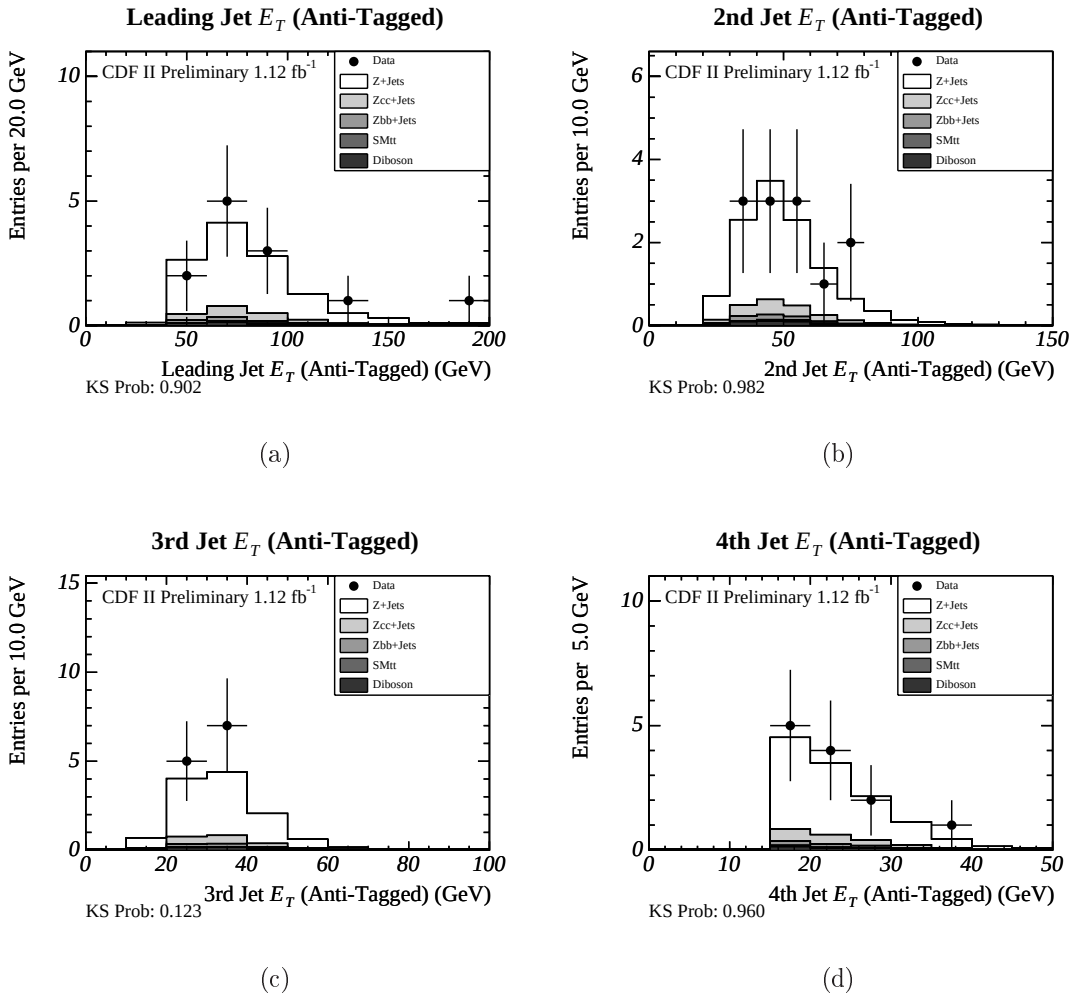
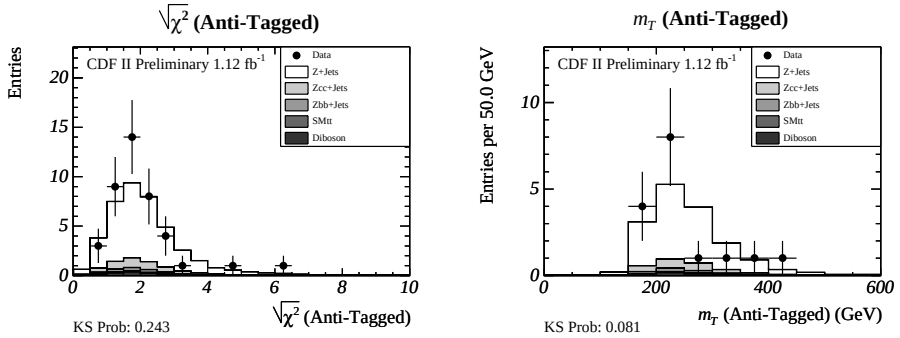
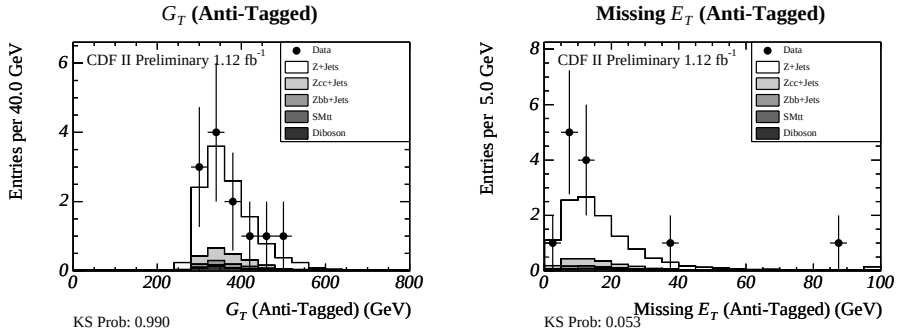


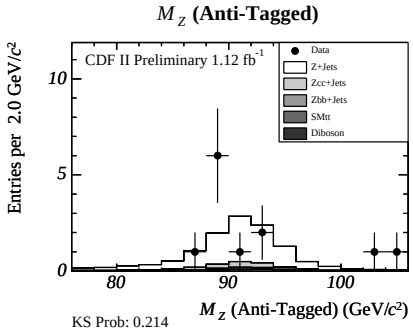
Figure 10.2: Transverse energy distributions of the four leading jets for the anti-tagged selection. The expected backgrounds are normalized to the data event yield. All distributions are $N-1$ distributions, (*i.e.*, all selection cuts are applied except for the kinematic variable shown).



(a) (b)



(c) (d)



(e)

Figure 10.3: Kinematic distributions for the anti-tagged selection: $\sqrt{\chi^2}$, transverse mass m_T, scalar sum of jet and lepton transverse momenta/energies G_T, missing transverse energy, and Z invariant mass M_Z. The expected backgrounds are normalized to the data event yield. All distributions are $N-1$ distributions.

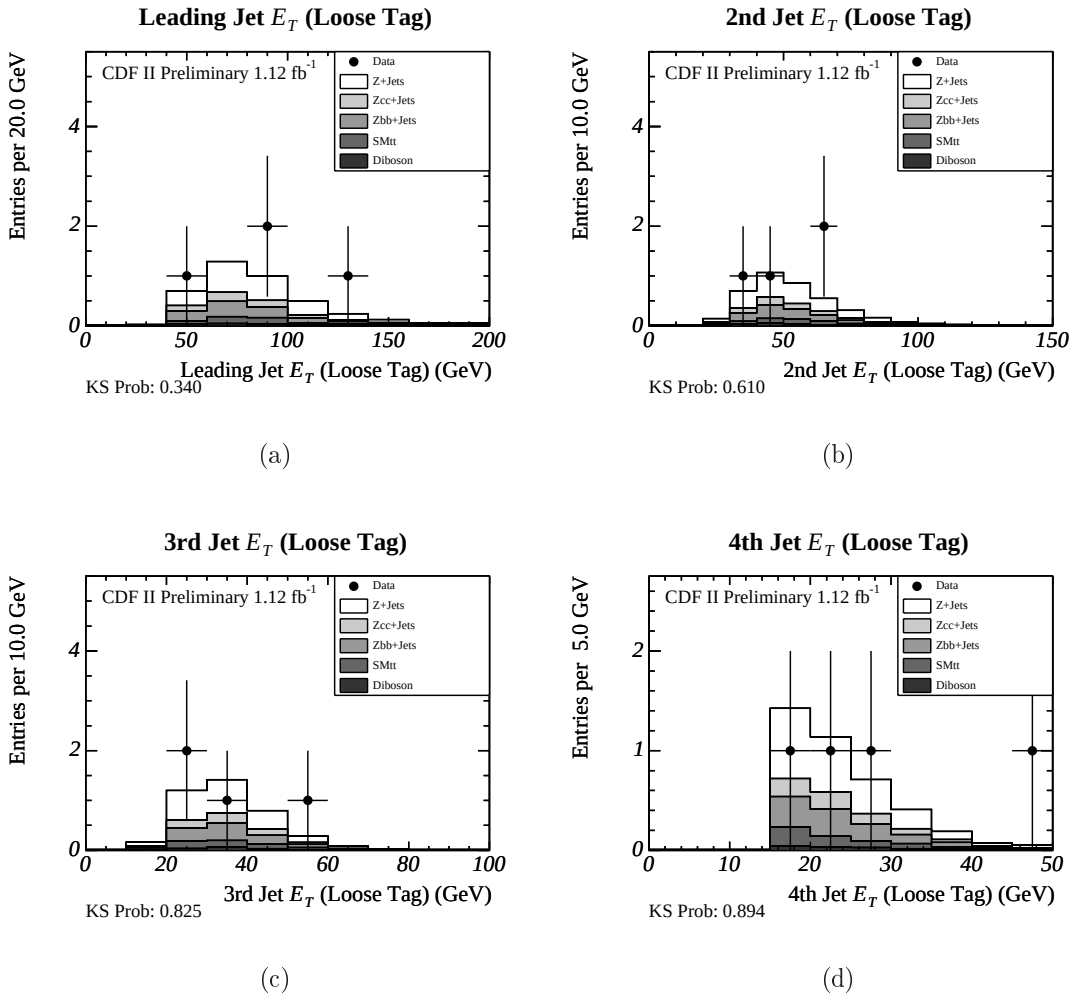
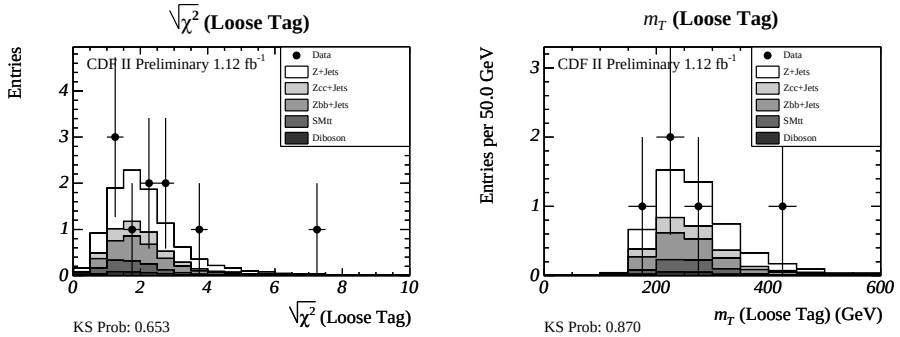
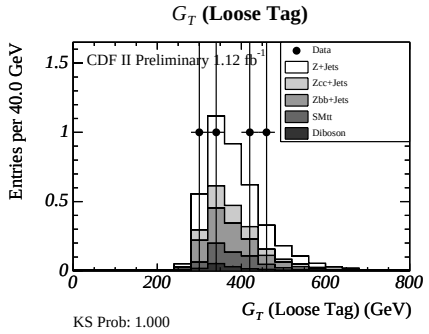


Figure 10.4: Transverse energy distributions of the four leading jets for the tagged selection. The expected backgrounds are normalized to the data event yield. All distributions are $N-1$ distributions.

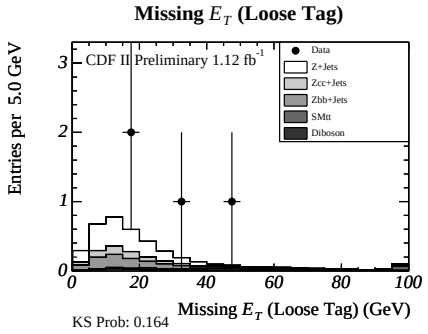


(a)

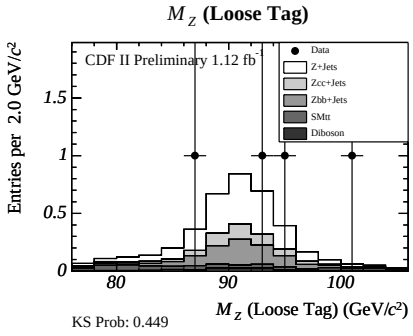
(b)



(c)



(d)



(e)

Figure 10.5: Kinematic distributions for the tagged selection: $\sqrt{\chi^2}$, transverse mass m_T, scalar sum of jet and lepton transverse momenta/energies G_T, missing transverse energy, and Z invariant mass M_Z. The expected backgrounds are normalized to the data event yield. All distributions are $N-1$ distributions.

Chapter 11

Conclusions

Yet all experience is an arch wherethro'
Gleams that untravell'd world, whose margin fades
For ever and for ever when I move.
Alfred Tennyson (from "Ulysses")

We have conducted a blind search for the top flavor changing neutral current decay $t \to qZ$ in $p\overline{p}$ collisions at $\sqrt{s} = 1.96$ TeV , using a data sample corresponding to 1.12 fb^{-1} of integrated luminosity collected by the CDF II detector. Our main signal of interest is $t\overline{t}$ events where one top decays via the FCNC mode to Z boson and a c or u quark while the other one decays via the SM mode to a W boson and a b quark, with the Z decaying to e^+e^- or $\mu^+\mu^-$ and W decaying to two quarks. Our signature is a Z reconstructed from two oppositely-charged electrons or muons and four or more jets. We account for extra acceptance from $t\overline{t}$ events where both tops decay to qZ and from other decays of the Z and W.

The main background is Z bosons produced in association with jets. There are also smaller contributions from standard model decays of $t\overline{t}$ pairs where a Z is falsely

reconstructed, from pairs of gauge bosons, ZZ and WZ, and from WW and W+jets events. We further refined our event selection criteria, exploiting the differences in the kinematics of signal and background events.

The data are consistent with the background prediction, and we set a limit for the $t \to qZ$ branching fraction, $\mathcal{B}(t \to qZ) < 11.3\%$. Although result is far above the standard model prediction of $\mathcal{B}(t \to qZ) \sim \mathcal{O}(10^{-14})$ [5] and also above the predictions from new physics models, $\mathcal{B}(t \to qZ) < \mathcal{O}(10^{-2})$ [6], it is the current world's best limit, improving the 13.7% set by L3 [8] by ∼20% and the CDF Run I limit of 33% [7] by almost a factor of three.

11.1 Future Directions

While we have set the world's best limit on the branching fraction for the top FCNC decay $t \to qZ$, further work can be done to improve it. Currently, CDF has collected 2 fb^{-1} of integrated luminosity and will collect 4 fb^{-1} or more by the end of Tevatron Run II. A simple extension of this analysis would be to repeat this counting experiment using the larger datasets. Figure 11.1 shows the projection of the expected limit from the current dataset with 1.12 fb^{-1} of integrated luminosity to the 4 fb^{-1} expected toward the end of the current Tevatron run. My collaborators at CDF are also planning to improve the limit by fitting the mass χ^2 distribution. A first attempt at this procedure is described in Appendix C.3 and already show promise. As more data is collected, one can also search for the "trilepton channel", $t\bar{t} \to qZWb$ with $Z \to \ell^+\ell^-$ and $W \to \ell\nu$, which is a rarer but has fewer background processes mimicking the signal.

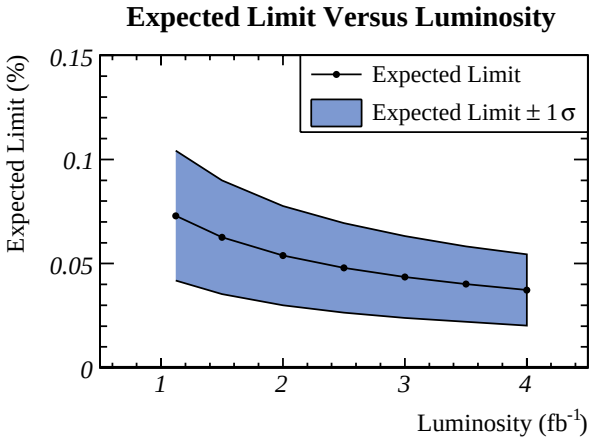


Figure 11.1: Projected expected limit versus luminosity for the current "counting experiment" analysis, for up to 4 fb^{-1} of integrated luminosity expected by the end of Tevatron Run II.

The high energy physics community is also eagerly awaiting the activation of the Large Hadron Collider (LHC) at CERN in Geneva, Switzerland. Colliding protons at $\sqrt{s} = 14$ TeV , the LHC will be a "top factory", since the $t\bar{t}$ production cross section at that energy is $\sigma_{t\bar{t}} = 833$ pb, roughly one hundred times bigger than the $t\bar{t}$ production cross section at the Tevatron. The LHC will also have higher instantaneous luminosity and, therefore, will collect a much larger dataset. Studies indicate that the LHC will be able to discover the FCNC decay $t \to qZ$ at 5σ down to branching fractions of $\mathcal{O}(10^{-4})$ [84] or place a 95% C.L. limit on the branching fraction of $\mathcal{O}(10^{-5})$ [85] with a dataset corresponding to 100 fb^{-1} of integrated luminosity. This will confirm or rule out many new physics models; or, perhaps, nature will surprise us with a $t \to qZ$ rate higher than any theory has predicted.

Appendix A

ALPGEN Tuning and Systematics

In the room women come and go
Talking of Michelangelo.
T. S. Eliot (from "The Love Song of J. Alfred Prufrock")

In this appendix, we will describe the work done to understand the ALPGEN Monte Carlo generator which we use to model Z+jets events, the dominant background in our analysis. At CDF, ALPGEN is also used to model other multijet processes, W+Jets and QCD multijets, which are backgrounds to other top analyses. While we will limit the discussion to the Z+jets samples except for some comparisons, the conclusions reached are applicable to other multijet processes.

A.1 ALPGEN Event Generator

ALPGEN [56] generates events using leading order matrix element calculations for the hard interaction. At CDF, ALPGEN is combined with Pythia or HERWIG which performs parton showers, generating initial and final state radiation for the

ALPGEN events. (Note that the events ALPGEN generates are leading order but the events from ALPGEN +PYTHIA (or HERWIG) are not a fixed order approximation.) The renormalization and factorization energy scale is set based on the mass of the electroweak boson and/or the p_T of the partons in the event. Additionally, the energy scale for determining the strong coupling, α_s is determined at each vertex.

Instead of generating an inclusive Z+jets (W+Jets) sample, ALPGEN generates different $Z(W)$+N parton ($Z(W)$+Np) samples and these samples must be added together to reproduce $Z(W)$+N jets events. At CDF, we have generated Z+0-4p, Z+$b\overline{b}$ +0-2p, and Z+$c\overline{c}$ +0-2p samples to model Z+jets events, and W+0-4p, W+$b\overline{b}$ +0-2p, W+$c\overline{c}$ +0-2p, and W+c+0-3p samples to model W+jets events. The Z bosons are forced to decay to e^+e^- or $\mu^+\mu^-$. The W bosons are forced to decay to $e\nu$ or $\mu\nu$.

There is some overlap in the phase space populated by the Z+Np samples. For example, a Z+2p event can gain an extra jet from the parton shower and contribute as a Z+3 jet event, overlapping with a Z+3p event with 3 jets. CDF uses the MLM [1] jet-parton matching procedure to remove this overlap.

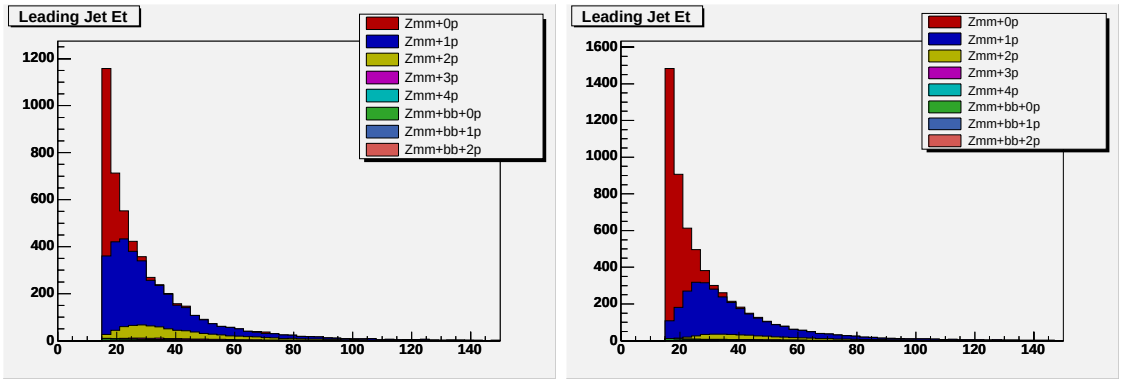
A.1.1 MLM Matching

In ALPGEN version 1, which was used in Gen 5 analyses, MLM matching was applied by hand after the samples were generated [86]. The individual $Z(W)$+Np samples could be used to reproduced the inclusive $Z(W)$+ $\geq$N jets events (e.g. $Z(W)$+3p events reproduce $Z(W)$+ $\geq$3 jets events). Adding the samples was necessary only for producing exclusive $Z(W)$+N jets samples, such as for predicting the W+jets

[1] Named after Michelangelo L. Mangano who is one of the authors of ALPGEN and developed the overlap removal scheme

backgrounds for top analyses. The disadvantage was that after matching, a large fraction of the generated events needed to be discarded, leading to a wasteful and computationally expensive procedure. In ALPGEN v2, MLM matching is built-in to the generation and events which fail the matching are discarded before the detector simulation, saving computational resources. However, the individual $Z(W)$+Np samples can no longer be used and the samples must always be combined, according to their generated cross sections.

MLM matching is a cone-based algorithm. The generated partons are matched to "theory jets" clustered from generator level (HEPG) particles, after the parton shower (by PYTHIA or HERWIG) but before hadronization. The p_T of the parton, E_T of the theory jet, and the cone size of theory jet can be set by the user. Note that the cone size is generally set to be the same as the minimum separation between the generated light flavor partons (there is no minimum separation between heavy flavor partons). The E_T requirement for the theory jet does not mean that jets below that E_T are not generated. The soft jets are generated by the parton showers. Changing the p_T and E_T thresholds changes which part of phase space is populated by each Np sample. This is illustrated in Figure A.1 where the distribution for the E_T of the leading (reconstructed) jet is shown for two different matching criteria are used. The plot on the left is from a combination of Z+Np samples where the p_T and E_T for matching were both set at 15 GeV/c (GeV). The plot on the left is from a combination of Z+Np samples where the p_T was set at 15 GeV/c and the E_T was set at 20 GeV . In the latter case, there was more contribution from the samples with smaller parton multiplicities. The default settings and the different settings



(a) (b)

Figure A.1: This is an illustration of the effects of MLM matching using leading jet E_T. (a) Shows the contribution from different Z+Np samples to the leading jet E_T when the p_T-E_T thresholds for MLM matching are set at 15 GeV/c-15 GeV . (b) Shows the contribution from different Z+Np samples to the leading jet E_T when the p_T-E_T thresholds for MLM matching are set at 15 GeV/c-20 GeV .

used to study systematic uncertainties will be discussed in Section A.2.4 and A.3, respectively.

There are two different matching modes. Exclusive matching requires that each parton is matched to a jet and vice versa. Inclusive matching requires that each parton is matched to a jet but there are allowed to be extra jets. Exclusive matching is used for all samples except the ones with the largest parton multiplicity (e.g. Z+4p, Z+$b\bar{b}$ +2p), where inclusive matching is used. Heavy flavor partons are not required to be matched.

A.2 ALPGEN Tuning

ALPGEN offers flexibility in a variety of its settings. We tune these settings to reproduce the kinematic distributions observed in data. In particular, we have studied whether using PYTHIA or HERWIG for the parton shower better models the data and we have determined what p_T and E_T criteria should be used in MLM matching.

A.2.1 Pythia vs. HERWIG Showers

We have generated a suite of Z+Np samples with ALPGEN v2 + PYTHIA and another with ALPGEN v2 + HERWIG. The Z bosons are forced to decay to $\mu^+\mu^-$. The cross sections for the two sets of Monte Carlo samples are given in Table A.2.1. Figures A.2 to A.4 show the kinematic distributions from the two sets of Monte Carlo samples compared to data. The variables plotted are: Z+N jets distribution (linear and log scale), Z p_T, G_T (the scalar sum of the p_T of the muons from the Z decay and the E_T of the jets in the event), the E_T and η of the three leading jets, and the ΔR, $\Delta\eta$, and $\Delta\phi$ between the Z and the leading jet and between the leading and second jets. Note that we have only examined $Z+ \leq$3 jets events because $Z+ \geq$4 jets events are blinded for the FCNC $t \to qZ$ search.

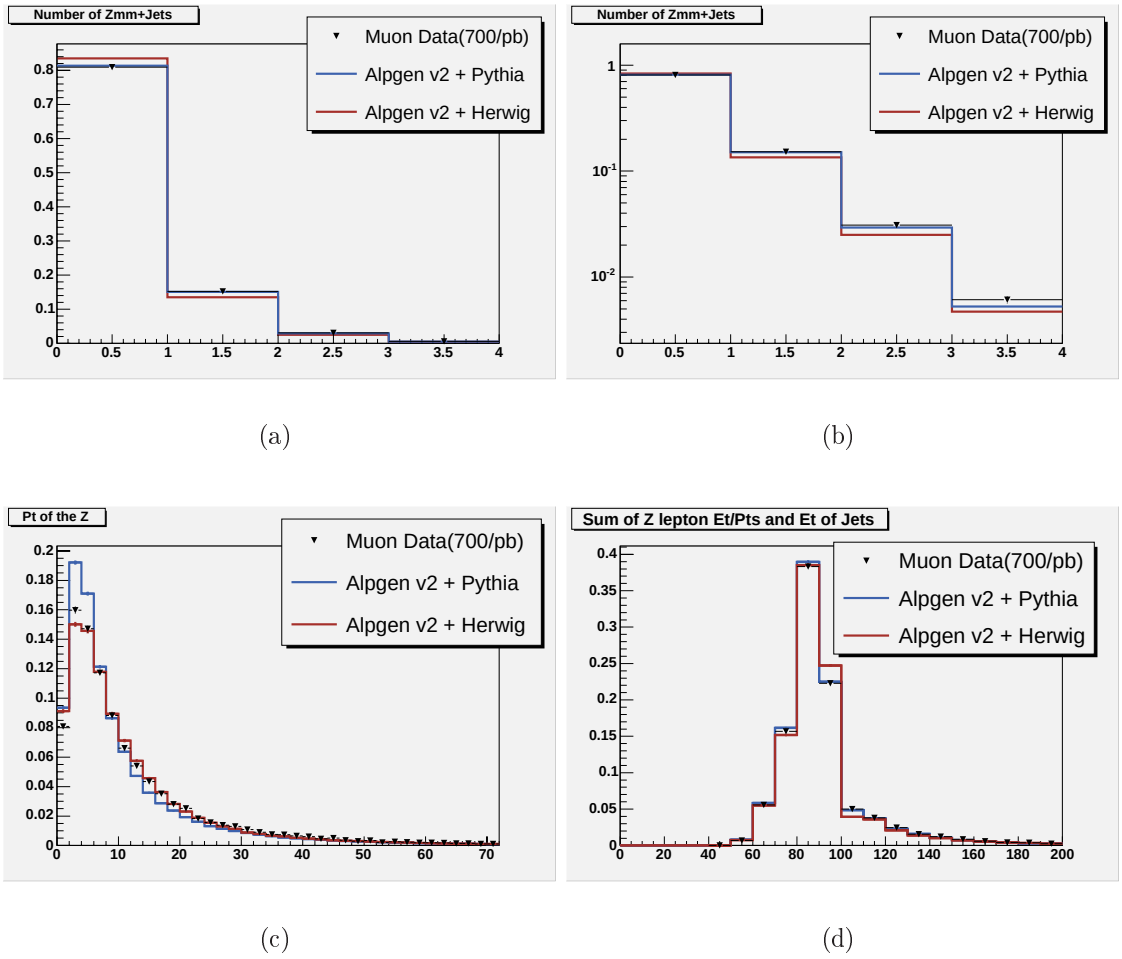
ALPGEN + PYTHIA reproduces data more accurately compared to ALPGEN + HERWIG for all distributions except for Z p_T and $\Delta\phi$ between the leading and second jets. The discrepancy in Z p_T is due to PYTHIA tuning. The discrepancy in $\Delta\phi$ between the leading and second jets is due to a tuning problem in the Z+0p sample which produces too many pencil-thin back-to-back pairs of jets. We have chosen to use PYTHIA since it is better at reproducing the kinematic distributions

Sample	ALPGEN + PYTHIA	ALPGEN + HERWIG (pb)
$Z \to \mu^+\mu^-$ +0p	158	152
$Z \to \mu^+\mu^-$ +1p	21.6	20.5
$Z \to \mu^+\mu^-$ +2p	3.47	3.88
$Z \to \mu^+\mu^-$ +3p	0.550	0.713
$Z \to \mu^+\mu^-$ +4p	0.0992	0.184
$Z \to \mu^+\mu^-$ +$c\bar{c}$ +0p	1.08	1.13
$Z \to \mu^+\mu^-$ +$c\bar{c}$ +1p	0.332	0.391
$Z \to \mu^+\mu^-$ +$c\bar{c}$ +2p	0.107	0.146
$Z \to \mu^+\mu^-$ +$b\bar{b}$ +0p	0.511	0.531
$Z \to \mu^+\mu^-$ +$b\bar{b}$ +1p	0.134	0.156
$Z \to \mu^+\mu^-$ +$b\bar{b}$ +2p	0.0385	0.0533
Total	185.92	179.68

Table A.1: Cross Sections for ALPGEN Matching Systematic Samples

in data, in general. Furthermore, top signal samples are produced with PYTHIA, so choosing PYTHIA keeps signal and background Monte Carlo simulations consistent.[2]

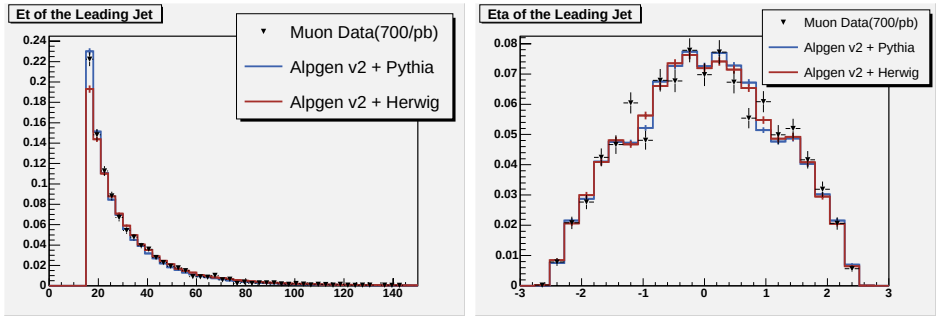
[2]In Gen 5, HERWIG was used to perform parton showers.



(a) (b)

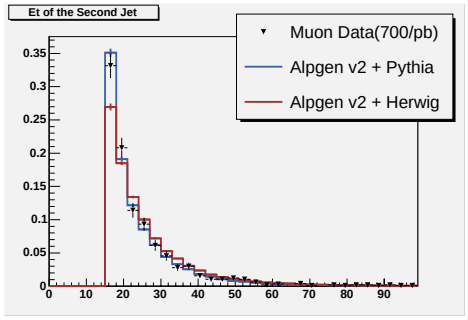
(c) (d)

Figure A.2: Kinematic distributions from ALPGEN v2 + PYTHIA (blue line) and ALPGEN v2 + HERWIG (red line), compared to data (black triangle). (a) Z+Njets (b) Z+Njets, log scale (c) Z p_T (d) G_T, the scalar sum of p_T of the muons from Z decay and E_T of the jets in the event

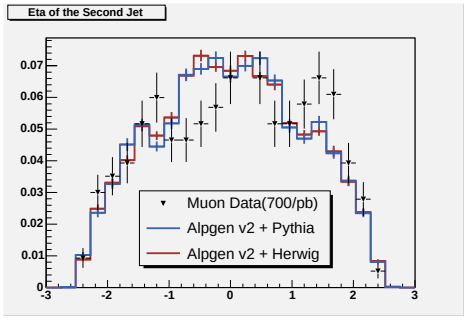


(a)

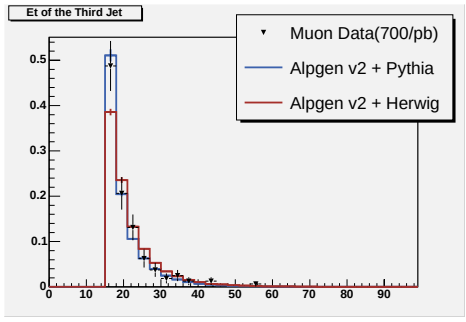
(b)



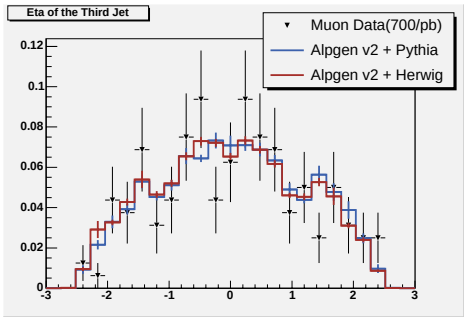
(c)



(d)



(e)



(f)

Figure A.3: Kinematic distributions from ALPGEN v2 + PYTHIA (blue line) and ALPGEN v2 + HERWIG (red line), compared to data (black triangles). (a) Leading Jet E_T (b) Leading Jet η (c) Second Jet E_T (d) Second Jet η (e) Third Jet E_T (f) Third Jet η

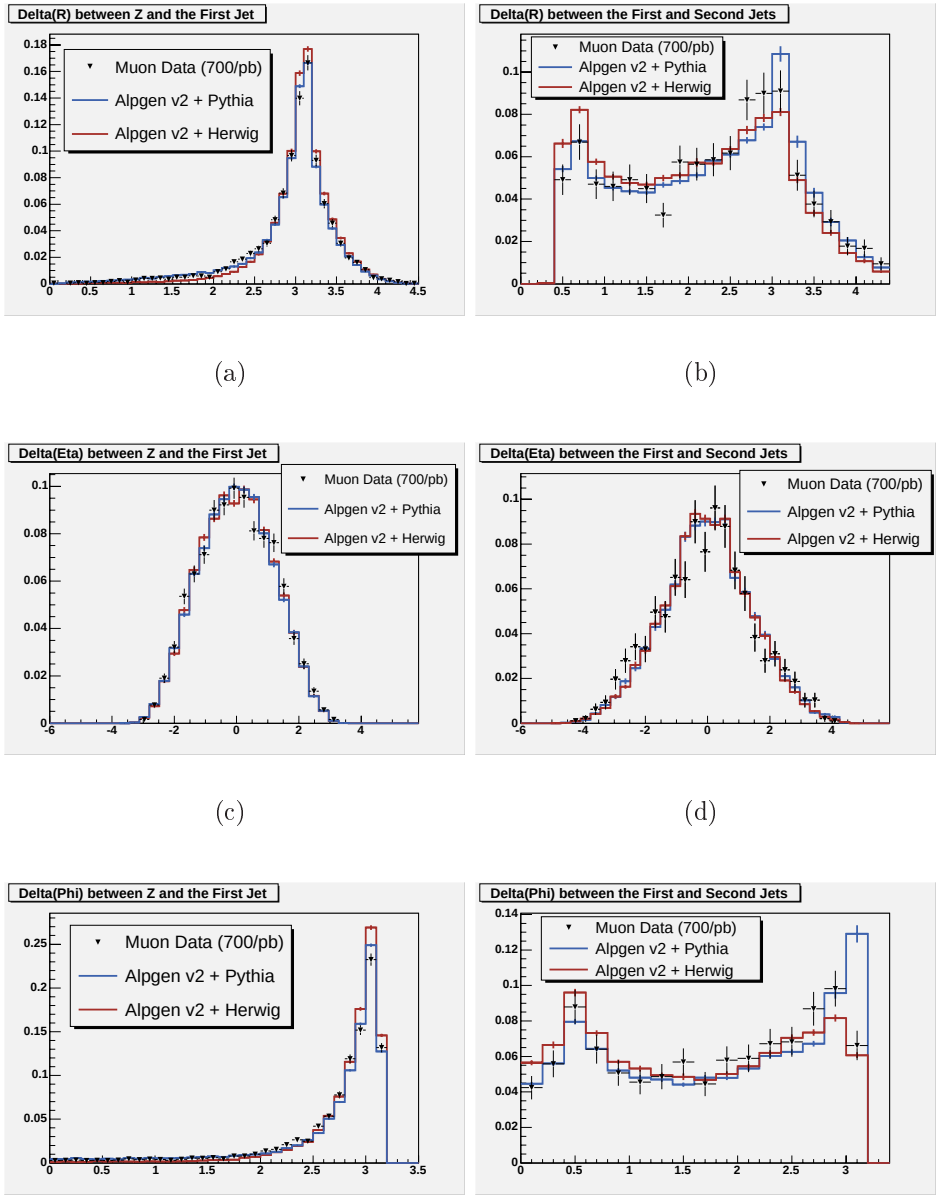
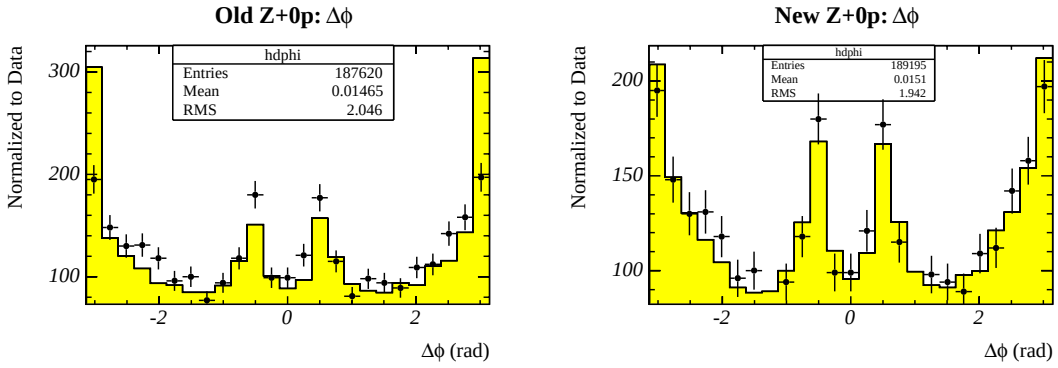


Figure A.4: Kinematic distributions from ALPGEN v2 + PYTHIA (blue line) and ALPGEN v2 + HERWIG (red line), compared to data (black triangles). (a) ΔR(Z-Jet1) (b) ΔR(Jet1-Jet2) (c) $\Delta\eta$(Z-Jet1) (d) $\Delta\eta$(Jet1-Jet2) (e) $\Delta\phi$(Z-Jet1) (f) $\Delta\phi$(Jet1-Jet2)



(a) (b)

Figure A.5: Z+0p bug fix. (a) $\Delta\phi$(Jet1-Jet2) before the Z+0p bug fix. There is an excess of back-to-back jets in MC. (b) $\Delta\phi$(Jet1-Jet2) after the Z+0p bug fix. Data and MC match well.

A.2.2 Z+0p Fix

The Z+0p sample was fixed to eliminate the over production of pencil-thin back-to-back jets. The $\Delta\phi$ distributions before and after the fix are shown in Figure A.5. Unfortunately, this fix was implemented only after the rest of the studies were performed and the $\Delta\phi$(jet1-jet2) plots shown in later sections will show the discrepancy.

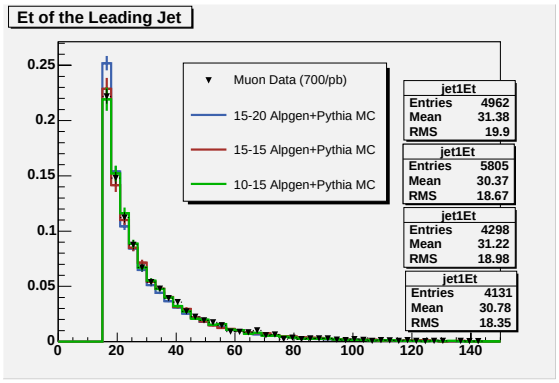
A.2.3 Matching Parameters

We also studied whether the parton p_T and jet E_T for matching should be the same or whether the jet E_T should be higher than the parton p_T. The concern is that the E_T from clustered jets has a resolution. Consequently, if the parton p_T and jet E_T matching criteria were the same, partons above the threshold can "slosh" to form jets with E_T below the cut but no partons below the cut can slosh to form

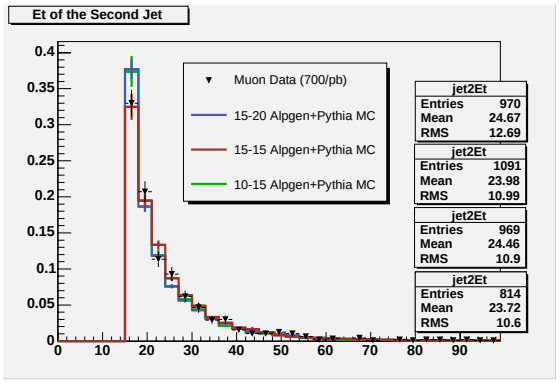
jets with E_T above the cut. Therefore, setting the E_T threshold above the p_T may be the correct choice. Figure A.6 show the E_T of the leading, second, and third jets with p_T-E_T matching thresholds set at 10-15, 15-15, and 15-20 compared to data. Surprisingly, the 15-15 requirement reproduces the jet E_T distributions the best (other kinematic distributions do not show a significant difference for between the three settings). The 15-15 setting is also faster to generate than the 10-15 and 15-20 settings because a larger fraction of the events are discarded when the p_T and E_T are set at different values. Therefore, we keep the 15-15 setting as our default. Note that the reconstructed jet E_T cut is set at 15 GeV after Level 5 jet energy corrections.

A.2.4 Default ALPGEN Parameters

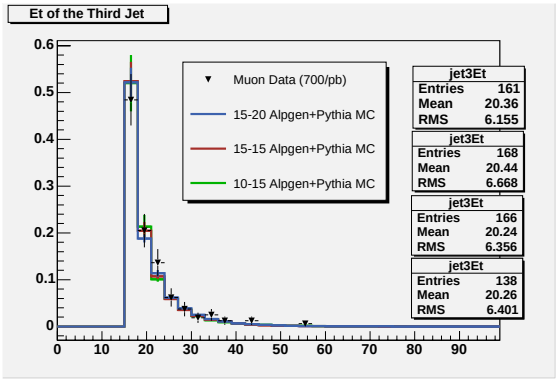
The parameters used for the main (default) CDF ALPGEN Z samples are listed in Table A.2.4. Note that settings which will not be varied for the systematic samples but are important for heavy flavor settings are also listed.



(a)



(b)



(c)

Figure A.6: Comparison of p_T-E_T thresholds of 15 GeV/c-15 GeV , 15 GeV/c-20 GeV , and 10 GeV/c-15 GeV . The 15-15 setting reproduces the data most closely (a) Leading Jet E_T (b) Second Jet E_T (c) Third Jet E_T. Other distributions do not show a significant difference between the different settings.

Parameter	Setting
ALPGEN_PROCESS	zjet
MLL_MIN	75.0
MLL_MAX	105.0
QFACTOR	1.0
KTFACTOR	1.0
QOPT	1
CLUOPT	1
PTMIN_L 1.0	
DRMIN_L 0.0	
ETAMAX_L 5.0	
PTMIN_JET	15
DRMIN_JJ	0.4
ETAMAX_JET 3.0	
MET_MIN 0.0	
ETMIN_CLUS_MATCH	15
DRMIN_CLUS_MATCH	0.4
PTMIN_B 0.0	
DRMIN_BB 0.0	
ETAMAX_B 10.0	

Table A.2: Default ALPGEN Parameters

A.3 ALPGEN Systematics

We discussed the default settings for ALPGEN in the section above. We vary these settings to study the systematic errors caused by our choice of a set of parameters. We vary the p_T and E_T matching thresholds, ΔR between light partons and matching cone size, the renormalization and factorization energy scale, and the energy scale for α_s at each vertex. The different settings for the systematics study are summarized in Table A.3.

Parameter	**Setting**
MLM Matching (E_T, p_T)	$(15\ \text{GeV}, 15\ \text{GeV}/c)$ (default) $(10\ \text{GeV}, 10\ \text{GeV}/c)$ $(20\ \text{GeV}, 20\ \text{GeV}/c)$
Renormalization and Factorization Scale	$Q = 1.0 \times \sqrt{M_Z^2 + \sum p_T^2(p)}$ (default) $Q = 1.0 \times M_Z^2$ $Q = 2.0 \times \sqrt{M_Z^2 + \sum p_T^2(p)}$ $Q = 0.5 \times \sqrt{M_Z^2 + \sum p_T^2(p)}$ $Q = 2.0 \times M_Z^2$ $Q = 0.5 \times M_Z^2$
Vertex Energy Scale	$Q = 1.0 \times p_T$ (default) $Q = 1.0 \times m_T$ $Q = 2.0 \times p_T$ $Q = 0.5 \times p_T$ $Q = 2.0 \times m_T$ $Q = 0.5 \times m_T$

Table A.3: Default settings for the ALPGEN MC generator and summary of systematic variations for samples with different ALPGEN settings. This table is also listed in Chapter 9.

A.3.1 Matching Parameters

Several parameters can be changed in MLM matching. The parton p_T and theory E_T can be set. In the section above, we showed that setting the p_T and E_T thresholds to be the same, at $(p_T, E_T) = (15\ \mathrm{GeV}/c, 15\ \mathrm{GeV})$ reproduced data well. We have generated systematic samples at $(10\ \mathrm{GeV}/c, 10\ \mathrm{GeV})$ and $(20\ \mathrm{GeV}/c, 20\ \mathrm{GeV})$ for systematic studies.

ALPGEN requires minimum separation between light partons and between light and heavy partons. There is no separation requirement for heavy partons. MLM matching also has a requirement for the cone size within which to cluster the energy deposition. These two ΔR requirements are set to be the same, $\Delta R = 0.4$ in the default scenario. We have generated a sample with the separation and clustering cone size set to $\Delta R = 0.7$.

Figures A.7 to A.9 shows the kinematic distributions for the default and systematics matching samples compared to data. The variables plotted are: Z+N jets distribution (linear and log scale), Z p_T, G_T (the scalar sum of the p_T of the muons from the Z decay and the E_T of the jets in the event), the E_T and η of the three leading jets, and the ΔR, $\Delta\eta$, and $\Delta\phi$ between the Z and the leading jet and between the leading and second jets. Overall, the kinematic distributions were similar between the different ALPGEN matching sessions and match data well. The generated cross sections for the different Np samples and the total cross section are given in Table A.3.1.

Sample	Default, 15-15, $\Delta R = 0.4$ (pb)	10-10 (pb)	20-20 (pb)	$\Delta R = 0.7$ (pb)
$Z \to \mu^+\mu^-$ +0p	158	141	169	155
$Z \to \mu^+\mu^-$ +1p	21.6	32.1	15.2	25.8
$Z \to \mu^+\mu^-$ +2p	3.47	6.67	2.01	4.51
$Z \to \mu^+\mu^-$ +3p	0.550	1.375	0.257	0.710
$Z \to \mu^+\mu^-$ +4p	0.0992	0.3579	0.0351	0.1140
$Z \to \mu^+\mu^-$ $+c\bar{c}$ +0p	1.08	0.95	1.15	1.40
$Z \to \mu^+\mu^-$ $+c\bar{c}$ +1p	0.332	0.418	0.258	0.358
$Z \to \mu^+\mu^-$ $+c\bar{c}$ +2p	0.107	0.194	0.065	0.116
$Z \to \mu^+\mu^-$ $+b\bar{b}$ +0p	0.511	0.439	0.551	0.492
$Z \to \mu^+\mu^-$ $+b\bar{b}$ +1p	0.134	0.167	0.106	0.147
$Z \to \mu^+\mu^-$ $+b\bar{b}$ +2p	0.0385	0.0708	0.0232	0.0424
Total	185.92	183.74	188.66	188.10

Table A.4: Cross Sections for ALPGEN Matching Systematic Samples

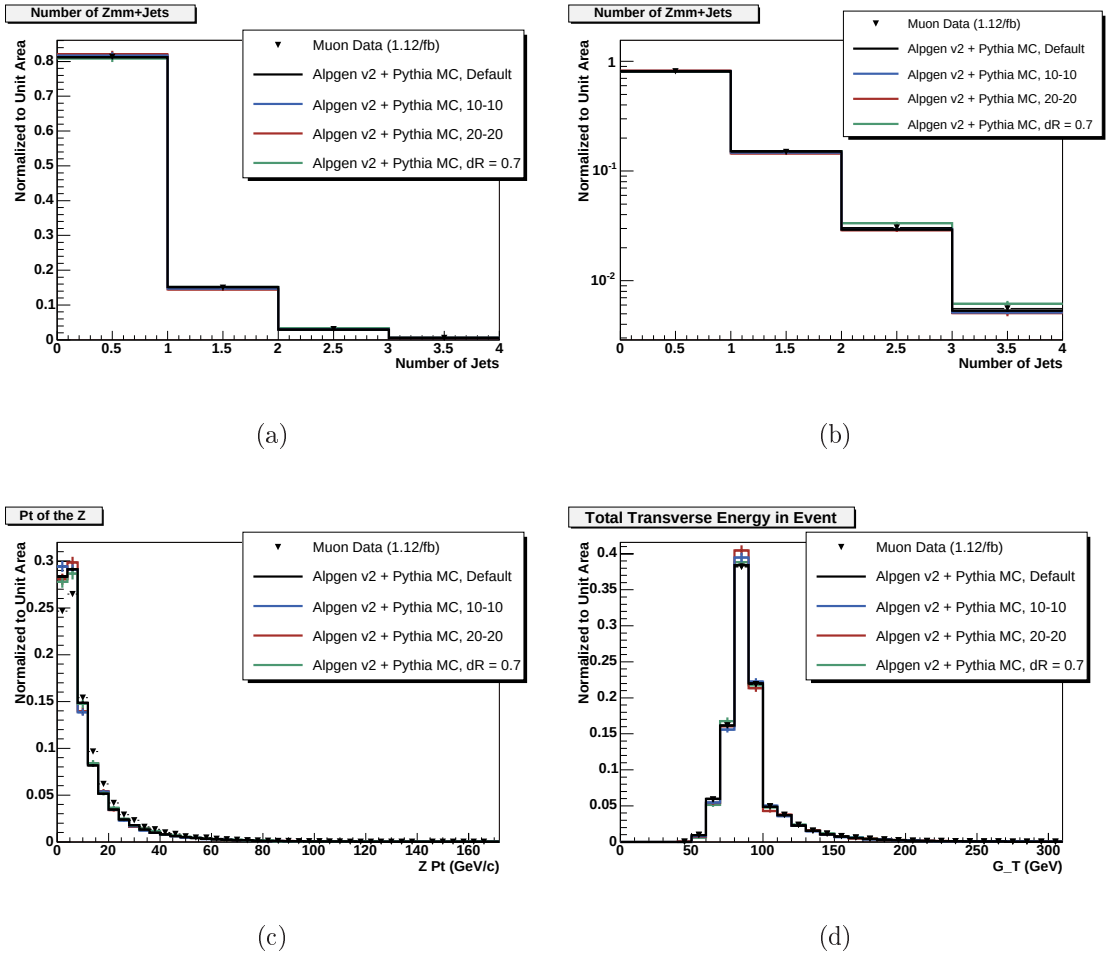


Figure A.7: Kinematic distributions from changing matching thresholds, the default, $(15\,\mathrm{GeV}, 15\,\mathrm{GeV}/c)$, $\Delta R = 0.4$ (black line), $(10\,\mathrm{GeV}, 10\,\mathrm{GeV}/c)$ (blue line), $(20\,\mathrm{GeV}, 20\,\mathrm{GeV}/c)$ (red line), and $DeltaR = 0.7$ (green line) compared to data (black triangles). (a) Z+Njets (b) Z+Njets, log scale (c) Z p_T (d) G_T, the scalar sum of p_T of the muons from Z decay and E_T of the jets in the event

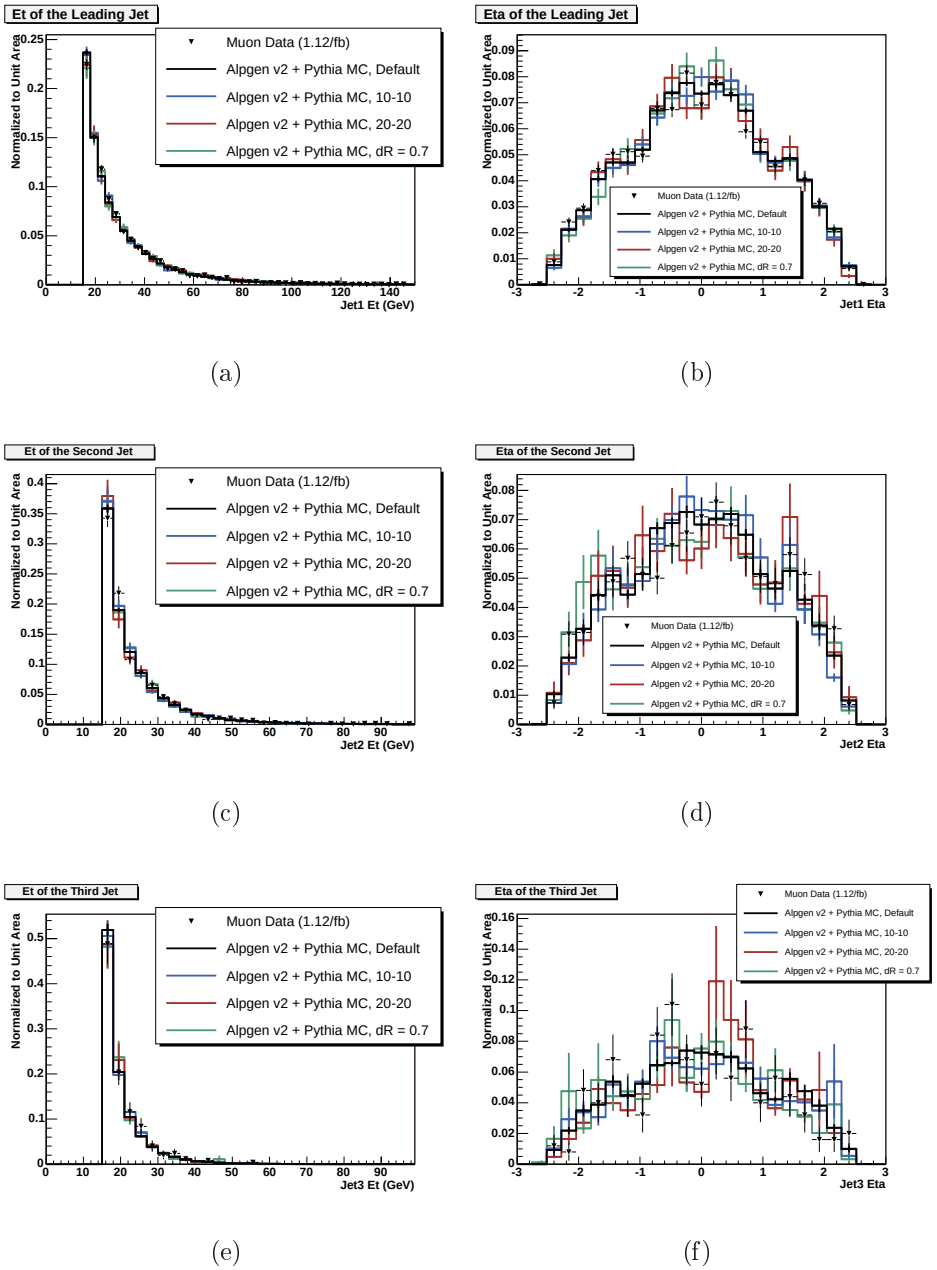


Figure A.8: Kinematic distributions from changing matching thresholds, the default, $(15\ \mathrm{GeV}, 15\ \mathrm{GeV}/c)$, $\Delta R = 0.4$ (black line), $(10\ \mathrm{GeV}, 10\ \mathrm{GeV}/c)$ (blue line), $(20\ \mathrm{GeV}, 20\ \mathrm{GeV}/c)$ (red line), and $DeltaR = 0.7$ (green line) compared to data (black triangles). (a) Leading Jet E_T (b) Leading Jet η (c) Second Jet E_T (d) Second Jet η (e) Third Jet E_T (f) Third Jet η

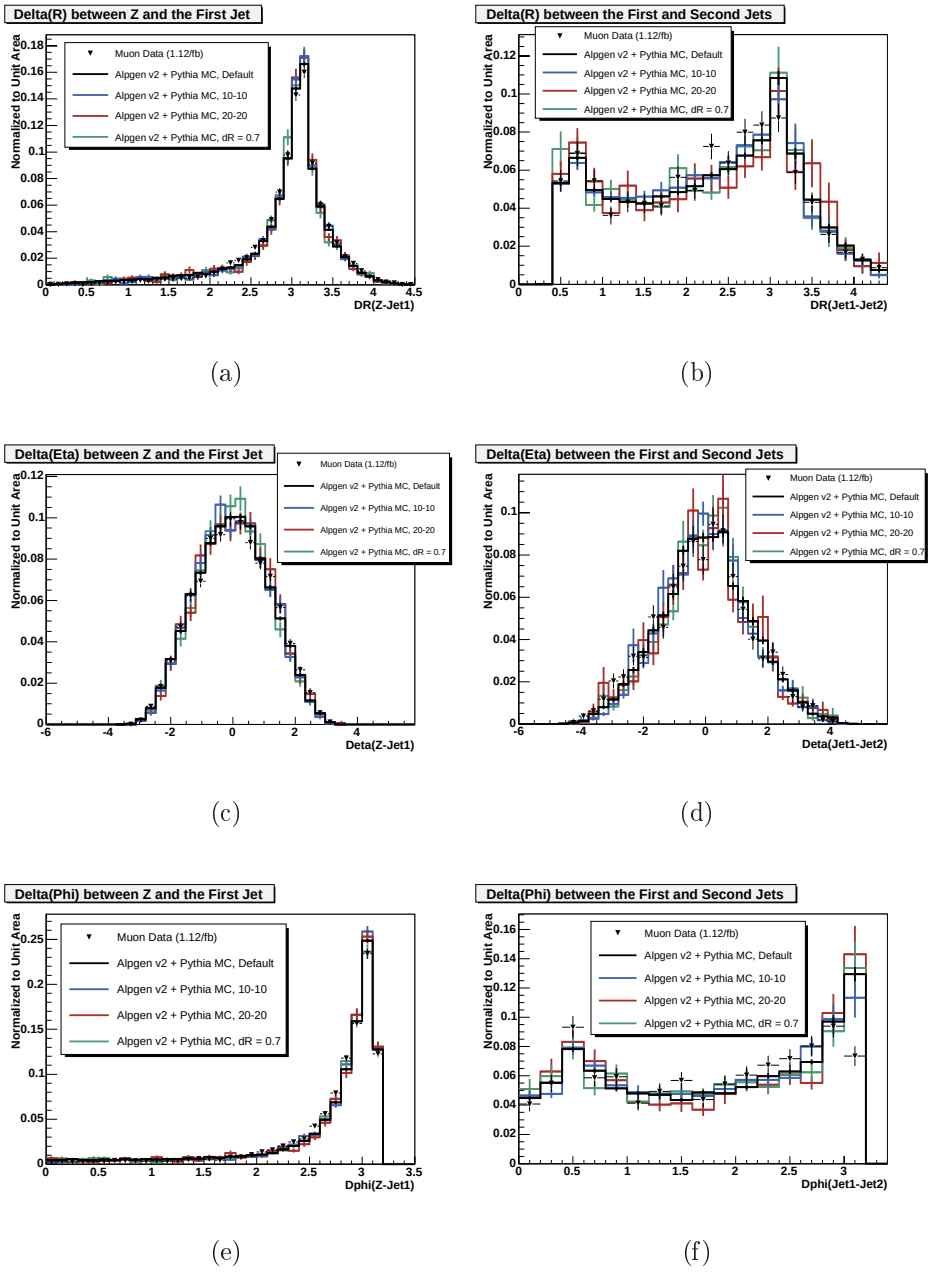


Figure A.9: Kinematic distributions from changing matching thresholds, the default, $(15\ \text{GeV}, 15\ \text{GeV}/c)$, $\Delta R = 0.4$ (black line), $(10\ \text{GeV}, 10\ \text{GeV}/c)$ (blue line), $(20\ \text{GeV}, 20\ \text{GeV}/c)$ (red line), and $DeltaR = 0.7$ (green line), compared to data (black triangles). (a) ΔR(Z-Jet1) (b) ΔR(Jet1-Jet2) (c) $\Delta\eta$(Z-Jet1) (d) $\Delta\eta$(Jet1-Jet2) (e) $\Delta\phi$(Z-Jet1) (f) $\Delta\phi$(Jet1-Jet2)

A.3.2 Renormalization and Factorization Energy Scale

The renormalization and factorization energy scale can be set in different ways. The default is to choose the scale Q = qfac $\times \sqrt{M_Z^2 + \sum p_T^2(p)}$, where qfac is a multiplicative factor set to 1.0. We have examined the effect of choosing the scale Q = qfac $\times \sqrt{M_Z^2}$ and varying qfac to 2.0 and 0.5 for both options.

Figures A.10 to A.12 shows the kinematic distributions for the default and systematics matching samples compared to data. The variables plotted are: Z+N jets distribution (linear and log scale), Z p_T, G_T (the scalar sum of the p_T of the muons from the Z decay and the E_T of the jets in the event), the E_T and η of the three leading jets, and the ΔR, $\Delta\eta$, and $\Delta\phi$ between the Z and the leading jet and between the leading and second jets. Varying the multiplicative factor, qfac has a bigger effect than changing the variable which sets the renormalization and factorization scale from $\sqrt{M_Z^2 + \sum p_T^2(p)}$ to $\sqrt{M_Z^2}$. The generated cross sections for the different Np samples and the total cross section are given in Table A.3.2.

Sample	Default, $\sqrt{M_Z^2+\sum p_T^2(p)}$ (pb)	Mz	$Q = 2.0 \times \sqrt{M_Z^2+\sum p_T^2(p)}$ (pb)	$Q = 0.5 \times \sqrt{M_Z^2+\sum p_T^2(p)}$ (pb)	$Q = 2.0 \times M_Z$ (pb)	$Q = 0.5 \times M_Z$ (pb)
$Z \to \mu^+\mu^-$ +0p	158	159	155	172	156	171
$Z \to \mu^+\mu^-$ +1p	21.6	22.1	20.1	25.4	20.1	25.9
$Z \to \mu^+\mu^-$ +2p	3.47	3.64	3.12	4.30	3.20	4.63
$Z \to \mu^+\mu^-$ +3p	0.550	0.606	0.479	0.709	0.505	0.803
$Z \to \mu^+\mu^-$ +4p	0.0992	0.1150	0.0850	0.1270	0.0923	0.1520
$Z \to \mu^+\mu^-$ +$c\bar{c}$ +0p	1.08	1.10	0.82	1.54	0.83	1.57
$Z \to \mu^+\mu^-$ +$c\bar{c}$ +1p	0.332	0.355	0.242	0.504	0.255	0.550
$Z \to \mu^+\mu^-$ +$c\bar{c}$ +2p	0.107	0.124	0.076	0.162	0.086	0.194
$Z \to \mu^+\mu^-$ +$b\bar{b}$ +0p	0.511	0.525	0.383	0.753	0.390	0.780
$Z \to \mu^+\mu^-$ +$b\bar{b}$ +1p	0.134	0.148	0.091	0.209	0.104	0.234
$Z \to \mu^+\mu^-$ +$b\bar{b}$ +2p	0.0385	0.0460	0.0272	0.8340	0.0311	0.0729
Total	185.92	187.72	180.23	206.10	181.33	205.96

Table A.5: Cross Sections for ALPGEN Renormalization and Factorization Scale Systematic Samples.

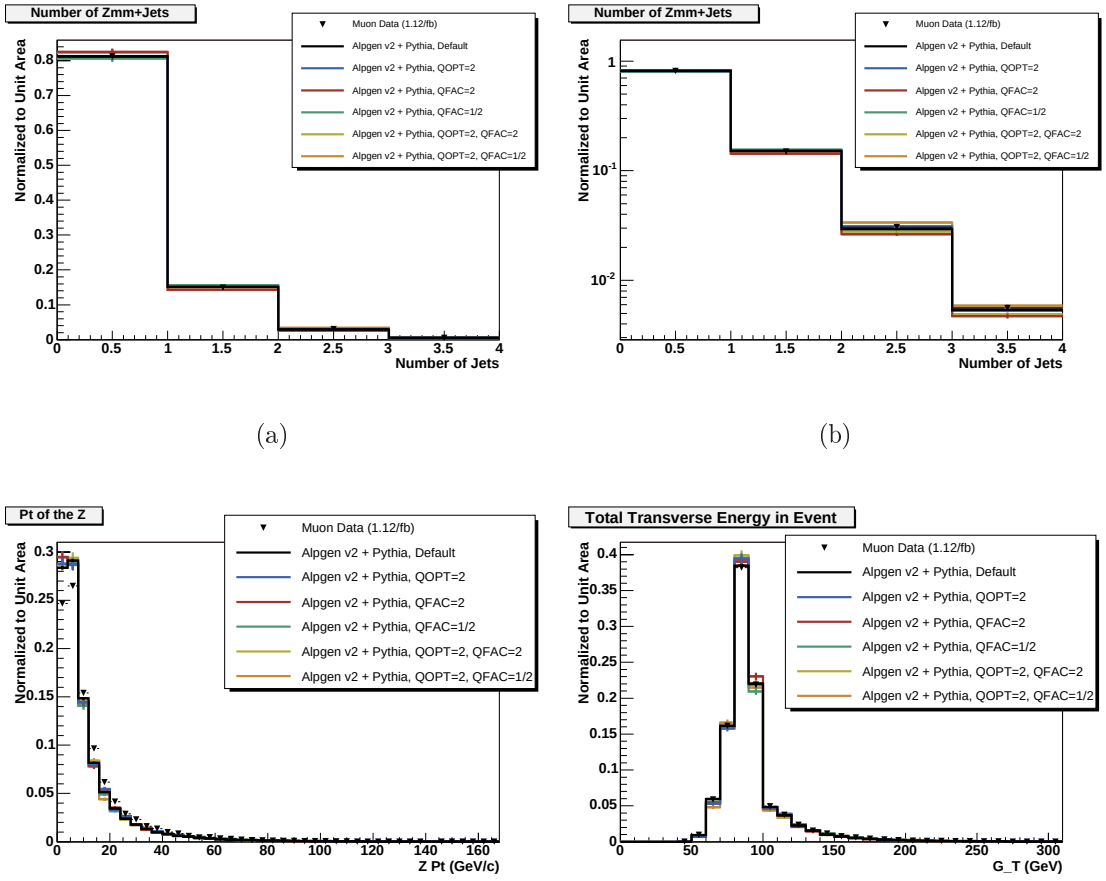


Figure A.10: Kinematic distributions from changing the renormalization and factorization energy scale, the default, $Q = 1.0 \times \sqrt{M_Z^2 + \sum p_T^2(p)}$ (black line), $Q = 1.0 \times M_Z$ (blue line), $Q = 2.0 \times \sqrt{M_Z^2 + \sum p_T^2(p)}$ (red line), $Q = 0.5 \times \sqrt{M_Z^2 + \sum p_T^2(p)}$ (green line), $Q = 2.0 \times M_Z$ (yellow line), and $Q = 0.5 \times M_Z$ (orange line), compared to data (black triangles). (a) Z+Njets (b) Z+Njets, log scale (c) Z p_T (d) G_T, the scalar sum of p_T of the muons from Z decay and E_T of the jets in the event

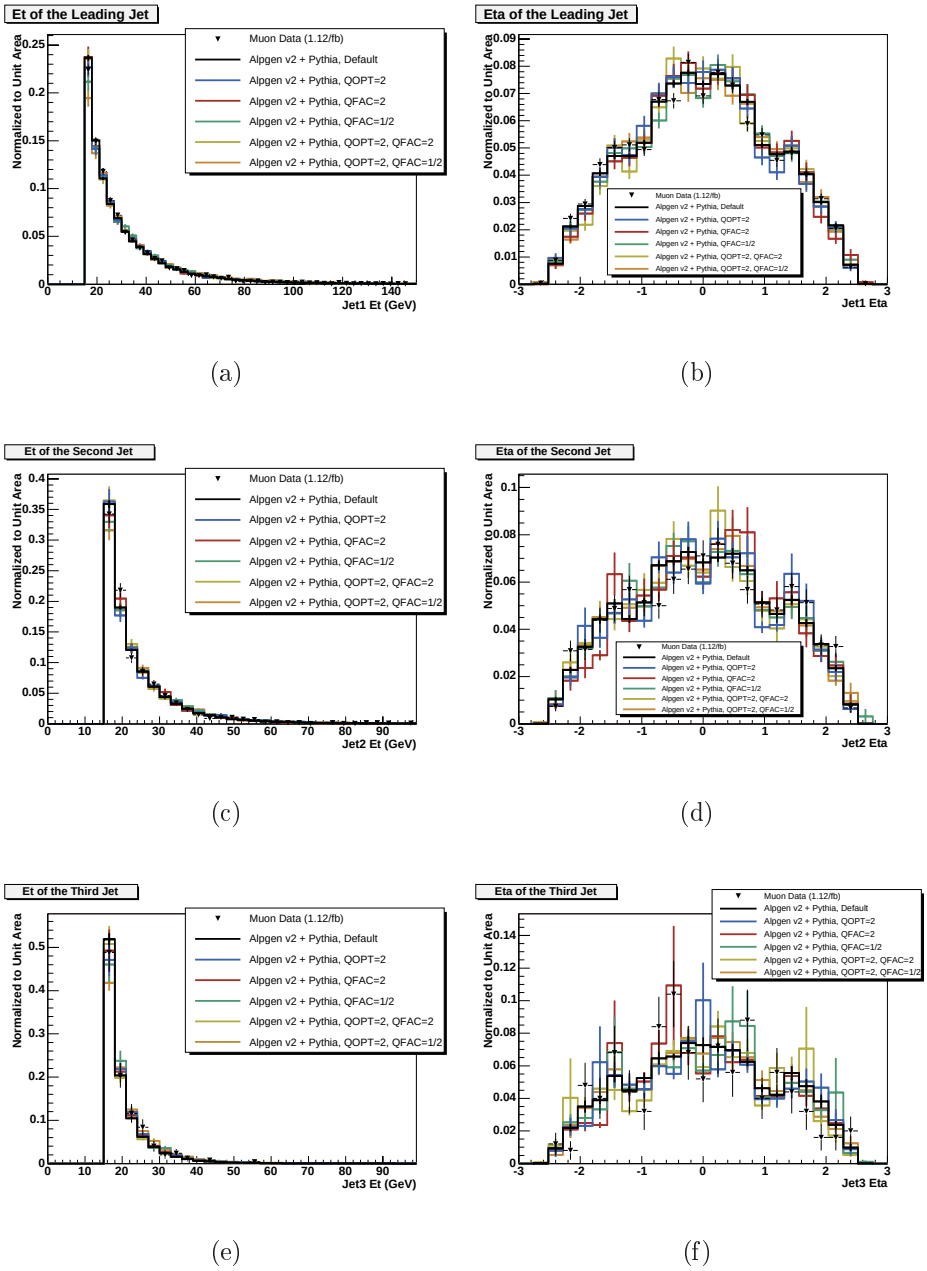
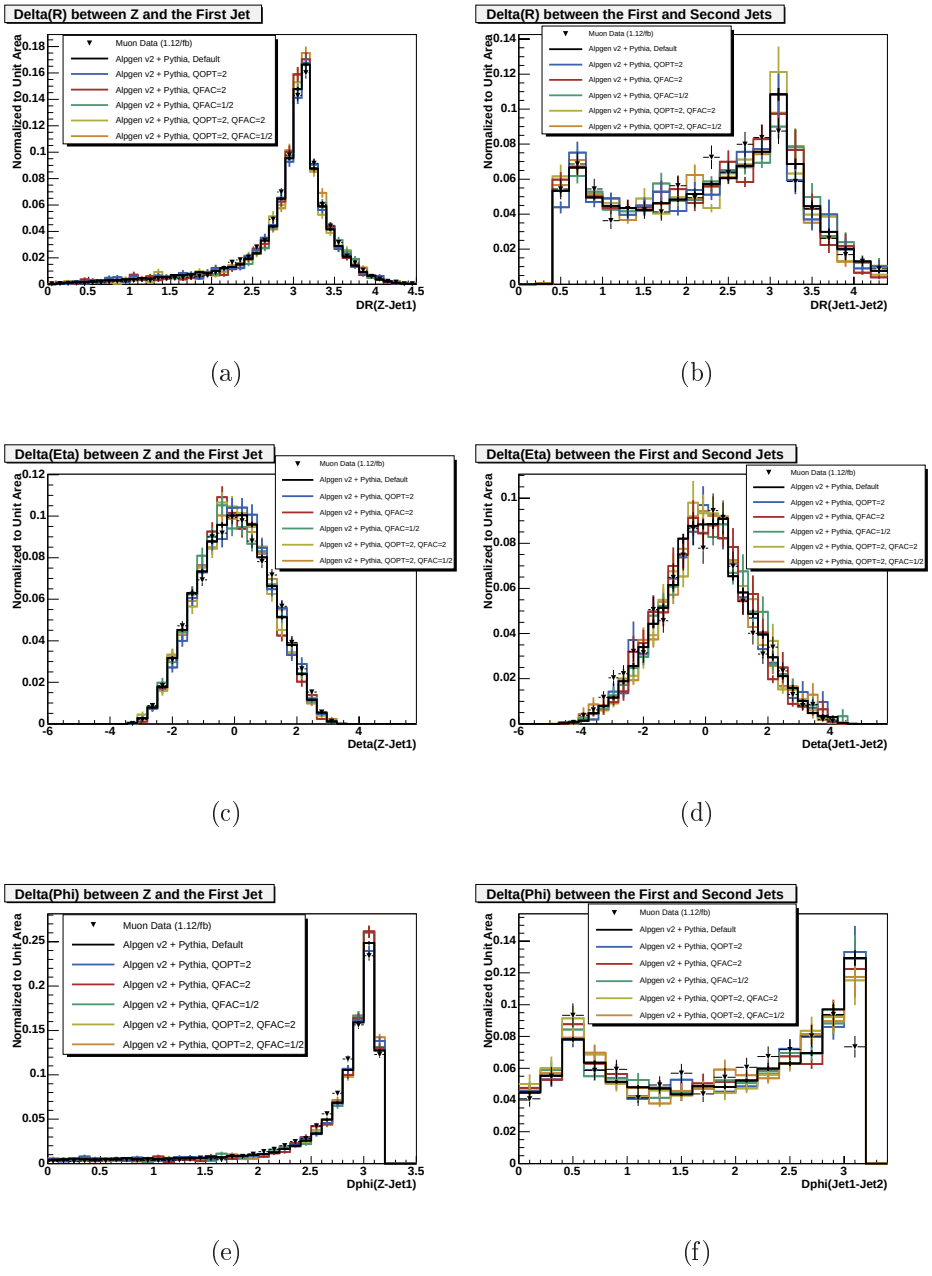


Figure A.11: Kinematic distributions from changing the renormalization and factorization energy scale, the default, $Q = 1.0 \times \sqrt{M_Z^2 + \sum p_T^2(p)}$ (black line), $Q = 1.0 \times M_Z$ (blue line), $Q = 2.0 \times \sqrt{M_Z^2 + \sum p_T^2(p)}$ (red line), $Q = 0.5 \times \sqrt{M_Z^2 + \sum p_T^2(p)}$ (green line), $Q = 2.0 \times M_Z$ (yellow line), and $Q = 0.5 \times M_Z$ (orange line), compared to data (black triangles). (a) Leading Jet E_T (b) Leading Jet η (c) Second Jet E_T (d) Second Jet η (e) Third Jet E_T (f) Third Jet η



(a) (b)

(c) (d)

(e) (f)

Figure A.12: Kinematic distributions from changing the renormalization and factorization energy scale, the default, $Q = 1.0 \times \sqrt{M_Z^2 + \sum p_T^2(p)}$ (black line), $Q = 1.0 \times M_Z$ (blue line), $Q = 2.0 \times \sqrt{M_Z^2 + \sum p_T^2(p)}$ (red line), $Q = 0.5 \times \sqrt{M_Z^2 + \sum p_T^2(p)}$ (green line), $Q = 2.0 \times M_Z$ (yellow line), and $Q = 0.5 \times M_Z$ (orange line), compared to data (black triangles). (a) ΔR(Z-Jet1) (b) ΔR(Jet1-Jet2) (c) $\Delta\eta$(Z-Jet1) (d) $\Delta\eta$(Jet1-Jet2) (e) $\Delta\phi$(Z-Jet1) (f) $\Delta\phi$(Jet1-Jet2)

A.3.3 Vertex Energy Scale

In ALPGEN, the value of the QCD coupling $\alpha_S(Q^2)$ is evaluated at every strong vertex individually. There are two options for the choice of the energy scale, $Q^2 =$ ktfac $\times\, {p_T}^2$ and $Q^2 =$ ktfac $\times\, {m_T}^2$, where p_T and m_T are the transverse momentum and the transverse mass of the vertex and ktfac is an additional scaling factor set to 1.0 by default. We have generated samples with both options, and also varied ktfac to 0.5 and 2.0 for both options.

Figures A.13 to A.15 shows the kinematic distributions for the default and systematics matching samples compared to data. The variables plotted are: Z+N jets distribution (linear and log scale), Z p_T, G_T (the scalar sum of the p_T of the muons from the Z decay and the E_T of the jets in the event), the E_T and η of the three leading jets, and the ΔR, $\Delta\eta$, and $\Delta\phi$ between the Z and the leading jet and between the leading and second jets. Varying the multiplicative factor, ktfac has a bigger effect than changing the variable which sets the vertex energy scale from p_T to m_T. The generated cross sections for the different Np samples and the total cross section are given in Table A.3.3.

Sample	Default, $Q = 1.0 \times p_T$ (pb)	$Q = 1.0 \times m_T$ (pb)	$Q = 2.0 \times p_T$ (pb)	$Q = 0.5 \times p_T$ (pb)	$Q = 2.0 \times m_T$ (pb)	$Q = 0.5 \times m_T$ (pb)
$Z \to \mu^+\mu^-$ +0p	158	160	158	159	159	159
$Z \to \mu^+\mu^-$ +1p	21.6	21.6	19.2	25.0	19.2	25.0
$Z \to \mu^+\mu^-$ +2p	3.47	3.49	2.74	4.59	2.73	4.61
$Z \to \mu^+\mu^-$ +3p	0.550	0.541	0.389	0.834	0.377	0.805
$Z \to \mu^+\mu^-$ +4p	0.0992	0.0962	0.0629	0.1700	0.0605	0.1630
$Z \to \mu^+\mu^- + c\bar{c}$ +0p	1.08	1.08	1.08	1.08	1.08	1.08
$Z \to \mu^+\mu^- + c\bar{c}$ +1p	0.332	0.322	0.274	0.381	0.266	0.377
$Z \to \mu^+\mu^- + c\bar{c}$ +2p	0.107	0.102	0.076	0.145	0.0728	0.14
$Z \to \mu^+\mu^- + b\bar{b}$ +0p	0.511	0.510	0.511	0.510	0.511	0.510
$Z \to \mu^+\mu^- + b\bar{b}$ +1p	0.134	0.130	0.115	0.163	0.111	0.156
$Z \to \mu^+\mu^- + b\bar{b}$ +2p	0.0385	0.0371	0.0290	0.0543	0.0279	0.0517
Total	185.92	187.73	182.46	191.59	183.40	191.83

Table A.6: Cross Sections for ALPGEN Vertex Energy Scale Systematic Samples

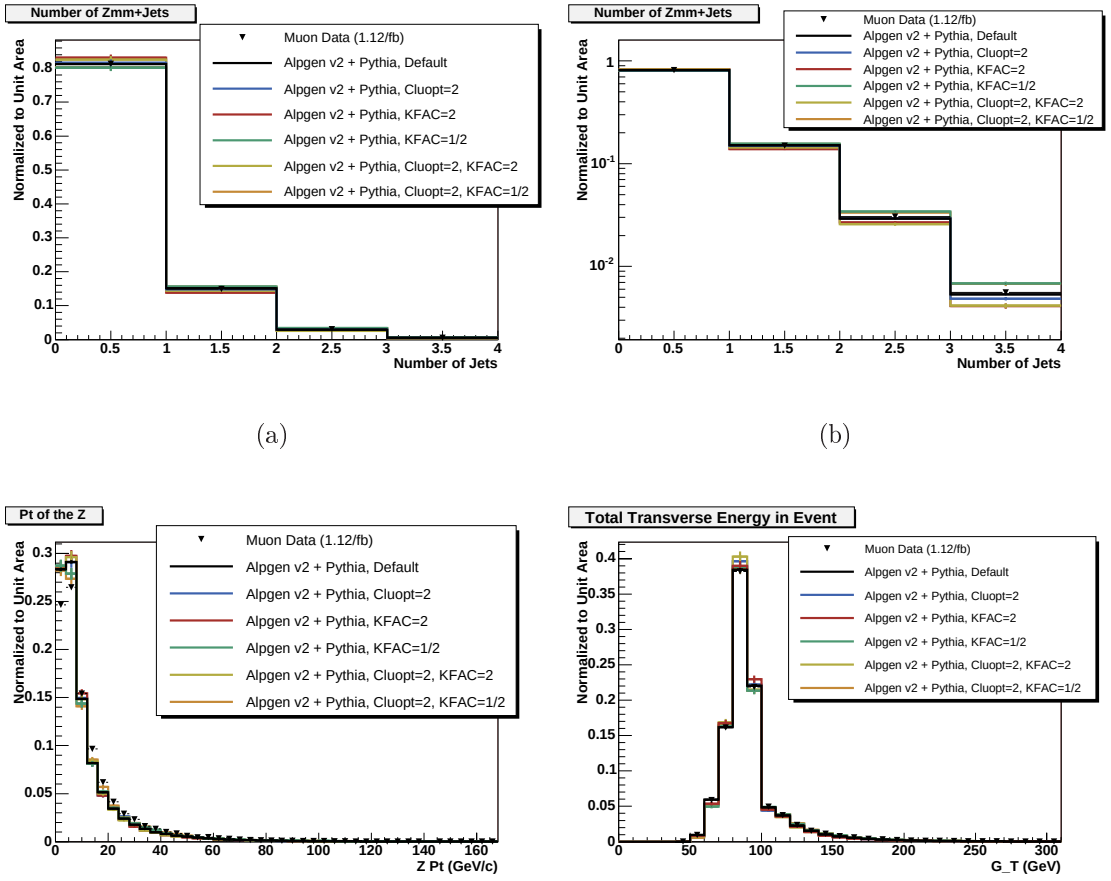
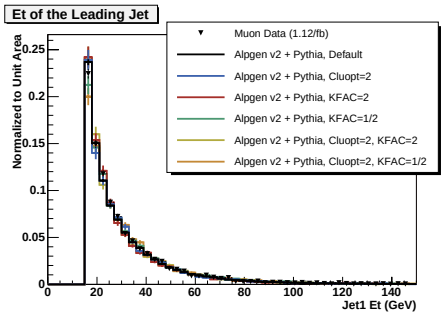
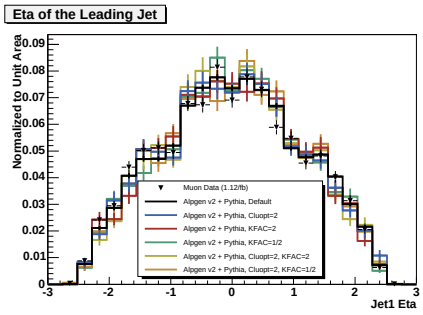


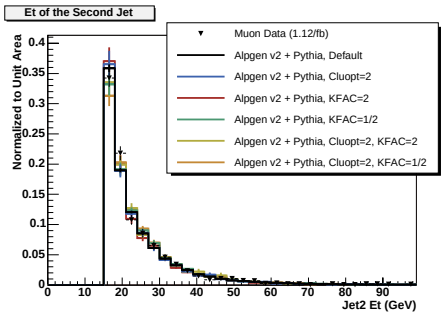
Figure A.13: Kinematic distributions from changing the vertex energy scale, the default, $Q = 1.0 \times p_T$ (black line), $Q = 1.0 \times m_T$ (blue line), $Q = 2.0 \times p_T$ (red line), $Q = 0.5 \times p_T$ (green line), $Q = 2.0 \times m_T$ (yellow line), and $Q = 0.5 \times m_T$ (orange line), compared to data (black triangles). (a) Z+Njets (b) Z+Njets, log scale (c) Z p_T (d) G_T, the scalar sum of p_T of the muons from Z decay and E_T of the jets in the event



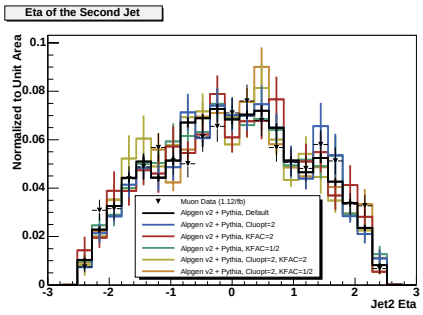
(a)



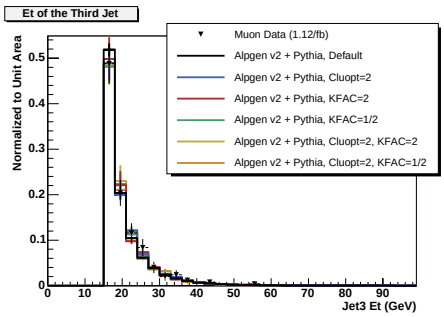
(b)



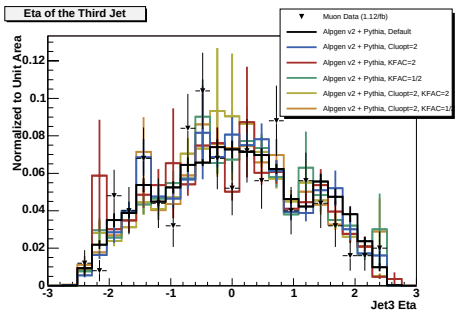
(c)



(d)

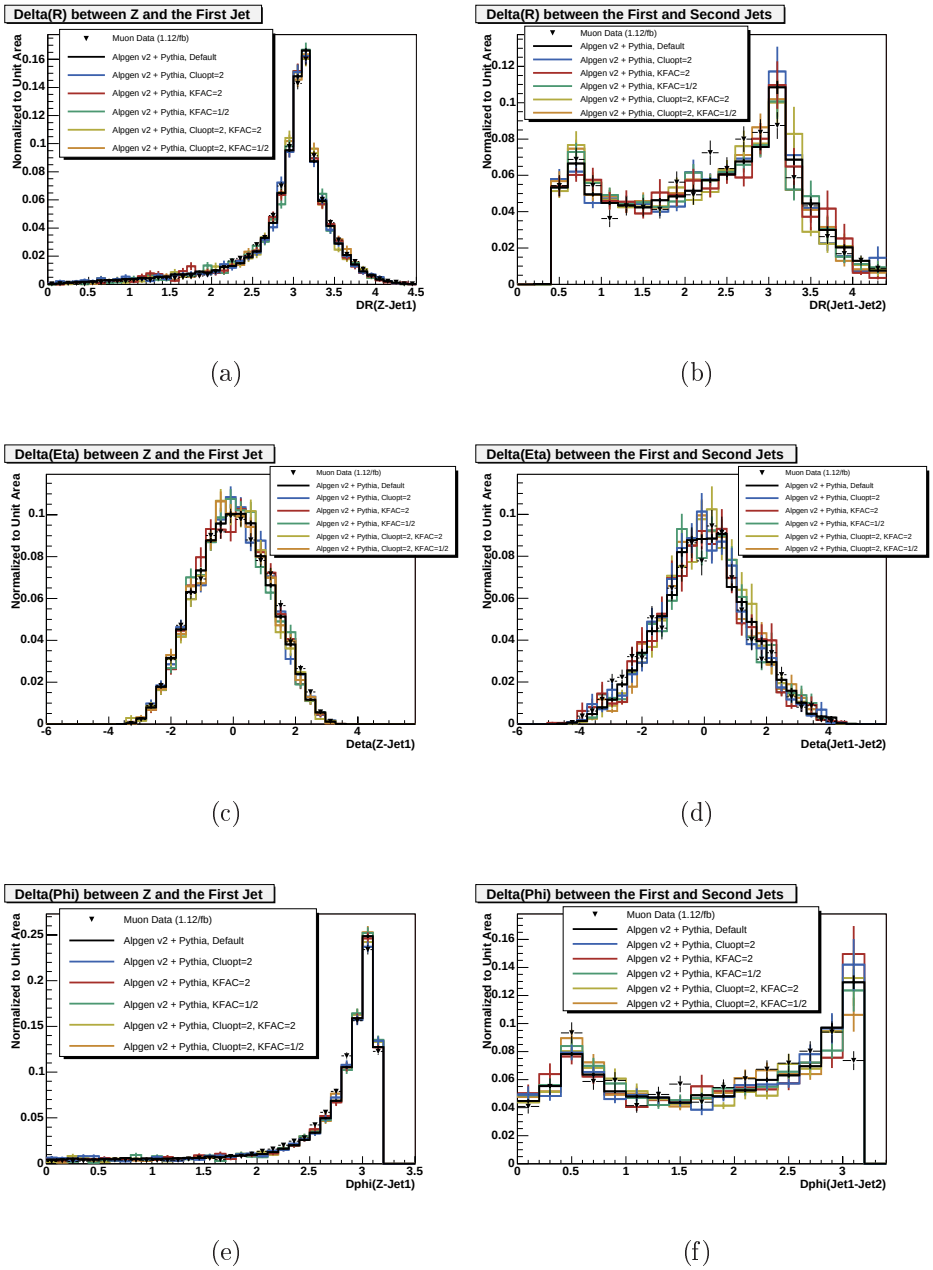


(e)



(f)

Figure A.14: Kinematic distributions from changing the vertex energy scale, the default, $Q = 1.0 \times p_T$ (black line), $Q = 1.0 \times m_T$ (blue line), $Q = 2.0 \times p_T$ (red line), $Q = 0.5 \times p_T$ (green line), $Q = 2.0 \times m_T$ (yellow line), and $Q = 0.5 \times m_T$ (orange line), compared to data (black triangles). (a) Leading Jet E_T (b) Leading Jet η (c) Second Jet E_T (d) Second Jet η (e) Third Jet E_T (f) Third Jet η



(a) (b)

(c) (d)

(e) (f)

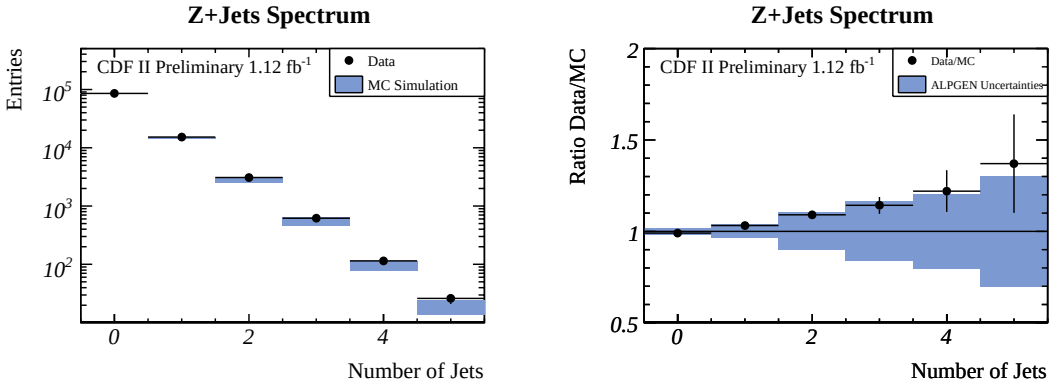
Figure A.15: Kinematic distributions from changing the vertex energy scale, the default, $Q = 1.0 \times p_T$ (black line), $Q = 1.0 \times m_T$ (blue line), $Q = 2.0 \times p_T$ (red line), $Q = 0.5 \times p_T$ (green line), $Q = 2.0 \times m_T$ (yellow line), and $Q = 0.5 \times m_T$ (orange line), compared to data (black triangles). (a) ΔR(Z-Jet1) (b) ΔR(Jet1-Jet2) (c) $\Delta\eta$(Z-Jet1) (d) $\Delta\eta$(Jet1-Jet2) (e) $\Delta\phi$(Z-Jet1) (f) $\Delta\phi$(Jet1-Jet2)

A.3.4 Systematics Summary

Several conclusions can be drawn from ALPGEN systematic studies. The cross sections for the individual samples vary by a wide range, 0% to a factor of twenty. However, the total (inclusive) cross sections remains stable to $\sim \pm 10\%$. Note that all the inclusive cross sections, 180-207 pb^{-1}, are low compared to the measured inclusive Z cross section, 254.9 ± 16.2pb [77]. The shapes of the kinematic distributions, shown in the previous sections, also remains stable to within $\sim$25%. In general, changing the multiplicative scale for the energy scales has a bigger effect than changing the variable which sets the scale.

Jet Multiplicities

After unblinding, we examined the systematic uncertainties due to ALPGEN settings for the Monte Carlo Z+N jets distribution. The uncertainty introduced by varying the ALPGEN parameters can be translated into an uncertainty of the ALPGEN prediction for jet multiplicities. Similar to the ALPGEN-induced χ^2 shape systematics discussed in Section 9.3.2, we assign the standard deviation of the systematic shifts of the number of events per jet bin as the uncertainty of the ALPGEN prediction. The result is shown in Figure A.16. The observed jet multiplicities are compatible with the ALPGEN prediction within the uncertainties; however, all central values tend to be too low.



(a) (b)

Figure A.16: Data-MC comparison of the jet multiplicity spectrum in events with a reconstructed Z. (a) Jet multiplicity spectrum. (b) Ratio of data over MC prediction. The background prediction contains Z+jets production, SM $t\bar{t}$ production, and diboson production. The colored band represents the standard deviation of the uncertainties expected from varying the ALPGEN parameters.

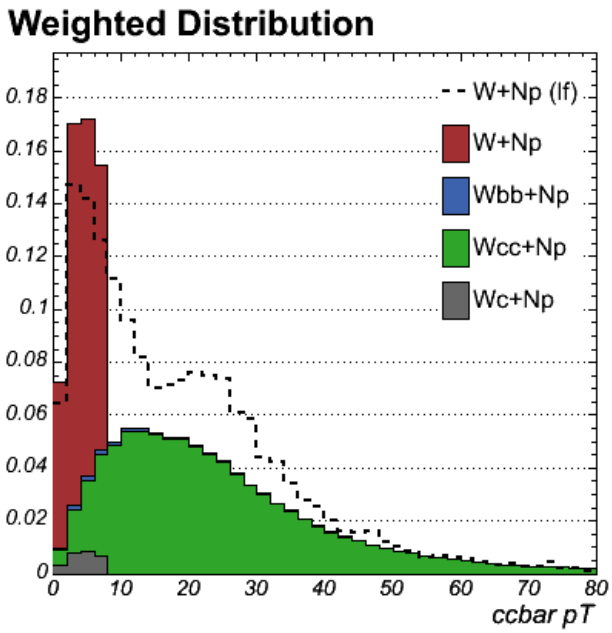
A.4 Heavy Flavor in ALPGEN

The work described previously in this appendix was done to tune and validate the pre-tag kinematic distributions in the Z+jets sample. We are also interested in the tagged background. To estimate it, we measure the fraction of events with b or c quarks, the heavy flavor fraction, in the ALPGEN samples. The heavy flavor in the ALPGEN samples come from both the Z+$b\bar{b}$ ($c\bar{c}$)+0-2p (heavy flavor) samples and the Z+0-4p (light flavor) samples where a gluon splits to a $b\bar{b}$ or $c\bar{c}$ pair. ALPGEN also contains massless charm from the matrix element in the light flavor samples. Therefore, there is an overlap in heavy flavor production which must be removed before combining light flavor and heavy flavor samples.

A.4.1 Overlap Removal

The simplest way to remove this overlap is to accept events with heavy flavor only from the heavy flavor sample and to discard the events with heavy flavor in the light flavor samples. The Z+jets samples were generated with no requirement on the p_T of the heavy flavor partons. The W+jets samples, however, have the requirement that the heavy flavor partons have a $p_T > 8$ GeV/c in the heavy flavor samples. Therefore, for W+jets samples, in this removal scheme, events with heavy flavor partons with p_T below 8 GeV/c were accepted. This scheme is called "p_T-based" overlap removal. The p_T-based overlap removal is problematic because the heavy flavor production rate in the Pythia parton showers is higher than the heavy flavor production rate in the matrix element. This leads to discontinuities in the kinematic contributions for the W+jets samples as shown in Figure A.17.

An alternative approach was developed for the SecVtx $t\bar{t}$ cross section analysis [71]. This approach, also used for our analysis, is based on the separation angle between the heavy flavor partons contained within reconstructed jets. The governing physics insight is that the ALPGEN matrix element calculations accurately model the production of well separated heavy flavor pairs (i.e. with large ΔR) while the Pythia showers are better at modelling the collinear heavy flavor pairs from gluon splitting. To exploit this, we keep the heavy flavor events in the light flavor samples if they fall within the same reconstructed jet, and we keep heavy flavor in heavy flavor samples if they fall in two different reconstructed jets. Note that in this analysis, we use jets with a cone size of $\Delta R = 0.4$. We keep or discard events in the light flavor and heavy flavor samples as follows:

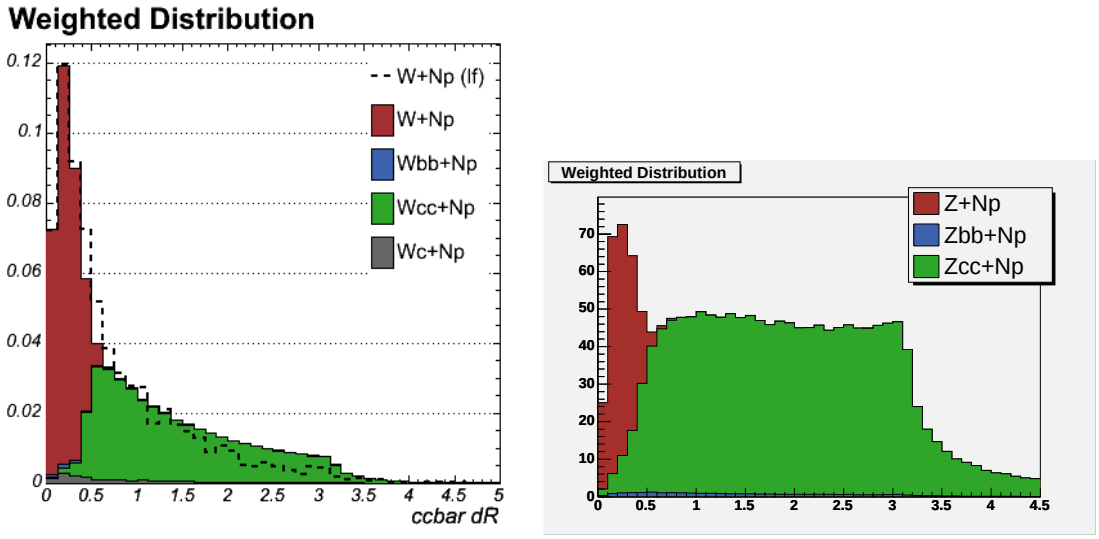


(a)

Figure A.17: $c\overline{c}$ p_T in W+jets sample with p_T based overlap removal (courtesy of Daniel Sherman). The solid histograms are the accepted events after overlap removal. The dashed histogram is the $c\overline{c}$ p_T distribution in W+Np samples.

- In light flavor samples:
  - Keep even if no heavy flavor is present
  - Keep if heavy flavor pairs from the parton shower (`STDHEP=2`) are in the same reconstructed jet
  - Discard events with matrix element (`STDHEP=3`) massless charm pairs
- In heavy flavor samples: Remove events with matrix element (`STDHEP=3`) heavy flavor pairs which fall within the same reconstructed jet.

This scheme is also summarized in Table 7.1.1, and is called "jet-based" overlap



(a) (b)

Figure A.18: $c\bar{c}$ ΔR in W+jets and Z+jets samples with jet based overlap removal. The solid histograms show the accepted events after overlap removal. (a) $c\bar{c}$ ΔR in W+jets samples (courtesy of Daniel Sherman) (b) $c\bar{c}$ ΔR in Z+jets samples.

removal (a.k.a. "Harvard" overlap removal). The distribution of the ΔR between $c\bar{c}$ pairs in W+jets and Z+jets samples after the jet-based overlap removal has been implemented are shown in Figure A.18. Z+jets events have more well separated charm pairs compared to W+jets due to an extra diagram in Z+HF production, shown in Figure A.19 (c). Since the c and $\overline{c}$ are from the splitting of two different gluons, they are more likely to be separated than the $c\bar{c}$ pairs from the splitting of one gluon as shown in Figure A.19 (a) and (b) which have analogous diagrams in W+HF production.

A.4.2 Z+HF Content

We measure the heavy flavor content of the ALPGEN Z+jets Monte Carlo sample. The Feynman diagrams for the production of Z + heavy flavor are shown in Figure A.19. We separate the sample by the number of (reconstructed) jets in the event and by whether they contain b or c hadrons. A jet is identified as a heavy flavor jet if a b or c hadron in the observed particles (OBSP) list falls within the jet cone ($\Delta R < 0.4$). We further categorize the events as having one heavy flavor jet or two or more heavy flavor jets, since events with two heavy flavor jets is more likely to be tagged than an event with only one heavy flavor jet. The heavy flavor fractions measured in the ALPGEN Z+jets Monte Carlo samples are given in Table A.7. Only statistical errors are shown.

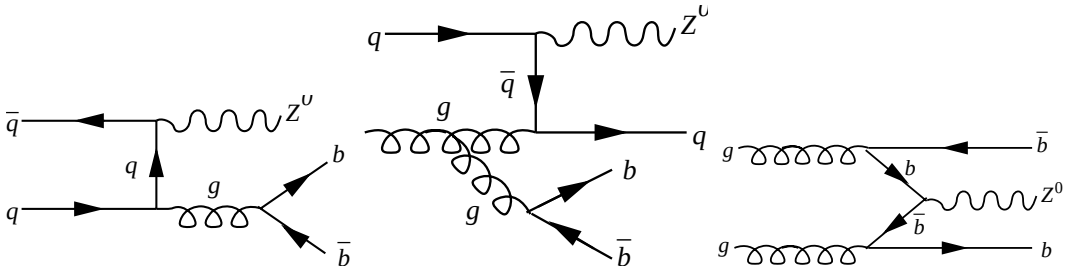


Figure A.19: Feynman diagrams for Z+HF production.

In Section 7.1.3, we discussed that the fraction of tagged events was smaller in MC than in data. We accounted for this difference in our analysis by taking the tag rates from data and assigning a systematic uncertainty to cover the under-prediction in Monte Carlo. One reason for the lower tag rate in MC might be that some charm

Sample	**1-jet (%)**	**2-jet (%)**	**3-jet (%)**	**4-jet (%)**
$Zb\bar{b}$, 1 b	0.83±0.01	1.54±0.03	2.37±0.05	3.30±0.10
$Zb\bar{b}$, 2 b	—	0.96±0.02	2.08±0.05	4.24±0.45
$Zc\bar{c}$, 1 c	1.90±0.03	3.65±0.07	5.48±0.14	7.66±0.81
$Zc\bar{c}$, 2 c	—	1.41±0.02	3.27±0.14	5.94±0.17

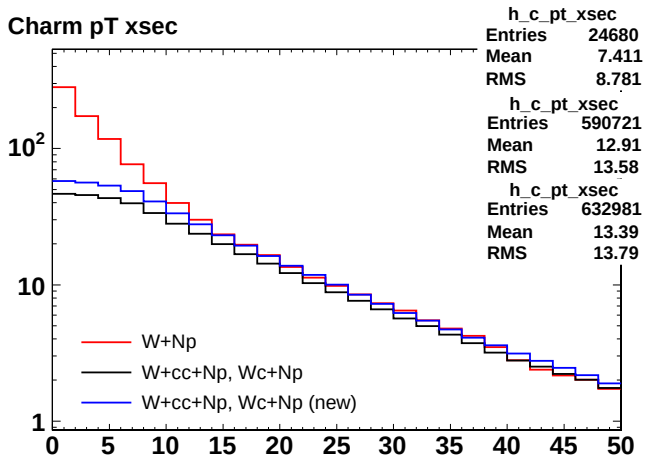
Table A.7: Z+Heavy Flavor fractions in ALPGEN Monte Carlo samples. This table gives the fraction of Z+jets events (in %) that contain heavy flavor jets, for each physical process, sorted by the amount of heavy flavor and number of jets. Only statistical errors are given.

content is missing.

Lack of Charm

It was found in that in the W+jets samples that the charm content in the W+Np samples was bigger than the charm content in the W+$c\bar{c}$ +np and W+c+np samples. This turned out to be due to a bug in the Wc generation. Figure A.20 shows the inclusive charm cross section as a function of p_T for the W+Np sample and the W+$c\bar{c}$ +Np combined with W + c+np samples before and after fixing the Wc bug. While there was more charm in the light flavor samples before the fix, the charm cross section in the heavy and light flavor samples match after the fix. Note that the discrepancy at low p_T is due to the requirement that the $c\bar{c}$ pairs in the W+HF samples have p_T > 8 GeV/c.

Figure A.21 (a) shows the inclusive charm cross section as a function of p_T in the Z+jets sample. We see that the charm cross section is higher in the Z+Np samples than the Z+$c\bar{c}$ +Np samples for all p_T. We, therefore, worry that the Z+jets samples do not contain enough charm. We explore the source of this discrepancy by comparing



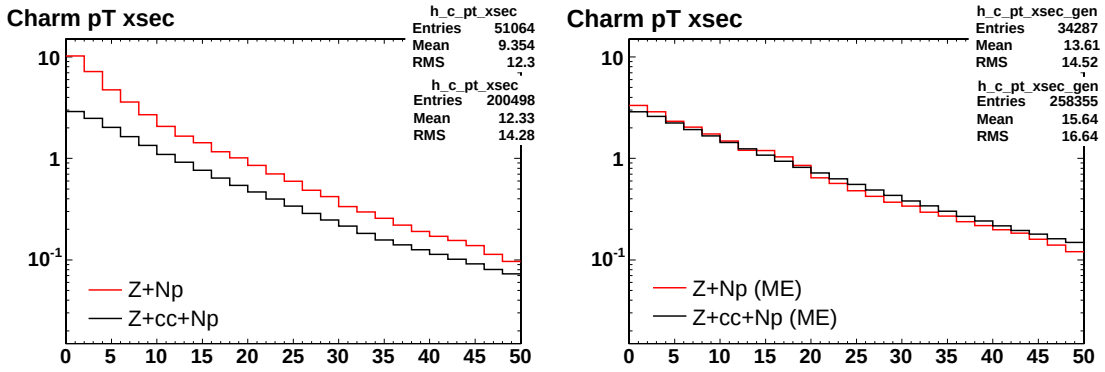
(a)

Figure A.20: Inclusive charm cross section in ALPGEN W+jets sample, before and after the Wc bug fix (courtesy of Sebastian Grinstein).

the charm content from the matrix element in both the light flavor and heavy flavor samples. Figure A.21 (b) shows the charm cross section from matrix element charm in Z+Np and Z+$c\bar{c}$ +Np samples. We see that the two distributions match well. We conclude that the discrepancy results from the parton shower. Since we take the parton shower contribution from the Z+Np samples and the matrix element contribution from the Z+$c\bar{c}$ +Np samples, the resulting charm content should be the same as in the Z+Np sample.

A.4.3 Z+HF and W+HF Comparisons

As a cross check, we compare the heavy flavor fractions in the Z+jets ALPGEN MC sample, given in Table A.7 to the W+HF fractions measured in W+jets ALPGEN MC sample [71], given in Table A.8. The first thing to note is that the



(a) (b)

Figure A.21: Inclusive charm cross section in ALPGEN Z+jets sample (courtesy of Sebastian Grinstein) (a) From matrix element and parton shower (b) Only from matrix element

$W+1c-$jet fractions are higher than the $Z+1c$-jet fractions. This is because the production of $Z + c$ requires the presence of a charm quark in the hard interaction (the charm content of the proton is small at $\sqrt{s} = 1.96$ TeV) and we do not have $Z + c$ production. The $b\bar{b}$ heavy flavor fractions are ∼30% higher for the Z+jets sample than the W+jets samples. This is because the process represented by the Feynman diagram in Figure A.19 (c) contributes only to Z+jets events and not W+jets events. However, the $c\bar{c}$ heavy flavor fractions are similar, perhaps indicating that there is missing charm in the Z+jets samples.

K Factor

We also check the heavy flavor content in QCD multijet ALPGEN samples and compare them to data. The ratio of data to MC is called the "K factor". This process is described in detail in CDF Note 8768 [88]. The K factor derived from this process,

Sample	**1-jet (%)**	**2-jet (%)**	**3-jet (%)**	**4-jet (%)**
$Wb\bar{b}$, 1 b	0.68±0.01	1.36±0.04	2.43±0.08	3.26±0.10
$Wb\bar{b}$, 2 b	—	0.86±0.02	1.84±0.07	3.28±0.11
$Wc\bar{c}$ or Wc, 1 c	5.5±0.1	8.7±0.2	11.7±0.3	13.3±0.4
$wc\bar{c}$, 2 c	—	1.38±0.03	3.27±0.13	5.99±0.14

Table A.8: W+Heavy Flavor fractions in ALPGEN Monte Carlo samples. This table gives the fraction of W+jets events (in %) that contain heavy flavor jets, for each physical process, sorted by the amount of heavy flavor and number of jets. Only statistical errors are given. [87]

with jet-based overlap removal, is 1.0 ± 0.3. This K factor is applied to W+HF in the background calculation for the $t\bar{t}$ cross section measurement [71]. However, we do not use the K factor in our FCNC search since the extra Z+HF process is not present for QCD multijet processes. As discussed in Sections 7.1.3 and 7.5. We take the Z+jets tag rates from data.

A.4.4 Summary of ALPGEN Heavy Flavor

In this section, we have discussed the measurement of the heavy flavor content in the Z+jets ALPGEN sample and compared it to the heavy flavor content in W+jets samples. The tag rates in Z+jets data are higher than the tag rates in Z+jets ALPGEN MC samples. While we account for this discrepancy in our analysis by taking the tag rates from data and assigning a systematic uncertainty to cover the under-prediction in MC, we suggested a possible reason for it in Section A.4.2. This is an issue which can be further investigated with help from the authors of the ALPGEN generator.

Appendix B

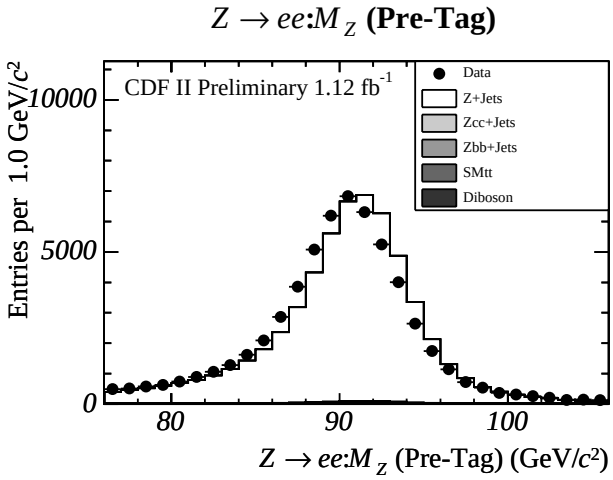
Cross Checks with Pre-Tag Events

We're all capable of mistakes, but I do not care to enlighten you on the mistakes we may or may not have made.
Dan Quayle

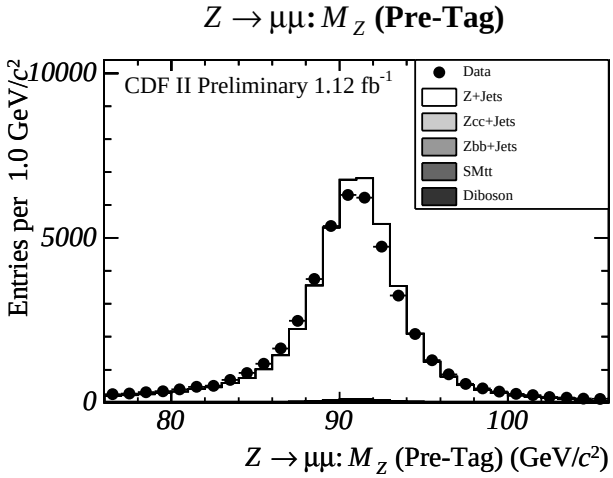
We have cross checked many distributions in the $Z+ \geq 4$ jets (pre-tag) and the full Z+jets samples. The results are shown in this appendix.

B.1 Z Mass

We compare the Z mass distributions in data and Monte Carlo simulations. We have separated the Z boson candidates into those reconstructed from electrons and those reconstructed from muons. We have separated them further into different combinations of the lepton types. These distributions are shown in Figures B.1 to B.3. The data and Monte Carlo simulations agree well, except for a small shift in the $Z \to e^+e^-$ invariant mass which is negligible due to our large Z mass window.

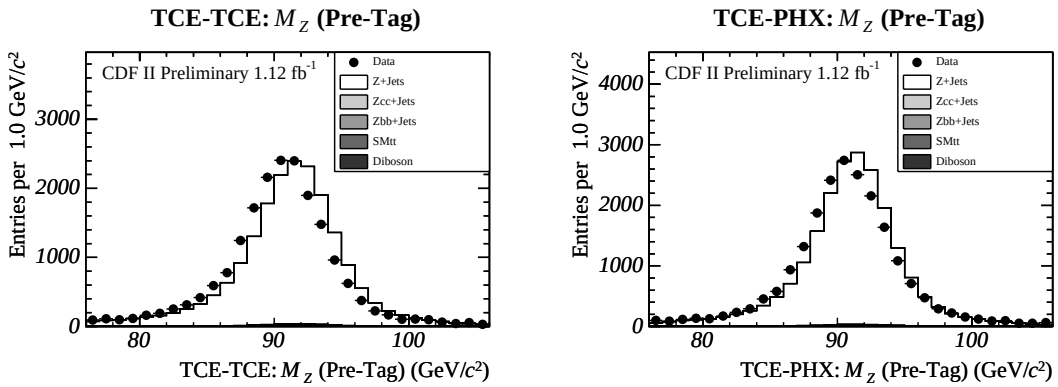


(a)



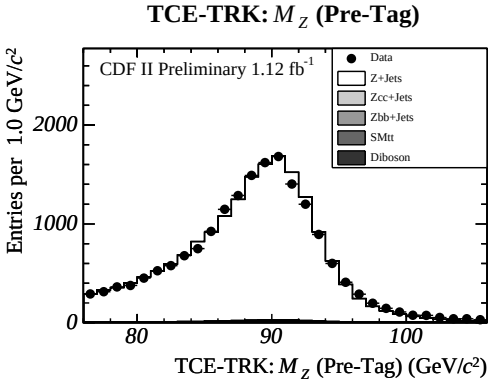
(b)

Figure B.1: Data-MC comparisons of the reconstructed Z mass. (a) $Z \rightarrow e^+e^-$. (b) $Z \rightarrow \mu^+\mu^-$. The Z+jets samples are normalized to the the data, and contributions from SM $t\bar{t}$ and diboson productions are added according to their predicted cross sections. Due to the large Z mass window, the small shift in between the Z masses in data and MC is negligible.



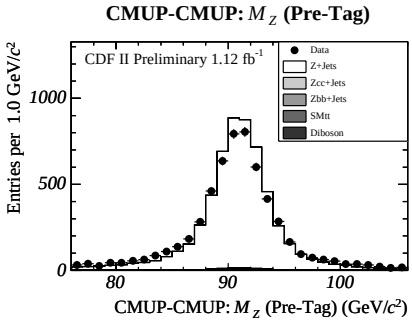
(a)

(b)

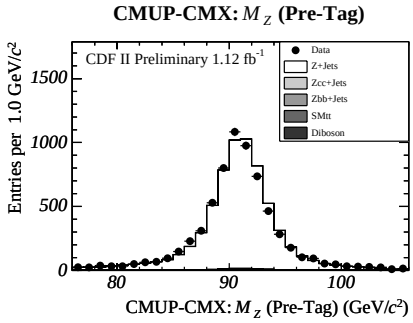


(c)

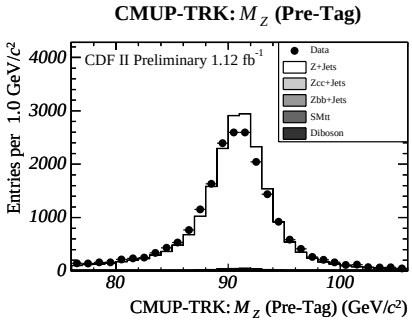
Figure B.2: Data-MC comparisons of the reconstructed Z mass. (a) $Z \to e^+e^-$ (both electrons: tight central electrons). (b) $Z \to e^+e^-$ (tight central electron paired with tight phoenix electron). (c) $Z \to e^+e^-$ (tight central electron paired with track lepton). The Z+jets samples are normalized to the the data, and contributions from SM $t\bar{t}$ and diboson productions are added according to their predicted cross sections.



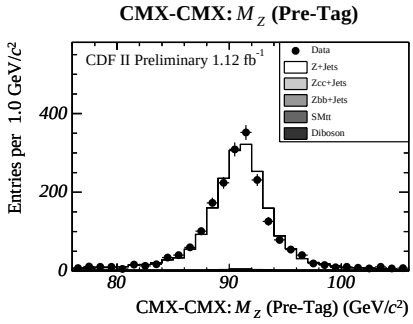
(a)



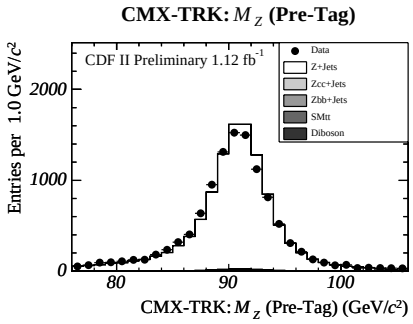
(b)



(c)



(d)



(e)

Figure B.3: Data-MC comparisons of the reconstructed Z mass. (a) $Z \to \mu^+\mu^-$ (both muons: CMUP). (b) $Z \to \mu^+\mu^-$ (CMUP paired with CMX). (c) $Z \to \mu^+\mu^-$ (CMUP paired with track lepton). (d) $Z \to \mu^+\mu^-$ (CMX paired with CMX). (e) $Z \to \mu^+\mu^-$ (CMX paired with track lepton). The Z+jets samples are normalized to the the data, and contributions from SM $t\bar{t}$ and diboson productions are added according to their predicted cross sections.

B.2 Z + N Jets

While we were vetoing events in the blinded region, we could only study the Z+N jet distribution with three or fewer jets. In Figure B.4, we show the full Z+N jet distribution including the $Z+ \geq 4$ jets events, with and without a loose SecVtx tag. These plots compare data to the ALPGEN +PYTHIA Monte Carlo simulations without our corrections. It confirms our previuos conclusion that the MC underpredicts the events in higher bins as well as the number of tags. Note that our background estimates correct for this from the lower jet multiplicity data before unblinding.

In Figure B.5, we have separated the Z+N jet distribution into $Z \to e^+e^-$ events and $Z \to \mu^+\mu^-$ events as a cross-check. The two distributions and their data-MC ratios are consistent.

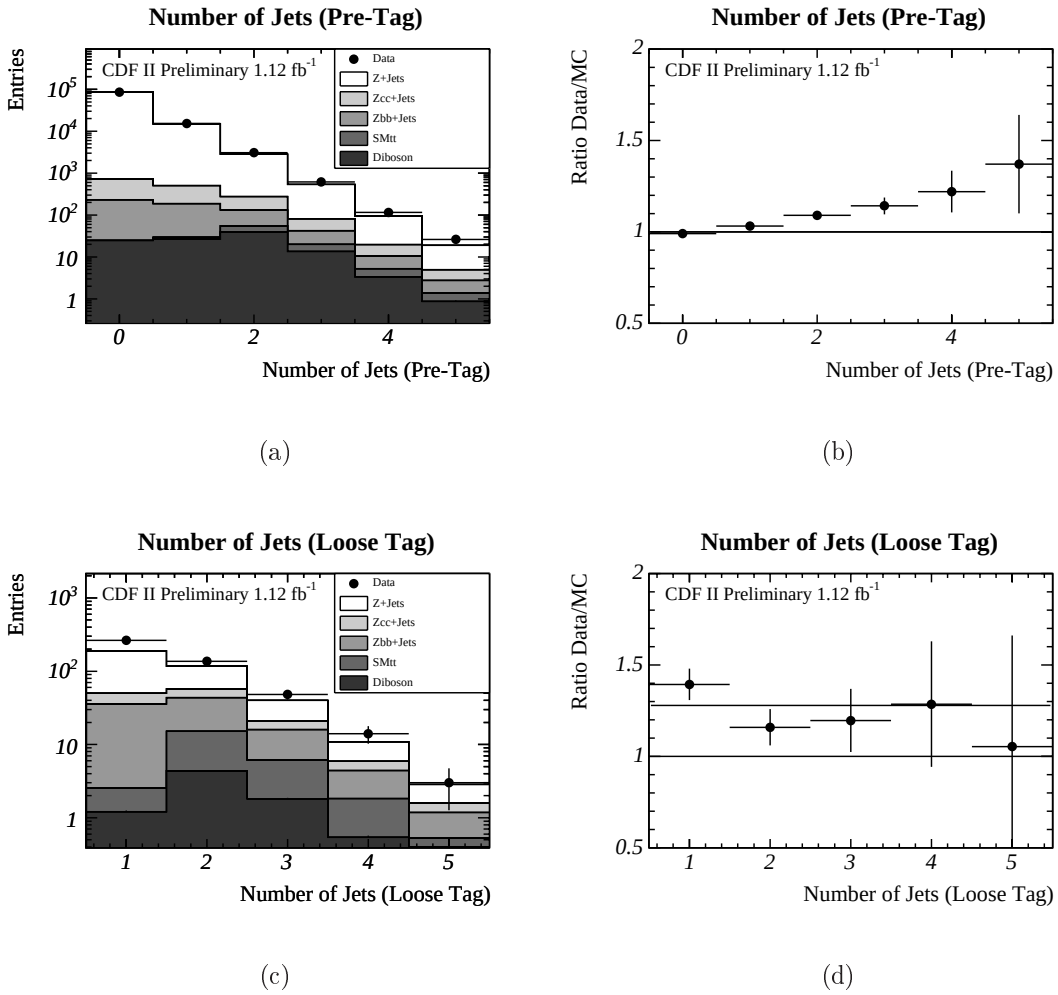


Figure B.4: Data-MC comparisons of the number of jets in events with a reconstructed Z. (a) Distribution of the number of jets before b-tagging. (b) Ratio of data over MC before b-tagging. (c) Distribution of the number of jets after b-tagging. (d) Ratio of data over MC after b-tagging. The Z+jets samples are normalized to the 0–3-jet bins before b-tagging, and contributions from SM $t\bar{t}$ and diboson productions are added according to their predicted cross sections.

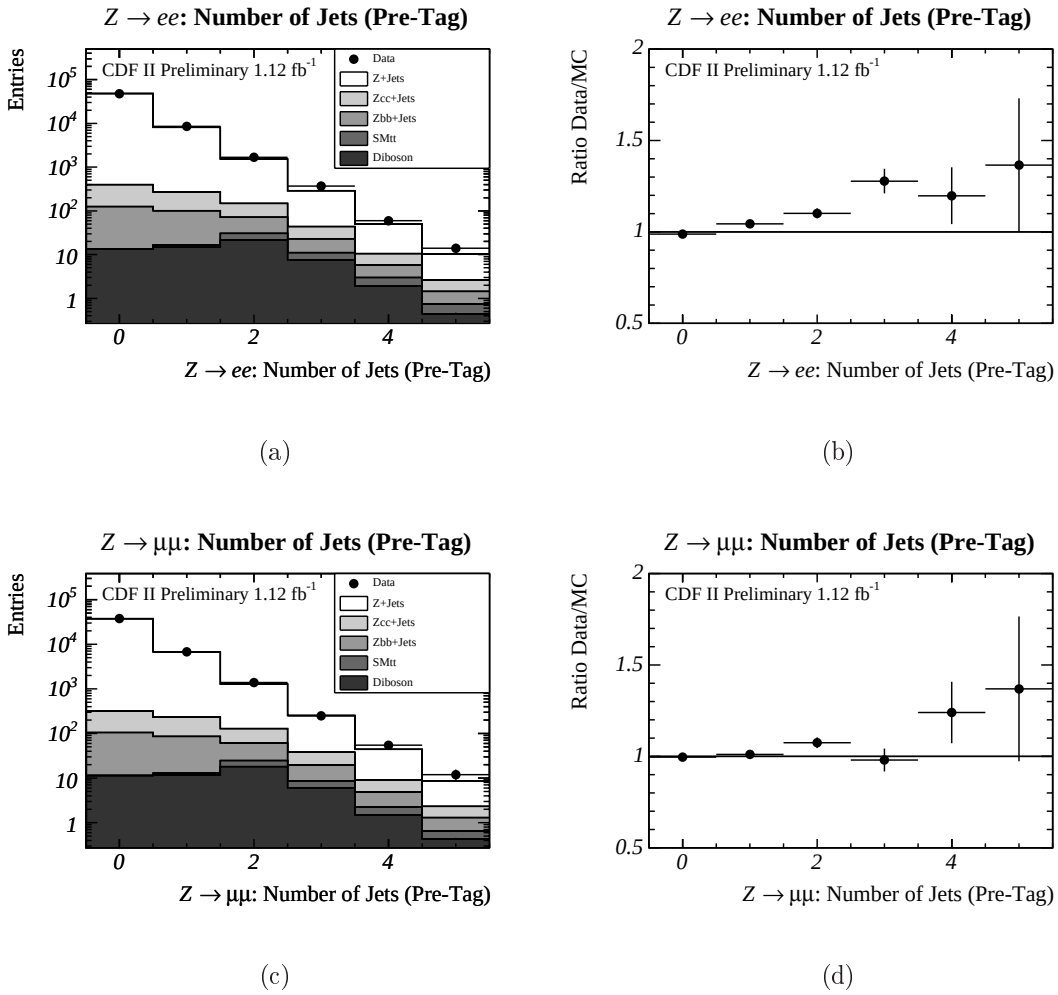


Figure B.5: Data-MC comparisons of the number of jets in events with a reconstructed Z before b-tagging. (a) Distribution of the number of jets for $Z \to e^+e^-$ only. (b) Ratio of data over MC for $Z \to e^+e^-$. (c) Distribution of the number of jets for $Z \to \mu^+\mu^-$ only. (d) Ratio of data over MC for $Z \to \mu^+\mu^-$. The Z+jets samples are normalized to the 0–3-jet bins before b-tagging, and contributions from SM $t\bar{t}$ and diboson productions are added according to their predicted cross sections.

B.3 Kinematics

In this section, we show the kinematic distributions for the full pre-tag ($Z+ \geq 4$ jets) sample to check for shape discrepancies. The data and Monte Carlo predictions are normalized to equal area. Figure B.6 shows the transverse energies of the four leading jets. Figure B.7 shows distributions of the event kinematic variables: mass χ^2, transverse mass m_T, total transverse energy G_T (without missing E_T), missing E_T, and the invariant mass of the reconstructed Z. All the kinematic distributions are consistent between data and Monte Carlo predictions, except for a slight deviation in the mass χ^2 distribution which will be discussed in the next section.

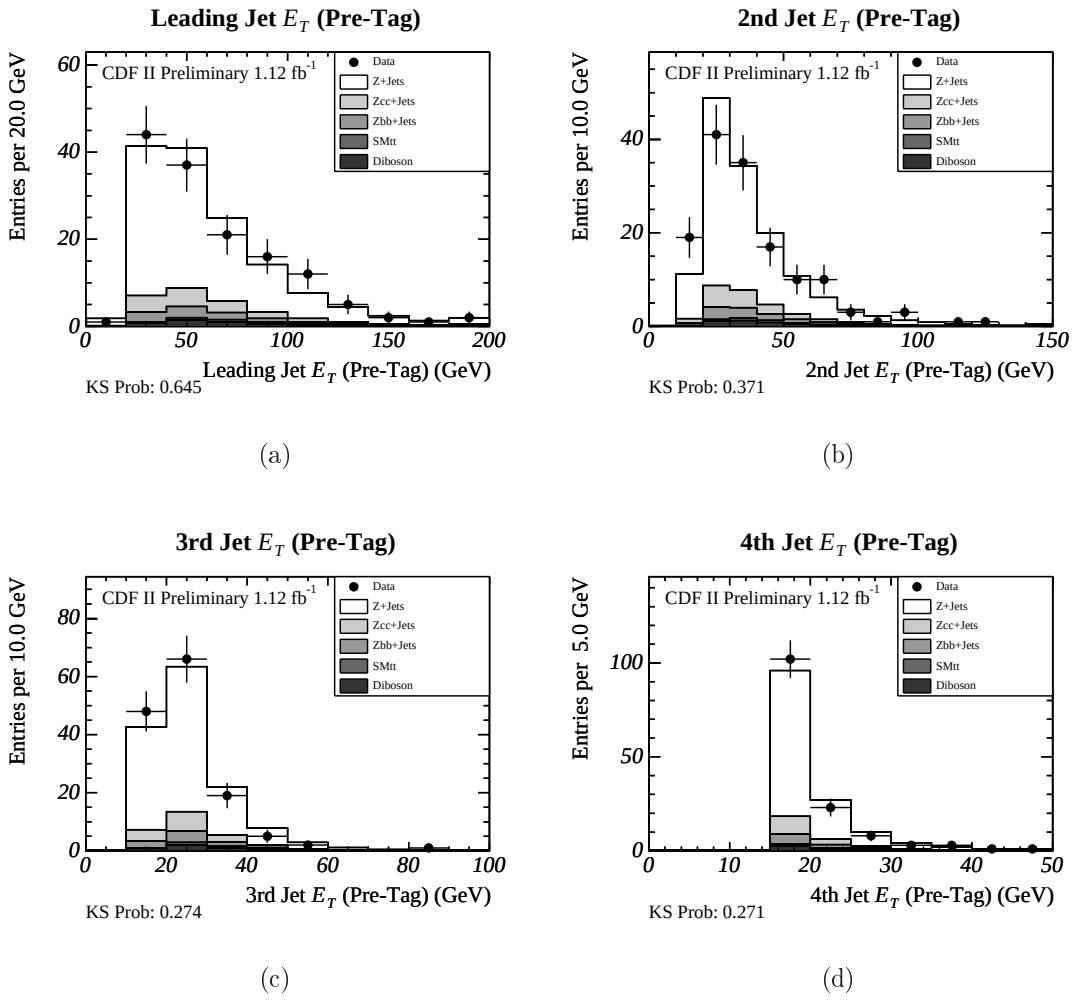
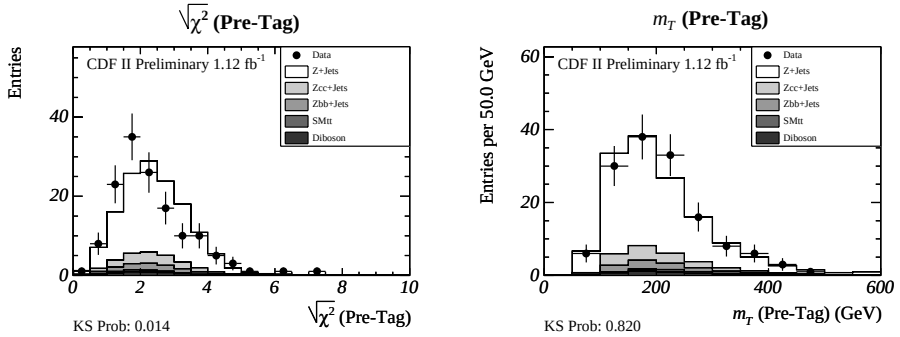
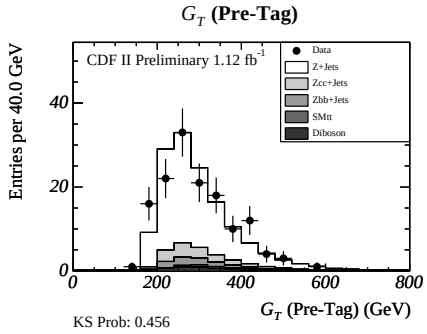


Figure B.6: Transverse energy distributions of the four leading jets for the base selection. The expected backgrounds are normalized to the data event yield.

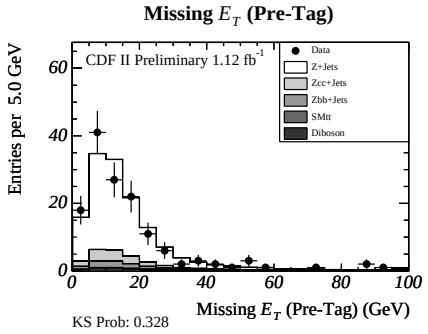


(a)

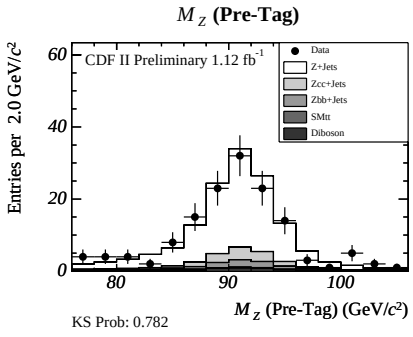
(b)



(c)



(d)



(e)

Figure B.7: Kinematic distributions for the base selection: $\sqrt{\chi^2}$, transverse mass m_T, scalar sum of jet and lepton transverse momenta/energies G_T, missing transverse energy, and Z invariant mass M_Z. The expected backgrounds are normalized to the data event yield. The largest difference between data and MC is observed in the $\sqrt{\chi^2}$ distribution, as will be further discussed in Fig. B.8.

B.4 Pre-Tag χ^2 Distribution

As indicated in the previous section, there is a slight discrepancy between the data and Monte Carlo shapes in the pre-tag χ^2 distribution. This section will describe the various cross-checks we have performed to understand this difference and demonstrate that it has a negligible effect on our background estimates and result.[1]

B.4.1 Soft Jets

It is widely suspect at CDF that the Monte Carlo simulations we use model soft jets poorly. For example, the lepton+jets $t\bar{t}$ cross section analysis (which we normalize to) require that all jets have $E_T > 20$ GeV . We have lowered this requirement to 15 GeV for our pre-tag selection criteria to increase the acceptance and have more events to study before optimizing. In Figure B.8 we demonstrate the effect of removing events with soft jets. Figure B.8 (a) shows the pre-tag χ^2 distribution with all events. In Figure B.8 (b), we show the E_T distribution of the second jet in events where the leading jet has $E_T < 40$ GeV . The data and MC do not agree well. This leads us to suspect that the discrepancy results from poor modeling of the soft jets in Monte Carlo. Figure B.8 (c) shows the mass χ^2 distribution for events with leading jet $E_T > 40$ GeV . We can see that the data and MC agree well. In Figure B.8 (d), the mass χ^2 distribution is shown for events with the leading jet $E_T < 40$ GeV , showing that these events have a large discrepancy between data and MC. Since one of the requirements in our optimized selection criteria is leading jet $E_T > 40$ GeV , this χ^2

[1] For the extension of this analysis where we fit the mass χ^2 distribution, we will have to better understand this discrepancy.

discrepancy is not present in the events in our signal region, as shown in Chapter 10.

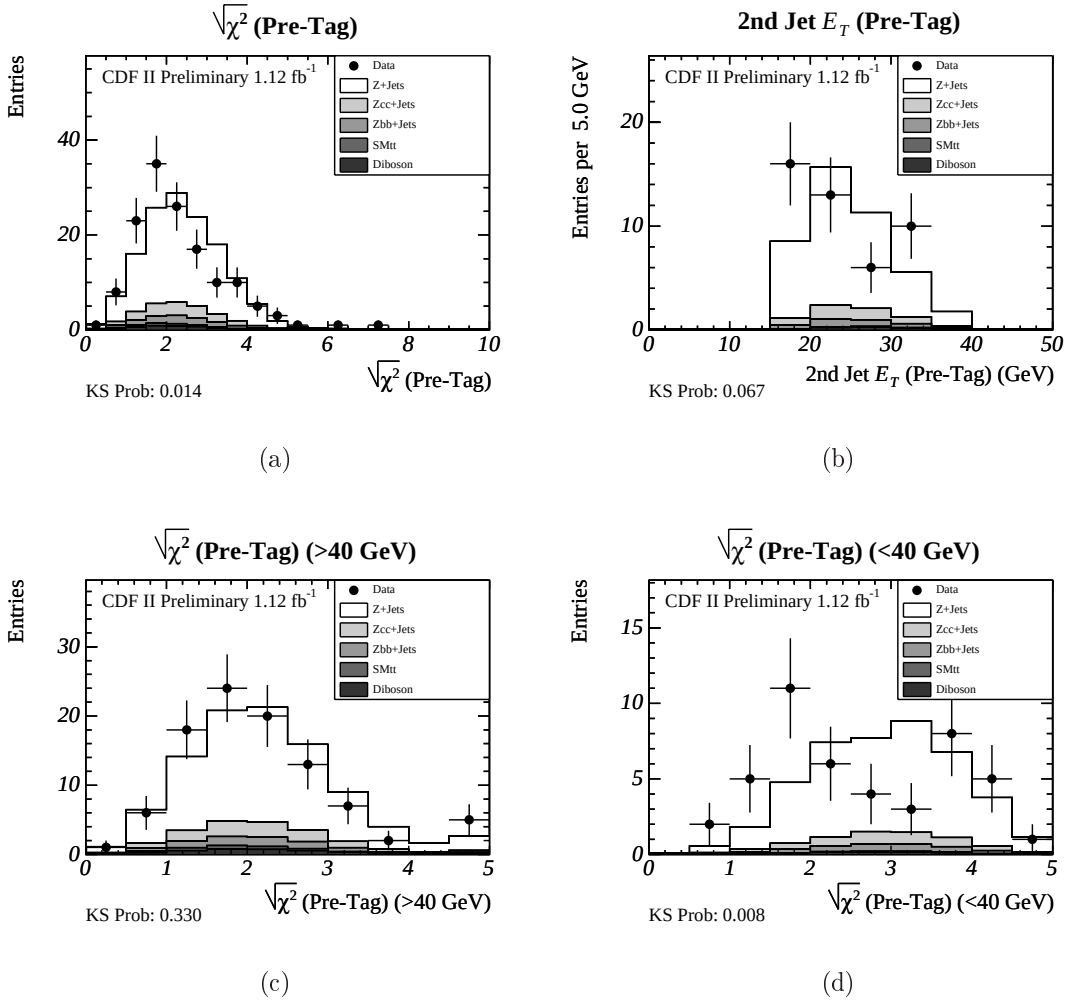


Figure B.8: The difference between data and MC in the pre-tag $\sqrt{\chi^2}$ distribution can be partially explained by bad MC modeling of events with four low E_T jets: (a) Pre-tag $\sqrt{\chi^2}$ distribution. (b) E_T distribution of second leading jet for events with leading jet E_T smaller than 40 GeV . (c) $\sqrt{\chi^2}$ distributionfor events with leading jet E_T greater than 40 GeV . (d) $\sqrt{\chi^2}$ distributionfor events with leading jet E_T smaller than 40 GeV .

B.4.2 Effect on Background Prediction

This analysis is a counting experiment. Therefore, we are not sensistive to the fine details of the mass χ^2 shape. However, we use the pre-tag χ^2 distribution in two ways:

1. The first place in which we use the χ^2 distribution directly is the estimate of the total pre-tag background. We predict 130 ± 28 events and observe 141. So the discrepancy is well covered by the systematic uncertainties.

2. The second place where we use the χ^2 distribution is in determining the cut efficiency for the optimized signal regions. The MC prediction for the cut efficiency is 8.3% 0.8% (syst.). In the data, the same cuts have an efficiency of $11.3\% \pm$ 2.7% (stat.). These numbers are consistent within one standard deviation.

Furthermore, in combining these two steps, we are senstive only to the ratio of events with $\sqrt{\chi^2} > 3.0$, the region from which we predicted the background, to the events with $\sqrt{\chi^2} < 1.6$, our optimized signal region. This is demonstrated explicitly in the following algebra on the next page.

The agreement between data and MC is good in the high χ^2 region (control) and the low χ^2 region (signal). Since we are not sensitive to the middle region where the discrepancy exists, our background estimates and result are not adversely affected by the pre-tag χ^2 discrepancy.

$$
\begin{aligned}
\mathcal{N}_{\text{tagged background}} &= \frac{\mathcal{N}_{\chi^2<1.6,\varepsilon_{(E_T,M_T)}} \cdot \varepsilon_{\text{loose}}}{\mathcal{N}_{\text{total}\chi^2}} \cdot \mathcal{N}_{Z+4\,\text{jets}} \\
&= \frac{\mathcal{N}_{\chi^2<1.6,\varepsilon_{(E_T,M_T)}} \cdot \varepsilon_{\text{loose}}}{\mathcal{N}_{\text{total}\chi^2}} \cdot \frac{31\ \text{events} \cdot \mathcal{N}_{\text{total}\chi^2}}{\mathcal{N}_{\chi^2>3.0}} \\
&= 31\ \text{events} \cdot \frac{\mathcal{N}_{\chi^2<1.6,\varepsilon_{(E_T,M_T)}} \cdot \varepsilon_{\text{loose}}}{\mathcal{N}_{\chi^2>3.0}}
\end{aligned}
$$

where

$\mathcal{N}_{\text{tagged background}}$ - expected tagged background

$\mathcal{N}_{\chi^2<1.6,\varepsilon_{(E_T,M_T)}}$ - number of weighted MC events with $\sqrt{\chi^2} < 1.6$ and passing E_T and M_T cuts

$\varepsilon_{\text{loose}}$ - loose SECVTX event tag rate

$\mathcal{N}_{\text{total}\chi^2}$ - total weighted MC events

31 events - number of events observed on high χ^2 tail

B.5 Pre-Tag Event Properties

We also provide the full event list of $Z+ \geq 4$ jets events. They are separated into $Z \to e^+e^-$ and $Z \to \mu^+\mu^-$ events. The kinematics of these events are also given.

# run	event	lepton1		lepton2		mZ	njets	Et1	Et2	Et3	Et4	nloose	chi	m_T
#														
166406	1417498	t15	TRK	e0	TCE	78.55	4	28.79	20.40	18.36	16.72	0	2.5583	88.15
168563	499220	t1	TRK	e0	TCE	91.57	4	87.33	78.30	40.89	27.61	0	2.3286	280.66
177628	2888858	e1	PHX	e0	TCE	96.93	4	33.18	17.22	15.55	15.48	0	1.8930	121.92
178886	847956	e1	TCE	e0	TCE	100.53	4	38.54	25.31	15.51	15.10	0	3.8621	135.32
183696	377838	e1	TCE	e0	TCE	84.67	4	22.64	19.57	18.60	15.86	0	1.8257	81.32
183696	3879681	t1	TRK	e0	TCE	89.63	4	30.78	23.09	20.82	18.89	0	2.2058	119.11
183965	5932573	e1	PHX	e0	TCE	90.63	5	30.85	22.85	16.42	16.26	0	3.2527	109.34
184234	1085615	e1	PHX	e0	TCE	79.47	4	47.22	23.32	18.31	15.48	0	2.8485	157.33
184237	3542281	e1	TCE	e0	TCE	85.28	5	116.87	97.51	83.67	18.97	0	1.9417	331.34
184291	4748871	e1	TCE	e0	TCE	100.01	4	33.93	26.80	24.07	20.47	0	1.6840	132.11
184802	2710370	t48	TRK	e0	TCE	85.94	5	35.15	22.14	21.86	16.76	0	2.4471	128.56
185725	373678	e1	TCE	e0	TCE	87.71	4	84.85	68.42	30.13	17.96	0	2.1137	267.00
185782	2053789	e1	TCE	e0	TCE	89.88	4	51.55	39.34	28.55	15.59	0	1.6018	177.75
186047	1910955	t1	TRK	e0	TCE	101.27	4	55.49	21.61	20.10	16.35	0	3.3573	181.91
186088	577377	e1	TCE	e0	TCE	85.29	4	22.14	17.17	15.74	15.14	0	3.4036	73.26
191596	563150	t1	TRK	e0	TCE	94.94	4	116.28	91.32	34.03	17.84	1	3.9051	294.33
191783	2615807	e1	TCE	e0	TCE	92.29	5	36.52	30.98	22.61	21.60	1	1.6157	179.50
192282	578612	e1	TCE	e0	TCE	92.66	4	45.06	20.90	17.08	16.46	0	1.9425	152.69
192348	7020043	e1	TCE	e0	TCE	94.82	4	47.93	36.05	31.58	16.67	0	1.8053	207.85
193030	230267	e2	TCE	e0	PHX	90.90	4	59.17	28.71	27.58	15.45	0	1.3861	194.33
193106	40122	e1	TCE	e0	TCE	92.27	4	33.64	24.25	20.04	15.17	0	1.0804	143.49
194029	1068143	e1	TCE	e0	TCE	93.55	4	19.27	18.47	17.36	16.62	0	3.7029	110.97
194460	1659144	e1	TCE	e0	PHX	94.40	4	102.94	51.85	25.76	17.85	0	2.6760	223.90
194460	16052561	e1	TCE	e0	TCE	93.58	4	106.62	40.25	34.79	15.50	0	2.0948	329.29
194590	4899694	e1	TCE	e0	TCE	90.25	4	54.02	34.59	19.61	17.53	0	1.8322	170.05
195408	867905	e1	PHX	e0	TCE	90.68	4	110.30	48.92	18.30	17.90	0	2.9287	272.73
196085	1298284	e1	PHX	e0	TCE	89.57	4	86.00	24.89	19.46	15.84	0	2.7960	230.50
196664	159168	e1	TCE	e0	TCE	92.15	4	65.92	23.98	23.36	20.09	0	2.3122	236.81
196946	7683410	e1	PHX	e0	TCE	87.66	5	42.63	38.92	25.85	19.86	0	0.9383	168.80
197289	13880641	t0	TRK	e0	TCE	90.15	4	25.26	17.49	16.30	15.49	0	4.3650	104.32
197321	1089982	t1	TRK	e0	TCE	87.67	4	190.80	54.06	24.92	24.70	0	2.6141	399.72
197405	2324976	e1	TCE	t0	TRK	95.11	4	59.19	29.73	22.50	20.62	0	2.5572	177.91
197990	734708	e1	TCE	e0	TCE	89.99	4	106.97	37.30	17.50	15.84	0	2.0596	334.31
198082	2320282	e1	TCE	e0	PHX	91.18	4	27.21	23.53	20.44	16.75	0	3.7061	103.57
198154	8821482	e1	TCE	e0	TCE	88.56	4	31.40	30.26	24.96	20.09	0	3.8516	147.42
198695	13706367	t0	TRK	e0	TCE	87.06	4	58.98	47.61	22.11	20.41	0	0.4854	168.94
199025	17558564	e1	TCE	e0	TCE	92.99	4	64.41	54.83	43.15	35.81	0	2.2663	330.05
199983	5248988	e1	PHX	e0	TCE	97.24	4	36.47	34.38	29.83	15.48	0	3.7760	230.74
200027	495680	e1	PHX	e0	TCE	87.17	4	20.89	19.06	15.56	15.55	0	2.1985	105.55
200536	2020853	e1	TCE	e0	PHX	92.27	4	81.36	36.70	23.77	18.60	0	2.3381	294.36
203190	3666112	e3	TCE	t0	TRK	86.96	4	120.86	97.83	19.65	18.32	0	5.2384	287.96
203190	4690217	e1	TCE	e0	TCE	91.98	4	34.61	32.18	26.05	17.76	0	1.4225	209.82
203335	806814	e0	TCE	t0	TRK	88.90	4	47.99	31.79	17.18	15.12	0	1.7074	186.90
204695	2539622	e2	TCE	e0	TCE	89.08	5	74.88	49.82	28.46	18.53	0	1.9966	221.66
205009	1907663	e1	TCE	e0	TCE	77.78	4	22.54	19.32	18.89	17.39	0	2.3269	91.20
205151	5256262	t36	TRK	e0	TCE	83.72	4	55.75	24.65	17.64	16.45	0	0.7981	211.20
206830	5018897	e1	TCE	e0	TCE	85.39	4	60.00	34.68	19.44	15.76	0	2.2259	234.54
209437	3698758	t0	TRK	e0	TCE	77.40	4	36.89	34.61	25.02	19.37	0	0.7898	171.10
209439	1281606	e1	PHX	e0	TCE	80.87	4	38.79	27.88	20.72	17.42	0	1.7067	181.42
209537	4331005	e0	TCE	e1	PHX	87.70	4	49.05	16.50	15.93	15.71	0	2.6551	111.17
209537	7092842	e1	TCE	e0	TCE	88.60	4	30.92	24.91	23.52	17.47	0	1.1388	160.00
210008	6886210	e1	PHX	e0	TCE	95.66	4	44.79	18.61	16.12	15.19	0	4.5989	193.19
219477	5104247	e1	TCE	e0	TCE	81.57	4	109.31	64.21	17.36	16.42	0	2.1638	268.12
219659	5769490	e1	PHX	e0	TCE	90.75	6	61.44	27.77	26.24	23.90	0	3.2910	183.67
220246	2862914	e1	TCE	e0	TCE	88.60	5	55.75	20.22	18.58	17.76	0	3.2570	186.22
220859	664459	e1	TCE	e0	TCE	91.46	4	38.96	23.33	20.48	16.77	0	3.2605	145.66
220844	1231419	e1	TCE	e0	TCE	93.01	4	50.76	39.03	26.44	16.15	0	2.8283	237.64
220847	1149831	e1	TCE	e0	TCE	95.40	6	23.27	18.34	18.31	18.23	0	2.9615	128.07
220847	1282490	e1	TCE	e0	TCE	90.63	4	24.59	15.88	15.64	15.45	0	4.4494	89.83
221604	6389386	e1	TCE	t0	TRK	92.50	5	70.76	46.19	20.48	15.32	0	1.8802	293.64
221751	726940	e1	TCE	e0	TCE	95.80	4	42.90	26.77	22.53	18.68	0	3.5947	228.10
221911	3599338	e1	TCE	t42	TRK	87.15	4	34.07	23.79	19.97	18.11	0	0.8457	182.33
222253	2548028	e0	TCE	t0	TRK	90.96	4	38.71	19.93	16.11	15.81	0	4.2101	116.08
222322	2329377	e1	TCE	e0	TCE	91.35	4	115.01	89.91	42.56	17.35	1	2.7607	303.35

Table B.1: List of the remaining 64 events with $Z \to e^+e^-$ and four or more jets. The table gives the run and event number, the types of the two leptons that form the Z, the Z mass, the number of jets, the transverse momenta of the four leading jets, the number of loose tags, $\sqrt{\chi^2}$, and the transverse mass.

# run	event	lepton1		lepton2		mZ	njets	Et1	Et2	Et3	Et4	nloose	chi	m_T
#														
147805	2869684	m1	CMUP	t0	TRK	98.70	4	76.04	20.44	17.58	17.37	0	1.7044	164.68
151974	194905	m1	CMUP	m0	CMX	94.79	4	46.55	26.56	17.73	15.70	1	2.0711	111.02
153389	5693738	m0	CMX	t0	TRK	95.10	4	30.14	24.78	16.35	15.91	0	1.0282	123.65
155130	1391882	m1	CMUP	t17	TRK	90.18	4	84.29	28.71	15.85	15.55	0	2.0211	178.68
155895	422689	m0	CMUP	t0	TRK	87.06	4	45.23	21.40	16.52	16.12	0	2.3577	119.72
160152	870415	t1	TRK	m0	CMUP	90.24	5	62.53	37.97	30.48	30.12	0	1.5634	223.17
161171	1163361	t1	TRK	m0	CMX	93.23	4	146.76	127.46	39.40	15.30	1	7.4685	388.64
161411	490024	m1	CMX	m0	CMX	77.97	4	61.83	17.21	16.20	15.45	0	2.4563	160.74
162480	6663699	t1	TRK	m0	CMUP	89.37	4	35.19	31.23	18.79	15.93	0	1.9054	174.06
162498	2822619	m1	CMUP	t0	TRK	87.16	4	33.82	19.67	18.44	15.73	0	1.3786	153.37
164352	57731	m0	CMUP	t14	TRK	93.97	4	94.80	28.13	21.90	16.15	1	1.9825	201.92
166367	2030787	t1	TRK	m0	CMUP	88.20	4	92.33	23.89	16.60	15.87	0	1.8515	263.40
166479	8406792	m1	CMUP	m0	CMUP	92.20	4	54.00	51.25	25.83	20.09	0	1.5628	228.61
166779	3467414	t1	TRK	m0	CMX	79.84	4	42.41	31.04	23.34	22.13	1	1.3930	156.52
168775	956114	m1	CMUP	t0	TRK	94.96	4	51.89	24.71	20.55	16.66	0	3.3774	209.77
178684	2540865	t1	TRK	m0	CMUP	86.67	4	79.43	41.69	29.49	16.89	1	2.0267	267.78
178758	2986179	m0	CMX	t0	TRK	90.82	5	102.47	67.40	21.82	17.31	1	2.8098	239.50
178785	1341596	m1	CMX	m0	CMUP	83.80	4	101.21	41.17	25.97	19.09	0	1.5707	280.04
184311	649387	t1	TRK	m0	CMUP	90.74	4	72.10	42.86	28.45	16.29	0	1.5706	233.64
184414	3419885	m1	CMX	m0	CMUP	91.21	4	73.97	68.00	49.57	43.25	0	6.4200	443.84
184832	9164981	t1	TRK	m0	CMUP	91.10	5	97.90	60.06	44.18	27.12	0	1.7800	311.34
185248	1240319	m0	CMUP	t43	TRK	84.09	4	31.81	17.59	17.12	16.30	0	3.7236	89.76
191640	809029	m1	CMUP	m0	CMUP	88.03	4	75.25	25.02	23.96	18.40	0	1.3977	157.09
192892	1789554	m1	CMUP	m0	CMX	86.67	5	52.52	40.27	39.67	37.18	0	1.8178	226.00
192987	336742	t0	TRK	m0	CMUP	94.27	4	47.02	33.68	24.23	19.54	1	2.2972	266.53
192988	5765614	t1	TRK	m0	CMUP	91.62	4	44.92	28.16	21.91	19.59	0	1.2351	199.66
194161	2470909	m1	CMUP	t0	TRK	92.32	4	158.23	38.45	18.56	17.31	0	1.8286	369.29
195758	4044389	m1	CMX	m0	CMX	91.77	4	37.41	25.55	21.18	19.55	0	2.3893	122.38
195759	763924	m0	CMUP	t0	TRK	90.45	4	25.87	21.84	18.21	17.65	0	4.0180	112.06
196368	115764	t1	TRK	m0	CMUP	90.29	4	40.58	24.99	24.62	20.35	1	3.4576	108.61
197657	3503621	t0	TRK	m0	CMUP	90.35	4	139.88	40.61	23.76	17.55	0	1.6342	357.15
198623	8399395	t40	TRK	m0	CMX	100.65	4	61.37	50.37	29.18	23.60	0	1.2408	183.53
198739	2419331	m1	CMX	t36	TRK	84.81	4	34.68	18.16	17.95	17.32	0	1.9278	125.52
198882	14643528	t32	TRK	m0	CMUP	92.32	4	59.74	32.74	21.31	20.64	0	1.9627	220.53
199655	1582724	t1	TRK	m0	CMUP	89.85	4	48.06	38.49	21.87	15.68	1	3.4704	138.88
199725	5297157	m0	CMUP	t0	TRK	92.52	5	34.27	30.99	24.15	23.15	0	1.6995	155.49
200570	5986514	m1	CMUP	m0	CMUP	89.97	4	76.78	43.50	39.16	32.24	0	2.3661	231.04
202771	5370397	m1	CMX	t0	TRK	90.91	4	28.43	19.85	16.12	15.10	0	2.4159	111.76
203265	2512685	m1	CMX	m0	CMX	93.18	6	113.86	56.04	29.53	28.44	0	1.6805	286.43
203437	303925	t1	TRK	m0	CMUP	92.43	4	32.21	22.09	17.69	15.17	0	3.8309	101.14
203797	4373580	m0	CMX	t0	TRK	91.66	4	47.37	36.77	31.72	20.51	0	2.2291	171.56
204257	6508508	m1	CMUP	t0	TRK	80.39	5	32.05	29.37	27.29	21.48	0	1.8396	152.67
204682	7124551	m1	CMX	t27	TRK	85.09	4	84.23	38.81	24.77	17.01	0	2.5080	220.69
205317	2479823	m1	CMUP	m0	CMUP	91.25	4	60.24	32.10	18.84	16.32	0	1.4905	175.42
205719	75411	m0	CMUP	t0	TRK	88.69	4	58.42	39.50	35.60	28.70	0	3.2945	234.77
209373	13152078	t27	TRK	m0	CMUP	81.75	4	44.32	29.07	21.11	16.70	0	2.7144	189.47
209402	3098550	m1	CMUP	t0	TRK	90.52	4	34.85	33.48	26.60	15.71	0	2.8991	184.80
209402	10058885	m1	CMX	m0	CMX	102.84	4	93.37	43.30	34.06	26.20	0	2.2786	313.62
209819	7699	t1	TRK	m1	CMX	91.53	5	56.27	55.71	22.60	18.70	0	1.6348	169.52
211265	4790879	m1	CMX	m0	CMX	90.83	5	87.18	69.58	25.54	20.82	0	1.5378	247.65
219095	9948579	m1	CMX	m0	CMUP	93.58	4	31.72	19.69	19.14	17.50	0	1.5071	120.42
219612	4359432	m1	CMUP	m0	CMX	77.73	5	163.36	114.16	53.88	33.92	0	4.5374	486.16
220067	121545	m1	CMUP	t0	TRK	92.27	4	111.26	68.81	23.34	20.45	0	2.1440	279.29
220227	2025089	m0	CMUP	t0	TRK	89.04	4	49.52	41.55	23.45	16.55	0	2.6398	156.73
220182	8623669	m1	CMUP	m0	CMUP	87.79	4	73.00	35.79	22.51	15.28	0	0.8659	171.91
220510	1885735	m1	CMUP	m0	CMUP	91.84	4	20.57	16.31	15.73	15.34	0	4.8525	110.44
220547	3074400	t1	TRK	m0	CMUP	94.27	4	37.34	30.17	26.24	15.86	0	2.5037	130.48
221464	1513862	m1	CMUP	t0	TRK	79.58	4	139.18	38.16	18.88	16.55	0	2.1038	357.02
221657	11733197	m1	CMX	m0	CMX	94.32	4	39.23	34.67	26.59	16.14	0	1.5416	192.73
222271	20319806	t1	TRK	m0	CMUP	97.15	4	37.54	23.97	18.08	17.00	1	4.2815	150.34
221528	145977	t1	TRK	m0	CMX	89.84	4	26.20	25.59	21.92	16.64	0	3.5150	137.51

Table B.2: List of the remaining 61 events with $Z \to \mu^+\mu^-$ and four or more jets. The table gives the run and event number, the types of the two leptons that form the Z, the Z mass, the number of jets, the transverse momenta of the four leading jets, the number of loose tags, $\sqrt{\chi^2}$, and the transverse mass.

Appendix C

Other Choices for the Limit

Two roads diverged in a yellow wood,
And sorry I could not travel both
And be one traveler, long I stood
And looked down one as far as I could
Robert Frost (from "The Road Not Taken")

C.1 Limit on $t \to qZ$ with $m_t = 170\,\mathrm{GeV}/c^2$

All signal MC samples employed in this analysis have been generated with top mass of 175 GeV/c^2. We have evaluated the top mass dependence of the limit on $\mathcal{B}(t \to qZ)$ by repeating the full analysis with a top mass of 170 GeV/c^2 to obtain the limit. We generated samples of approximately 110,000 events each, covering the full 1.12fb^{-1} run range, for the main $Z(ll)W(qq)$ sample, the "trilepton" $Z(ll)W(l\nu)$ sample, and the "double FCNC" $Z(ll, q\overline{q})Z(ll, q\overline{q})$ sample (see Table 6.1 for the nomenclature).

To illustrate the effect of the top mass on the mass χ^2, Figure C.1 shows a com-

parison of the mass χ^2 shape for our base selection for top masses of 175 GeV/c^2 and 170 GeV/c^2. We observe a small shift of the χ^2 distribution, as expected for a 5 GeV/c^2 shift within a mass resolution of 21–24 GeV/c^2.

From the signal MC samples at 170 GeV/c^2, we have re-calculated all acceptances required for the limit calculation, assuming that the systematic uncertainties remain unchanged. The acceptances are compared to the acceptances at 175 GeV/c^2 in Table C.1. The limit calculation yields an upper limit of $\mathcal{B}(t \to qZ) < 11.9\%$ at 95% C.L., compared to 11.3% at 175 GeV/c^2.

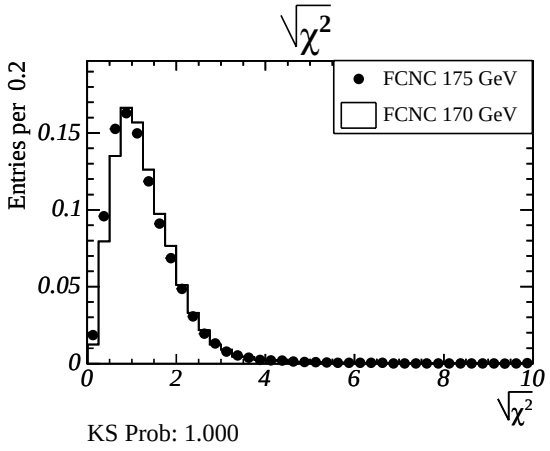


Figure C.1: Comparison of top mass χ^2 distributions for top masses of 175 GeV/c^2 and 170 GeV/c^2 after the base event selection criteria.

C.2 Normalizing to Loose Single Tag SecVtx $t\bar{t}$ Cross Section

We chose to normalize our FCNC signal to the measured double tag loose SecVtx $t\bar{t}$ cross section. Our choices were the tight single tag, loose single tag, tight double

Acceptance	**170 GeV/c^2**	**175 GeV/c^2**
$\mathcal{R}_{WZ/WW}$ (%)	34.18	32.10
$\mathcal{R}_{ZZ/WW}$ (%)	2.27	1.96
Baseline $\mathcal{A}_{WZ}$ (%)	0.778	0.778
Anti-Tagged $\mathcal{A}_{WZ}$ (%)	0.184	0.190
Loose Tag $\mathcal{A}_{WZ}$ (%)	0.278	0.292
Baseline $\mathcal{A}_{ZZ}$ (%)	0.249	0.251
Anti-Tagged $\mathcal{A}_{ZZ}$ (%)	0.794	0.823
Loose Tag $\mathcal{A}_{ZZ}$ (%)	0.627	0.665

Table C.1: Comparison of acceptances for signal MC generated with a top mass of 170 GeV/c^2 and 175 GeV/c^2.

tag, and loose double tag cross sections and we made our choice based on which one had the smallest FCNC to SM lepton+jets acceptance ratio, 32%, and, therefore, the biggest enhancement factor in the equation below:

$$\mathcal{N}_{\text{signal}} = \mathcal{B}_Z \cdot (\mathcal{N}_{LJ} - B_{LJ}) \cdot \frac{\mathcal{A}_{WZ}}{\mathcal{A}_{LJ\text{ww}}} \cdot \underbrace{\frac{\left(2 \cdot (1 - \mathcal{B}_Z) + K_{ZZ/WZ} \cdot \mathcal{B}_Z\right)}{(1 - \mathcal{B}_\text{z})^2 + 2 \cdot \mathcal{B}_\text{z}(1 - \mathcal{B}_\text{z}) \cdot \mathcal{R}_{\text{wz/ww}} + \mathcal{B}_\text{z}^2 \cdot \mathcal{R}_{\text{zz/ww}}}}_{\text{Full Running Acceptance Correction}} \tag{C.1}$$

The full derivation of this can be found in Section 6.4. The price for this choice was that, due to the difference in tagging requirements in the two selections, our systematics did not completely cancel, especially for the our anti-tag selection. Our systematic uncertainties were 11% for the base (pre-tag) selection, 17% for the anti-tagged selection criteria, and 6% for the tagged selection criteria.

If we normalize to the loose *single* tag cross section, the differences are as follows. We get a much better systematic error cancellation from taking the acceptance ratio. The systematic uncertainties are 5% for the pre-tag selection, 10% for the anti-tagged selection, and 2% for the tagged selection. However, the FCNC to SM lepton+jets

acceptance ratio is 45%, instead of 31%. This means that more of the FCNC signal would have fallen within the cross section measurement's acceptance and we would not get as big an enhancement from normalizing to the measured cross section (instead of the theoretical cross section). Figure C.2 shows the Bayesian expected limit distribution for a 32% acceptance ratio compared to a 45% acceptance ratio. It can be seen that a 32% acceptance ratio give a better expected limit than a 45% acceptance ratio. Our limit is more sensitive to the acceptance ratio than to systematic uncertainties. Therefore, we chose to normalize to the measured loose double tag SecVtx $t\bar{t}$ cross section.

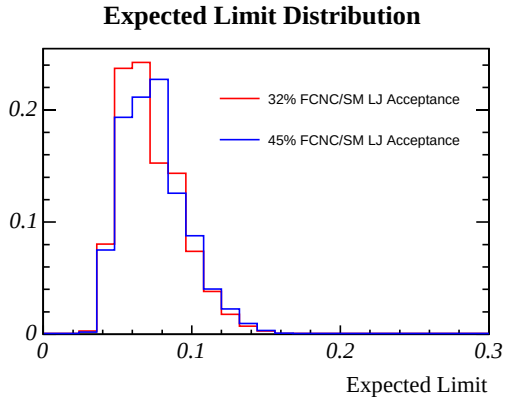


Figure C.2: This figure shows the Bayesian expected limit distribution for the 32% acceptance ratio (from normalizing to the loose double tag cross section) compared to the 42% acceptance raito (from normalizing to the loose single tag cross section). The 32% acceptance ratio gives a better expected limit.

C.3 Limit from Preliminary Mass χ^2 Fit

We have performed a counting experiment in this analysis. An alternative is fit the mass χ^2 distribution. This method would be more sensitive since it takes the

difference is χ^2 shape between signal and background into account better. Figure C.3 shows the χ^2 fit with our current data sample.

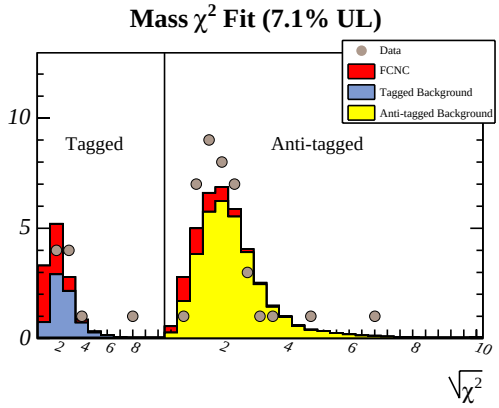


Figure C.3: Bayesian fit to the mass χ^2 template. Systematic errors are not included. The fit is consistent with no FCNC signal present. The signal in the plots corresponds to the 95% C.L. upper limit we set, $\mathcal{B}\,(t \to qZ) < 7.1\%$.

We have used the Bayesian limit technique assuming a flat prior in FCNC branching fraction between 0.0 and 1.0, as described in Section 8.1.2, except that systematic errors have not been taken into account. All optimized selection criteria have been applied except for the χ^2 requirement. The tagged and anti-tagged backgrounds are allowed to float independently but we have only used the Z+jets shape for the background. The acceptance used is the full acceptance, including the running acceptance correction. The best fit is consistent with having no FCNC signal present. We extract an upper limit on the branching fraction, $\mathcal{B}\,(t \to qZ) < 7.1\%$ at 95% C.L. (compared to 10.6%).

The fit method is promising and my collaborators are pursuing this technique for the next iteration of this analysis with 2.0 fb^{-1} of integrated luminosity. However, this fit is very preliminary and much work still needs to be done. The affect of systematic

uncertainies need to be studied and understood. A Feldman-Cousins method needs to be implemented. More importantly, the concerns with the pre-tag χ^2 distribution discussed in the last appendix must be fully dispelled before a final result can be obtained with the mass χ^2 fit.

C.4 Tighter Transverse Mass Cut

We have seen in our results that there is a slight ($\sim 1\sigma$) excess in our data compared to the background expectation. We find that most of the excess is in one bin of the anti-tagged transverse mass distribution (there were no significan excesses in any other kinematic variable distribution), as shown in Figure C.4. We believe this to be a statistical fluctuation. Raising the transverse mass requirement to 220 GeV from 200 GeV would have made our data yield much closer to the background expectation, as given in Table C.2.[1] Unfortunately, this is a blind analysis and we cannot retune our selection criteria after examining the data, in order to prevent bias.

Selection	**Observed (Expected) Events**	
	$m_T > 200$ GeV	$m_T > 220$ GeV
Anti-Tagged	12 (7.7)	7 (6.4)
Tagged	4 (3.2)	3 (2.8)
Total	16 (10.8)	10 (9.2)
Cut Efficiency (%)	11.3 (8.3)	7.1 (7.1)

Table C.2: This table shows the event yields for a transverse mass requirement of 200 GeV (our selection criterion) and a transverse mass requirement of 220 GeV (beyond the bin with an excess).

[1] We believe that the transverse mass is responsible for our limit being higher than the expected limit, instead of the mass χ^2.

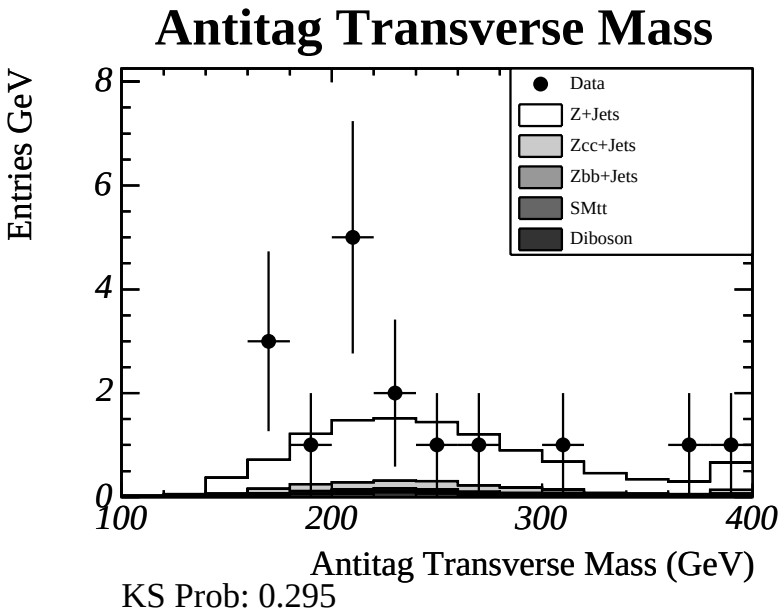


Figure C.4: Anti-tag transverse mass distribution in data showing an excess over background expectation in one bin. We believe this to be a statistical fluctuation.

Bibliography

[1] F. Abe et al. (CDF Collaboration), *Observation of Top Quark Production in $p\overline{p}$ Collisions with the Collider Detector at Fermilab*, Physical Review Letters **74** (1995), 2626–2631.

[2] W.-M. Yao et al. (Particle Data Group Collaboration), *Review of particle physics*, J. Phys. **G33** (2006), 1–1232.

[3] S. Eidelman and others (Particle Data Group), *Review of Particle Physics*, Physics Letters B **592** (2004), 1–4.

[4] R. Ammar et al. (CLEO Collaboration), *Evidence for penguin-diagram decays: First observation of $B\to K^*\gamma$*, Physical Review Letters **71** (1993), 674–678.

[5] J. A. Aguilar-Saavedra, *Top flavour-changing neutral interactions: Theoretical expectations and experimental detection*, Acta Phys. Polon. **B35** (2004), 2695–2710.

[6] F. Larios, R. Martínez, and M. A. Pérez, *New Physics Effects in the Flavor-Changing Neutral Couplings of the Top Quark*, International Journal of Modern Physics A **21** (2006), 3473–3493.

[7] F. Abe et al. (CDF Collaboration), *Search for Flavor-Changing Neutral Current Decays of the Top Quark in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV*, Phys. Rev. Lett. **80** (1998), 2525–2530.

[8] P. Achard et al. (L3 Collaboration), *Search for Single Top Production at LEP*, Phys. Lett. **B549** (2002), 290–300.

[9] S. F. Novaes, *Standard Model: An Introduction*, hep-ph/0001283, 2000.

[10] D. Griffiths, *Introduction to Elementary Particles*, Introduction to Elementary Particles, by David Griffiths, pp. 400. ISBN 0-471-60386-4. Wiley-VCH , March 1987., March 1987.

[11] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, Westview Press, 1995., 1995.

[12] I. J. R. Aitchison and A. J. G. Hey, *Gauge Theories in Particle Physics*, Institute of Physics Publishing, 1993. 2nd ed., 1993.

[13] M. Kobayashi and T. Maskawa, *CP violation in the renormalizable theory of weak interaction*, Prog. Theor. Phys. **49** (1973), 652–657.

[14] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963), 531–533.

[15] T. Inami and C. S. Lim, *Effects of Superheavy Quarks and Leptons in Low-Energy Weak Processes $K_L \to \mu^+\mu^-$, $K^+ \to \pi^+\nu\overline{\nu}$ and $K^0 \leftrightarrow \overline{K}^0$*, Prog. Theor. Phys. **65** (1981), 297.

[16] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. **D2** (1970), 1285–1292.

[17] D. Ambrose et al. (E871 Collaboration), *Improved Branching Ratio Measurement for the Decay $K_L^0 \to \mu^+\mu^-$*, Physical Review Letters **84** (2000), 1389–1392.

[18] A. J. Schwartz, *Overview of Rare and Forbidden Charm Decays*, Modern Physics Letters A **8** (1993), 967–977.

[19] K. Abe et al. (Belle Collaboration), *Observation of the Decay $B \to K\ell^+\ell^-$*, Physical Review Letters **88** (2002), no. 2, 021801.

[20] B. Aubert et al. (BaBar Collaboration), *Evidence for the Rare Decay $B \to K^*\ell^+\ell^-$ and Measurement of the $B \to K\ell^+\ell^-$ Branching Fraction*, Physical Review Letters **91** (2003), no. 22, 221802.

[21] D. Acosta et al. (CDF Collaboration), *Measurement of $\frac{B(t\to Wb)}{B(t\to Wq)}$ at the Collider Detector at Fermilab*, Phys. Rev. Lett. **95** (2005).

[22] B. A. Arbuzov and M. Y. Osipov, *Enhancement of the neutral tc transition in the model of dynamical electroweak-symmetry breaking*, Physics of Atomic Nuclei **62** (1999), 485–490.

[23] F. T. Division, *Tevatron Rookie Books*, http://www-bdnew.fnal.gov/operations/rookie_books/rbooks.html, 2006.

[24] C. Hojvat, M. Joy, and R. C. Webber, *Stripping Foils for Multiturn Charge Exchange Injection into the Fermilab Booster*, http://lss.fnal.gov/archive/test-tm/0000/fermilab-tm-0871.pdf.

[25] D. Capista and I. Kourbanis, *Main Injector Stacking Bunch Rotation Tune Up*, http://www-fmi.fnal.gov/Brotation/Brotation.html, 2000.

[26] S. Bogacz, *Main Injector - Coherent Instability Limits*, http://Iss.fnal.gov/archive/fn/FN-0519.pdf, 1989.

[27] J. Slaughter et al., *Tevatron Run II Luminosity, Emittance, and Collision Point Size*, 2003 Particle Accelerator Conference, 2003.

[28] *CDF Luminosity Webpage*, http://www-cdfonline.fnal.gov/ops/opshelp/stores/.

[29] C. I. Collaboration, *The CDF II Detector Technical Design Report*, FERMILAB-Pub-96/390-E, 1996.

[30] C. S. Hill, *Operational experience and performance of the CDFII silicon detector*, Nuclear Instruments and Methods in Physics Research A **530** (2004), 1–2.

[31] A. Sill, *CDF Run II silicon tracking projects*, Nuclear Instruments and Methods in Physics Research A **447** (2000), 1–2.

[32] A. Affolder et al., *Intermediate silicon layers detector for the CDF experiment*, Nuclear Instruments and Methods in Physics Research A **453** (2000), 84–88.

[33] T. Affolder et al., *CDF Central Outer Tracker*, Nuclear Instruments and Methods in Physics Research A **526** (2004), 249–299.

[34] L. Balka et al., *The CDF central electromagnetic calorimeter*, Nuclear Instruments and Methods in Physics Research A **267** (1988), 272–279.

[35] T. Moulik, *Offline Central PreRadiator Reconstruction in Run IIa*, CDF/DOC/CALORIMETRY/CDFR/6192, 2002.

[36] M. Albrow et al., *The CDF plug upgrade electromagnetic calorimeter: test beam results*, Nuclear Instruments and Methods in Physics Research A **480** (2002), 524–546.

[37] G. Ascoli, L. E. Holloway, I. Karliner, U. E. Kruse, R. D. Sard, V. J. Simaitis, D. A. Smith, and T. K. Westhusing, *CDF central muon detector*, Nuclear Instruments and Methods in Physics Research A **268** (1988), 33–40.

[38] P. Schlabach, *Central Muon Extension Detector (CMX)*, http://www-cdfonline.fnal.gov/ops/cdf_muon/chambers.html#CMX.

[39] G. Gómez-Ceballos et al., *Event Builder and Level 3 at the CDF experiment*, Nuclear Instruments and Methods in Physics Research A **518** (2004), 522–524.

[40] D. Acosta et al., *The CDF Cherenkov luminosity monitor*, Nuclear Instruments and Methods in Physics Research A **461** (2001), 540–544.

[41] D. Acosta et al., *The performance of the CDF luminosity monitor*, Nuclear Instruments and Methods in Physics Research A **494** (2002), 57–62.

[42] C. Hays, P. Tamburello, A. Kotwal, P. Wittich, and R. Snider, *The COT Pattern Recognition Algorithm and OfflineCode*, CDF/DOC/TRACKING/CDFR/6992, 2004.

[43] P. Azzi, G. Busetto, P. Gatti, and A. Ribon, *Histogram Tracking in the COT*, CDF/DOC/TRACKING/CDFR/5562, 2001.

[44] K. Bloom and W.-M. Yao, *Outside-In Silicon Tracking at CDF*, CDF/DOC/TRACKING/CDFR/5991, 2002.

[45] T. Nelson, R. Snider, and D. Stuart, *Forward Electron Tracking with the Phoenix-Mods Package*, CDF/DOC/TRACKING/PUBLIC/6278, 2003.

[46] J. Bellinger, K. Bloom, D. Dagenhart, A. Korn, S. Krutelyov, V. Martin, and M. Schmitt, *A Guide to Muon Reconstruction for Run 2*, CDF/DOC/COMP_UPG/CDFR/5870, 2002.

[47] R. G. Wagner, *Understanding and Using Lshr*, CDF/DOC/ELECTRON/CDFR/6249, 2003.

[48] J. Goldstein, S. Harper, B. Heinemann, G. Manca, and P. Renton, *Reconstructing the Plug Electron Energy in 5.3.3*, CDF/DOC/ELECTRON/CDFR/7687, 2005.

[49] B. Flaugher and J. Mueller, *A Guide to JetClu: The CDF Jet Clster Algorithm*, CDF/DOC/JET/CDFR/1814, 1992.

[50] A. Bhatti and F. Canelli, *Jet Energy Corrections at CDF*, CDF/PUB/JET/PUBLIC/7543, 2005.

[51] J. Guimarães da Costa and D. Sherman, *Measurement of the Top Pair Production Cross Section in 700/pb with Loose SecVtx*, CDF/ANAL/TOP/CDFR/8120, 2006.

[52] S. Rappoccio, *Measurement of the $t\bar{t}$ Production Cross Section*, CDF/THESIS/TOP/PUBLIC/7894, 2005.

[53] F. Gerberson, S. Grinstein, J. Guimarães da Costa, J. Incandela, C. Neu, and D. Sherman, *Combination of the SecVtx 1.2/fb b-Tagging Scale Factors*, /CDF/ANAL/SEC_VTX/CDFR/8666, 2007.

[54] S. Budd, T. Junk, T. Liss, and C. Neu, *Tight Loose and Ultratight SECVTX Tag Rate Matrix in 1.2* fb^{-1} , CDF/DOC/SEC_VTX/GROUP/8519, 2006.

[55] T. Sjöstrand et al., *High-energy physics event generation with* PYTHIA *6.1*, Comput. Phys. Commun. **135** (2001), 238–259.

[56] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **07** (2003), 001.

[57] R. Brun, R. Hagelberg, M. Hansroul, and J. C. Lassalle, *GEANT: Simulation program for particle physics experiments. User guide and reference manual*, CERN-DD-78-2-REV, 1993.

[58] CDF Collaboration, *Good Run List v16*, http://www-cdf.fnal.gov/internal/dqm/goodrun/good.html.

[59] D. Allspach, D. Ambrose, M. Binkley, K. Burkett, R. Kephart, R. Madrak, T. Miao, A. Mukherjee, R. Roser, and R. L. Wagner, *Aging in the Large CDF Axial Drift Chamber*, CDF/PUB/TRACKING/PUBLIC/7406, 2004.

[60] Y. Ishizawa and J. Nielsen, *Trigger Efficiencies for High E_T Electrons*, CDF/DOC/ELECTRON/CDFR/7401, 2004.

[61] U. Grundler, L. Lovas, and A. Taffard, *High-Pt muons recommended cuts and efficiencies for Winter 2007*, CDF/ANAL/TOP/CDFR/8618, 2006.

[62] CDF Collaboration, *Baseline High-p_T Electron Selection Criteria and Efficiencies: Gen5 and Gen6*, http://www-cdf.fnal.gov/internal/physics/joint_physics/instructions/electron_cuts_gen6.html.

[63] CDF Collaboration, *Baseline High-p_T Muon (CMUP, CMX, BMU) Selection Criteria: Gen5 and Gen6*, http://www-cdf.fnal.gov/internal/physics/joint_physics/instructions/muon_cuts_gen6.html.

[64] J. Incandela, C. Mills, P. Savard, T. Spreitzer, and J. Thom, *A Measurement of the $t\bar{t}$ Dilepton Cross Section in the 1.1* fb^{-1} *Lepton + Isolated Track Sample*, CDF/DOC/TOP/CDFR/8696, 2007.

[65] B. Han and V. Boisvert, *Trigger Efficiencies for the High Et Central Electron in the Gen6 data*, CDF/DOC/ELECTRON/CDFR/8629, 2006.

[66] D. Hare, E. Halkiadakis, and T. Spreitzer, *Electron ID Efficiency and Scale Factors for Winter 2007 Analyses*, CDF/DOC/ELECTRON/CDFR/8614, 2006.

[67] V. Martin, *High p_T muons, recommended cuts and efficiencies for release 5.3.1*, CDF/DOC/MUON/CDFR/7031, 2004.

[68] J. Adelman et al., *Top Quark Mass Measurement Using the Template Method in the Lepton+Jets Channel at CDF II*, CDF/PHYS/TOP/CDFR/7532, 2005.

[69] S. Grinstein, J. Guimarães da Costa, and D. Sherman, *Electron-Method SecVtx Scale Factor for Winter 2007*, /CDF/DOC/TOP/SEC_VTX/CDFR/8625, 2006.

[70] F. Garberson, J. Incandela, and C. Neu, *SECVTX b-Tag Efficiency Measurement Using Muon Transverse Momentum for 1.2/fb Analyses*, /CDF/DOC/SEC_VTX/CDFR/8640, 2007.

[71] M. Franklin, S. Grinstein, J. Guimarães da Costa, and D. Sherman, *Measurement of the Top Pair Cross Section in Lepton+Jets in 1.12* fb^{-1} , /CDF/ANAL/TOP/CDFR/8767, 2007.

[72] T. Sjöstrand, L. Lönnblad, S. Mrenna, and P. Skands, PYTHIA *6.3: Physics and Manual*, hep-ph/0308153, 2003.

[73] T. Tait, private communication.

[74] J. Guimarães da Costa and S. Rappoccio, *SecVtx Tag Matrices for 5.3.3_nt*, CDF/PUB/SECVTX/PUBLIC/7326, 2004.

[75] J. Guimarães da Costa, S. Rappoccio, and D. Sherman, *Measurement of the SecVtx Mistag Asymmetry in 5.3.3*, /CDF/ANAL/SEC_VTX/CDFR/7585, 2005.

[76] S. Grinstein, J. Guimarães da Costa, and D. Sherman, *SecVtx Mistag Asymmetry for Winter 2007*, /CDF/ANAL/SEC_VTX/CDFR/8626, 2007.

[77] D. Acosta et al. (CDF Collaboration), *First measurements of inclusive W and Z cross sections from Run II of the Tevatron collider*, Phys. Rev. Lett. **94** (2005), 091803.

[78] J. M. Campbell and R. K. Ellis, *An update on vector boson pair production at hadron colliders*, Phys. Rev. **D60** (1999), 113006.

[79] A. Ivanov, *Top Group Gen 6 MC Production*, http://www-cdf.fnal.gov/internal/physics/top/RunIIMC/topmc6/ (retrieved: June 3, 2007).

[80] G. Feldman and R. Cousins, *Unified approach to the classical statistical analysis of small signals*, Phys. Rev. D **57(7)** (1998), 3873–3889.

[81] R. Cousins and V. Highland, *Incorporating systematic uncertainties into an upper limit*, Nucl. Inst. Meth. **A320** (1992), 331.

[82] C. Plager and K. Bloom, *Using The Feldman-Cousins Construction to Measure* $\mathcal{R} = \frac{B(t->Wb)}{B(t->Wq)}$, CDF/PHYS/TOP/CDFR/7400, 2004.

[83] CDF Collaboration, *Jet Energy Correction Web Page*, http://www-cdf.fnal.gov/internal/physics/top/jets/ (retrieved: December 14, 2006).

[84] L. Chikovani and T. Djobava, *ATLAS Sensitivity to the Flavour-Changing Neutral Current Decay* $t \to Zq$, hep-ex/0205016, 2002.

[85] J. Carlvalho, N. Castro, L. Chikovani, T. Djobava, J. Dodd, S. McGrath, A. Onofre, J. Parsons, and F. Veloso, *Study of ATLAS sensitivity to FCNC top decays*, SN-ATLAS-2007-059, 2007.

[86] H. Bachacou, C. Feretti, J. Nielsen, and W.-M. Yao, *Heavy Flavor Contributions to the SECVTX-tagged W+jets Sample*, CDF/ANAL/TOP/CDFR/7007, 2004.

[87] M. Franklin, S. Grinstein, K. Guimarães da Costa, J. Lannon, T. Schwarz, A. Taffard, D. Sherman, and I. Zaw, *Heavy Flavor Content of the W+Jets Sample*, /CDF/ANAL/TOP/CDFR/8765, 2007.

[88] M. Franklin, S. Grinstein, J. Guimarães da Costa, D. Sherman, and I. Zaw, *Calibration of Heavy-Flavor Production in QCD Data*, /CDF/ANAL/TOP/CDFR/8768, 2007.