



Fermi National Accelerator Laboratory

FERMILAB-Conf-95/277-E

CDF

Kinematics of the $t\bar{t}$ Events in W^+ Jets at CDF

A. Beretvas

For the CDF Collaboration

*Fermi National Accelerator Laboratory
P.O. Box 500, Batavia, Illinois 60510*

August 1995

Published Proceedings from the *International Symposium on Particle Theory and Phenomenology, Workshop on Top Quark Physics*, Iowa State University, Ames, Iowa, May 25-26, 1995

Disclaimer

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Kinematics of the $t\bar{t}$ events in W + Jets at CDF

Andrew Beretvas

for the CDF Collaboration

Fermi National Accelerator Laboratory, Batavia Illinois 60510, USA

We compare the CDF W + n \geq 3 Jets data with predictions using the standard model. Herwig is used to simulate $t\bar{t}$ production and VECBOS is used to simulate QCD W + n \geq 3 Jets. We look at four different data sets with $t\bar{t}$ content varying from 20% to 75%. We examine several kinematic variables. We conclude that the data is consistent with the standard model.

1 Introduction

This talk is about $t\bar{t}$ kinematics in W + n \geq 3 jets events where the W decays leptonically. The topics covered will be the H analysis¹, comparisons of directly measured quantities², kinematics using mass fitting³, and a search for resonances decaying into $t\bar{t}$. The emphasis in this talk will be on showing that the kinematics of the CDF W + n \geq 3 jet events agrees with Monte Carlo predictions for top production (Herwig Monte Carlo)⁴ plus QCD W + jets background (VECBOS)⁵. The last two topics are presented in outline form as they will be presented in much greater depth later this year.

2 H Analysis

For this analysis of the CDF W + n \geq 3 jet data we use a variable called H. H is defined as the scalar sum of the lepton transverse energy, the neutrino transverse energy (measured by the \cancel{E}_T in the event) and the transverse energy of the jets ($E_T > 8$ GeV, and $|\eta| < 2.4$). We will see that the H variable is a good way to separate the top signal from the QCD W + jets background. We first define the requirements for a W sample. The event must contain an isolated electron or muon with transverse momentum greater than 20 GeV/c and pseudorapidity $|\eta| < 1.0$. Events with an ee or $\mu\mu$ pair with an invariant mass between 75 and 105 GeV/c² are removed since they are likely from a Z decay. Electrons from converted photons are removed with high efficiency by using tracking information. The \cancel{E}_T must be greater than 20 GeV. The signal sample requires that the W sample contain 4 or more jets. Three of the jets must pass a high threshold cut ($E_T > 15$ GeV and $|\eta| < 2.0$) and at least one additional jet must pass a low threshold cut ($E_T > 8$ GeV and $|\eta| < 2.4$).

The control sample called "3 Jet LOW" contains 814 events. The jet requirement is that there be 3 jets with E_T greater than 8 GeV and that there be no fourth jet with E_T greater than 8 GeV. The jets are only considered if $|\eta| < 2.4$. We have added to the VECBOS Monte Carlo data a very small component of top (1%)

corresponding to the expected top contamination in this sample. We find very good agreement between data and VECBOS when the Q^2 scale is chosen so that it is equal to the square of the average P_T of the jets ($\langle P_T \rangle^2$). The agreement is not quite as good when the Q^2 scale is chosen to be the square of the W boson mass.

At first glance this good agreement may be surprising because the background to $t\bar{t}$ production also includes $WW + \text{jets}$, where one W decays into an $e\nu$ or $e\mu$ pair and $W + \text{jets}$ where $W \rightarrow \tau\nu$ and $\tau \rightarrow l\nu_l\nu_\tau$. There are also non-W backgrounds. These are QCD multijet where one jet fakes an electron or muon, $b\bar{b} + \text{multijet}$ production where one of the b quarks decays semileptonically, $ZZ + \text{jets}$ with Z decaying leptonically but only one of the leptons being found, $Z \rightarrow \tau\tau$ followed by $\tau \rightarrow l\nu_l\nu_\tau$, and Drell-Yan production of lepton pairs along with extra QCD jets. The H distribution for these backgrounds are very similar to that of the main background of $W + \text{jets}$ (direct production of a W which recoils against light quarks and gluons). Thus we will consider only the $W + \text{jets}$ background (VECBOS).

The signal sample consisting of 99 events is given in Fig. 1. It is clear that the data is not the same as the VECBOS background. We show the plots for both choices of Q^2 scale. The H distribution for the data is at a higher average H than for VECBOS. The next figure (2) shows how the data can be expressed in terms of two

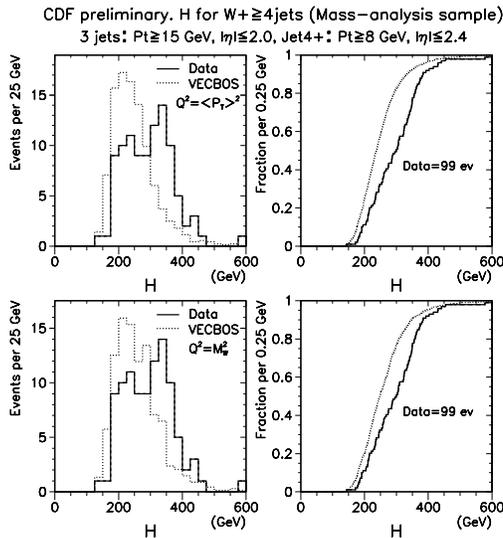


Figure 1: H distribution for the $W + \geq 4$ -jet events passing the high threshold cuts.

components. The first component is from $t\bar{t}$ production simulated using the Herwig Monte Carlo with $M_{\text{top}} = 180$ GeV/c^2 . The second component is from QCD $W + \text{jets}$ background (VECBOS). We see that the area corresponding to two components are about equal. When the two components are combined the result is a good fit to the data.

The signal events are fit with a binned likelihood fit to a linear combination of $t\bar{t}$ (Herwig) and background $W + \text{jets}$ (VECBOS). After fitting the data points as a function of top quark mass with a cubic polynomial we find the top quark mass

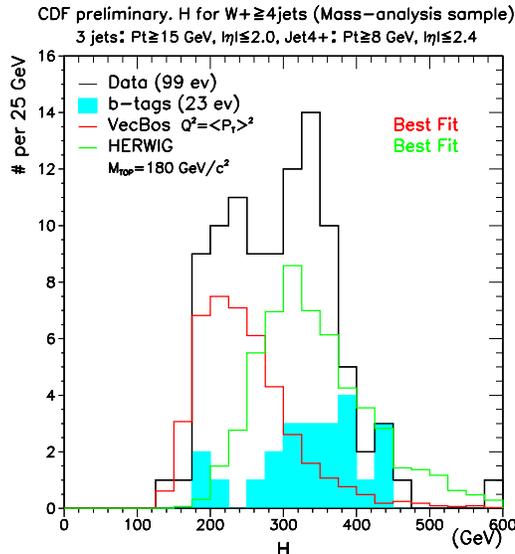


Figure 2: H distribution for the Top Monte Carlo (Herwig with $M_{\text{TOP}} = 180 \text{ GeV}/c^2$) and H distribution for VECBOS (QCD $W + \text{jets}$).

is $180 \pm 12 \text{ GeV}$. This value of the top mass is in excellent agreement with our previously published value³.

The most important source of systematic uncertainty is due to the jet energy scale. Other important considerations are the Q^2 scale in VECBOS, the underlying events in VECBOS and initial state radiation in Herwig. The final answer we obtain from this analysis is 180 ± 12 (stat) $^{+19}_{-15}$ (syst) GeV/c^2 . Both the statistical and systematic errors are a little larger than our previously published value³. Because this sample is also correlated with the b-tag sample used in our published result the result is quoted just to show consistency.

We have established that the total transverse energy distribution of $W + n \geq 4$ jets is not consistent with known backgrounds (deviates by more than 3.8σ using the Kolmogorov test). The best fit to the H distribution is obtained using a linear combination of Herwig with a top mass of $180 \text{ GeV}/c^2$ and VECBOS. We also find a large fraction of the b-tagged events in the high H region.

3 Directly Measured Kinematics

We will show that the $W + \geq 3$ Jet data (where the W decays leptonically) agrees with Monte Carlo predictions. The Monte Carlo predictions consist of $t\bar{t}$ events using Herwig and background consisting of QCD $W + n \geq 3$ jet production using VECBOS. Herwig is a Monte Carlo program based on the leading order QCD matrix elements for the hard process, followed by coherent parton shower evolution, hadronization, and an underlying event model based on data. VECBOS is a parton-level Monte Carlo program based on tree-level matrix element calculations.

We will show that the data agrees with the Monte Carlo predictions for many data sets, many kinematic variables and for a range of cuts allowed by statistics.

Four different data sets are used with $t\bar{t}$ content varying from about 20% to 75%. The first data set is the standard W + 3 jet sample. For this sample there must be 3 jets with $E_T > 15$ GeV. The second data set (used in our mass fits) requires the standard 3 jets plus a fourth jet with $E_T > 8$ GeV. The third set is called the high threshold sample and requires $E_T > 15$ GeV for the fourth highest energy jet. This sample also has a dijet separation cut ($\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} > 0.6$). The fourth sample requires the standard 3 jets and that there be an SVX-btag.

The kinematic variables are put into two classes which have different characteristics for the $t\bar{t}$ signal and the QCD W +jets background. Our first class is the energy variables which are used by most analyses. All the energy variables listed use only the transverse components of the energy since they give better separation between the $t\bar{t}$ signal and the QCD W + jets background. These variables all have the property that the mean of the distribution is greater for the $t\bar{t}$ signal than for the QCD W + jets background. The other class is that of angular variables. The polar angle variables separate the signal from the background because top is more centrally produced than the QCD W + jets background. The distribution in ϕ can also be used to separate top from background because top events are more circular than VECBOS. Also useful are combinations of the angular variables(θ, ϕ) like the aplanarity (top production is more aplanar than VECBOS).

We next show several of the kinematic variables for data set II. The jets are ordered in transverse energy (E_T) with the notation being $E_T(2)$ refers to the transverse energy of the second highest E_T jet. The distribution for $E_T(3) + E_T(4)$ is given in Fig. 3. This is one of the best variables for separating top from QCD W + jets. We see that the mean of the distribution for TOP 170 is 79.5 GeV and that for VECBOS is 56.8 GeV. The data has a mean of 64.3 GeV, much as expected. We have estimated the $t\bar{t}$ contribution to data set II to be 30%. This estimate is based on the number of SVX-btags in data set I and their estimated background. We need to know the efficiency of the SVX tagging ($42 \pm 5\%$) and the efficiency of data set II for top relative to data set I(86%). Fig. 4 shows the comparison of the data and a mixture of 30% TOP 170 and a 70% VECBOS. As expected good agreement is obtained between data and Monte Carlo.

To make a comparison of four different data sets we show what we call an ‘‘Overview Plot’’. This is an integral plot that shows deviations of the data from VECBOS predictions in units of statistical uncertainty. Each figure shows a different variable for the 4 data sets. One feature of these plots is that only the shapes of the distributions are compared (no absolute normalization is used). The horizontal axis is the fraction of TOP 170 Monte Carlo events passing the cut. The vertical axis is the standard deviation of the fraction of events above the cut from the predictions of a VECBOS template:

$$\frac{(f_{\text{data}} - f_{\text{VECBOS}})}{\sigma}$$

$$\sigma = \sqrt{\frac{(f_{\text{VECBOS}} + 1/n)(1 - f_{\text{VECBOS}} + 1/n)}{n}}$$

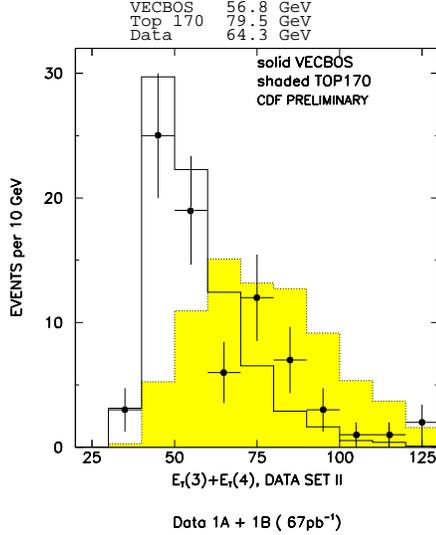


Figure 3: The distribution of $E_T(3) + E_T(4)$ for data set II. Monte Carlo distributions for VECBOS and TOP 170 for data set II.

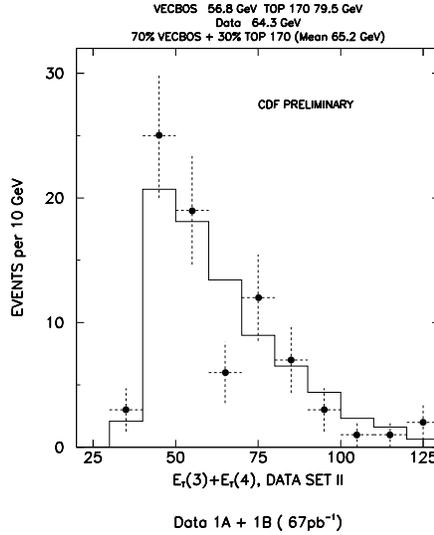


Figure 4: The histogram corresponds to a mixture of 30% TOP 170 and 70% VECBOS for data set II.

n = number of data events

The vertical deviation is by definition zero for no cut. As cuts are made on the variable, the cut efficiency for TOP 170 decreases from 1.0 as one moves along the horizontal axis from right to left. At first, the deviation from the VECBOS template typically increases as the cut efficiency decreases. At some cut efficiency, the curves typically turn over because the deviation from the long tails of the VECBOS template decreases in terms of expected statistical uncertainty. Below the cut efficiency when the VECBOS template predicts one event (0.5 for data set IV), all curves are set to zero. If the data is assumed distributed like the indicated mixture of $t\bar{t}$ and VECBOS, the expected statistical uncertainty is typically 1.0 to 1.5 vertical units (error bars on selected data points are shown). The error bars on adjacent points are correlated because it is an integral plot.

Before going to the Overview plots we show a Deviation plot (Fig. 5). This plot compares the predictions for several variables for data set II assuming there are 80 events of which 33% are TOP 170. All the curves start at 0.86 because this is the efficiency of the data set II cuts with respect to data set I. The solid triangles correspond to the variable $E_T(3) + E_T(4)$ which is the variable that has the best predicted discriminating power. The variable $(E_T(2) - 20) \times (E_T(3) - 20)$ is indicated by an open square. This variable also has very good discriminating power and is very similar to the variable used in the likelihood analysis⁶. The next

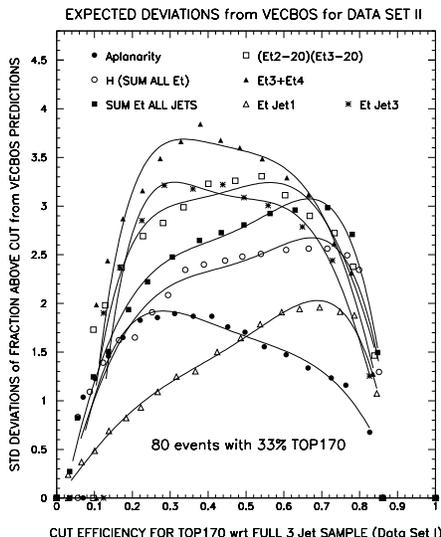


Figure 5: Deviation plot comparing variables for data set II. This is a purely Monte Carlo plot using 33% TOP 170.

best variable (indicated by an asterisk) is $E_T(3)$. Next in predicted discriminating power is the $\sum E_T$ of all of the jets above threshold (solid squares). Next comes the variable H (sum of all E_T 's in the event) which we discussed in detail in section 2 (open circles). The final two variables $E_t(1)$ and the aplanarity have relatively little predicted discriminatory power.

The data in the overview plot is represented by solid points. The hatched band is the expectation for 100% VECBOS; the width of the hatched region corresponds to two extreme q^2 scales for $\alpha(s)$: $\langle P_T \rangle^2$ and M_W^2 . For the energy variables the lower edge of the hatched region corresponds to $q^2 = \langle P_T \rangle^2$ and the upper edge to M_W^2 . The shaded (striped) band is the indicated mixture of TOP 170 (190). The width of the bands represents the uncertainty in the q^2 scale for VECBOS plus any indicated uncertainty in the percentage of $t\bar{t}$ events.

The overview plot for data set I is about what we expected while data sets II and IV indicate a higher top mass, while data set III indicates a lower top mass. High data points may also be an indication that the percentage of top has been underestimated because of the correlation between the energy and the percentage of $t\bar{t}$ in the Monte Carlo predictions. The general impression for all four data sets is much as expected for this given variable. In Fig. 6 we show the overview plot for $E_T(3) + E_t(4)$. Remember this is the variable with best discriminatory power. The behavior for data set I, II and III seems much as expected. Data set IV favors a lower value of top mass. The overview plot of H for data set III seems much as expected. Data set I, II and IV all favor a higher top mass than 170. Data set II especially seems to have fluctuated to higher values of top mass. The overview plot for the $P_T(\text{electron})$ shows no deviation from VECBOS. This as expected is exactly what we see for all four data sets. The aplanarity is a variable that does not depend

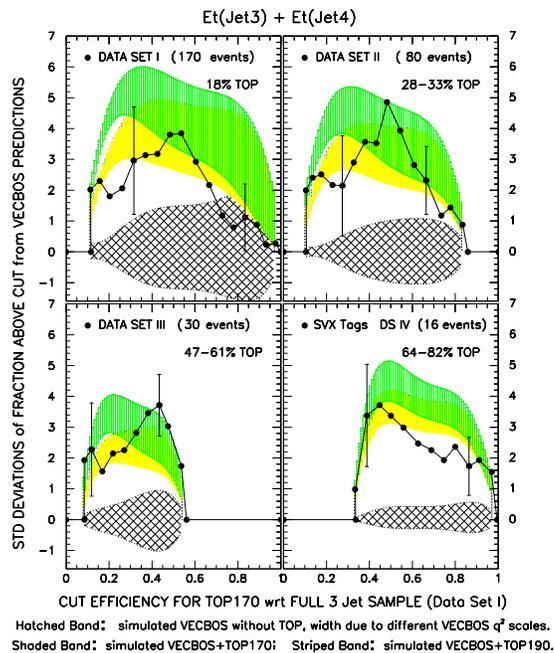


Figure 6: Overview plot of the variable $E_T(3) + E_T(4)$. Points are the data; bands are Monte Carlo predictions.

on the energy thus the curves for TOP 170 and TOP 190 are essentially identical. Data set III agrees with our expectations. Data set I, II and IV show slightly larger deviations from VECBOS than expected. For this and all angular plots the upper edge corresponds to $q^2 = \langle P_T \rangle^2$. The fact that the deviations are on the high side is what we would expect from the H analysis presented in section 2.

In this section we have compared four data sets and five variables to Monte Carlo predictions. We have looked² at a much larger set of variables that include $E_T(2)$, $P_T(W)$, η_{MAX} , $\cos(\Theta^*)_{MAX}$, and circularity. Our conclusion is that the data is well fit by the mixture of VECBOS and TOP 170 indicated by the SVX tagging rate. This mixture has varied from about 20% TOP 170 for data set I to 75% for data set IV. There is a slight tendency of the data to fit a slightly higher value of the top mass (or a larger percentage of top).

4 Kinematics using mass fitting

Now that the existence of the top quark has been established^{3 7}, other properties of the $t\bar{t}$ system need to be investigated. To proceed further we must fit the events to the $t\bar{t}$ hypothesis. We use the measured energy and angle of each of the four leading jets to infer the 4-momentum of the primary partons³. The constraints of the mass fit improve the resolution of the kinematic quantities. The situation is complicated because of gluon radiation. It is further complicated in that the assignments may

not be correct (clearly b-tagging helps, and tagging both b 's helps more). Perhaps the most interesting quantity is the mass of the $t\bar{t}$ system. This is shown in Fig. 7 for the same data set used in the H analysis (section 2) and very similar to data set II used in the directly measured kinematics analysis (section 3).

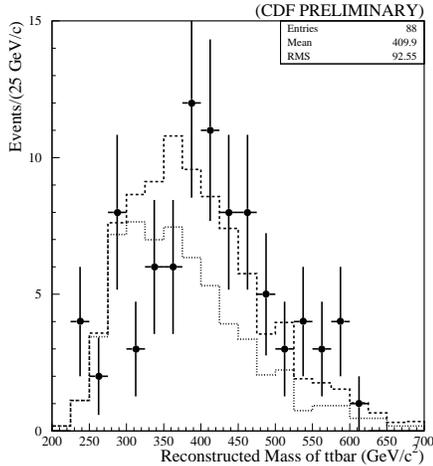


Figure 7: Invariant mass distribution of the $t\bar{t}$ for pretagged events (points). The background (VECBOS) is shown as a dotted curve. A fit to the data using a mixture of VECBOS and TOP 175 is shown as a dashed curve.

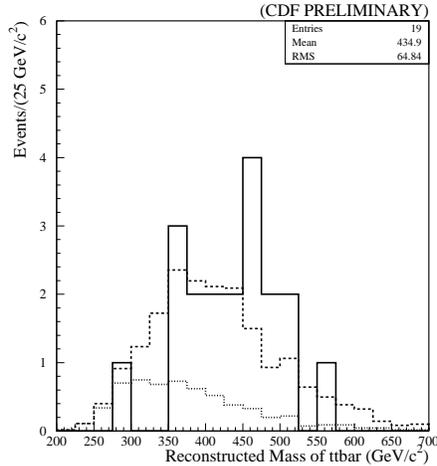


Figure 8: Invariant mass distribution of the $t\bar{t}$ for b-tagged events (solid curve). The background (VECBOS) is shown as a dotted curve. A fit to the data using a mixture of VECBOS and TOP 175 is shown as a dashed curve.

The distribution is only for events that have a good mass fit ($\chi^2 < 10$). The data consists of 88 events. The background (QCD W + jets) is constrained to 61.4 events. The data is well fit by the mixture of TOP 175 and background (VECBOS). We now show the $t\bar{t}$ mass distribution for the b-tagged sample (Fig. 8). The data consists of 19 events (both SVX and SLT tags have been used). The background (QCD W + jets) is constrained to 6.2 events. Again the data is well fit by the mixture of of TOP 175 and background(VECBOS).

We will soon present data on a 100 pb^{-1} sample. At that time we plan to present data including the $P_T(\text{top})$ and $\eta(\text{top})$.

5 Search for resonances decaying into $t\bar{t}$.

Physics beyond the standard model could appear as structure in the $t\bar{t}$ mass distribution. One such non standard model has been proposed by C.T. Hill⁸. In this model a technicolor Z' could decay to $t\bar{t}$ or $b\bar{b}$. We will use this model for purposes of illustration. In section 4 we improved the kinematics by requiring a mass fit. In this section we further improve the kinematics by constraining the mass of the top to our measured value of $176 \text{ GeV}/c^2$.

If the Z' existed then there would be three components to the $t\bar{t}$ mass distri-

bution. The first would be the standard model top (Herwig TOP 175), the second would be the QCD W +jets background (VECBOS) and the third would be new physics (Z'). For purposes of illustration we have chosen Z' 's of mass 400, 500 and 600 GeV/c^2 (Fig. 9). Note that as the Z' mass gets heavier the corresponding yield

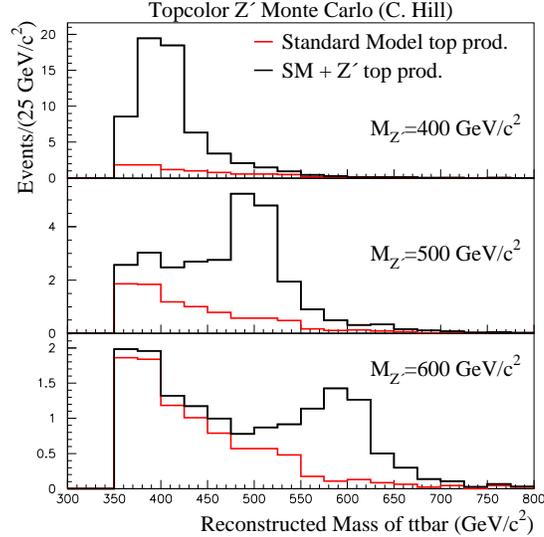


Figure 9: Predicted $t\bar{t}$ mass distributions reconstructed with top mass constraint. Standard model and Z' top production are shown ($175/c^2 \text{ GeV top}$).

of events gets smaller. The contribution of the QCD W + jet background is not included in the figure. However, the presence of the QCD W + jets background should not stop us from observing a Z' signal.

The mass constrained fit for the b-tagged data sample is shown in Fig. 10. The b-tagged sample consists of 19 events, but when the top mass constraint is added ($\chi^2 < 10$), 2 events were lost. The black curve shows the b-tagged events. The dotted curve shows the standard model prediction (Herwig TOP 175 and VECBOS). The excess of the data over the standard model would be a signal for new physics. The data seem to be consistent with the standard model, but clearly adding more data will be interesting. The present run should yield over 100 pb^{-1} . Following this run the CDF detector will be upgraded. This upgrade will allow us to use the higher luminosity provided by a new main injector. It is hoped that a luminosity of $10^{32} \text{ cm}^{-2} \text{ sec}^{-1}$ can be achieved for our next run (II). The goal for run II is to have an integrated luminosity of 2000 pb^{-1} .

Acknowledgements

I would like to thank the organizers of this conferences for their hospitality. I would also like to thank my collaborators for help with this paper, especially G. Bellettini, M. Binkley, J. Skarha, and G.P. Yeh. Thanks also go to R. Herber for help with PostScript.

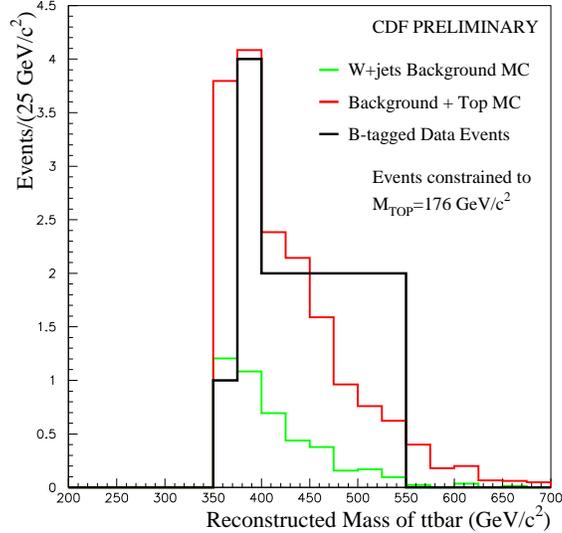


Figure 10: The \bar{t} mass with the top mass constrained to 176 GeV. The solid histogram is the 17 b-tagged events. The dotted curve is VECBOS and the dashed curve is a mixture of TOP 175 and VECBOS.

References

1. “Study of $t\bar{t}$ Production in $p\bar{p}$ Collisions Using Total Transverse Energy”, Fermilab-PUB-95/149E, submitted to Phys. Rev. Lett.
2. Morris Binkley, “ $t\bar{t}$ Kinematics in $W + \geq 3$ Jet Events”, 10th Topical Workshop on Proton-Antiproton Collider Physics, Fermilab May 1995, Fermilab-PUB-95/176E.
3. F. Abe *et al.*, Phys. Rev. Lett. **74**, 2626 (1995).
F. Abe *et al.*, Phys. Rev. Lett. **73**, 225 (1994).
F. Abe *et al.*, Phys. Rev. D **50**, 2966 (1994).
4. G. Marchesini and B.R. Webber, Nucl.Phys. **B310**, 461,(1988).
G. Marchesini *et al.*, Comput. Phys. Commun. **67**, 465 (1992).
5. F.A. Berends, W.T. Giele, H. Kuijf and B. Tausk, Nucl. Phys. **B357**, 32 (1991).
6. F. Abe *et al.*, Phys. Rev. D **51**, 4623 (1995).
“Kinematical Evidence for top pair production in $W +$ Multijet events in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV”, Fermilab-PUB-94/411E, submitted to Phys. Rev. Lett.
7. S. Abachi *et al.*, Phys. Rev. Lett. **74**, 2632 (1995).
8. Christopher T. Hill, Phys. Lett. **B345**, 483 (1995).