CDF/PUB/TOP/PUBLIC/8912

A Measurement of the $p\bar{p} \to t\bar{t} \to b\bar{b}\ell^+\nu\ell^-\nu$ Cross-Section in 1.0 fb^{-1} Using Lepton plus Isolated Track Selection with Identified b-Jets at $\sqrt{s} = 1.96$ TeV

The CDF Collaboration
URL http://www-cdf.fnal.gov
(Dated: July 23, 2007)

Using 1.0 fb^{-1} of CDF data from Run II of the Fermilab Tevatron, we complete the first measurement of the $t\bar{t}$ cross section in the dilepton channel using lepton plus isolated track selection and identified b-jets. A secondary vertex-finding algorithm is used to reconstruct heavy flavour decay vertices within jets. We perform the measurement using an estimate of the number of expected background events and the number of observed events with two leptons, significant missing transverse energy, and ≥ 2 jets with at least one jet identified as a b-jet. Use of an isolated track as the second lepton significantly increases the dilepton acceptance, at the price of some increase in background, particularly from W + jets events where one of the jets is identified as a lepton. The requirement that at least one of the jets is identified as a b-jet significantly reduces the background. The addition of the identified b-jet requirement improves the observed signal : background ratio from 1:1 for the before requiring identified b-jets to $\approx 5.5 : 1$ for the tagged measurement. Assuming a branching ratio of $BR(W \to \ell\nu) = 10.8$ % and a top mass of $m_t = 175$ GeV/c^2, the measured cross-section is $\sigma(p\bar{p} \to t\bar{t}) = 10.1 \pm 1.8(\text{stat.}) \pm 1.1(\text{syst.}) \pm 0.6(\text{lumi.})$ pb.

I. INTRODUCTION

This note describes a measurement of the $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV with the CDF detector at the Fermilab Tevatron. The most recent theoretical calculations of the cross-section have uncertainties of less than 15% [1, 2]. As CDF continues to collect data, more stringent tests of this prediction are becoming possible. Discrepancies between the predicted and observed cross-section could indicate new physics influencing or imitating the top signature. Many models for physics beyond the Standard Model predict discrepancies in the production or decay of the top quark [3, 4]. Even in the absence of indications for new physics, measurements of top quark properties are crucial to understanding the details of the Standard Model and provide perhaps the best testing ground currently available for the study of complex, high-p_T final states of the type expected for many new physics processes (particularly supersymmetry) at the Large Hadron Collider.

The dilepton channel, where both top quarks decay $t \to Wb \to \ell\nu b$, is of particular interest due to its high purity. This analysis takes advantage of that relative purity by using an isolated track for the second lepton to increase the acceptance. Purity is recovered by identifying at least one of the jets as having originated from a b-quark, called "b-tagging". The b-identification algorithm (SecVtx) [10] takes advantage of the long lifetime of bottom quarks to explicitly reconstruct decay vertices displaced from the primary interaction point. We use a high efficiency version of this algorithm at the expense of sample purity to increase the statistical power of the sample.

The results described here are the first to require b-jet identification with the lepton+isolated track selection in the dilepton channel. The CDF collaboration has published results of this analysis without imposing the b-jet identification in 200 pb^{-1}of data [5]. An update with improved background estimation methods is described in [6]. The analysis without the b-jet identification requirement is the "pretag" analysis, and shares selection and acceptance with the b-jet identified version. The pretag background estimates also form an input to the tagged analysis background estimation.

II. DATA SAMPLE & EVENT SELECTION

The top dilepton signature consists of two charged leptons, two b quarks which hadronize and are detected as jets, and two neutrinos which escape the detector without interacting. To isolate this signature, we select events from the data with two reconstructed leptons, two or more jets, and significant missing transverse energy ($\not\!\!E_T$) from the neutrinos. We use a higher efficiency configuration of the SecVtx tagger to identify at least one jet in the event as a b-jet, known as"b-tagging". Additional event selection constraining the $\not\!\!E_T$ in the event is made with the goal of reducing the background from Drell-Yan ($q\bar{q} \to Z/\gamma^\star \to \ell\ell$) events.

A. Data Sample

This analysis is based on an acceptance-weighted integrated luminosity of 1.0 fb^{-1} collected with the CDF detector [7]. The candidate event sample in data is based on those events identified online by the trigger system as having at least one central ($|\eta| \lesssim 1$) electron or muon candidate with $p_T > 18$ GeV/c or one forward ($1 \lesssim |\eta| \lesssim 2$) electron with $E_T > 20$ GeV in an event with $\not\!\!E_T > 15$ GeV [8]. The rest of the event selection takes place offline, after full event reconstruction.

B. Lepton Identification

The primary lepton candidate must be fully reconstructed and have an isolated energy signature in the calorimeter. To qualify, electrons must have a well-measured track pointing at an energy deposit in the calorimeter with $E_T >$ 20 GeV. The electromagnetic shower profile of the electron candidate must be consistent with expectations. Muon candidates are selected from well-measured tracks with $P_T > 20$ GeV/c which point at a track segment in the muon detectors which lie outside the rest of the detector. In the region of the calorimeter crossed by the track, the hadronic and electromagnetic energy deposited must be consistent with the signature of a minimum ionizing particle.

The second lepton candidate, the "track lepton", is a well-measured, isolated track with $P_T > 20$ GeV/c. A track is considered isolated if the total transverse momentum of all other tracks with $P_T > 0.5$ GeV/c and within a cone $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \leq 0.4$ around the candidate track is $\leq 10\%$ of the lepton candidate P_T. To ensure quality reconstruction of the lepton track, we require the impact parameter in the transverse plane to be within 250 μm of the beam axis. The track lepton criteria refers only to tracking information. By avoiding reference to other subdetectors, we maintain acceptance for leptons which have crossed uninstrumented regions of the detector. This also allows us to accept single prong hadronic as well as leptonic decays of the τ lepton from $W \to \tau\nu$, which includes 85% of τ decays.

The track lepton must have the opposite charge from the primary lepton and have a z_0 within 5 cm of the z_0 of the primary lepton.

C. Jet Reconstruction

We consider jets with $E_T > 20$ GeV in the range $|\eta| < 2$. The jets are defined using a fixed-cone clustering algorithm with a cone size of $R = 0.4$. The E_T values of the jets are corrected for the effects of jet fragmentation, calorimeter non-uniformities and calorimeter absolute energy scale [9]. Events with two or more jets are used to calculate the cross-section.

D. $\not\!\!E_T$ Corrections and Selection

Significant $\not\!\!E_T$ is one of the most distinctive features of the top dilepton signature, so all candidate events are required to have corrected $\not\!\!E_T > 25$ GeV. This rejects the majority of the background from the Drell-Yan ($Z/\gamma^\star \to \ell\ell$) process, which would otherwise overwhelm the top dilepton signature due to its much larger cross-section. To reduce the contribution from events where the $\not\!\!E_T$ is produced by mismeasurement of the energy of objects in the event, the $\not\!\!E_T$ is corrected for all known cases where the calorimeter information is not the best measure of an object's energy. The first correction is for all well-measured tracks pointing at a 3 by 3 block of calorimeter towers where the E_T measured is less than 70% of that track's p_T. This accounts for minimum-ionizing particles such as muons which do not deposit all of their energy in the calorimeter. Then we correct the $\not\!\!E_T$ for the jet energy calibrations, subtracting off the difference between the corrected and uncorrected jet energies.

However, Drell-Yan events with false $\not\!\!E_T$, which can occur when the p_T of the lepton candidate track and/or the E_T of a jet are mismeasured, can still pass the selection. To reduce the contamination from such events, we make several event selections in addition to the 25 GeV requirement. No physics object in the event should be pointing in the φ direction of the $\not\!\!E_T$. Specifically, we require that the primary lepton not point within 5° of the $\not\!\!E_T$, and that the track lepton not be within 5° of parallel or anti-parallel to the $\not\!\!E_T$. All jets in the event must be more than 25° away from the direction of the $\not\!\!E_T$, unless the event has $\not\!\!E_T > 50$ GeV. Also, if the dilepton invariant mass is within 15 GeV/c^2 of the Z boson resonance, the $\not\!\!E_T$ threshold is raised 40 GeV.

E. b-Jet Tagging

The algorithm used to identify b-quarks is called the SecVtx tagging algorithm. SecVtx works by vertexing the decay daughters of the b hadron. The algorithm was developed in RunI and is described in detail in [10], although we provide a short description here. SecVtx starts from the primary interaction vertex of each event. The primary vertex position resolution is $\approx$ 10-20 μm. SecVtx considers each jet in an event, examining those tracks in a jet which pass basic quality cuts. The algorithm then attempts to find a secondary vertex by vertexing tracks with a large impact parameter significance with respect to the primary vertex. We calculate the distance between primary and secondary vertices in the plane perpendicular to the beamline, and project the vector onto the jet axis to obtain L_{2D}, the two dimensional decay length along the jet axis. "Displaced" tracks have $|L_{2D}/\sigma_{L_{2D}}| > 3.0$. The sign of L_{2D} is taken from the dot product of the vector with the jet direction. If L_{2D} is positive, we consider this a positive tag, if $L_{2D} < 0$ the jet is negatively tagged. Long-lived quarks (b's and c's) predominantly have positive tags, while negative tags are primarily mistakes due to resolution effects and material tags, referred to as "mistags".

III. $t\bar{t}$ DILEPTON ACCEPTANCE

We determine geometric and kinematic acceptance for the $t\bar{t}$ signal before imposing the b-tagging requirement using PYTHIA Monte Carlo calculations [11] followed by a full simulation of the CDF detector. This version of PYTHIA uses the CTEQ5L parton distribution functions [12] to model the momentum distribution of the initial state partons and assumes a leptonic branching fraction for the W boson of 0.108.

The acceptance before the b-tagging requirement is defined as the number of events passing all selection, which are identified as dilepton events at generator level, divided by the *total* number of $t\bar{t}$ events in the sample. The total lepton + track acceptance, corrected for all known discrepancies between data and Monte Carlo, is 0.85 $\pm$ 0.03 %.

68% of the acceptance comes from events reconstructed with an electron and an isolated track, and the rest from events with a muon and isolated track. In terms of the generated leptons, the acceptance is 20% from ee events, 42%

Source	Uncertainty on acceptance
Primary/track lepton	1.1%/1.1%
Jet energy scale	1.3%
Initial/Final-state radiation	1.6%/.5%
Parton distribution functions	0.5%
Monte Carlo generator	1.5%

TABLE I: Summary table of systematic uncertainties on the pretag signal acceptance.

from $e\mu$, 19% from $\mu\mu$, and the remaining 19% from events with at least one τ. 13% of the total acceptance is from events where the lepton from the W is a τ which then decays hadronically, which is acceptance gained by using an isolated track for the second lepton rather than identifying it strictly as a muon or electron.

The dominant uncertainty on the signal acceptance comes from the dependence of the acceptance on the production of extra jets from QCD radiation. A complete list of the systematic uncertainties on the acceptance is shown in Table I.

IV. EVENT TAGGING EFFICIENCY

The event tagging efficiency is defined as the efficiency to tag at least one jet which passes the jet selection requirements. This efficiency is calculated separately from the geometric and kinematic acceptance to take advantage the work already done in the pretag lepton plus isolated track analysis. Calculation of the jet tagging efficiency as a separate quantity also makes validation and comparison with other similar measurements easier.

The event tagging efficiency is based on the observed fraction of events with at least one b-tagged jet in simulated top dilepton events. We correct for three factors, the data/Monte Carlo scale factor, the rate of tagging light quark jets in top dilepton events, and the efficiency to reconstruct a b hadron as a jet. The scale factor is determined in a heavy flavour enhanced independent sample, and is defined as the ratio of the efficiency in data, to that in Monte Carlo. The scale factor is 0.94 ± 0.06. We apply the scale factor to the efficiency measured in Monte Carlo to get the event tagging efficiency in data. We use Monte Carlo to find the efficiency for tagging light quark jets in top events to be $0.013\pm$ 0.001 per jet. We find the efficiency to reconstruct a jet matched in $\eta - \varphi$ space to a decaying b hadron as 0.9889 ± 0.0004 per jet. We correct the event tagging efficiency to account for these effects. The final event tagging efficiency is found to be 0.669 ± 0.037.

Uncertainty on the event tagging efficiency is 5.5% from the data/Monte Carlo scale factor. This uncertainty arises from our limited knowledge of the charm and light flavour contamination in our data samples used to derive the scale factor.

V. BACKGROUNDS

The backgrounds for the pretag sample is estimated using different methods than the tagged background. But, the tagged backgrounds are an input to the tagged background estimate. We outline here the pretag background estimation methods, details are found in [6].

A. Pretag Backgrounds

The dominant backgrounds in the pretag top dilepton sample all arise from the leptonic decay of one or more massive vector bosons. If enough jets are produced by extra energetic partons emerging from the interaction to substitute for the jets from the b quarks, these events can pass the selection if they have, or can mimic the existence of, two leptons and significant missing transverse energy ($\not\!\!E_T$). These sources include diboson (WW, WZ and ZZ) production, $W \to \ell\nu$ + jets where one of the jets is misidentified as a second lepton, and Drell-Yan ($Z/\gamma^\star \to \ell\ell$) events where $\not\!\!E_T$ is produced by the decay of τ leptons or the mismeasurement of objects in the event.

1. *Diboson and $Z/\gamma^\star \to \tau\tau$*

Background from diboson (WW,WZ,ZZ) and $Z/\gamma^\star \to \tau\tau$ events are predicted based on the acceptance as measured in simulation and normalized to the theoretical prediction for the cross-section. These processes are relatively difficult to isolate in data, so that estimates based on data are as yet unfeasible, but because there are real leptons and neutrinos in the final state, Monte Carlo calculations are expected to be reliable.

2. *W + jets with a Fake Lepton*

W events with extra jets from QCD radiation and events from the decay of top in the lepton+jets channel can pass the event selection if one of the jets fluctuates down to a single track and is misidentified as a second lepton. We estimate this background using data by calculating the rate at which jets produce such a "fake" lepton and applying that fake lepton rate to the jets in $W \to \ell\nu$ data. We calculate fake rates for the primary lepton types as well as the track leptons, although the fake rates for the fully-reconstructed leptons are over an order of magnitude smaller and the corresponding contribution to the signal sample is only 7% of the total.

3. *Drell-Yan ($Z/\gamma^\star \to ee/\mu\mu$)*

To estimate the Drell-Yan contribution from $Z/\gamma^\star \to ee$ and $Z/\gamma^\star \to \mu\mu$ events to the candidate sample, we use the data in combination with PYTHIA Monte Carlo.

Z boson candidates are selected in the data, using all of the criteria described above except for the Z veto, inside a mass range of 76 GeV/$c^2 < M <$ 106 GeV/c^2. The number of events with $\not\!\!E_T > 25$ GeV and with $\not\!\!E_T > 40$ GeV are counted. After subtraction of the expected non-Drell-Yan contribution from diboson, $t\bar{t}$, $Z/\gamma^\star \to \tau\tau$, and W events with a fake lepton, these two numbers are the normalization for the estimate.

A sample of simulated $Z/\gamma^\star \to ee/\mu\mu$ events is selected with the same criteria as the data. The ratio of the number of events with $\not\!\!E_T > 25$ GeV outside the mass window 76 GeV/$c^2 < M <$ 106 GeV/c^2 to the number inside the window, measured as a function of the number of jets in the event, is used to predict the number of events outside the mass window, given the number of observed events in data inside the mass window.

B. Tagged Backgrounds

We use a data driven background technique to estimate the tagged backgrounds. The background can be organized into two parts. The first is made up of signals with the same decay signature, such as $Z + b\bar{b}$ events, and mistags, events which are falsely tagged. The second is fakes, where a jet is identified falsely as an isolated track.

1. *Physics and Mistag Backgrounds*

Positive and negative tag rates of generic QCD jets are parameterized in five quantities, jet E_T, the number of tracks in the jet, jet η, the number of primary vertices in the event, and the total scalar sum of the E_T of all the jets. These parameterizations are termed "tag matrices". The positive tag matrix is used to estimate the number of tagged events in our dataset from physics backgrounds and mistags based on the jets in pretag events.

Many pretag events are actually top decays, and as such, should not contribute to the background estimation. Top content of the tag matrix background is estimated by finding the tag rate from the matrix for top Monte Carlo events and applying this tag rate to the number of predicted top events from the pretag analysis.

2. *Fake Lepton Background*

The pretag fake lepton background consists of two processes. The dominant is W events with extra jets from QCD production. These events have a tag rate equal to that of generic QCD jets, and are estimated well by the tag matrices. The second process contributing to the background is top decay in the lepton+jets channel. We consider only top events in the dipelton channel as part of our acceptance, and thus lepton+jets top decays are classified as

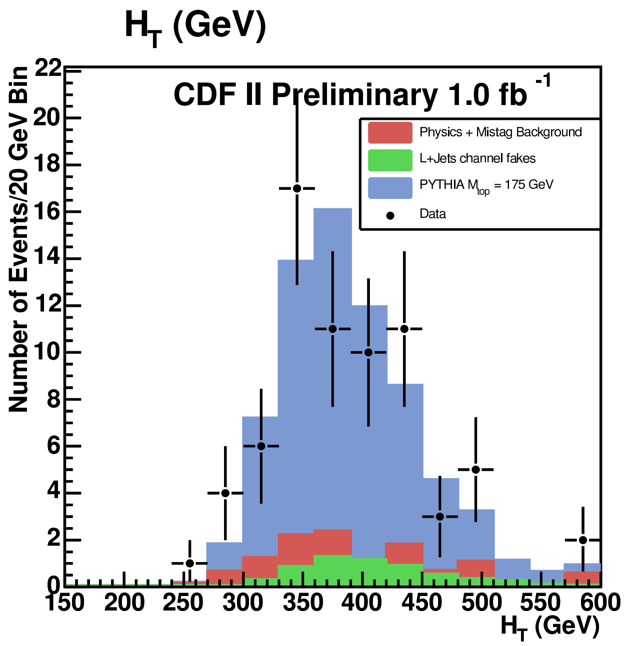

FIG. 1: Predicted and observed number of events as a function of H_T, the scalar sum of jet E_T, lepton E_T/P_T and $\not\!\!E_T$.

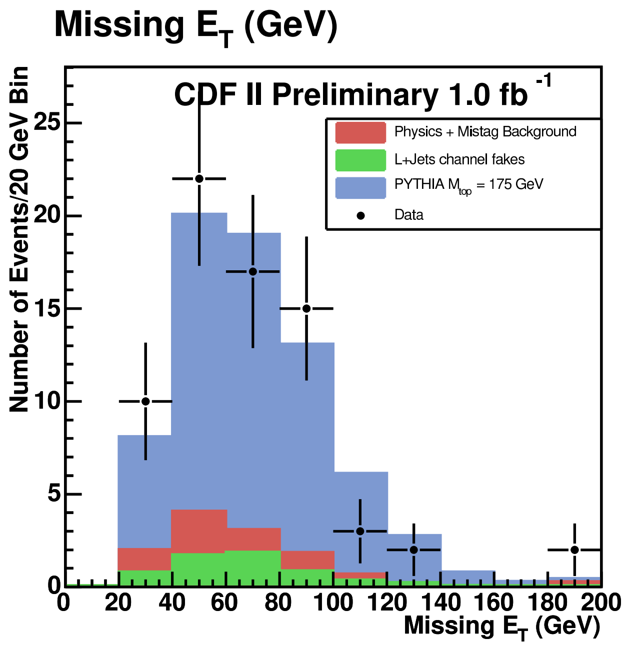

FIG. 2: $\not\!\!E_T$ for the $t\bar{t}$ signal and the backgrounds (≥ 2 jets).

backgrounds. Top lepton+jets decays have a higher tag rate than generic QCD events, and their tag rates need to be estimated separately from the tag matrix estimates.

First, we subtract the top lepton+jets contribution to the tag matrix background because the top decays have a higher tagging rate than the generic QCD jets used to build the matrices. This subtraction is done the same way as for dilepton top events: we estimate the tagging rate for top events in the matrix by applying the matrix tag rates to top lepton+jets decays in top Monte Carlo.

The top lepton+jets fake lepton background is estimated using both data and Monte Carlo. We use the number of pretag fakes from top as the overall normalization. Using data we obtain the fraction of events passing W+jets selection which are tagged. We use data and Monte Carlo to estimate the fraction of tagged W+jets events which are top lepton+jets decays. These two fractions are multiplied by the number of pretag fake leptons from top lepton+jets decays to arrive at the total number of tagged top lepton+jets decays which pass the lepton+isolated track selection.

C. Background Uncertainties

Uncertainty on the background consists of uncertainty on the inputs from the pretag backgrounds, and the uncertainty on tagging predictions made by the tag rate matrices. Tag rate predictions from the matrices are subject to an 8% uncertainty. Uncertainties from the pretag background estimates are added in quadrature with the 8% from the tag rate matrix. We find a total uncertainty on the background prediction to be 34%.

VI. RESULTS

The cross section is calculated using

$$\sigma = \frac{N_{\rm obs} - N_{\rm bkgnd}}{A \cdot \epsilon \cdot \int \mathscr{L} dt}$$

The number of observed events is 69, the background is calculated to be 11.9 $\pm$ 4.1 events, the acceptance times efficiency is 0.85 $\pm$ 0.03%, the event tagging efficiensy is 0.669 $\pm$ 0.037 and the total acceptance-weighted integrated luminosity is 1000 $\pm$ 60 pb^{-1}. Assuming $m_t = 175$ GeV/c^2 and $BR(W \to \ell\nu) = 10.8$ %,

$$\sigma_{t\bar{t}} = 10.1 \pm 1.8(\text{stat.}) \pm 1.1(\text{sys.}) \pm 0.6(\text{lum.})\text{pb} \ ,$$

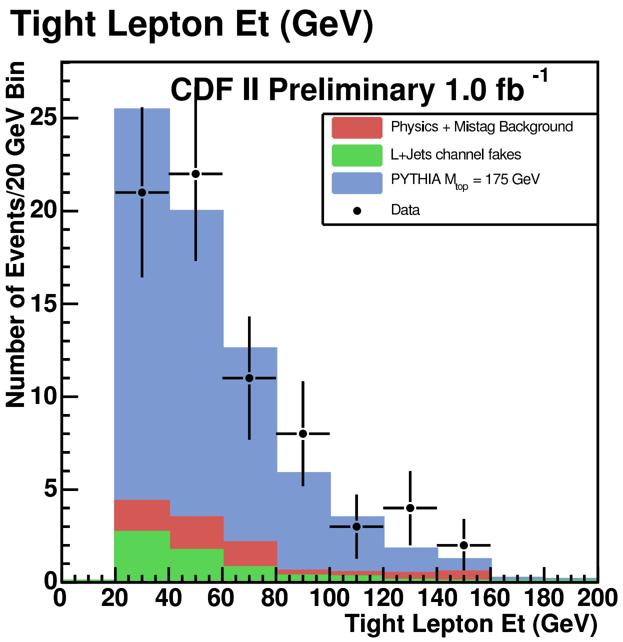

FIG. 3: Tight lepton E_t for the $t\bar{t}$ signal and the backgrounds ($\geq$ 2 jets).

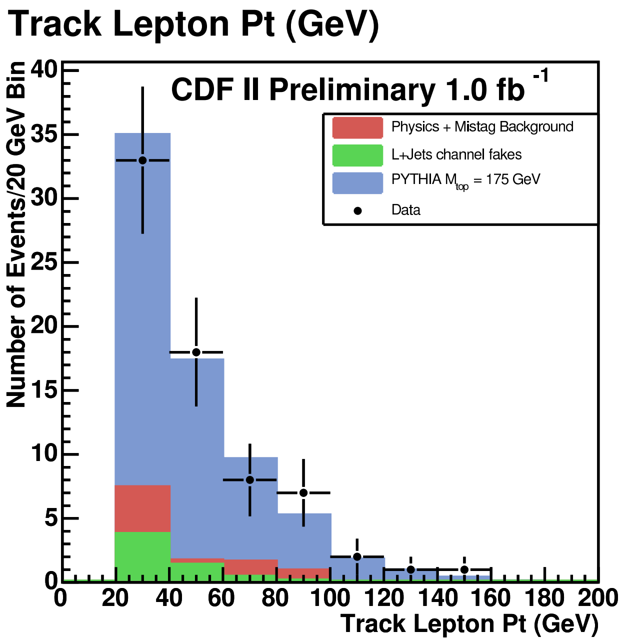

FIG. 4: Track lepton p_T for the $t\bar{t}$ signal and the backgrounds ($\geq$ 2 jets).

compared to the Standard Model prediction of $6.7^{+0.7}_{-0.9}$ pb [1, 2].

VII. KINEMATIC FEATURES OF CANDIDATE EVENTS

Figures 1-5 compare the various kinematic quantities such as lepton transverse momentum, dilepton invariant dilepton mass, and jet quantities of the candidate events in the data to the expectation based on the Standard Model $t\bar{t}$ signal and background. Note that the measured cross-section is used to normalize the $t\bar{t}$ distribution, so that the integral number of observed events is equal to the number predicted, and the comparison is between shapes only. We find good agreement between the distribution in the data and the Standard Model expectation.

Acknowledgments

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto Nazionale DP Fiske Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the Bundesministerium für Bildung und Forschung, Germany; the Korean Science and Engineering Foundation and the Korean Research Foundation; the Particle Physics and Astronomy Research Council and the Royal Society, UK; the Institut National de Physique Nucleaire et Physique des Particules/CNRS; the Russian Foundation for Basic Research; the Comisión Interministerial de Ciencia y Tecnología, Spain; the European Community's Human Potential Programme; the Slovak R&D Agency; and the Academy of Finland.

[1] M. Cacciari, S. Frixione, G. Ridolfi, M. Mangano, and P. Nason, J. High Energy Phys. 04 (2004) 68.
[2] N. Kidonakis and R. Vogt, Phys. Rev. D **68**, 114014 (2003).
[3] C. T. Hill and S. J. Parke, Phys. Rev. D **49**, 4454 (1994).
[4] H. P. Nilles, Phys. Rep. **110**, 1 (1984); H. E. Haber and G. L. Kane, *ibid.* **117**, 75 (1985).

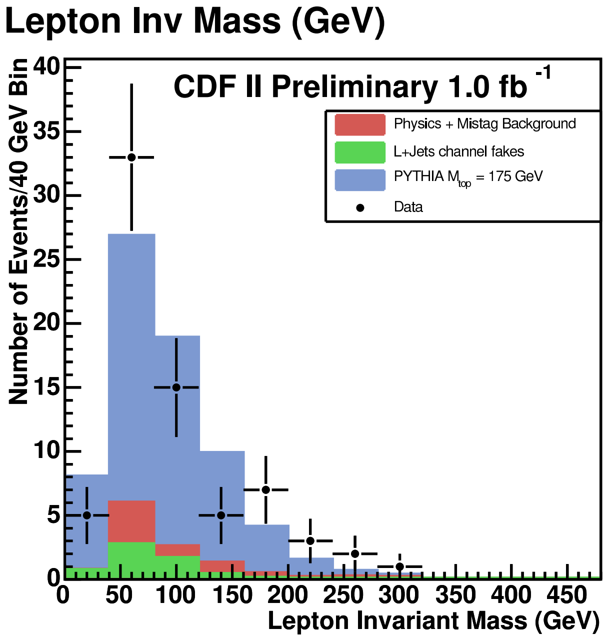

FIG. 5: Lepton invariant mass for the $t\bar{t}$ signal and the backgrounds (≥ 2 jets).

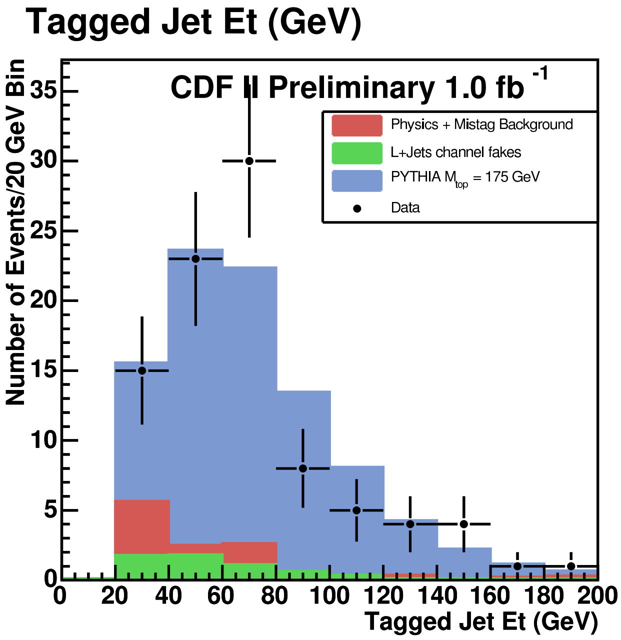

FIG. 6: Tagged jet E_t for the $t\bar{t}$ signal and the backgrounds.

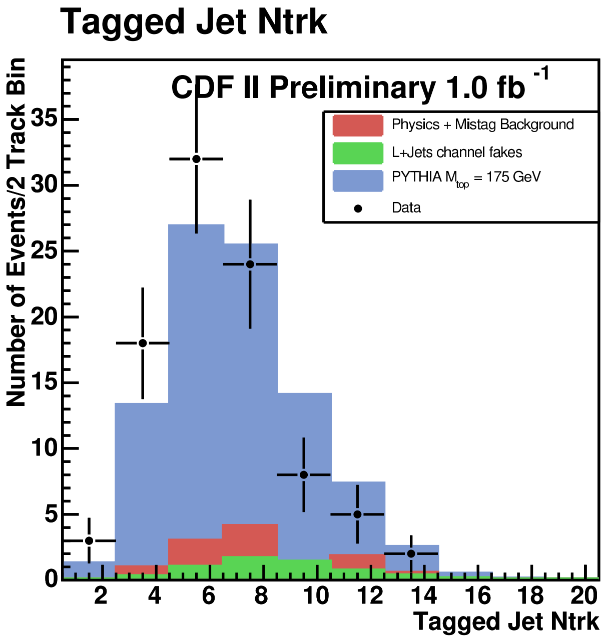

FIG. 7: Number of tracks in tagged jets for the $t\bar{t}$ signal and the backgrounds.

[5] CDF Collaboration, D. Acosta*et al*, Phys. Rev. Lett. **93**, 142001 (2004).
[6] http://www-cdf.fnal.gov/physics/new/top/confNotes/cdf8770_ltrk_1invfb_pub_v2.ps
[7] CDF Collaboration, FERMILAB-PUB-96/390-E.
[8] We use a cylindrical coordinate system where θ is the polar angle to the proton beam, ϕ is the azimuthal angle about the beam axis, and pseudorapidity is defined $\eta = -\ln\tan(\theta/2)$. We define $E_T = E\sin\theta$ and $p_T = p\sin\theta$ where E is an energy measured in the calorimeter and p is a momentum measured by the spectrometer. Missing transverse energy, $\not\!\!E_T$, is $-|\sum_i E^i_T\mathbf{n}_i|$, where $\mathbf{n}_i$ is the unit vector in the azimuthal plane that points from the beamline to the ith calorimeter tower.

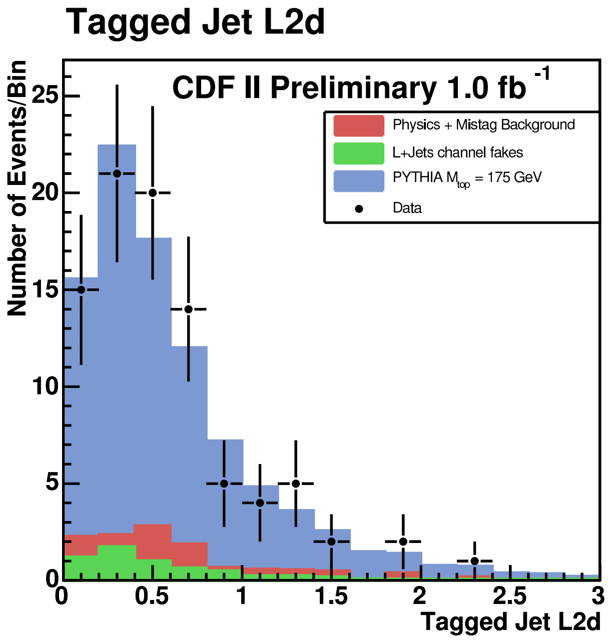

FIG. 8: Tagged jet 2-D decay length for the $t\bar{t}$ signal and the backgrounds.

[9] A. Bhatti et al, *Determination of the Jet Energy Scale at the Collider Detector at Fermilab*, 2005

[10] CDF Collaboration, T. Affolder*et al*, Phys. Rev. D **64**, 032002 (2001).

[11] T. Sjostrand et al., High-Energy-Physics Event Generation with PYTHIA 6.1, Comput. Phys. Commun. **135**, 238 (2001).

[12] H. L. Lai *et al*, Eur. Phys. J. **C12**, 375 (200).

[13] J. M. Campbell and R. K. Ellis, Phys. Rev. D**60**, 113006 (1999).

[14] P. Sutton, A. Martin, R. Roberts, and W. Stirling, Phys. Rev. D **45**, 2349 (1992); P. Rijken and W. van Neerven, Phys. Rev. D **51**, 44 (1995); R. Hamberg, W. van Neerven and T. Matsuura, Nucl. Phys. **B359**, 343, (1991); R. Harlander and W. Kilgore, Phys. Rev. Lett. **88**, 201801 (2002); W. van Neerven and E. Zijlstra, Nucl. Phys. **382**, 11, (1992).