Search for a $Z'$-Like Resonance Decaying to $t\bar{t}$ Pairs in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV

Michael Kagan
makagan@umich.edu
University of Michigan

In collaboration with

Dante Amidei
amidei@umich.edu
University of Michigan

Stephen Miller
jsmiller3@gmail.com
University of Michigan

Tom Schwarz
schwarzt@umich.edu
University of Michigan

Abstract

A search for a heavy resonance ($Z' \to t\bar{t}$) in the $t\bar{t}$ invariant mass spectrum has been conducted in 682 pb\(^{-1}\) of CDF Run II data. For $b$-tagged $t\bar{t}$ candidates in the lepton+jets decay channel, we use the standard $\chi^2$ based mass-fit algorithm to reconstruct the $M_{t\bar{t}}$ spectrum. Expected spectra from Standard Model top, non-top Standard Model backgrounds, and a candidate resonance of $Z' \rightarrow t\bar{t}$ are modeled with Monte Carlo. A binned likelihood fit to the three components is performed for $Z'$ masses from 400 GeV/$c^2$ to 900 GeV/$c^2$, and we establish upper limits for $\sigma_{Z'}BR\{Z' \to t\bar{t}\}$ at the 95% CL, with statistical errors only. We also investigate the effect of the Jet Energy Scale systematic uncertainties.
A senior thesis submitted in partial fulfillment
of the requirements for the degree of
Honors Bachelor of Science in Physics
University of Michigan
2006

Written by

Michael Kagan

Research Advisor

Dante Amidei
Contents

1 Introduction ................................. 5
  1.1 The Standard Model .......................... 5
  1.2 The Tevatron and the CDF-II Detector ............... 5
  1.2.1 SecVtx b-Tagging .......................... 8
  1.2.2 Modeling Particle Interactions And The Detector .... 8

2 Top Quark Theory ................................ 8
  2.1 SM Top Production ............................ 8
  2.2 SM Top Decay ................................ 9
    2.2.1 All Hadronic Channel ....................... 10
    2.2.2 Dilepton Channel .......................... 10
    2.2.3 Lepton+Jets Channel ........................ 10
  2.3 Z' Top Production ............................ 10

3 Data Samples and Event Selection .................. 11

4 Signal and Background Models ........................ 13

5 Method ......................................... 14
  5.1 Overview .................................... 14
  5.2 Event Reconstruction .......................... 15
  5.3 Top Mass Resolution ........................... 16
    5.3.1 The Reconstructed $t\bar{t}$ Mass Spectrum .............. 17
    5.3.2 The "$M_{tt}$" Spectra of the Backgrounds .......... 18
    5.3.3 The $Z'$ Signal Reconstruction .................. 19
  5.4 Likelihood Fit Method .......................... 20

6 Analysis ...................................... 23
  6.1 Pseudo Experiments ............................ 23
  6.2 Likelihood Function Linearity Studies ............... 24
  6.3 Choice of Reconstruction and $\chi^2$ Cut .............. 25
  6.4 Expected Limit Without Systematic Errors ............. 27

7 Shape Systematic Errors .......................... 27
  7.1 Shifted Template Shapes and Acceptances .............. 28
  7.2 Cross section Shift Parameterization .................. 29
  7.3 Marginalization Method ........................ 30
  7.4 Shifted Expected Limits ........................ 31

8 $M_{tt}$ Distributions in CDF-II Data .................. 32
1 Introduction

1.1 The Standard Model

The fundamental constituents of matter and the forces which govern their interactions are described by the Standard Model of Particle Physics (SM), a quantum field theory which includes descriptions of the electromagnetic, weak and strong interactions [1]. Within the SM, the electromagnetic and weak interactions have been unified through the Glashow-Salam-Weinberg (GSW) theory.

In the SM, all matter is described by spin-1/2 Fermions while interactions are mediated by the exchange of spin-1 gauge Bosons. The Fermions are split into three generations of particles. These generations are identical with respect to their quantum numbers but vary with respect to mass. Each generation consists of one charged lepton and one neutrino, which partake in the electroweak interaction, and one up-type and one down-type quark, which partake in the electroweak and strong interactions. In addition, there are four Bosons in the SM: the photon mediates the electromagnetic interaction, $W^\pm$ and $Z^0$ mediate the weak interactions, and the gluon mediates the strong interaction. The SM Fermions and their masses can be found in table 1, while the SM gauge Bosons and their masses can be found in table 2.

The SM has seen unprecedented success in its agreement with experiment, yet there are observed phenomenon which are not explained by the SM (such as neutrino oscillations, the origin of mass, and electroweak symmetry breaking) [9]. Furthermore, the SM does not include a quantum description of gravity. Thus, the SM is not a complete fundamental theory and new theories which extend beyond the SM could play an important role in our understanding of the nature of the universe.

One possible avenue for new signs of physics which extend beyond the SM is in top quark decays. The top quark was discovered at the Tevatron accelerator at Fermilab, and has a much larger mass than all other known particles, as seen in table 1. In the SM, top quark pair ($t\bar{t}$) production is a strong process dominated by gluon exchange. However, many new physics schemes suggest the existence of a new neutral massive gauge boson, often called the $Z'$, whose decay could appear as an additional source of top quarks. As the $Z'$ could be responsible for producing $t\bar{t}$, one could examine the invariant mass of the $t\bar{t}$ system ($M_{t\bar{t}}$) and the $Z'$ signal would appear as a larger than expected top production rate and some kind of resonant structure in the $M_{t\bar{t}}$ spectrum. Thus, we have performed a search for such a $Z'$ particle. However, because the only mass requirement is that the $Z'$ mass is greater than $2 \cdot m_{\text{top}}$, we search for a $Z'$ of unknown mass in the region $400 \text{ GeV}/c^2 \leq M_{t\bar{t}} \leq 900 \text{ GeV}/c^2$.

1.2 The Tevatron and the CDF-II Detector

The Tevatron synchrotron at the Fermi National Accelerator Laboratory, located in Batavia, Illinois, produces proton-anti-proton ($p\bar{p}$) collisions with a center of mass energy of $\sqrt{s} = 1.96 \text{ TeV}$ every 396 nanoseconds. At this energy, SM top-anti-top quark
1 INTRODUCTION

<table>
<thead>
<tr>
<th>First Generation</th>
<th>Second Generation</th>
<th>Third Generation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quarks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>up (u)</td>
<td>1.5-4 MeV</td>
<td></td>
</tr>
<tr>
<td>down (d)</td>
<td>4-8 MeV</td>
<td></td>
</tr>
<tr>
<td>charm (c)</td>
<td>1.15 - 1.35 GeV</td>
<td></td>
</tr>
<tr>
<td>strange (s)</td>
<td>80 - 130 MeV</td>
<td></td>
</tr>
<tr>
<td>top (t)</td>
<td>172.5±2.3 GeV</td>
<td></td>
</tr>
<tr>
<td>bottom (b)</td>
<td>4.1-4.4 GeV</td>
<td></td>
</tr>
<tr>
<td>Leptons</td>
<td></td>
<td></td>
</tr>
<tr>
<td>electron (e)</td>
<td>0.511 MeV</td>
<td></td>
</tr>
<tr>
<td>electron neutrino ($\nu_e$)</td>
<td>$&lt;3$ eV</td>
<td></td>
</tr>
<tr>
<td>muon (\mu)</td>
<td>105.7 MeV</td>
<td></td>
</tr>
<tr>
<td>muon neutrino ($\nu_\mu$)</td>
<td>$&lt;0.19$ MeV</td>
<td></td>
</tr>
<tr>
<td>tau (\tau)</td>
<td>1.78 GeV</td>
<td></td>
</tr>
<tr>
<td>tau neutrino ($\nu_\tau$)</td>
<td>$&lt;18.2$ MeV</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: The three generations of Standard Model fermions and their masses.

<table>
<thead>
<tr>
<th>Particle (symbol)</th>
<th>Mass (GeV/c^2)</th>
<th>Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Photon (\gamma)</td>
<td>0</td>
<td>Electromagnetic</td>
</tr>
<tr>
<td>W Boson (W^{\pm})</td>
<td>80.4</td>
<td>Weak (Charged)</td>
</tr>
<tr>
<td>Z Boson (Z^0)</td>
<td>91.2</td>
<td>Weak (Neutral)</td>
</tr>
<tr>
<td>Gluon (g)</td>
<td>0</td>
<td>Strong</td>
</tr>
</tbody>
</table>

Table 2: The Standard Model Gauge bosons and their masses

Pair (t\bar{t}) production will occur through strong quark-anti-quark (q\bar{q}) annihilation and gluon fusion. However, the t\bar{t} production cross section is extremely small relative to the p\bar{p} collision cross section ($\sigma(p\bar{p} \rightarrow t\bar{t})/\sigma(p\bar{p} \rightarrow \text{anything}) \approx 10^{-10}$), and thus most collisions result in non-top events [9].

The CDF-II detector is a general purpose solenoid detector designed for charged particle tracking, lepton identification, and fast energy measurements (calorimetry) [9]. In this way, the detector can identify and measure the energy and momentum of electrons, muons, photons and jets (collimated bunches of charged and uncharged particles resulting from the decay of a high energy quark or gluon). A polar coordinate system is used at CDF: the \hat{z} axis points along the direction of the proton beam (z = 0 corresponds to the center of the detector); the polar angle \theta is measured from the \hat{z} axis; the azimuthal angle \phi is measured from the vertical. The detector is azimuthally symmetric as well as symmetric about the z = 0 plane. It is frequently convenient to change the polar angle to pseudorapidity, defined as $\eta = -\log (\tan (\theta/2))$, as the number of high energy particles in p\bar{p} collisions has a nearly uniform distribution in \eta.

The central region of the detector, $|\eta| < 1.0$, is immersed in a 1.4 Tesla magnetic field oriented parallel to the \hat{z} axis, produced by a superconducting solenoid. Within the solenoid sits the tracking system for the CDF detector, which consists of silicon microstrip detectors and the Central Outer Tracker (COT). The silicon detectors provide radial tracking from approximately 1.2 cm to 30 cm from the collision point (primary vertex) corresponding to the fiducial region of $|\eta| < 2.0$. The silicon detectors can locate charged particle tracks with a spatial resolution of $\approx 60 \mu m$ and can identify tracks which originate from a secondary decay vertex which is sufficiently displaced from the primary vertex. The COT is a cylindrical open-cell drift chamber designed to provide three-dimensional measurements of charged particle trajectories. The COT extends 3.1 meters along the \hat{z} axis and has an inner and outer radius of 40 cm and 138 cm, respectively, corresponding to the fiducial region of $|\eta| < 1.0$. 

The electromagnetic and hadronic calorimetry systems, which sit outside of the solenoid and surround the tracking systems, measure the energy of particles which pass through the calorimeters and produce showers of particles. The calorimeters cover the fiducial region of $|\eta| < 3.6$. This system can be used to measure the energies of electrons, photons, hadrons, and jets. However, the muon does not leave much energy in the calorimeter due to its large mass and relativistic speed. Muon drift tubes are located outside of the calorimeter system and are used to identify muons (essentially the only particles that will traverse the rest of the detector and arrive at these drift tubes). The muon detectors cover the region of $|\eta| < 1.0$.

Finally, gas Čerenkov detectors, which are located at $3.7 < |\eta| < 4.7$, are used to obtain luminosity information by measuring the number of inelastic pp collisions per bunch crossing. A cross section view of the CDF detector can be found in figure 1 and further description of the CDF-II detector can be found in reference [3].

Figure 1: The longitudinal cross section view of the CDF detector
1.2.1 SecVtx b-Tagging

The Secondary Vertex Detector (SVX), one of the silicon sub-detectors, is capable of identifying tracks which originate from a secondary decay vertex, which is particularly important for identifying b quarks. Because the b quark has a relatively long lifetime ($\approx 10^{-12}$ seconds), a b quark in a $t\bar{t}$ event with a typical momentum of 70 GeV/$c$ will travel nearly 500 $\mu$m before decaying into a jet. By locating tracks in a jet which originate from a secondary vertex, we can identify b quark jets (b-jets). Once the secondary vertex is identified, the jet is said to have a SecVtx tag (or simply a b-tag). This method for identifying b quarks is vital for identifying $t\bar{t}$ events and rejecting backgrounds.

1.2.2 Modeling Particle Interactions And The Detector

In order to understand the complexities of the data measured at CDF, one must have accurate predictions of the physical processes taking place. This is done with Monte Carlo (MC), which is a computer simulation of complex particle interactions and their signatures in the CDF detector. Thus, MC is a simulation of CDF data that we use to describe the expected physics at the Tevatron. Furthermore, MC stores truth information, or the precise masses, energies, and momenta of every particle in an MC event before it has been run through the CDF simulation and distorted by our experimental resolution. In this way, we may always check the reliability and predictive power of our analysis by comparing results with MC truth information.

2 Top Quark Theory

2.1 SM Top Production

In the SM, $t\bar{t}$ is produced through the strong interaction in $p\bar{p}$ collisions. The main processes which produce $t\bar{t}$ are $q\bar{q}$ annihilation, as seen in figure 2. While the weak interaction can produce single top quarks, the production cross section for this process is much smaller than $t\bar{t}$ production [14]. Furthermore, single top production is plagued by large backgrounds, while most background to $t\bar{t}$ can be rejected. This analysis is interested in $t\bar{t}$ production and single top production will only be considered as a background process.

The production cross section for top quark pairs is [9],

$$\sigma(p\bar{p} \rightarrow t\bar{t}) = \sum_{i,j} \int dx_i dx_j F_p(x_i, Q^2) F_{\bar{p}}(x_j, Q^2) \sigma_{ij}(Q^2, m_t)$$

(1)

where $i$ and $j$ correspond to quarks and gluons, or partons, within the proton and anti-proton respectively, $x$ is the momentum of a given parton, and $Q^2$ is an arbitrary parameter with units of energy that is introduced for the purpose of renormalization. The function $F_p(x_i, Q^2)$ is the empirical probability distribution for parton $i$ within
2 TOP QUARK THEORY

<table>
<thead>
<tr>
<th>W Decay Channel</th>
<th>Branching Ratio [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>qq</td>
<td>67.39 ± 0.35</td>
</tr>
<tr>
<td>e\nu</td>
<td>10.72 ± 0.16</td>
</tr>
<tr>
<td>µ\nu</td>
<td>10.57 ± 0.22</td>
</tr>
<tr>
<td>τ\nu</td>
<td>10.74 ± 0.27</td>
</tr>
</tbody>
</table>

Table 3: Branching ratios for W boson decays

the proton to have a momentum $x_i$, and similarly for $F_{\bar{p}}(x_j, Q^2)$. Finally, $\sigma_{ij}(Q^2, m_t)$ is the cross section for an interaction between parton i in the proton and parton j in the anti-proton, which depends on the top quark mass $m_t$. The SM theoretical value for $\sigma(p\bar{p} \rightarrow t\bar{t})$ at $\sqrt{s} = 1.96$ TeV and $m_t = 175$ GeV/c$^2$ has been calculated using QCD to be $\sigma(p\bar{p} \rightarrow t\bar{t}) = 6.7^{+0.7}_{-0.9}$ [18]. As $\sigma_{ij}(Q^2, m_t)$ depends explicitly on the top mass, $\sigma(p\bar{p} \rightarrow t\bar{t})$ will change depending on the value of $m_t$.

2.2 SM Top Decay

In order to experimentally measure the $t\bar{t}$ production cross section, we must understand the decay channels of the top quark and observe the final states of the decay products at CDF. The SM predicts that the top quark will decay almost exclusively to a W Boson and a b quark ($t \rightarrow Wb$). Thus a $t\bar{t}$ pair will decay as $t\bar{t} \rightarrow W^+ + b + W^- + \bar{b}$ as seen in figure 2. The high energy b quark will form a jet which is subsequently measured by CDF. The W is known to decay in two ways: leptonically to an electron, muon or tau and a neutrino; hadronically to two quarks. The branching ratios for each type of W decay have been measured with high precision [8], and are found in table 3. Quarks resulting from hadronic W decay will form jets, whereas the electrons and muons from leptonic W decay will be directly measured by CDF. Because $\tau$'s decay quickly, CDF can only measure the decay products of the $\tau$ in the detector. Therefore, $\tau$ decays will not be used in this analysis. Thus, the decay of $t\bar{t} \rightarrow W^+ + b + W^- + \bar{b}$ can be sorted into three possible final states: all hadronic channel; dilepton channel; lepton+jets channel.

Figure 2: Feynman diagram for $t\bar{t} \rightarrow W^+ + b + W^- + \bar{b}$. 
2.2.1 All Hadronic Channel

This decay channel is \((t\bar{t} \rightarrow W^+ + b + W^- + \bar{b} \rightarrow 2 \text{ b jets} + 4 \text{ light flavor jets})\). This channel has the largest branching ratio, and also the lowest signal to background ratio (S/B). This is due to the fact that there are many non-top processes which can result in events with high jet multiplicity. Furthermore, it is necessary to properly assign jets to the partons from which they decayed in order to properly reconstruct the events. Even requiring one b-tag in the event leaves 240 possible jet to parton assignments, i.e. determining the correct combination is highly non-trivial. Thus, we do not use this decay channel due to the large backgrounds and difficult reconstruction.

2.2.2 Dilepton Channel

This decay channel is \((t\bar{t} \rightarrow W^+ + b + W^- + \bar{b} \rightarrow l + \nu + l + \nu + 2 \text{ b jets})\). This channel has the highest S/B, but is plagued by low statistics due to the small branching ratio. Furthermore, the event contains two neutrinos which are not measured by CDF. These neutrinos carry an unknown amount of energy, called the missing transverse energy \((\not{E}_T)\), in an unknown direction. For each event we can only measure the total \(\not{E}_T\) and cannot identify the components of missing energy which should be allotted to each neutrino. Essentially, reconstruction of the top mass becomes an extremely difficult task and thus our analysis does not use events in this channel.

2.2.3 Lepton+Jets Channel

This decay channel is \((t\bar{t} \rightarrow W^+ + b + W^- + \bar{b} \rightarrow l + \nu + 2 \text{ b jets} + 2 \text{ light flavor jets})\) as seen in figure 3. This channel has a higher S/B than the all hadronic channel, but lower S/B than the dilepton channel. This channel also has a smaller branching ratio than the all hadronic channel and does suffer from relatively low statistics. However, the fact that there is only one neutrino in the final decay products of \(t\bar{t}\) implies that we can obtain a unique solution for \(E_T\) and are only left with the task of determining the longitudinal momentum \(p_z\) of the neutrino. While this channel has 4 jets, the requirement of at least one b-tag leaves only 12 possible jet to parton assignments. Combined with the two possible \(p_z\) directions of the neutrino, there are a total of 24 combinations to be considered in these events. While it is difficult to identify the proper combination, this channel has the most promising features for reconstructing the top mass. Therefore, we will use this channel to reconstruct the top mass and top 4-vectors. By adding the 4-vectors of the two top quarks we can obtain the 4-vector of \(t\bar{t}\) system \((V_i)\). We can calculate the invariant mass of the \(t\bar{t}\) system \((M_{t\bar{t}})\) by taking the inner product \(V^iV_i = M^2_{t\bar{t}}c^2\).

2.3 \(Z'\) Top Production

The SM of particle physics has done exceeding well at predicting outcomes of particle interactions, and yet several questions still remained unanswered. Furthermore, with
the unification of the electromagnetic and weak forces having been verified, the question of unifying the strong and electroweak interactions has become increasingly relevant. Several extensions of the SM have been able to achieve this unification, including some superstring models [15]. Many of these theories predict the existence of a new neutral gauge boson, often called the $Z'$. This $Z'$ boson could have properties much like the SM $Z^0$ boson, but with a larger mass, and could be produced in extremely high energy interactions [15].

Often, the $Z'$ is expected to decay leptonically in reactions such as $p\bar{p} \rightarrow Z' + ... \rightarrow (e^+e^- + ... \mu^+\mu^- + ... )$. Searches for a resonance of this type have been conducted at CDF, but have seen no significant evidence for this signal [16]. However, these searches do not rule out the existence of a $Z'$ resonance as it may have suppressed decays to leptons, in which case the $Z'$ is called “leptophobic”. In such a case, several models suggest that $Z'$ would decay predominantly to quark-anti-quark pairs, as seen in figure 4, and would have a larger branching ratio for decaying to up-type quark pairs rather than down-type quark pairs [15]. If the $Z'$ mass is larger than $2m_{top}$, the decay channel to top quark pairs will be open and the $Z'$ could appear as an additional source of $t\bar{t}$ beyond the SM $t\bar{t}$ production from gluons [15]. Unfortunately, quark-anti-quark decays are plagued by QCD background processes and make $Z'$ detection difficult [17]. That is, SM $t\bar{t}$ is dominated by production from gluons, but deviations from the expected physics in $t\bar{t}$ production could be signs for $t\bar{t}$ production from a $Z'$ Boson. The signal of this would be a larger than expected top production rate and some kind of resonant structure in the $M_{t\bar{t}}$ spectrum. Searches for a leptophobic $Z'$ decaying to $t\bar{t}$ have been conducted at CDF in Run I [5] and recently at CDF in Run II with 682 pb$^{-1}$ [4]. Our search is conducted in 682 pb$^{-1}$ of CDF Run II data in the Lepton+Jets decay channel of $t\bar{t}$ and can considered complimentary to the recent results of reference [4].

## 3 Data Samples and Event Selection

In this analysis, we use CDF Run II data taken from July, 2001 through July, 2005. The amount of recorded data is measured in luminosity (in units of inverse barns$^1$) so that our data sample corresponds to a total luminosity of 682 pb$^{-1}$. For a given process

---

$^1$1 barn = $10^{-24}$ cm$^2$
Figure 4: Feynman diagram for $Z' \rightarrow t\bar{t}$.

(which we call process X), the number of events $N_X$ in a given data sample is given by

$$N_X = \sigma_X \cdot A_X \cdot \int \mathcal{L} \, dt$$

where $\sigma_X$ is the production cross section for this interaction, $A_X$ is our selection efficiency, or "acceptance", for this process, and $\int \mathcal{L} \, dt$ is the total luminosity of the data set. For example, for a process with a cross section of 2 pb with an acceptance of 10% there will be approximately $N=2 \times 0.1 \times 682=13.64$ events present in the data sample.

We will conduct our search in the lepton + jets decay channel. Since the $t\bar{t}$ production cross section is so much smaller than the $p\bar{p}$ collision cross section, it is vital that we properly identify the relatively small sample of $t\bar{t}$ events in the data. Several background processes can fake $t\bar{t}$ events and may be present in our data sample, which will be discussed in further detail in Section 4. Thus, we apply several cuts to the kinematic properties measured by the CDF-II detector in order to isolate $t\bar{t}$ events and reject backgrounds. For both data and Monte Carlo, a standard top selection [11] has been optimized to select a relatively pure $t\bar{t}$ sample without reducing our event statistics to an unreasonably low level. The cuts are:

1. A high-$E_t$ ($\geq 20\text{GeV}$) isolated muon or electron.
2. Missing $E_t > 20\text{GeV}$
3. Three tight jets with $E_t > 15\text{GeV}$ and $|\eta| < 2.0$
4. A fourth jet with $E_t > 8\text{GeV}$ and $|\eta| < 2.4$
5. One tight SecVtx tagged jet

Because top quark pairs are quite heavy, we expect the final state constituents in $t\bar{t}$ events to have large energies. The first cut ensures that a lepton is present in the event and is not located near a jet (which could contain or be mis-identified as a lepton). The second cut requires the event to have a large missing energy which is associated with the neutrino. The combination of the first two cuts helps to greatly reduce data sample contamination from QCD background processes. The third and fourth cuts ensure that
we have the four jets which must be present in the Lepton+Jets \( t\bar{t} \) decay channel. We have allowed one of the four jets to have a lower energy in order to increase event statistics. The fifth cut ensures that we have an event which contains an identified b-quark. Of course a \( t\bar{t} \) event should contain two b-jets, but the tag rate for b-jets is only \( \approx 60\% \), and the requirement of one b-jet is sufficient for rejecting a large portion of the backgrounds.

The acceptances for selecting \( t\bar{t} \) events is found in table 4 for SM \( t\bar{t} \) and for \( t\bar{t} \) produced by \( Z' \) of various masses. We do not present the acceptances for the backgrounds in this paper because the expected number of events in the 682 pb\(^{-1} \) \( t\bar{t} \) sample for each background process have been calculated in reference [6]. We use these predicted background normalizations for our backgrounds and only rely on the event selection of background MC to model the background shapes. Background normalizations will be further discussed in section 5.3.2.

We will refer to this b-tagged lepton + 3.5 jet sample as the Ljb3.5 sample. With the specified cuts we can get a signal to background ratio of approximately \( S/B \approx 4 \). After we properly reconstruct the \( t\bar{t} \) events, an additional cut will be applied to the data and MC in order to improve signal sensitivity and to eliminate improperly reconstructed outlier events which may not be well modeled by our Monte Carlo. This additional cut will be discussed in Section 6.3.

### 4 Signal and Background Models

The expected \( t\bar{t} \) signal is modeled with an official CDF MC sample produced with \( M_{t\bar{t}} = 175 \text{GeV}/c^2 \). This sample contains approximately 1.1 million events of which 47 thousand pass the Ljb3.5 cuts.

The \( Z' \) signal model was produced with PYTHIA version 6.216 process 141, quark-
5 METHOD

anti-quark annihilation into a heavy neutral gauge boson with the same couplings as the Standard Model $Z^0$, but of course with mass above the $t\bar{t}$ threshold. Signal samples were produced with masses of $M_{Z'} = \{400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900\}$ GeV/c$^2$. Each sample contains about 60 thousand events of which about 3 thousand pass our cuts.

There are several SM processes which can fake a $t\bar{t}$ event, even with our cuts. These are called the non-top SM backgrounds (or simply the non-top backgrounds). The main contributions to the non-top backgrounds are:

- W + heavy flavor quark production in which the initial $q\bar{q}$ interaction does not decay to top quarks but decays directly to a W Boson along with heavy flavored quarks (bottom and charm). The main W + heavy flavor quark processes are $q\bar{q}$ decaying to $W + b\bar{b}$, $W + c\bar{c}$, and $Wc$. This is the largest contribution to our backgrounds.

- Non-top QCD interactions which result in events with four jets, a lepton and missing energy. Often the jets in these events are produced from an initial state or final state radiation of one or several gluons which produce jets.

- Electroweak production in which the $q\bar{q} \rightarrow WW$ or $q\bar{q} \rightarrow WZ$. The presence of W bosons in these events implies that their decay products could look similar to $t\bar{t}$.

- SM single top production. As these events do have one top quark, they can have a similar signature to $t\bar{t}$ events in the detector.

The expected contributions of each process is discussed in Section 5.3.2.

5 Method

5.1 Overview

Before a search for a $Z'$ resonance can be conducted in the Ljb3.5 sample, we must reconstruct the events. By reconstruction, we refer to the process of determining the longitudinal component of the neutrino $p_z$ and determining which parton each jet is associated with. This reconstruction is done with a $\chi^2$ based algorithm, which allows energy and momentum measurements of the CDF detector to vary within experimental resolution and uses the known masses of the quarks, electrons and W Bosons, in order to best fit each possible combination of jet combinations and the neutrino $p_z$. The lowest $\chi^2$ combination is used as our fit. The result is a full kinematic reconstruction of the 4-vectors for every particle which was present in the event decay chain, including the $t\bar{t}$ system.

We use the standard tools developed for the lepton+jets top mass and b-tag cross section analysis [11] to model and normalize the expected Standard Model contributions
(SM $t\bar{t}$ and non-top SM backgrounds) to the $M_{t\bar{t}}$ invariant mass spectrum, along with a vanilla version of a PYTHIA $Z'$ to model the signal resonance. The $\chi^2$ fit is applied to the models of SM $t\bar{t}$, non-top SM backgrounds, and the $Z'$ signal for various masses. The reconstructed $M_{t\bar{t}}$ spectrum for each process is a distribution of the $t\bar{t}$ fit masses of each event placed into a single histogram, called a "template". Using the $M_{t\bar{t}}$ templates for these samples, we do a 3 parameter binned likelihood fit to measure or derive limits on the apparent $Z'$ production cross section. The likelihood fit is essentially a shape fitting algorithm which determines the most probable composition of the three input model distributions in a given data sample.

5.2 Event Reconstruction

In the lepton+jets channel both top quarks decay to $W + b$, where each b quark produces a hadronic jet. One of the W’s decays leptonically to a lepton and a neutrino, while the other W decays hadronically and produces two jets. Therefore, an event in the lepton+jets decay channel will have four jets, a tight lepton, and a missing transverse energy. To determine the parton assignments of the jets, fit the missing longitudinal neutrino momentum, and reconstruct the $t\bar{t}$ kinematics, we use the $\chi^2$ minimization algorithm (also called the kinematic fit) to compare the event with the $t\bar{t}$ hypothesis. The $\chi^2$ function measures the goodness of fit between the fit values and measured values of the parameters in the function. Some parameters, such as the mass of the W boson, are known very precisely and we can constrain the fit value of the parameter to the known value. By trying each combination of jet-parton assignment and neutrino $p_z$ directions, and varying all parameters in the fit simultaneously, we can find the solution with the lowest $\chi^2$ value.

The $\chi^2$ definition is:

$$\chi^2 = \frac{(M_{jj} - M_W)^2}{\Gamma_W^2} + \frac{(M_{l\nu} - M_W)^2}{\Gamma_W^2} + \frac{(M_{bjj} - M_{Fit})^2}{2.5^2} + \frac{(M_{bl\nu} - M_{Fit})^2}{2.5^2}$$
$$+ \sum_{i=l,jets} \frac{(P_{i,meas} - P_{i,fit})^2}{\sigma_i^2} + \sum_{j=x,y} \frac{(P_{UE,meas} - P_{UE,fit})^2}{\sigma_j^2}$$

(3)

The input parameters to the $\chi^2$ function are:

- The transverse momentum of the lepton and the jets, $P_{t,i}^i$, which vary within experimental resolution, $\sigma_i$.

- The mass of the leptonic W boson, $M_{l\nu}$, which is calculated by adding the 4-vectors of the lepton and neutrino.

- The mass of the hadronic W boson, $M_{jj}$, which is calculated by adding the 4-vectors of the two non-b jets.
• The mass of the leptonically decaying top quark, $M_{bl\nu}$, which is calculated by adding the 4-vectors of the leptonic W and a b quark.

• The mass of the hadronically decaying top quark, $M_{bjj}$, which is calculated by adding the 4-vectors of the hadronic W and a b quark.

• The unclustered energy of all particles in the event not associated with the jets or lepton, $P_{jUE}$, which varies within the experimental resolution, $\sigma_j$.

In the standard fit, the leptonic and hadronic top masses are required to be identical within $2.5 \text{ GeV}/c^2$, and the W masses are constrained to $80.4 \text{ GeV}/c^2$, weighted by the W width of $\Gamma_W = 2.1 \text{ GeV}$. In addition to the kinematic constraints, b-tagged jets are required to be used as b’s. We choose the solution with the lowest $\chi^2$.

5.3 Top Mass Resolution

![Figure 5: Left: Reconstructed top mass spectrum in MC. Right: The un-normalized pull from $M_{top} = 175 \text{ GeV}/c^2$.](image)

We assess the reliability of our reconstruction by studying the SM $t\bar{t}$ signal Monte Carlo. Events are produced with $M_t = 175 \text{ GeV}/c^2$, and passed through the standard selection and kinematic fit. By examining the fitted top mass spectrum, we can check that the kinematic fit is properly reconstructing the known mass of the top quark. Furthermore, since the top quark is produced with a width of $2.1 \text{ GeV}/c^2$, we can measure our resolution by examining how spread out our fit top mass spectrum has become. We measure the un-normalized pull distribution, $M_{fit} - M_{true}$, where $M_{true}$ is the MC true mass at which the top quark was generated. This true mass also correspond to the mass of the MC top quark before it was put through detector simulation.

The fitted top mass is shown on the in figure 5a. The mean is $173.4 \text{ GeV}/c^2$ indicating that we have a small bias of $\approx 1.6\text{GeV}/c^2$. Thus, we have reconstructed the
5. METHOD

Top mass to within 1% of the mass at which it was generated. The un-normalized pull distribution \( M_{\text{fit}} - M_{\text{true}} \), shown in Fig. 5b, has a mean of \( \approx 1.5 \text{GeV}/c^2 \) (as expected) and RMS of \( \approx 35 \text{GeV}/c^2 \). In other words, our resolution on measuring the top mass is \( \approx 35 \text{GeV}/c^2 \). Since \( M_{\text{tt}} \approx 2 \times M_t \) we might expect an \( M_{\text{tt}} \) resolution of approximately 70 GeV/c².

5.3.1 The Reconstructed \( t\bar{t} \) Mass Spectrum

Next we check how well our reconstruction reproduces the invariant mass of the top pair. Using the same events as above we calculate \( M_{tt} \) from MC truth information and compare to the result from the reconstruction.

The standard reconstruction requires \( M_t = M_{\bar{t}} \), but of course the mass is to be determined. In a measurement of \( M_{tt} \) we can constrain both quantities to the known top mass \( M_t = 175 \text{GeV}/c^2 \) (weighted by the top width \( \Gamma_W = 2.1 \text{ GeV} \)) and improve our resolution. We call this the “constrained fit”. The standard mass reconstruction is the “unconstrained fit”.

The reconstructed \( M_{tt} \) spectrum for the standard fit is shown on the left in Fig. 6. At the lower masses the reconstruction slumps below the pair threshold. On the right, we show the difference between the true pair mass and the fitted pair mass for the lowest \( \chi^2 \) solution. We can see that pair mass spectrum has tails at the low and high masses and the difference fit - true (the un-normalized pull) is biased, with a mean at -16 GeV/c², and broad, with an RMS of 71.8 GeV/c², as expected.

The \( M_{tt} \) spectrum for the constrained fit is shown in Fig. 7. The pair mass has a sharp edge at the \( t\bar{t} \) threshold, and the falling part of the distribution is reasonably reproduced, although the reconstruction is still slightly low. The pair mass difference has a smaller bias, and although the RMS does not differ significantly from the unconstrained case, the bulk of the distribution is much narrower.
5 METHOD

Figure 7: Left: Reconstructed pair mass in MC with additional top mass constraint. Right: The difference between fit and true pair mass.

<table>
<thead>
<tr>
<th>Process</th>
<th>$N_{\text{events}}$</th>
<th>$\delta N_{\text{events}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>W+Jets</td>
<td>20.12</td>
<td>$\pm 3.19$</td>
</tr>
<tr>
<td>QCD</td>
<td>13.08</td>
<td>$\pm 2.57$</td>
</tr>
<tr>
<td>Mistags</td>
<td>18.35</td>
<td>$\pm 1.55$</td>
</tr>
<tr>
<td>Single Top T</td>
<td>1.27</td>
<td>$\pm 0.18$</td>
</tr>
<tr>
<td>Single Top S</td>
<td>0.95</td>
<td>$\pm 0.13$</td>
</tr>
<tr>
<td>Total Method 2</td>
<td>53.77</td>
<td>$\pm 4.39$</td>
</tr>
</tbody>
</table>

Table 5: The predicted normalizations of the non $t\bar{t}$ backgrounds from reference [14].

strained fit as our default reconstruction (The relative sensitivity of the constrained vs. unconstrained fit will be discussed in Section 6.3).

5.3.2 The “$M_{t\bar{t}}$” Spectra of the Backgrounds

The reconstructed shape of the effective $M_{t\bar{t}}$ spectra for the main backgrounds are displayed in Fig. 8. These are derived from the samples discussed in Section 4. Events passing all cuts are reconstructed with the constrained fit.

Since all of the distributions are quite similar, it is convenient to create a single “non-$t\bar{t}$ background template” template by superposing them after normalizing them to their predictions from reference [6]. The predicted number of events for each background process is shown in Table 5. We use these as relative weights to superimpose the distributions in Fig. 8 to create the single non-$t\bar{t}$ background template shown in Fig. 9.
5.3.3 The Z’ Signal Reconstruction

The $Z' \rightarrow t\bar{t}$ MC samples were produced such that a heavy neutral $Z'$ boson was forced to decay into $t\bar{t}$. This was the only restriction on the MC. In the subsequent $t\bar{t}$ decays,
the W bosons were allowed to decay leptonically or hadronically. Signal acceptances for each \( Z' \to t\bar{t} \) mass are found in table 4.

Figure 10 shows the \( Z' \to t\bar{t} \) templates for \( M_{Z'} = \{500, 700, 900\} \) GeV/c\(^2\), reconstructed with the constrained fit in the left plots and with the unconstrained fit in the right plots. The mean and RMS for all \( Z' \to t\bar{t} \) templates for both fits is found in table 6. In the low mass range, the mean and RMS for both fits are quite similar, while in the high mass region the unconstrained fit appears to do slightly better. However, there is relatively no backgrounds present in this region and thus we do not believe there will be much difference in sensitivity between the constrained and unconstrained fits at high \( Z' \) mass. We explore the sensitivity of the two fits further in section 6.3.

The true mass from the MC is also seen in the \( M_{Z'} = 700 \) GeV/c\(^2\) plot for both fits and we see that the natural width of the \( Z' \) is much smaller than our resolution. Further, the templates have almost no tail in the high mass region, whereas there is a long low mass tail for both constrained and unconstrained fits. This long low mass tail is an artifact of the kinematic fitter. The kinematic fitter attempts to minimize the \( \chi^2 \) value for each jet to parton combination and it is possible that a wrong combination (incorrect leptonic b assignment) results in a lower \( \chi^2 \) value than the correct combination. The reconstruction with an incorrect combination will result in an incorrect invariant \( M_{t\bar{t}} \). The \( Z' \to t\bar{t} \) template for \( M_{Z'} = 700 \) GeV/c\(^2\) is found in Fig. 11, which displays the correct and incorrect combinations present in the signal template for no \( \chi^2 \) cut on the left and a cut of \( \chi^2 < 5.0 \) on the right.

As seen in Fig. 11, the main contribution to the long low mass tail is events reconstructed with incorrect jet to parton assignments. However, the properly reconstructed events form a peak around the true mass of \( M_{Z'} \). We would like to discard events with incorrect combinations and isolate the peak around the true mass. While the incorrect combination events can be discarded with a \( \chi^2 \) cut, there is a large loss in event statistics. For example, a \( \chi^2 < 5.0 \) cut causes a 68% loss in event statistics. We determine the \( \chi^2 \) cut which allows for maximum sensitivity to observing an \( Z' \to t\bar{t} \) resonance in section 6.3.

5.4 Likelihood Fit Method

Our search for \( (Z' \to t\bar{t}) \) is based upon a shape-fitting analysis, whereby we measure a production cross section on \( Z' \to t\bar{t} \). As we do not know the branching ratio (BR) for \( Z' \to t\bar{t} \) we measure \( BR(Z' \to t\bar{t}) \cdot \sigma_{Z'} \), but, for simplicity, we will represent this as just \( \sigma_{Z'} \). A previous attempt to measure \( \sigma_{Z'} \) with a similar method was performed with very low statistics in the "top discovery" dataset and is discussed in [5]. Using this method, we may also set cross section limits as a function of \( M_{Z'} \). We assume that the \( M_{t\bar{t}} \) spectrum contains only \( Z' \to t\bar{t} \), SM \( t\bar{t} \), and non-\( t\bar{t} \) backgrounds, and perform a three parameter binned likelihood fit to the data.

The \( M_{t\bar{t}} \) values are organized in a binned histogram, called a template. If the total number of events is \( N_d \), then the expected number of events in the \( i^{th} \) bin is
\[ \mu_i = N_d \cdot P_{t,i} = \sigma_{Z'} \cdot \text{BR}\{Z' \to t\bar{t}\} \cdot A \cdot \int \mathcal{L} dt \cdot P_{Z',i} + N_d \cdot \beta_1 P_{tt,i} + N_d \cdot \beta_2 P_{bkg,i} \tag{4} \]

where \( A \) is the \( Z' \) signal acceptance and \( \int \mathcal{L} dt \) is the total integrated luminosity. The variables \( P_{Z',i} \), \( P_{tt,i} \), and \( P_{bkg,i} \) are the probabilities of observing a signal event, \( t\bar{t} \) event and non-\( t\bar{t} \) background event, respectively, in the \( i^{th} \) bin. For each template, these are simply the fraction of the template found in the \( i^{th} \) bin. The parameters \( \beta_1 \) and \( \beta_2 \) are the SM \( t\bar{t} \) fraction and the non-\( t\bar{t} \) background fraction respectively, i.e. they are the fraction of the total number of events which are background. The probability of observing \( n_i \) events in the \( i^{th} \) bin when \( \mu_i \) events are expected is given by the Poisson distribution as each bin of the distribution essentially acts as an individual counting experiment. The Poisson distribution is given by,

\[ P_i(n_i, \mu_i) = \frac{\mu_i^{n_i} e^{-\mu_i}}{n_i!} \tag{5} \]

The likelihood function \( L \) is then defined as

\[ L = \left( \prod_i P_i(n_i, \mu_i) \right) \cdot \left( \frac{1}{\delta c \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{\beta_2 - c}{\delta c} \right)^2} \right) \tag{6} \]

Assuming that the background estimations are accurate, the contributions of the non-\( t\bar{t} \) backgrounds in the likelihood function can be constrained. Thus, we constrain the parameter \( \beta_2 \) to \( c \) weighted by the error on the constrained value, \( \delta c \). This constraint ensures that non-\( t\bar{t} \) background fraction does not deviate far from the background predictions. As we are constraining a background fraction, we see that \( c = (\sum \text{background predicted normalizations})/N_d \) and \( \delta c = (\text{total background prediction error})/N_d. \)

Table 6: Constrained and unconstrained mean and RMS for each \( Z' \to t\bar{t} \) template. All number are in units of GeV/c^2.

<table>
<thead>
<tr>
<th>( M_{Z'} )</th>
<th>Mean (CON)</th>
<th>RMS (CON)</th>
<th>Mean (UNC)</th>
<th>RMS (UNC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>400</td>
<td>413</td>
<td>52</td>
<td>402</td>
<td>74</td>
</tr>
<tr>
<td>450</td>
<td>439</td>
<td>56</td>
<td>438</td>
<td>77</td>
</tr>
<tr>
<td>500</td>
<td>461</td>
<td>61</td>
<td>468</td>
<td>73</td>
</tr>
<tr>
<td>550</td>
<td>498</td>
<td>78</td>
<td>509</td>
<td>83</td>
</tr>
<tr>
<td>600</td>
<td>531</td>
<td>94</td>
<td>546</td>
<td>92</td>
</tr>
<tr>
<td>650</td>
<td>569</td>
<td>107</td>
<td>581</td>
<td>101</td>
</tr>
<tr>
<td>700</td>
<td>602</td>
<td>120</td>
<td>619</td>
<td>109</td>
</tr>
<tr>
<td>750</td>
<td>641</td>
<td>131</td>
<td>650</td>
<td>117</td>
</tr>
<tr>
<td>800</td>
<td>674</td>
<td>145</td>
<td>687</td>
<td>124</td>
</tr>
<tr>
<td>850</td>
<td>707</td>
<td>158</td>
<td>715</td>
<td>140</td>
</tr>
<tr>
<td>900</td>
<td>740</td>
<td>172</td>
<td>750</td>
<td>152</td>
</tr>
</tbody>
</table>
Figure 10: Constrained fit (left plots) and unconstrained fit (right plots) $Z' \rightarrow t\bar{t}$ templates for $M_{t\bar{t}} = \{500, 700, 900\}$.

We then use a minimization algorithm to minimize $-\ln L$ with respect to the three parameters $\sigma_{Z'}$, $\beta_1$ and $\beta_2$, thereby maximizing the likelihood function $L$. The fit cross section is defined as the cross section which maximizes the likelihood. In order to establish an upper limit on the cross section for $Z' \rightarrow t\bar{t}$ at the 95% CL, the likelihood
Figure 11: Constrained $Z' \rightarrow t\bar{t}$ shape showing correct and incorrect leptonic $b$ assignment for Left: No $\chi^2$ cut. Right: $\chi^2 < 5.0$ cut.

function is integrated over $\sigma_{Z'}$ and the upper limit is determined by the value of the cross section which corresponds to 95% of the likelihood function area, $\sigma_{Z'}^{95\%}$.

6 Analysis

We now assess the accuracy and reliability of our likelihood search machinery, optimize our sensitivity to accurately measure $\sigma_{Z'}$, and derive the SM expected limits. To perform all three of these tasks, we must have data simulations of the $M_{t\bar{t}}$ spectrum, called pseudo data. The process for creating pseudo data samples is outlined in Section 6.1. We then perform pseudo experiments (PE’s) by applying the likelihood fit machinery to the pseudo data samples and measuring $\sigma_{Z'}$ (or $\sigma_{Z'}^{95\%}$). By measuring the $\sigma_{Z'}$ for various $Z'$ input cross sections, $\sigma_{input}$, we can test the linearity of our system, i.e. we test how accurately we can measure a given input cross section. Furthermore, we can optimize our sensitivity by choosing the reconstruction method (constrained vs. unconstrained) and the $\chi^2$ cut. Once the system is optimized, we measure upper limits in pseudo data samples which only contain SM processes (i.e. $\sigma_{input} = 0$) in order to derive our expected limits given that only SM process are present in the data (i.e. the SM expected limits).

6.1 Pseudo Experiments

In order to perform reliable pseudo experiments, it is necessary to create pseudo data that will model our actual data sample. Therefore, it is necessary that the SM $t\bar{t}$, non-top backgrounds, and $Z' \rightarrow t\bar{t}$ templates are properly normalized before they are combined to form pseudo data sample. This will ensure that the pseudo experiments produce valid predictions. Thus, all pseudo experiments were performed with an total
luminosity of $\int L \, dt = 682 \text{pb}^{-1}$. In addition, we assume that the presence of a new production mechanism for $t\bar{t}$ would reduce the number of SM $t\bar{t}$ events while leaving the non-$t\bar{t}$ events unaffected. The template normalizations are as follows:

- The total $Z'$+SM $t\bar{t}$ sample is scaled from the CDF average top cross section of 7.3 pb [10] such that $\sigma_{t\bar{t}}^{CDF} = \sigma_{t\bar{t}}^{SM} + \sigma_{Z'}^{input}$.
- The $Z'$ template is normalized to $N_{Z'}^{input} = \sigma_{Z'}^{input} \cdot A_{Z'} \cdot \int L \, dt$.
- The SM $t\bar{t}$ template is normalized to $N_{t\bar{t}}^{SM} = N_{t\bar{t}}^{CDF} - N_{Z'}^{input}$ where $N_{t\bar{t}}^{CDF} = \sigma_{t\bar{t}}^{CDF} \cdot A_{t\bar{t}} \cdot \int L \, dt$.
- Non-$t\bar{t}$ backgrounds are normalized to the background predictions of [14].

In an analysis such as this, where each bin of the distribution acts like an individual counting experiment, each bin is subject to statistical fluctuations. Thus, once the normalizations are determined, each sample is individually Poisson fluctuated by bin, and then combined to form a pseudo data sample. Each pseudo data sample is run through the likelihood fit in order to measure a fit $Z'$ cross section, $\sigma_{fit}$, and an upper limit on the $Z'$ cross section at the 95% CL, $\sigma_{95\%fit}$. A typical fit to pseudo data and the fit likelihood function can be found in Fig. 12. In this typical fit, we have input an $M_{Z'} = 700 \text{ GeV}/c^2$ signal of 1 pb and run the likelihood fit. On the left we see the fit, and can see the fit contribution of the $Z'$ template in purple as a bump at approximately 700 GeV/c^2. On the right, we see the likelihood function, where the maximum value is the fit value and the upper limit at the 95% CL is displayed with an arrow.

### 6.2 Likelihood Function Linearity Studies

We test the response and reliability of our likelihood fit method for each $Z'$ mass by inputting pseudo data samples created with a $Z'$ input cross section from 0 pb to 4 pb and measuring $\sigma_{fit}$. 1000 pseudo experiments are performed for each $Z'$ input cross section, the median $\sigma_{fit}$ is used as our measured value and the RMS of the distribution is used as an estimate of the statistical error. For this linearity study, the constrained fit and a $\chi^2$ cut of $\chi^2 < 50$ were used (both motivated by sensitivity studies from section 6.3). The result of this linearity test can be found in figure 13 along with black dashed lines representing a $\pm 10\%$ deviation from a perfect response (slope of 1). The likelihood fit machinery is seen to have a linear response to a given input cross section, and is seen to measure a given input cross section to within 10% of the input value. When we investigate systematic errors, the response curves in figure 13 will be compared with our response curves shifted due to systematic uncertainties. For those purposes, the curves in figure 13 will be referred to as the unshifted measurements, $\sigma_{fit}^{unshifted}$. 
6.3 Choice of Reconstruction and $\chi^2$ Cut

Sensitivity for measuring $\sigma_{Z'}$ is optimized by investigating how well we measure upper limits on pseudo data samples which do not contain a $Z'$ signal, i.e. we test to see
how well we can measure a null signal. We first compare the effectiveness of the unconstrained kinematic fit to that of the constrained fit (that is when we force $M_t = 175$ GeV/$c^2$). Once we have chosen a fit type, we then optimize our $\chi^2$ cut, where $\chi^2$ is the "goodness of fit" value calculated by the reconstruction for each event.

For the comparison of unconstrained to constrained fits, pseudo data samples are created with $\sigma_{Z'^{input}} = 0$ and the upper limit cross section at the 95% CL, $\sigma_{95\%fit}$, is calculated for each event. 1000 pseudo experiments are performed for each $M_{Z'} = \{500, 600, 700, 800, 900\}$ GeV/$c^2$ using both the unconstrained and constrained fits. The median $\sigma_{95\%fit}$ of the pseudo experiments are shown in figure 14. The upper limits placed at each mass were seen to be similar. This is not surprising as the first and second moments of the constrained and unconstrained $Z'$ distribution were similar at each $Z'$ Mass, as seen in table 6. The constrained fit was chosen, as it appears to measure slightly lower limits when considering the entire mass range, but the difference is obviously minor.

![Constrained vs Unconstrained Upper Limits](image)

Figure 14: Upper limits calculated with the constrained and unconstrained fits.

The $\chi^2$ cut is the value at which we only use events which have a $\chi^2$ lower than the chosen cut. We suspect that events which have not been reconstructed well by the fitter may have a large value of the $\chi^2$, and thus it may be possible to eject these poorly constructed events from our sample with a $\chi^2$ cut. Thus, we expect that a cut will help to remove the non-$t\bar{t}$ background contamination in our sample. However, a cut will also reduce the number of good events in our samples and lower the statistical sensitivity.

To optimize the $\chi^2$ cut, the constrained fit was used and pseudo data samples were generated with $\sigma_{Z'^{input}} = 0$. For $M_{Z'} = \{500, 700, 900\}$ GeV/$c^2$ and a $\chi^2$ cut of $\{10, 20, 30, 50, \text{no cut}\}$, the median $\sigma_{95\%fit}$ was measured as a function of $Z'$ mass (and
RMS used as a measure of statistical error), as seen in figure 15. A cut of \( \chi^2 < 50 \) appears to be optimal considering the entire range of \( Z' \) masses. This cut does not greatly affect our upper limit measurements, but does reduce any high \( \chi^2 \), poorly reconstructed events that may be present in our data but not properly modeled by MC.

### 6.4 Expected Limit Without Systematic Errors

SM expected limits are obtained by fitting 682pb\(^{-1}\) pseudo experiments with no \( Z' \) signal present. The constrained fit and \( \chi^2 < 50 \) cut were used and \( \sigma_{95\%}^{\text{fit}} \) was measured for 1000 pseudo experiments. The median of this distribution is used as the central value, and the RMS is used as an estimate of a \( \pm 1\sigma \) deviation from the central value. These pseudo experiments were done for each \( Z' \) signal template. The resulting expected limits can be found in figure 16.

### 7 Shape Systematic Errors

Since our analysis depends on the reliability of the \( M_{tt} \) shape from the MC, we investigate the systematic uncertainties that depend on the shapes of our template distributions. It should be noted that these shape systematics also change the acceptances for each distribution. The prior analysis with small statistics showed that the effects of systematic uncertainties are small [5]. We validate that the systematic uncertainties are small by investigating the largest source of shape systematic uncertainty, the Jet Energy Scale (JES) systematic error. The JES systematic error is the uncertainty associated with possible mis-measurements of jet energies in the CDF calorimeter. We
study the JES systematic error for $M_{Z'} = \{500, 650, 800\}$ GeV/$c^2$, and show that these uncertainties are small compared to statistical errors over the entire mass range.

To investigate the JES systematic error, a $\pm 1\sigma$ shift in the JES was applied to each jet energy measurement in each event of each of our templates before event selection. The shifted templates are then run through the event selection and reconstruction with the kinematic fitter. The effect of the JES systematic on template shapes and acceptances is investigated in section 7.1.

The shifted $Z'$ and SM $t\bar{t}$ templates were then used to create PE’s with shifted shapes and acceptances. Using the likelihood fit, the shifted PE’s were fit with the unshifted templates in order to parameterize the change in fit cross section, $\Delta(\sigma_{Z'})$, as a function of the cross section itself. This parameterization of the change between the unshifted and shifted likelihood response curves gives us direct information on how much the JES systematic will change our cross section measurements. Thus, $\Delta(\sigma_{Z'})$ is taken as our systematic uncertainty on the fit cross section and is used as a Gaussian like error to smear, or marginalize, the likelihood function.

All systematic studies are done using the constrained fit and with a cut of $\chi^2 < 50$.

7.1 Shifted Template Shapes and Acceptances

The $\pm 1\sigma$ JES shifted $Z'$ and SM $t\bar{t}$ templates have been plotted along with the unshifted templates in figure 17. There appears to only be a small change in shape between the shifted and unshifted templates for both $Z'$ and SM $t\bar{t}$. Thus we do not believe the change in shape will have a large effect on expected limits. The largest effect of the
Figure 17: The unshifted and ±1σ JES shifted shapes of the SM $t\bar{t}$ (left) and $Z'$ with a mass of 700 GeV/$c^2$ (right).

Table 7: Shifted acceptances due to JES uncertainties for SM $t\bar{t}$ and $Z' \rightarrow t\bar{t}$ templates

<table>
<thead>
<tr>
<th>Template</th>
<th>Acceptance [%] Unshifted</th>
<th>Acceptance [%] JES +1</th>
<th>Acceptance [%] JES -1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_{Z'}=400$</td>
<td>3.26</td>
<td>3.44</td>
<td>3.1</td>
</tr>
<tr>
<td>$M_{Z'}=450$</td>
<td>3.48</td>
<td>3.65</td>
<td>3.31</td>
</tr>
<tr>
<td>$M_{Z'}=500$</td>
<td>3.64</td>
<td>3.82</td>
<td>3.44</td>
</tr>
<tr>
<td>$M_{Z'}=550$</td>
<td>3.76</td>
<td>3.93</td>
<td>3.58</td>
</tr>
<tr>
<td>$M_{Z'}=600$</td>
<td>3.72</td>
<td>3.79</td>
<td>3.48</td>
</tr>
<tr>
<td>$M_{Z'}=650$</td>
<td>3.77</td>
<td>3.92</td>
<td>3.63</td>
</tr>
<tr>
<td>$M_{Z'}=700$</td>
<td>3.70</td>
<td>3.84</td>
<td>3.57</td>
</tr>
<tr>
<td>$M_{Z'}=750$</td>
<td>3.70</td>
<td>3.81</td>
<td>3.57</td>
</tr>
<tr>
<td>$M_{Z'}=800$</td>
<td>3.88</td>
<td>4.02</td>
<td>3.74</td>
</tr>
<tr>
<td>$M_{Z'}=850$</td>
<td>3.72</td>
<td>3.84</td>
<td>3.59</td>
</tr>
<tr>
<td>$M_{Z'}=900$</td>
<td>3.91</td>
<td>4.03</td>
<td>3.77</td>
</tr>
<tr>
<td>SM $t\bar{t}$</td>
<td>3.4</td>
<td>3.55</td>
<td>3.25</td>
</tr>
</tbody>
</table>

JES uncertainties is from the change in template acceptances, as seen in table 7. This change in acceptance will alter the normalizations of the templates in the PE’s, and will be the leading source of any shifts in measured cross sections.

### 7.2 Cross section Shift Parameterization

For input $Z'$ cross sections between 0pb and 4pb, PE’s with the shifted templates are fit with the unshifted templates in order to measure the shifted fit cross section, $\sigma_{\text{shifted}}^{\text{fit}}$. The median $\sigma_{\text{shifted}}^{\text{fit}}$ of 1000 PE’s is used as the measured value. The previously mentioned unshifted fit cross section measurements, $\sigma_{\text{fit}}^{\text{unshifted}}$, are found in section 6.2.
7 SHAPE SYSTEMATIC ERRORS

The raw change in fit cross section, $\delta \sigma_{Z'} = \sigma_{fit}^{shifted} - \sigma_{fit}^{unshifted}$, plotted as a function of the cross section itself is found in left plot of figure 18. As $\delta \sigma_{Z'}$ will be used as a Gaussian like error to smear the likelihood function, we symmetrize $\delta \sigma_{Z'}$ for the $\pm 1\sigma$ JES shift by using the larger absolute value of $\delta \sigma_{Z'}$ for each value of $\sigma_{Z'}^{input}$, as seen in the right plot of figure 18. By using the larger absolute value of $\delta \sigma_{Z'}$ we are taking the largest possible shift in our cross section measurements. Finally, we need a smooth shift parameterization so that we may calculate cross section shifts at any cross section value. Thus, a linear fit to the symmetrized $\delta \sigma_{Z'}$ is used as our systematic uncertainty parameterization, $\Delta(\sigma_{Z'})$, such that $\Delta(\sigma_{Z'}) = \alpha_1 \cdot \sigma_{Z'} + \alpha_2$, as seen in figure 19. The measured values of $\alpha_1$ and $\alpha_2$ can be found in table 8.

### 7.3 Marginalization Method

Now that we have obtained a smooth parameterization for the shift in cross section, as a function of cross section, we want to input this information into our likelihood fit. Since
\[ L'(\sigma_{Z'}) = \int_0^\infty L(x) \cdot \frac{1}{\Delta(x)\sqrt{2\pi}} \exp \left( -\frac{1}{2} \left( \frac{\sigma_{Z'} - x}{\Delta(x)} \right)^2 \right) dx \]  

The marginalized likelihood function, \( L'(\sigma_{Z'}) \), now accounts for the for the \( \pm 1\sigma \) JES uncertainty when measuring data samples. The new fit cross section, \( \sigma_{fit} \), is defined as the \( \sigma_{Z'} \) which maximizes \( L'(\sigma_{Z'}) \). \( L'(\sigma_{Z'}) \) is integrated to 95\% of the total area to establish the upper limit on the cross section for \( Z' \rightarrow t\bar{t} \) at the 95\% CL, \( \sigma_{fit}^{95\%} \), with JES systematic uncertainties. A typical likelihood function and marginalized likelihood function are shown in figure 20. The effect of the marginalization is to spread the original likelihood function over a broader range of fit cross sections.

### 7.4 Shifted Expected Limits

In order to understand the effect of the systematic uncertainties on our upper limit measurements, we calculate the SM expected limits with the marginalized likelihood function and compare them to the un-marginalized SM expected limits of Section 6.4. For \( M_{Z'} = \{500, 650, 800\} \) GeV/c\(^2\), we performed 1000 PE’s with no input cross section and measured the median \( \sigma_{fit}^{95\%} \) with both the likelihood function and the marginalized likelihood function, as seen in figure 21. The JES systematic uncertainties appear to
have a small effect on our expected limit, especially for heavy $Z'$ and especially when compared to the large statistical errors inherent in our upper limit measurements (seen in figure 16). It should also be noted that the maximum deviations were used to symmetrize $\delta \sigma_{Z'}$, and thus these shifted limits represent the maximum expected limit shift due to JES uncertainties. Thus, while systematic uncertainties will slightly change our final upper limits, with the idea that uncertainties are dominated by statistics, we present our result in $682 \text{pb}^{-1}$ of CDF Run II data with statistical error only.

8 $M_{\text{tt}}$ Distributions in CDF-II Data

For the Ljb3.5 sample with no $\chi^2$ cut, the $319 \text{pb}^{-1}$ of pre-2005 data containing 121 candidate events is shown in the left plot of figure 22. In this early data set the structure at $\approx 500 \text{GeV}/c^2$ seemed interesting. The $362 \text{pb}^{-1}$ of data with no $\chi^2$ cut from January to August of 2005 containing 136 candidate events is shown in the right plot in figure 22. It can easily be seen that the peak at $500 \text{GeV}/c^2$ seen in the initial $319 \text{pb}^{-1}$ is not present in the latter $362 \text{pb}^{-1}$ of data. The full $682 \text{pb}^{-1}$ data sample after the $\chi^2 < 50$ cut is found in the left plot of figure 23, and contains 243 candidate events. Thus, 17 candidate events were discarded from our data sample due to the $\chi^2 < 50$ cut. Furthermore, while a small peak still persists at $500 \text{GeV}/c^2$ in the full data sample, it is no longer as prevalent as it was in the pre-2005 data. The right plot of figure 23 shows the $682 \text{pb}^{-1}$ data sample passing the $\chi^2 < 50$ cut together with the SM expectation. While the mean and RMS of the data and SM predictions agree fairly well, there also appears to be deviations between the shapes of the data and SM expectation.
9 Results

9.1 Most Likely Values

The most likely values (MLV) of the cross section \( \sigma_{Z'} \) for \( Z' \rightarrow t\bar{t} \) in the data can be found in table 9. We also include the MLV measurements separately for the first 319 pb\(^{-1}\) of data and the latter 362 pb\(^{-1}\) of data. At \( M_{Z'} = 500 \text{ GeV}/c^2 \) in the first 319 pb\(^{-1}\) of data, there appeared to be a 2 sigma deviation from zero cross section measurement. This is no longer present in the full 682 pb\(^{-1}\) data sample, and we see
Figure 23: Left: Total data sample with cut. Right: Total data sample with cut plotted along with the SM expectation.

<table>
<thead>
<tr>
<th>$M_{Z'}$</th>
<th>MLV (682pb$^{-1}$)</th>
<th>MLV (319pb$^{-1}$)</th>
<th>MLV (362pb$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>400</td>
<td>$0.0^{+1.04}_{-0.0}$</td>
<td>$0.0^{+1.00}_{-0.0}$</td>
<td>$2.06^{+1.87}_{-1.90}$</td>
</tr>
<tr>
<td>450</td>
<td>$0.0^{+1.20}_{-0.0}$</td>
<td>$1.30^{+1.92}_{-1.30}$</td>
<td>$0.0^{+0.87}_{-0.0}$</td>
</tr>
<tr>
<td>500</td>
<td>$0.57^{+0.92}_{-0.57}$</td>
<td>$2.57^{+1.40}_{-1.32}$</td>
<td>$0.0^{+0.47}_{-0.0}$</td>
</tr>
<tr>
<td>550</td>
<td>$0.51^{+0.51}_{-0.51}$</td>
<td>$1.16^{+1.05}_{-0.94}$</td>
<td>$0.0^{+0.18}_{-0.0}$</td>
</tr>
<tr>
<td>600</td>
<td>$0.0^{+0.29}_{-0.0}$</td>
<td>$0.0^{+0.39}_{-0.0}$</td>
<td>$0.0^{+0.2}$</td>
</tr>
<tr>
<td>650</td>
<td>$0.0^{+0.12}_{-0.0}$</td>
<td>$0.0^{+0.12}_{-0.0}$</td>
<td>$0.0^{+0.2}$</td>
</tr>
<tr>
<td>700</td>
<td>$0.0^{+0.19}_{-0.0}$</td>
<td>$0.0^{+0.19}_{-0.0}$</td>
<td>$0.0^{+0.1}$</td>
</tr>
<tr>
<td>750</td>
<td>$0.0^{+0.22}_{-0.0}$</td>
<td>$0.0^{+0.22}_{-0.0}$</td>
<td>$0.0^{+0.12}$</td>
</tr>
<tr>
<td>800</td>
<td>$0.11^{+0.19}_{-0.11}$</td>
<td>$0.31^{+0.35}_{-0.32}$</td>
<td>$0.0^{+0.17}_{-0.0}$</td>
</tr>
<tr>
<td>850</td>
<td>$0.09^{+0.17}_{-0.09}$</td>
<td>$0.22^{+0.32}_{-0.21}$</td>
<td>$0.0^{+0.18}$</td>
</tr>
<tr>
<td>900</td>
<td>$0.05^{+0.13}_{-0.05}$</td>
<td>$0.07^{+0.27}_{-0.07}$</td>
<td>$0.0^{+0.15}_{-0.0}$</td>
</tr>
</tbody>
</table>

Table 9: Most likely values of the cross section of $\sigma_{Z'} \cdot BR\{Z' \rightarrow t\bar{t}\}$ as a function of $M_{Z'}$.

no significant deviation from a $Z'$ cross section of zero for any mass. The likelihood fit to the 682 pb$^{-1}$ sample with a $Z'$ of mass 500 GeV/c$^2$ is found in figure 24. Since there appears to be no significant deviation from a $Z'$ cross section of zero, we proceed to calculated the upper limits on the $Z'$ cross section as a function of $Z'$ mass.

### 9.2 Upper Limits

The upper limit on the cross section for $Z' \rightarrow t\bar{t}$ at the 95% CL with statistical error only in the data along with the SM expectations are found in figure 25. Again, all limits are calculated with 682 pb$^{-1}$, a $\chi^2 < 50$ cut, and with statistical error only. The
9 RESULTS

![Graphs](image)

Figure 24: Left: Total data sample fit with a 500 GeV/c² Z'. Right: Likelihood function for the 500 GeV/c² Z' fit to data.

<table>
<thead>
<tr>
<th>M_{Z'}</th>
<th>Expected Limit (682pb⁻¹)</th>
<th>Data Limit (682pb⁻¹)</th>
<th>Data Limit (319pb⁻¹)</th>
<th>Data Limit (362pb⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>400</td>
<td>2.47 ± 0.86</td>
<td>2.49</td>
<td>2.47</td>
<td>5.22</td>
</tr>
<tr>
<td>450</td>
<td>2.31 ± 0.80</td>
<td>2.55</td>
<td>4.74</td>
<td>2.84</td>
</tr>
<tr>
<td>500</td>
<td>1.70 ± 0.61</td>
<td>2.31</td>
<td>4.98</td>
<td>1.83</td>
</tr>
<tr>
<td>550</td>
<td>1.18 ± 0.41</td>
<td>1.69</td>
<td>3.12</td>
<td>1.74</td>
</tr>
<tr>
<td>600</td>
<td>0.94 ± 0.38</td>
<td>0.91</td>
<td>1.77</td>
<td>1.20</td>
</tr>
<tr>
<td>650</td>
<td>0.71 ± 0.25</td>
<td>0.54</td>
<td>1.14</td>
<td>0.79</td>
</tr>
<tr>
<td>700</td>
<td>0.61 ± 0.22</td>
<td>0.56</td>
<td>1.16</td>
<td>0.74</td>
</tr>
<tr>
<td>750</td>
<td>0.48 ± 0.17</td>
<td>0.52</td>
<td>1.16</td>
<td>0.57</td>
</tr>
<tr>
<td>800</td>
<td>0.43 ± 0.15</td>
<td>0.54</td>
<td>1.09</td>
<td>0.59</td>
</tr>
<tr>
<td>850</td>
<td>0.37 ± 0.13</td>
<td>0.48</td>
<td>0.96</td>
<td>0.56</td>
</tr>
<tr>
<td>900</td>
<td>0.29 ± 0.10</td>
<td>0.37</td>
<td>0.74</td>
<td>0.48</td>
</tr>
</tbody>
</table>

Table 10: Results for the Limit on the cross section of σ_{Z'} as a function of M_{Z'}

limits in the data show no significant deviation from SM expectation. The measured values for the upper limits for the data and SM expectation can be found in table 10. In this table, we have also included the upper limits calculated in the 319 pb⁻¹ and 362 pb⁻¹ data samples separately.

As a final check that uncertainties are dominated by statistics, the upper limits
Figure 25: Upper limits on the production cross section for $Z' \rightarrow t\bar{t}$ in data plotted along with the SM expected values.

with JES systematic uncertainties for $M_{Z'} = \{500, 650, 800\}$ GeV/$c^2$ in the full 682 pb$^{-1}$ data sample and SM expectation is plotted along with the statistical error only limits in figure 26. The upper limits change slightly due to JES uncertainties, but the total uncertainty is statistics dominated.

Figure 26: Upper limits in data and SM expectations show with and without the effect of JES systematic errors.
10 Summary

We have reconstructed $M_{tt}$ for events in the b-tagged Lepton+Jets decay channel using the kinematic fitter. Using the signal, SM $tt$ and non-$tt$ background samples, we performed a 3 parameter binned likelihood fit and measured most likely cross section on $Z' \rightarrow tt$ as well as the upper limit on the cross section for $Z' \rightarrow tt$ at the 95% CL. We saw no evidence for a new resonance of $Z' \rightarrow tt$ over the mass range $400\text{GeV}/c^2 \leq M_{tt} \leq 900\text{GeV}/c^2$. 
11 References

References


[7] “Combination of CDF top pair production cross section measurements.” CDF note 7700


[12] “Z’ generation with PYTHIA” CDF note 7755

[13] R. Harris, C. Hill, S. Parke, “Cross Section for Topcolor $Z’$ decaying to $t\bar{t}$” CDF note 5064.


[16] A. Abulencia et al., The CDF Collaboration, ”Search for $Z’ \rightarrow e^+e^-$ Using Dielectron Mass and Angular Distribution” CDF note 7992.
