Measurement of the $t\bar{t}$ Production Cross Section in $pp$ Collisions at $\sqrt{s}=1.96$ TeV using Soft Electron $b$-Tagging

We present a measurement of the top quark pair production cross section in $p\bar{p}$ collisions at $\sqrt{s}=1.96$ TeV using a data sample corresponding to 1.7 fb$^{-1}$ of integrated luminosity collected with the Collider Detector at Fermilab. We reconstruct $t\bar{t}$ events in the lepton+jets channel, consisting of $e\nu$+jets and $\mu\nu$+jets final states. The dominant background is the production of $W$ bosons in association with multiple jets. To suppress this background, we identify electrons from the semileptonic decay of heavy-flavor jets ('soft electron tags'). From a sample of 2196 candidate events, we obtain 120 tagged events with a background expectation of 51$\pm$3 events, corresponding to a cross section of $\sigma_{t\bar{t}} = 7.8 \pm 2.4$ (stat)$\pm 1.6$ (syst)$\pm 0.5$ (lumi) pb. We assume a top-quark mass of 175 GeV/$c^2$. This is the first measurement of the $t\bar{t}$ cross section with soft electron tags in Run II of the Tevatron.
I. INTRODUCTION

The top quark is the most massive fundamental particle observed to date, and has been studied by the CDF and D0 collaborations since its discovery in 1995 \[1\]. The \(t\bar{t}\) production cross section has been measured in each of the three canonical final states: \(q\bar{q}′b\bar{q}′\bar{b}\) \[2\], \(q\bar{q}′b\ell\bar{b}\) \[3, 4\], and \(\ell\bar{b}\ell\bar{b}\) \[5\] \((\ell = e, \mu,\) and \(q = u, d, c, s)\). In these measurements, different combinations of b-quark identification (‘tagging’) and kinematic information \[3\] have been used to suppress backgrounds. Tagging of b-quarks has been accomplished by identifying the long lifetime of the hadron with secondary vertex reconstruction or with displaced tracks \[4\] or through soft muons from semileptonic decay \[5\]. Along with measurements of the top-quark mass \[7\] and many other properties of the top quark, a consistent picture of the top quark as the third generation standard model (SM) isospin partner of the bottom quark emerges.

The Fermilab Tevatron produces top quarks, typically in pairs, by colliding \(p\bar{p}\) at \(\sqrt{s} = 1.96\) TeV. The \(t\bar{t}\) production cross section calculated at next-to-leading order is \(6.7 \pm 0.8\) pb \[3\] assuming \(m_t = 175\) GeV/c\(^2\), where the uncertainty is dominated by the choice of renormalization and factorization scales. At the Tevatron, approximately 85% of \(t\bar{t}\) production is via quark-antiquark annihilation and 15% is via gluon-gluon fusion. The measurement of the production cross section is important first as a test of perturbative QCD, but also as a platform from which the production cross section is calculated at next-to-leading order.

The measurement of the three canonical final states: \(tq\bar{q}\) and \(b\ell\nu\) has been accomplished by identifying the long lifetime of the hadron with secondary vertex reconstruction or with displaced tracks \[4\] or through soft muons from semileptonic decay \[5\]. Along with measurements of the top-quark mass \[7\] and many other properties of the top quark, a consistent picture of the top quark as the third generation standard model (SM) isospin partner of the bottom quark emerges.

The Fermilab Tevatron produces top quarks, typically in pairs, by colliding \(p\bar{p}\) at \(\sqrt{s} = 1.96\) TeV. The \(t\bar{t}\) production cross section calculated at next-to-leading order is \(6.7 \pm 0.8\) pb \[3\] assuming \(m_t = 175\) GeV/c\(^2\), where the uncertainty is dominated by the choice of renormalization and factorization scales. At the Tevatron, approximately 85% of \(t\bar{t}\) production is via quark-antiquark annihilation and 15% is via gluon-gluon fusion. The measurement of the production cross section is important first as a test of perturbative QCD, but also as a platform from which to study other top-quark properties. Moreover, measuring the \(t\bar{t}\) cross section in its various final states is an important consistency test of the SM and might highlight contributions to a particular decay channel from new physics.

In this paper, we present a measurement of the \(t\bar{t}\) production cross section in the lepton plus \(\geq 3\) jets final state. The dominant background in this channel is the production of a \(W\) boson associated with several jets. To suppress this background, we use a soft electron tagger (SLT\(_e\)) to identify the semileptonic decay of heavy-flavor (HF). Heavy-flavor refers to the product of the fragmentation of a bottom or charm quark.

Soft electron tagging is a challenging method of identifying b-jets because the semileptonic branching fraction (BF) is approximately \(20\% - \text{BF}(b \rightarrow e\nu X)\) and \(\text{BF}(b \rightarrow c \rightarrow \nu X)\) each contribute approximately \(10\%\) and because electron identification is complicated by the presence of a surrounding jet. The algorithm is able to distinguish electromagnetic showers from hadronic showers by using a shower-maximum detector embedded in the electromagnetic calorimeter. This detector has a high enough resolution that it can determine the transverse shape and position of electron showers and yet be unaffected by nearby activity. Additionally, \(\gamma \rightarrow e^+e^-\) conversions due to material interactions provide a significant background, which we suppress using a combination of geometric and kinematic requirements. Nevertheless, the soft electron technique is interesting because it is complementary to other b-tagging techniques and because it is a useful technique for other analyses.

This is the first measurement of the \(t\bar{t}\) cross section with soft electron tags in Run II of the Tevatron. A previous measurement at \(\sqrt{s} = 1.8\) TeV combined secondary vertex tagging, soft muon, and soft electron tagging \[9\].

We organize this paper as follows: Sec. III describes aspects of the CDF detector salient to this analysis. Section III describes the implementation of the SLT\(_e\). We discuss the SLT\(_e\) tagging efficiency in \(t\bar{t}\) events in Sec. IV. Section V describes the calculation of the background to tagged electrons in HF jets, including conversion electrons and hadrons. In Sec. VI we tune the SLT\(_e\) tagger in a \(b\bar{b}\) control sample. This ensures the tagger’s validity in high-momentum b-jets, such as those found in \(t\bar{t}\) events. Section VII reports the cross section measurement, including the event selection and signal and background estimation. Finally, in Sec. VIII we present our results and conclusions.

II. THE CDF DETECTOR

CDF II is a multi-purpose, azimuthally and forward-backward symmetric detector designed to study \(p\bar{p}\) collisions at the Tevatron. An illustration of the detector is shown in Fig. I. We use a cylindrical coordinate system where \(z\) points along the proton direction, \(\phi\) is the azimuthal angle about the beam axis, and \(\theta\) is the po-
lar angle to the proton beam direction. We define the pseudorapidity \( \eta \equiv - \ln \tan(\theta/2) \).

The tracking system consists of silicon microstrip detectors and an open-cell drift chamber immersed in a 1.4 T solenoidal magnetic field. The silicon microstrip detectors provide precise charged particle tracking in the radial range from \( 1.5 - 28 \) cm. The silicon detectors are divided into three different subcomponents, comprised of eight total layers. Layer00 (L00) \([10]\) is a single-sided silicon detector mounted directly on the beampipe. The silicon vertex detector (SVXII) \([11]\) consists of five double-sided sensors with radial range up to 10.6 cm. The intermediate silicon layer (ISL) \([12]\) is composed of two layers of double-sided silicon, extending coverage up to \( |\eta| < 2.0 \).

The drift chamber, referred to as the central outer tracker (COT) \([13]\), consists of 96 layers of sense wires grouped in 8 alternating superlayers of axial and stereo wires, covering a radial range from 40 to 140 cm. The reconstructed trajectories of COT tracks are extrapolated into the silicon detectors, and the track is refit using the additional hits in the silicon detectors. In combination, the COT and silicon detectors provide excellent tracking up to \( |\eta| \leq 1.1 \). The transverse momentum \( (p_T) \) resolution, \( \sigma(p_T)/p_T \), is approximately \( 0.07\%p_T \ (\text{GeV}/c)^{-1} \) when hits from the SVXII and ISL are included.

Beyond the solenoid lie the electromagnetic and hadronic calorimeters, with coverage up to \( |\eta| \leq 3.6 \). The calorimeters have a projective geometry with a segmentation of \( \Delta \eta \approx 0.1 \) and \( \Delta \phi \approx 15^\circ \) in the central \( (|\eta| \leq 1.1) \) region. The central electromagnetic calorimeter (CEM) \([14]\) consists of \( > 18 \) radiation lengths \( (X_0) \) of lead-scintillator sandwich and contains wire and strip chambers embedded at the expected shower maximum \( (\sim 6X_0) \). The wire and strip chambers are collectively referred to as the central shower-maximum (CES) chambers and provide measurements of the transverse electromagnetic shower shape along the \( r - \phi \) and \( z \) directions with a resolution of 1 and 2 mm, respectively. The central hadronic calorimeter (CHA) \([15]\) consists of \( \sim 4.7 \) interaction lengths of alternating lead-scintillator layers at normal incidence. Measured in units of GeV, the CEM has an energy resolution, \( \sigma(E)/E = 13.5\%/\sqrt{E} \sin(\theta) + 2\% \), and the CHA has an energy resolution, \( \sigma(E)/E = 50\%/\sqrt{E} \).

Muon chambers \([16]\) consist of layers of drift tubes surrounding the calorimeter. The central muon detector (CMU) is cylindrical and covers a pseudorapidity range \( |\eta| < 0.63 \). The central muon upgrade (CMP) is a box-shaped set of drift chambers located beyond the CMU and separated by more than three interaction lengths of steel. Muons which produce hits in both the CMU and CMP are called CMUP. The central muon extension (CMX) extends the muon coverage up to \( |\eta| < 1 \).

Gaseous Cherenkov luminosity counters (CLC) \([17]\) provide the luminosity measurement with a \( \pm 6\% \) relative uncertainty.

CDF uses a three-level trigger system to select events to be recorded to tape. The first two levels perform a limited set of reconstruction with dedicated hardware, and the third level is a software trigger performing speed-optimized event reconstruction algorithms. The triggers used in this analysis include electron, muon, and jet triggers at different transverse energy thresholds. The electron triggers require the coincidence of a track with an electromagnetic cluster in the central calorimeter. The muon triggers require a track that points to hits in the muon chambers. The jet triggers require calorimeter clusters with uncorrected \( E_T \) above a specified threshold.

### III. SOFT ELECTRON TAGGING

The SLT\(_e\) algorithm uses the COT and silicon trackers, central calorimeter and, in particular, the central shower maximum chambers to identify electrons embedded in jets from semileptonic decays of HF quarks. The tagging algorithm is ‘track-based’ - as opposed to ‘jet-based’ - in that we consider every track in the event that meets certain criteria as a candidate for tagging. Such tracks are required to be well-measured by the COT and extrapolated to the CES. This requirement forces the track to have \( |\eta| \) less than 1.2. We require that the track \( p_T \) is greater than 2 GeV/c. We consider only tracks that originate close to the primary vertex: \( |d_0| < 0.3 \) cm, \( |z_0| < 60 \) cm, and \( |z_0 - z_{\text{vtx}}| < 5 \) cm, where \( d_0 \) is the impact parameter, that is the distance of closest approach in the transverse plane, with respect to the beamline. The \( z \) position of the track at closest approach to the beamline is \( z_0 \), and \( z_{\text{vtx}} \) is the reconstituted \( z \) position of the primary vertex. Tracks must also pass a jet-matching requirement, which is that they are within \( \Delta R \equiv \sqrt{\Delta \eta^2 + \Delta \phi^2} < 0.4 \) from the axis of a jet with transverse energy \( E_T \) greater than 20 GeV. Jets are clustered with a fixed-cone algorithm with a cone of size \( \Delta R \leq 0.4 \). Jet energies are corrected for detector response, multiple interactions, and un-instrumented regions of the detector \([18]\). Finally, tracks must also pass a conversion filter described in Sec. V A. Although we have not explicitly required tracks to have silicon hits, the conversion filter insists that tracks with a high number of ‘missing’ silicon hits must be discarded. We consider tracks which meet all of the above criteria as SLT\(_e\) candidates.

Candidate tracks are passed through the SLT\(_e\) algorithm which uses information from both the calorimeter and CES detectors. The algorithm is designed to identify low-\( p_T \) electrons \([19]\) embedded in high-\( E_T \) jets while still maintaining a high identification efficiency for high-\( p_T \) electrons. This is particularly important for tagging \( t\bar{t} \) events, although the SLT\(_e\) algorithm is not specific to this final state. Figure 2 shows the \( p_T \) shape of candidate SLT\(_e\) electrons in the CDF II detector from a bottom quark, charm quark, and photon conversions in PYTHIA \([20]\) \( t\bar{t} \) Monte Carlo (MC) simulated events. Even in \( t\bar{t} \) events, the electron spectrum from \( b\)-jets peaks at low \( p_T \) but extends more than a decade in scale. Electrons from
FIG. 1: Illustration of the CDF II detector.

FIG. 2: Transverse momentum distribution of candidate SLT_e tracks in jets in Pythia t\bar{t} MC simulated events. Distributions for electrons from a bottom quark, charm quark, and photon conversions are normalized to unity to emphasize the relative difference in the shapes.

charm decay in t\bar{t} events are principally due to cascade decays, but some direct charm production occurs through the hadronic decay of the W boson.

The SLT_e candidate tracks are extrapolated to the front face of the calorimeter to seed an electromagnetic cluster in the CEM. The two calorimeter towers adjacent in \eta-space closest to the extrapolated point are used in the cluster. A candidate SLT_e must have an electromagnetic shower that satisfies 0.6 < E_{EM}/p < 2.5 and E_{Had}/E_{EM} < 0.2, where E_{EM} and E_{Had} are the total electromagnetic and hadronic energies in the cluster, respectively, and p is the momentum of the electron track. The E_{EM}/p requirement selects electromagnetic showers which have approximately the same energy as the track (as expected from electrons), while the E_{Had}/E_{EM} requirement suppresses late-developing (typically hadronic) showers. These requirements were tuned in simulated t\bar{t} events and are looser than for typical high-E_T electrons because the presence of photons and hadrons from the nearby jet distorts the energy deposition. Figure 3 shows the calorimeter variables for candidate SLT_e tracks in Pythia tt simulation.

Next, the SLT_e algorithm uses the track extrapolation to seed a wire cluster and strip cluster in the CES. We limit the number of strips and wires in the clusters to seven each in order to minimize the effects of the sur-
rounding environment. At least two wires (strips) with energy above a 80 (120) MeV threshold must be present, or the track is not tagged. This requirement suppresses low-$p_T$ hadrons that have a late developing shower in the CEM. Two discriminant quantities determined from the CES are used to distinguish electrons from hadrons. One is a $\chi^2$ comparison between the transverse shower profile of the SLT$_e$ candidate and the profile measured with test-beam electrons. The other is the distance $\Delta$, measured in cm, between the extrapolated track and the position of the cluster energy centroid. Each type of discriminant is determined for the wire and strip chambers separately.

We construct a likelihood-ratio discriminant by using the $\chi^2$ and $\Delta$ distributions from pure samples of electrons and hadrons as templates. The electron sample is selected by triggering on a $E_T > 8$ GeV electron from a photon conversion ($\gamma \rightarrow e^+ e^-$) and using the partner electron. For this sample, the conversion filter require-
Tagging Efficiency

The distributions for the CES wire chamber and strip chamber discriminants from each sample are shown in Fig. [2]. The relative difference in shapes between the wire and strip distributions is due to the different energy thresholds used, and the slightly different resolution due to the differing technology.

The likelihood-ratio is formed by binning the electron and hadron templates in a normalized 4-dimensional histogram to preserve the correlations between the four variables, $\chi^2_{\text{wire}}$, $\Delta_{\text{wire}}$, $\chi^2_{\text{strip}}$, and $\Delta_{\text{strip}}$, creating probability distribution functions for both signal and background. We use them to derive a likelihood ratio according to the formula:

$$L \equiv \frac{S_i}{S_i + B_i}$$

where $S_i$ and $B_i$ are the value of the probability distribution functions in the $i^{\text{th}}$ bin of signal and background templates, respectively. We tag a candidate track if $L > 0.55$. Two other operating points ($> 0.65$ and $> 0.75$) were also studied for this analysis, but the former point was found to give the best expected combined statistical and systematic uncertainty on the $t\bar{t}$ cross section. Table I summarizes the requirements for a candidate SLT$_e$ track to be tagged.

We measure the tagging efficiency – that is, the number of tracks that are tagged divided by the number of all candidate tracks – with the combination of calorimeter, wire/strip, and $L$ requirements in various samples. Figure 5 shows this tagging efficiency for the electron sample ($\sim 60\%$ at $L > 0.55$) and the hadron sample ($\sim 1.1\%$ at $L > 0.55$) as a function of the likelihood-ratio requirement. Note that because the hadron sample has not been corrected for the small contamination by electrons, the hadron tagging efficiency should only be considered an upper bound. This correction is discussed later in Sec. [VIB]. Also note that value of the likelihood-ratio does not extend to 1.0. This is an artifact of the four variables chosen for the likelihood. Hadrons occupy the entire phase-space of possible values for $\chi^2_{\text{wire/strip}}$ and $\Delta_{\text{wire/strip}}$, so that the background probability distribution function is never zero.

### IV. SLT$_e$ Tagging Efficiency in Jets

An important feature of the SLT$_e$ algorithm is the tagging efficiency dependence on the environment. In the previous section we described the per-track tagging efficiency for a sample of isolated conversion electrons where each leg is incident on a different calorimeter tower. However, the tagging efficiency for electrons from semileptonic $b$ decay with the same kinematic characteristics as the conversion electrons is markedly lower. This is due to the nearby jet which distorts the electromagnetic shower detected in the calorimeter. In general, the calorimeter variables, $E_{\text{EM}}/p$ and $E_{\text{Had}}/E_{\text{EM}}$, are strongly affected by the jet, whereas the CES variables - that is the $\chi^2$ and $\Delta$ variables as well as the number of wires and strips in the CES cluster - have a much weaker dependence.

For the SLT$_e$ algorithm, we introduce the isolation variable, $I_{\text{SLT}}$, defined as the scalar sum of the $p_T$ of tracks which point to the calorimeter cluster divided by the candidate track $p_T$: $\Sigma_{\text{clus}}p_T/p_T$. This variable is useful at quantifying the degree to which the local environment should affect the electron’s electromagnetic shower, and hence the identification variables. An isolated SLT$_e$ track has $I_{\text{SLT}}$ identically equal to 1.0, whereas for a non-isolated track, $I_{\text{SLT}} > 1.0$.

In order to measure the SLT$_e$ tagging efficiency of soft electrons in jets, we rely on a combination of MC simula-
tion and data-driven techniques. We study the calorimeter and the CES discriminants, which both enter the SLT\textsubscript{e} algorithm, separately. Although the calorimeter variables have a strong dependence on the local environment, they are well-modeled in the MC simulation. However, the CES variables, on the whole, are poorly modeled in the simulation due to the presence of early overlapping hadronic showers.

We study the modeling of the SLT\textsubscript{e} calorimeter-based discriminants in a sample of conversion electrons reconstructed in jets. This sample is constructed by identifying an electron and its conversion partner while both are close to a jet ($\Delta R \leq 0.4$). We select such conversions in data triggered on a 50 GeV jet and a kinematically comparable dijet MC simulation sample. We use the missing silicon layer variable, described in Sec. \ref{sec:conv}, to enhance the conversion electron content in the sample. This is done by requiring that the track associated with the conversion partner is expected to have, but does not have, hits in at least three silicon layers. The conversion partner is used as a probe to compare the efficiency of the combined calorimeter requirements in both data and simulated samples as a function of $p_T$ and $I_{\text{SLT}}$. We see very good agreement in the general trend between both samples, as shown in Fig. \ref{fig:conv_electron} from which we derive a 2.5\% relative systematic uncertainty (integrating over all bins) to cover the difference between data and simulation. The comparison between kinematically and environmentally similar samples is important to validate the behavior of the simulation modeling.

To account for the mis-modeling of the CES-based discriminants, we measure the tagging efficiency of candidate SLT\textsubscript{e} tracks directly in data and apply it to candidate SLT\textsubscript{e} tracks in the simulation that have already passed the calorimeter-based requirements. The efficiency is parameterized as a three-dimensional matrix in $p_T$, $\eta$, and $I_{\text{SLT}}$ to account for the correlations between the three variables. This matrix is constructed out of the pure conversion electron sample used to create the likelihood-ratio templates. The validity of the tag matrix is then verified in a sample of electrons from $Z$ boson decays and in a $b\bar{b}$ sample, as described in Sec. \ref{sec:conv}. A 3\% relative systematic uncertainty — derived from the agreement within the conversion sample and with the $Z \rightarrow e^+e^-$ sample — is applied to the tag-matrix prediction.

Applying the matrix as a weight on each candidate SLT\textsubscript{e} track identified in the simulated events, we find that the tagging efficiency for electrons from HF jets in $t\bar{t}$ events is approximately 40\% per electron track (See Sec. \ref{sec:conv}). This is calculated by identifying candidate SLT\textsubscript{e} tracks in $t\bar{t}$ events matched to electrons from HF jets in the simulation. For those electrons which pass the calorimeter requirements, the tag matrix determines the expected tagging probability.

V. SLT\textsubscript{e} Tagging Backgrounds

The two principal backgrounds to SLT\textsubscript{e} tagging are real electrons from photon conversions and misidentified electrons from charged hadrons (e.g. $\pi$, $K$, $p$). Although the tagging probability is very low for hadrons, the high multiplicity of such tracks makes their contribution non-negligible. Conversion electrons are much more abundant than electrons from HF jets. In $t\bar{t}$ events, three times as many candidate tracks are due to conversion electrons than to electrons from semileptonic decay of HF. Their removal is essential to effective $b$-tagging. Additionally, there is a small contribution from Dalitz decays of $\pi^0$, $\eta$, and $J/\psi$. In this section we discuss the estimation of the conversion electron and hadronic backgrounds.

A. Conversions

The primary procedure for conversion electron rejection relies on identifying the partner leg. We identify an SLT\textsubscript{e} tagged track as a conversion if, when combined with another nearby track in the event, the pair has the geometric characteristics of a photon conversion. In particular, the $\Delta \cot(\theta)$ between the tracks as well as the distance between the tracks when they are parallel in the $r-\phi$ plane must be small. However, for low-$p_T$ conversion electrons in jets, this requirement fails to identify the partner leg more than 40\% of the time. The primary reason for this is that the track reconstruction algorithms begin to fail at very low $p_T \sim 500$ MeV/c. The asymmetric energy sharing between conversion legs exacerbates this effect.

To recover conversion electrons when the partner leg is not found, we use the fact that conversions are produced through interactions in the material. We extrapolate the candidate track’s helix through the silicon detectors and identify silicon detector channels where no hit is found. If a track is missing hits on each side of more than three double-sided silicon layers, then it is identified as a conversion (at most six missing layers is possible \cite{21}). Figure \ref{fig:conv_angle} shows the reconstructed radius of conversion, $R_{\text{conv}}$, versus the number of missing silicon layers for conversion electrons with both legs tagged by the SLT\textsubscript{e} in an inclusive sample of $E_T > 8$ GeV electrons. Although high $R_{\text{conv}}$ values are suppressed because of the impact parameter requirement, there is a clear correlation between missing silicon layers and the $R_{\text{conv}}$. For SLT\textsubscript{e} candidates, we combine both the standard partner-track-finding algorithm with the missing silicon layer algorithm so that if a tag fails either, we reject it as a conversion electron.

We measure the conversion ID efficiency in data by decomposing the algorithm into a partner-track-finding component and a missing silicon layer component. We use the missing silicon layer templates to measure the partner-track-finding component efficiency, and we use a sample of conversions with both legs SLT\textsubscript{e} tagged to mea-
FIG. 6: Efficiency of the calorimeter requirements on an untagged conversion electron as a function of the track $p_T$, for both isolated (a) and non-isolated (b) tracks. Error bars reflect statistical uncertainties from both data and MC. We use the overall agreement to derive a 2.5% relative systematic uncertainty on the calorimeter requirements of the SLT_e tagger.

FIG. 7: Number of missing silicon layers versus the reconstructed radius of conversion for conversion electrons found in an inclusive ($E_T > 8$ GeV) electron sample. Tracks tagged by the SLT_e algorithm are rejected as conversions if they have more than three missing silicon layers.

sure the missing silicon layer component efficiency. We combine the efficiencies, accounting for their correlation.

We use an in situ process of building templates for the missing silicon layer variable for conversions and prompt tracks directly within the sample of interest to fit for the total conversion content before and after rejection. The in situ nature of the template construction is important because conversion identification depends strongly on kinematics and geometry that can vary across different samples. The conversion template is constructed from conversions where both legs are tagged by the SLT_e, and the prompt track template is constructed from tracks where the SLT_e requirements have been inverted, resulting in a nearly 100% pure hadronic sample.

A fit for the conversion component of SLT_e tags in events triggered on a 20 GeV jet is shown in Fig. 8. In the fit, only those electrons with hits in six expected silicon layers is considered. Those tracks with fewer than six expected layers are used as a consistency check, and a systematic uncertainty is assigned to the geometric bias incurred from this requirement. The dearth of tracks with four or five missing layers is an artifact of the CDF track reconstruction algorithm, which requires that at least three silicon hits must be added to any track or none will be added. The goodness-of-fit is limited by systematic biases in the template construction which contribute the dominant systematic uncertainties to the efficiency measurement. Such biases include correlations between track finding and missing silicon layers, modeling of prompt electrons (including HF decay) by prompt hadrons, geometric dependencies, and sample contamination.

We find that the conversion identification efficiency is overestimated in MC simulation relative to data. We characterize the difference by a multiplicative scale factor (SF), defined as the ratio of efficiencies measured in data and simulation. Because the conversion identification efficiency depends strongly on the underlying photon energy spectrum, it is important for the SF measurement to compare energetically similar samples. Therefore, we measure the SF in events triggered by a jet with $E_T > 20$, 50, 70, and 100 GeV and compare to MC simulated dijet
Efficiency/Scale Factor
Tags/Missing Layer

FIG. 8: Fit for the conversion and prompt component of SLT\(_e\) tags before conversion removal in events triggered on a \(E_T > 20\) GeV jet. The goodness-of-fit is limited by systematic biases in the template construction, and is accounted for in the final SF measurement.

FIG. 9: Conversion identification scale factor measurement in events triggered by a \(E_T > 20\) GeV jet. Shown also is the efficiency measured in data and the efficiency measured in MC simulation. The consistency across \(p_T\), among other variables, demonstrates the validity of the SF approach.

events which pass the same requirements. We measure a conversion efficiency SF of \(0.93 \pm 0.01\) (stat) \(\pm 0.02\) (syst). The dominant uncertainties are systematic effects related to the accuracy of the template models. We find that the SF behaves consistently as a constant correction across a variety of different event and track variables in multiple datasets. Figure 9 shows the SF as a function of track \(p_T\) in a sample of events triggered by a \(E_T > 20\) GeV jet. The gray band shows the value of the SF with statistical and systematic uncertainties for combined SF across jet 20, 50, 70, and 100 datasets.

We also measure a conversion ‘misidentification efficiency’ — defined as the efficiency to misidentify a non-conversion track as a conversion — multiplicative SF of \(1.0 \pm 0.3\) between data and simulation. This is done by measuring the efficiency to identify prompt tracks as conversions. The large systematic uncertainty accounts for the variation found across different kinematic variables, jet triggers, and particle types (such as the difference between a HF electron and a pion from \(K_s\) decay). In \(t\bar{t}\) events, the complete algorithm is approximately 70% efficient at rejecting candidate SLT\(_e\) tracks that are conversions. Only 7% of non-conversions are misidentified as conversions. Since the misidentification efficiency is an order of magnitude lower than the efficiency, the total contribution of systematic uncertainties from each is comparable.

B. Hadrons

We measure the tagging efficiency of hadrons in MC by defining a three-dimensional fake matrix out of tracks in jet-triggered events. The matrix parameterizes the probability that the CES discriminants (\(\mathcal{L}\) plus the number of wires and strips) can tag a hadron. We remove jets where a large fraction of energy is deposited by a single track, in order to reduce the contamination of hard electrons that are also reconstructed as jets. We find that the use of the track \(p_T\), \(\eta\), and \(I_{SLT}\) are sufficient to describe the dependence of the tagging efficiency on other variables as well. This is demonstrated in Fig. 10 which shows the measured and predicted tags in events triggered on a jet with \(E_T > 100\) GeV as a function of the \(E_T\) of the jet closest in \(\Delta R\) to the SLT\(_e\) track. We also cross-check the fake-matrix prediction in a distinct sample of tracks in jets triggered on a high-\(E_T\) photon. We find that the agreement is good within \(\pm 5\%\).
Before tagging, the tracks in the jet samples are almost purely hadronic; however, after tagging, we must correct for the electron contamination when we estimate the total efficiency to tag a hadronic track. Three classes of electrons are present in the sample: conversion electrons, HF electrons, and other sources (primarily Dalitz decay of π0). The conversion electron contamination is estimated by measuring the efficiency and misidentification efficiency of the conversion filter in the jet samples. Using this information in combination with the number of tracks before and after conversion removal determines the remaining conversion content. The HF electron contamination is estimated using correlations between the SLT_e tags and b-tags from a secondary vertex algorithm, SecVtx \cite{22}. We use SecVtx to enhance the HF content of the jet sample. Using MC simulation to estimate the expected size of this enhancement, we can extrapolate back to the original, pre-SecVtx tag HF component. The remaining contribution of electrons from other sources is small and estimated with the MC simulation.

We find that 35 ± 3% of the tags in the jet-triggered sample are electrons. This estimate is verified by measuring the SLT_e tagging efficiency for charged pions from K_s decay. By subtracting the electron contamination from the fake-matrix prediction, we find that, on average, 0.5% of hadronic tracks in t\bar{t} events produce a fake SLT_e tag.

VI. TUNING IN b\bar{b} SAMPLE

As a validation of the measured efficiency of the tagger, we measure the jet tagging efficiency in a highly enriched sample of b\bar{b} events. Events are selected through an 8 GeV electron or muon trigger, and we require that both the jet close to the lepton (∆R ≤ 0.4) and the recoiling (away) jet have a SecVtx tag. We measure the per-jet efficiency to find at least one SLT_e tag in the away jet. This efficiency is measured to be 4.4 ± 0.1 (stat) (%) in simulation and 4.3 ± 0.1 (stat) (%) in data.

The efficiency is calculated in simulation by taking all of the candidate tracks in the jet that pass the calorimeter requirements and using either the tag matrix for electrons or the fake matrix for hadrons to determine a tagging probability. If a track is identified as a conversion, then the tagging probability is rescaled according to the conversion efficiency or mis-identification efficiency SF. We tune the tag matrix with a multiplicative factor of 0.98 ± 0.03 to get the simulation to agree with data, where the systematic uncertainty is assigned to cover a jet-\(E_T\) dependence in the difference. The difference in the prediction and measurement is due to isolation effects of the jet environment not already accounted for by the \(J_{SLT}\) parameterization, specifically the presence of neutral hadrons. Figure 11 shows the predicted and measured tags in the combined 8 GeV electron and muon trigger samples as a function of the \(p_T\) of the SLT_e tag after the tuning. Statistical uncertainties from the data and simulation are added in quadrature and shown on the data points.

By combining the tag matrix, fake matrix, conversion identification and mis-identification efficiency SFs, and the correction for the jet environment, we estimate the tagging efficiency of data from simulation. Figure 12 shows the efficiency to tag a HF electron and a hadron in simulated t\bar{t} events as a function of the track \(p_T\) and the jet \(E_T\). While the tagging efficiency for electrons is steady as a function of the track \(p_T\), it decreases as a function of the jet \(E_T\) because of the decreasing isolation at high-\(E_T\).

VII. CROSS SECTION MEASUREMENT

The t\bar{t} production cross section is determined with the equation

\[
\sigma = \frac{N - B}{\epsilon_{t\bar{t}} A_{t\bar{t}} \int Ldt}
\]

where \(N\) is the number of tagged events, \(B\) is the expected background, \(\epsilon_{t\bar{t}}\) and \(A_{t\bar{t}}\) are the signal efficiency and acceptance, respectively and \(\int Ldt\) is the integrated luminosity. In this section, we describe the measurement of each of these quantities.

A. Event Selection and Expectation

We select t\bar{t} events in the lepton+jets decay channel through an inclusive lepton trigger which requires an electron (muon) with \(E_T > 18\) GeV (\(p_T > 18\) GeV/c).
After triggering, we further require that events contain an isolated electron (muon) with $E_T > 20 \text{ GeV}$ ($p_T > 20 \text{ GeV/c}$) in the central region ($|\eta| < 1.1$). We refer to this lepton as the primary lepton, to distinguish it from the soft lepton tag. The isolation of the primary lepton is defined as the transverse energy in the calorimeter surrounding the lepton in a cone of $\Delta R \leq 0.4$ but not including the lepton $E_T$ itself—divided by the electron (muon) $E_T$ ($p_T$). The lepton is considered isolated if the isolation is less than 0.1. Note that this isolation definition is different than the isolation variable, $I_{\text{SLT}}$, which is used with the SLT$_e$ algorithm.

We reject cosmic ray muons, conversion electrons, and $Z$ bosons. Only one primary lepton is allowed to be reconstructed in the lepton+jets sample, and the flavor of that lepton must be consistent with the trigger path. More details regarding this event selection can be found in Ref. [5]. An inclusive $W$ boson sample is constructed by requiring high missing transverse energy, $E_T > 30 \text{ GeV}$. We suppress background events by requiring $H_T > 250 \text{ GeV}$ when three or more jets are present. We define $H_T$ as the scalar sum of the transverse energy of the primary lepton, jets, and $E_T$.

In total, using events collected from February 2002 through March 2007 corresponding to an integrated luminosity of $\int L \, dt = 1.7 \pm 0.1 \text{ fb}^{-1}$, we find 2196 ‘pretag’ events with $\geq 3$ jets after the event selection described. We apply the SLT$_e$ algorithm to this sample and find 120 ‘tag’ events with $\geq 3$ jets with at least one SLT$_e$, of which five have two SLT$_e$ tags. Out of 120 events, 48 have a SecVtx tag present, in agreement with the expected 45 such double tags.

We use PYTHIA MC simulation with $m_t = 175 \text{ GeV/c}^2$ to simulate top-quark pair-production. By default, all MC simulated samples are generated with the CTEQ5L [23] parton distribution functions (PDF), and the program EvtGen [24] is used to decay the particle species. We measure $A_{E_{T\text{slt}}}$ by counting the number of events that pass the lepton+jets event selection described above divided by the total number of events generated. We do not restrict the decay channel at the generator level, so it is possible for some signal from other decay channels [25] to be reconstructed and categorized as lepton+jets. We then correct the acceptance with various scale factors to account for differences between simulation modeling and data. These scale factors result from differences in modeling of the lepton identification and isolation components, as well as corrections for requirements imposed on data but not the simulation, including the trigger efficiency, the position of the primary vertex along $z$, and the quality of the lepton track. The total acceptance for $tt$ events after corrections is 6.2%, comparable with the acceptance of other analyses in this final state [3,4,5]. A breakdown of the corrected accep-

<table>
<thead>
<tr>
<th>Lepton</th>
<th>1 jet</th>
<th>2 jets</th>
<th>3 jets</th>
<th>4 jets</th>
<th>$\geq$ 5 jets</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEM</td>
<td>$0.163 \pm 0.003$</td>
<td>$0.862 \pm 0.011$</td>
<td>$1.403 \pm 0.017$</td>
<td>$1.493 \pm 0.018$</td>
<td>$0.519 \pm 0.007$</td>
</tr>
<tr>
<td>CMUP</td>
<td>$0.089 \pm 0.002$</td>
<td>$0.477 \pm 0.009$</td>
<td>$0.788 \pm 0.015$</td>
<td>$0.826 \pm 0.015$</td>
<td>$0.284 \pm 0.006$</td>
</tr>
<tr>
<td>CMX</td>
<td>$0.042 \pm 0.001$</td>
<td>$0.220 \pm 0.005$</td>
<td>$0.353 \pm 0.008$</td>
<td>$0.381 \pm 0.008$</td>
<td>$0.130 \pm 0.003$</td>
</tr>
<tr>
<td>Total</td>
<td>$0.295 \pm 0.005$</td>
<td>$1.559 \pm 0.024$</td>
<td>$2.543 \pm 0.039$</td>
<td>$2.700 \pm 0.041$</td>
<td>$0.932 \pm 0.015$</td>
</tr>
</tbody>
</table>

TABLE II: Corrected $\bar{t}t$ acceptance in the lepton+jets decay channel. We have required $H_T > 250 \text{ GeV}$ for events with $\geq 3$ jets and $E_T > 30 \text{ GeV}$. Combined statistical and systematic uncertainties are shown.
tance by jet multiplicity and W lepton type is shown in Table II. Scaling the acceptance by the $t\bar{t}$ production cross section (assumed here to be 6.7 pb) and integrated luminosity yields a total pre-tag event expectation of 716.7 ± 44.4 events, where the dominant uncertainties result from the uncertainty on the luminosity and the acceptance corrections.

Finally, we measure the efficiency to find at least one SLT tag in events that pass the event selection by applying the calorimeter requirements, tag matrix, fake matrix, and conversion efficiency scale factors to candidate tracks. Assuming $\sigma_{t\bar{t}} = 6.7$ pb, and $\int L dt = 1.7$ fb$^{-1}$, we expect 59.2 ± 5.0 events after tagging in the $\geq 3$ jet region. This corresponds to a per-event tagging efficiency of $\epsilon_{t\bar{t}} = 8.3\%$.

B. Background Estimation and Sample Composition

We consider three categories of background in the identification of $t\bar{t}$ events. The first category, whose contribution is derived from MC simulation, includes the production of $WW$, $WZ$, $ZZ^*$ (where one $Z$ can be produced off-shell), single top quark production, $Z$ in association with jets, and Drell-Yan in association with jets. These backgrounds have a small uncertainty on the production cross section or contribute sufficiently little to the total background that a large uncertainty has little effect. For diboson production, we use PYTHIA generated samples scaled by their respective theoretical cross sections to estimate their contribution to the pretag and tag samples. The estimate for single top quark production uses a combination of MADEVENT [26] for generation and PYTHIA for showering, and is calculated separately for $s$- and $t$-channel processes again using the theoretical cross sections. $Z$+jets and Drell-Yan+jets use an ALPGEN [27] and PYTHIA combination, where ALPGEN is used for the generation and PYTHIA is used for the showering. The cross section is scaled to match the measured $Z$+jets cross section with an additional 1.2 ± 0.2 correction to match the measured jet multiplicity spectrum. Table II lists the cross sections used for each process.

The second category consists of background from multijet production, called QCD. We estimate the QCD contribution by releasing the $E_T$ requirement and fitting the total $E_T$ distribution to templates for the backgrounds and signal. To model the QCD $E_T$ spectrum, we use two samples: a PYTHIA $bb$ dijet sample, and a data sample with an $E_T > 20$ GeV electron candidate that fails at least two electron ID requirements. This sample is principally composed of multijet events with a similar topology to those that fake a high-$E_T$ electron. We fit for the fraction of QCD events in the sample by fixing the $t\bar{t}$ and MC simulation-driven background normalizations, and varying the $W$+jets and QCD template normalizations separately. The total QCD contribution has virtually no dependence on the assumed $t\bar{t}$ cross section.

We also include a 15% systematic uncertainty due to the real electron contamination in the electron-like sample. Table IV shows the measured fits for the fraction of pretag events with $E_T > 30$ GeV that are due to pretag and tag QCD events, $F_{\text{pre}}^{\text{QCD}}$ and $F_{\text{tag}}^{\text{QCD}}$, respectively. The result of the fit in the pretag region for $\geq 3$ tags is shown in Fig. 13.

![QCD Template Fit](image)

**FIG. 13: QCD fit for pretag events with $\geq 3$ jets.**

$W$+jet and QCD templates are allowed to float.

The third category and largest background is the production of $W$ bosons in association with multiple jets. We use a combination of simulation and data-driven techniques to measure this background. We use ALPGEN as the generator of the $W$+multijet datasets and PYTHIA for fragmentation and showering.

The $W$+jet normalization is determined by assuming that all pretag data events, not already accounted for by $t\bar{t}$ or by the first two background categories, must be $W$+jets. The tag estimate is derived from the pretag estimate by assuming that the tagging efficiency measured in MC simulation for separate HF categories is accurate and only the relative amount of HF needs adjustment. The equations below elucidate this procedure:

$$N_{\text{pre}}^{\text{tag}} = N_{\text{data}}^{\text{pre}} - N_{\text{MC}}^{\text{pre}} - N_{\text{QCD}}^{\text{pre}} - N_{\text{QCD}}^{\text{pre}}$$

(3)

$$N_{\text{W}+b\bar{b}}^{\text{tag}} = N_{\text{W}}^{\text{pre}}(\epsilon_2F_{2b} + \epsilon_1F_{1b})$$

(4)

$$N_{\text{W}+c\bar{c}}^{\text{tag}} = N_{\text{W}}^{\text{pre}}(\epsilon_2F_{2c} + \epsilon_1F_{1c})$$

(5)

$$N_{\text{W}+\text{LF}}^{\text{tag}} = N_{\text{W}}^{\text{pre}}\epsilon_3(1 - F_{2b} - F_{1b} - F_{2c} - F_{1c})$$

(6)

where, $N_{\text{tag}}$ and $N_{\text{pre}}$ are the number of tag and pretag events for various signal and background components. The tagging efficiencies, $\epsilon$, are measured in separate HF categories, where the subscript designates the number of reconstructed jets in an event identified as a $b$− or $c$− jet with information from the generator. For bookkeeping purposes, the presence of a $b$-jet supersedes the presence of a $c$-jet. The HF fractions, $F$, designate the fraction of $W$+jet events for each HF category.

While both the HF efficiencies and HF fractions are measured in MC simulation, the fractions are calibrated...
K
Q
tainty is dominated by the contribution from varying the K
by the agreement of the
SecVtx
rived from a data/MC comparison of multijet events with
by a single, multiplicative
K
-scale. All numbers are shown in units of %.

<table>
<thead>
<tr>
<th>Process</th>
<th>Cross Section×BF (pb)</th>
<th>Generator</th>
</tr>
</thead>
<tbody>
<tr>
<td>WW</td>
<td>12.4 ± 0.25 [28]</td>
<td>PYTHIA</td>
</tr>
<tr>
<td>WZ</td>
<td>3.96 ± 0.06 [28]</td>
<td>PYTHIA</td>
</tr>
<tr>
<td>ZZ*</td>
<td>2.12 ± 0.15 [28]</td>
<td>PYTHIA</td>
</tr>
<tr>
<td>single top (s-channel)</td>
<td>0.29 ± 0.02 [29]</td>
<td>MADEVENT+PYTHIA</td>
</tr>
<tr>
<td>single top (t-channel)</td>
<td>0.66 ± 0.03 [29]</td>
<td>MADEVENT+PYTHIA</td>
</tr>
<tr>
<td>Z+Jets</td>
<td>308 ± 51 [30]</td>
<td>ALPGEN+PYTHIA</td>
</tr>
<tr>
<td>Drell-Yan+Jets</td>
<td>2882 ± 480 [30]</td>
<td>ALPGEN+PYTHIA</td>
</tr>
</tbody>
</table>

TABLE III: Cross sections and generators used for the MC-simulation-derived backgrounds. The production of single top, Z+jets and Drell-Yan+jets is constrained to decay (semi-)leptonically at generator level. The cross sections for these processes are multiplied by the leptonic branching fraction. The decay of the diboson simulation, however, remains unconstrained and the full production cross section is quoted.

<table>
<thead>
<tr>
<th>Process</th>
<th>Cross Section×BF (pb)</th>
<th>Generator</th>
</tr>
</thead>
<tbody>
<tr>
<td>WW</td>
<td>12.4 ± 0.25 [28]</td>
<td>PYTHIA</td>
</tr>
<tr>
<td>WZ</td>
<td>3.96 ± 0.06 [28]</td>
<td>PYTHIA</td>
</tr>
<tr>
<td>ZZ*</td>
<td>2.12 ± 0.15 [28]</td>
<td>PYTHIA</td>
</tr>
<tr>
<td>single top (s-channel)</td>
<td>0.29 ± 0.02 [29]</td>
<td>MADEVENT+PYTHIA</td>
</tr>
<tr>
<td>single top (t-channel)</td>
<td>0.66 ± 0.03 [29]</td>
<td>MADEVENT+PYTHIA</td>
</tr>
<tr>
<td>Z+Jets</td>
<td>308 ± 51 [30]</td>
<td>ALPGEN+PYTHIA</td>
</tr>
<tr>
<td>Drell-Yan+Jets</td>
<td>2882 ± 480 [30]</td>
<td>ALPGEN+PYTHIA</td>
</tr>
</tbody>
</table>

TABLE IV: Summary of the fraction of the pretag and tag QCD events for different jet multiplicities.

<table>
<thead>
<tr>
<th>Fraction</th>
<th>1 Jet</th>
<th>2 Jets</th>
<th>≥ 3 Jets</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F^\text{QCD}_{\text{pt}}$ (%)</td>
<td>3.7 ± 6.0</td>
<td>4.6 ± 0.6</td>
<td>9.2 ± 1.5</td>
</tr>
<tr>
<td>$F^\text{QCD}_{\text{tag}}$ (%)</td>
<td>0.045 ± 0.011</td>
<td>0.10 ± 0.02</td>
<td>0.28 ± 0.14</td>
</tr>
</tbody>
</table>

TABLE V: Heavy-flavor fractions multiplied by the K-factor for W+jets events. Uncertainties are dominated by the agreement of the K-factor across jet bins and the $Q^2$ scale. All numbers are shown in units of %.

<table>
<thead>
<tr>
<th>Fraction</th>
<th>1 Jet</th>
<th>2 Jets</th>
<th>3 Jets</th>
<th>≥ 4 Jets</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_{1b}$</td>
<td>0.8 ± 0.3</td>
<td>1.6 ± 0.6</td>
<td>3.0 ± 1.1</td>
<td>3.7 ± 1.4</td>
</tr>
<tr>
<td>$F_{2b}$</td>
<td>—</td>
<td>1.0 ± 0.4</td>
<td>2.2 ± 0.8</td>
<td>3.5 ± 1.3</td>
</tr>
<tr>
<td>$F_{1c}$</td>
<td>5.8 ± 1.6</td>
<td>9.1 ± 2.6</td>
<td>10.2 ± 3.3</td>
<td>12.1 ± 3.9</td>
</tr>
<tr>
<td>$F_{2c}$</td>
<td>—</td>
<td>1.5 ± 0.6</td>
<td>3.4 ± 1.3</td>
<td>6.3 ± 2.3</td>
</tr>
</tbody>
</table>

TABLE VI: SLT_e tagging efficiency for different classes of HF in W+jet events. Uncertainties shown include all SLT_e tagging systematic uncertainties. All numbers are shown in units of %.

by a single, multiplicative K-factor, $K = 1.0 ± 0.4$, derived from a data/MC comparison of multijet events with HF enhanced by a SecVtx tag. The systematic uncertainty is dominated by the contribution from varying the $Q^2$ of the samples and the agreement of the K-factor across jet multiplicities. Phase-space overlap of jets simulated by ALPGEN and PYTHIA is accounted for by allowing ALPGEN to simulate those HF jets well-separated in $\eta-\phi$ space and allowing PYTHIA to simulate the rest [31]. Tables [IV] and [VI] show the measured values for the HF fractions and efficiencies, respectively.

C. Measurement and Uncertainties

Although the W+jets background depends explicitly on the assumed value of $\sigma_{t\bar{t}}$ (see Eq. [4]), we can solve algebraically for the cross section, resulting in a central value of $7.8 ± 2.4$ pb, where the statistical uncertainty is determined through error propagation and is verified with pseudo-experiments. The final sample composition is shown in Table [VII] and is shown graphically in Fig. [14]. This table shows the $t\bar{t}$ expectation for the measured cross section along with the background estimates corrected for the signal contribution. The observed number of pretag events and the expected number of pretag $t\bar{t}$ events is also presented.

The combined systematic uncertainties due to the luminosity, acceptance, background cross sections, SLT_e tagging, K-factor, and QCD fit are given in the table. Note that some of the background contributions—in particular, the W+jets components—are negatively correlated with each other, and this is reflected in the systematic uncertainties presented.

Figure [15] show the SLT_e tag $p_T$ distribution and the event $H_T$ distribution in the ≥ 3 jet region.

In the previous sections, we have described systematic uncertainties related to the SLT_e tagger and the background estimations. The tagger uncertainties derive from the calorimeter variable modeling, the tag- and fake-matrix predictions, the conversion (mis-)identification scale factors, and the jet environment correction from the b-jet tuning. The background uncertainties are derived from the theoretical or experimental production cross sections, the W+jets HF K-factor, the QCD fit, and the acceptance modeling. Here we discuss the uncertainties arising from the jet energy scale (JES) [15] and the
Events
120599
19695
1358
645
193


Pretag $t\bar{t}$ ($\sigma = 7.84$ pb)


<table>
<thead>
<tr>
<th>Process</th>
<th>1 jet</th>
<th>2 jets</th>
<th>3 jets</th>
<th>4 jets</th>
<th>$\geq$ 5 jets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretag</td>
<td>39.82 ± 2.11</td>
<td>211.2 ± 11.2</td>
<td>345.4 ± 18.3</td>
<td>366.6 ± 19.4</td>
<td>126.67 ± 6.71</td>
</tr>
<tr>
<td>$WW$</td>
<td>12.87 ± 1.27</td>
<td>12.36 ± 1.14</td>
<td>1.53 ± 0.14</td>
<td>0.64 ± 0.06</td>
<td>0.25 ± 0.02</td>
</tr>
<tr>
<td>$WZ$</td>
<td>1.37 ± 0.13</td>
<td>3.04 ± 0.26</td>
<td>0.41 ± 0.04</td>
<td>0.21 ± 0.02</td>
<td>0.06 ± 0.01</td>
</tr>
<tr>
<td>$ZZ$</td>
<td>0.16 ± 0.02</td>
<td>0.17 ± 0.02</td>
<td>0.05 ± 0.01</td>
<td>0.02 ± 0.00</td>
<td>0.01 ± 0.00</td>
</tr>
<tr>
<td>Single Top ($s$)</td>
<td>0.55 ± 0.06</td>
<td>2.31 ± 0.23</td>
<td>0.46 ± 0.05</td>
<td>0.17 ± 0.02</td>
<td>0.05 ± 0.01</td>
</tr>
<tr>
<td>Single Top ($t$)</td>
<td>1.88 ± 0.17</td>
<td>2.67 ± 0.25</td>
<td>0.36 ± 0.03</td>
<td>0.09 ± 0.01</td>
<td>0.01 ± 0.00</td>
</tr>
<tr>
<td>$Z$ +Jets</td>
<td>46.3 ± 10.1</td>
<td>19.52 ± 4.02</td>
<td>2.44 ± 0.44</td>
<td>1.09 ± 0.20</td>
<td>0.28 ± 0.05</td>
</tr>
<tr>
<td>Drell-Yan+Jets</td>
<td>10.01 ± 2.27</td>
<td>6.32 ± 1.42</td>
<td>1.11 ± 0.25</td>
<td>0.33 ± 0.07</td>
<td>0.09 ± 0.02</td>
</tr>
<tr>
<td>QCD</td>
<td>53.9 ± 14.1</td>
<td>20.20 ± 4.65</td>
<td>3.75 ± 1.92</td>
<td>1.78 ± 0.91</td>
<td>0.53 ± 0.27</td>
</tr>
<tr>
<td>$W+bb$</td>
<td>28.2 ± 10.9</td>
<td>22.74 ± 8.70</td>
<td>2.43 ± 0.94</td>
<td>1.04 ± 0.43</td>
<td>0.23 ± 0.10</td>
</tr>
<tr>
<td>$W+c\bar{c}$, $W+c$</td>
<td>104.2 ± 30.2</td>
<td>47.1 ± 14.6</td>
<td>3.80 ± 1.31</td>
<td>1.66 ± 0.62</td>
<td>0.36 ± 0.15</td>
</tr>
<tr>
<td>$W+$Light Flavor</td>
<td>960.7 ± 90.8</td>
<td>281.0 ± 22.9</td>
<td>18.56 ± 2.10</td>
<td>5.60 ± 1.14</td>
<td>1.22 ± 0.32</td>
</tr>
<tr>
<td>Total $W$+Jets</td>
<td>1093 ± 101</td>
<td>350.8 ± 24.0</td>
<td>24.78 ± 2.05</td>
<td>8.30 ± 1.38</td>
<td>1.81 ± 0.43</td>
</tr>
<tr>
<td>Backgrounds</td>
<td>1220.0 ± 94.8</td>
<td>417.4 ± 25.5</td>
<td>34.89 ± 2.36</td>
<td>12.64 ± 1.32</td>
<td>3.09 ± 0.41</td>
</tr>
<tr>
<td>$t\bar{t}$ ($\sigma = 7.84$ pb)</td>
<td>1.41 ± 0.10</td>
<td>13.25 ± 0.96</td>
<td>26.27 ± 1.94</td>
<td>30.70 ± 2.16</td>
<td>12.41 ± 0.86</td>
</tr>
<tr>
<td>Tags</td>
<td>1312</td>
<td>427</td>
<td>56</td>
<td>45</td>
<td>19</td>
</tr>
</tbody>
</table>

TABLE VII: Sample composition of lepton+jet events with $\geq 1$ SLT$_e$ tag corrected for the measured signal contribution. Uncertainties include effects from luminosity, acceptance corrections, cross section uncertainties, SLT$_e$ tagger modeling, $K$-factor, and the QCD fit.

FIG. 14: Jet multiplicity of SLT$_e$ tagged events in the lepton+jets dataset. The embedded plot is the $\geq 3$ jet subsample. Dashed lines represent the combined systematic uncertainties, while the data shows only the statistical uncertainty.

modeling of the $t\bar{t}$ signal. The systematic uncertainties are summarized in Table VII.

The effect of the JES uncertainty is calculated by adjusting the jet energy corrections that are applied to the MC simulation by $\pm 1\sigma$ and remeasuring the cross section. The central value for the cross section is 7.2 pb with $+1\sigma$ JES and 8.5 pb with $-1\sigma$ JES, so we assign a $\pm 8.6\%$ relative systematic uncertainty due to the JES.

We also determine the uncertainty from initial state radiation (ISR) and final state radiation (FSR) by remeasuring with the PYTHIA MC simulation tuned with more or less ISR and FSR. We take the mean deviation as a systematic uncertainty.

Uncertainties related to top quark kinematic modeling and the jet fragmentation model are considered by replacing PYTHIA with HERWIG [22] as the event generator for the $t\bar{t}$ sample. The result is a 2.2% relative difference in the $t\bar{t}$ acceptance, which we take as a systematic uncertainty.

The uncertainty from PDFs is considered from three sources. The first source is the difference in $t\bar{t}$ acceptance when the CTEQ5L PDF set is reweighted within its own uncertainties. The second source is the difference between the CTEQ5L and an MRST98 [23] set. The third source is calculated by varying $\alpha_S$ within the same PDF set. The final PDF uncertainty is calculated by taking the larger of the two first uncertainties and combining it in quadrature with the $\alpha_S$ uncertainty. This results in a 0.9% uncertainty on the cross section.

The final result is

$$\sigma_{t\bar{t}} = 7.8 \pm 2.4 \text{(stat)} \pm 1.6 \text{(syst)} \pm 0.5 \text{(lumi)} \text{ pb}$$

where we separate the luminosity uncertainty from the other systematic uncertainties. Although we have assumed for this analysis a top quark mass of 175 GeV/$c^2$, the world average for the top quark mass is now approximately 172.4 GeV/$c^2$. This moves the theoretical value of the cross section to approximately 7.2 pb. The systematic uncertainty on the $t\bar{t}$ cross section due to the error on the top-quark mass is small, and leaves the result unchanged.
FIG. 15: (a) $p_T$ distribution of SLT$_e$ tags in lepton+jet events with $\geq 3$ jets. (b) $H_T$ distribution of SLT$_e$ tagged events with $\geq 3$ jets.

<table>
<thead>
<tr>
<th>Source</th>
<th>Relative Uncertainty on $\sigma_{\bar{t}t}$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jet Energy Scale</td>
<td>8.4</td>
</tr>
<tr>
<td>QCD Fit</td>
<td>5.0</td>
</tr>
<tr>
<td>$K$-factor</td>
<td>3.0</td>
</tr>
<tr>
<td>HERWIG/PYTHIA</td>
<td>2.2</td>
</tr>
<tr>
<td>Acceptance Corrections</td>
<td>1.6</td>
</tr>
<tr>
<td>Background Cross Section</td>
<td>0.6</td>
</tr>
<tr>
<td>PDFs</td>
<td>0.9</td>
</tr>
<tr>
<td>FSR</td>
<td>0.6</td>
</tr>
<tr>
<td>ISR</td>
<td>0.5</td>
</tr>
<tr>
<td>Conversion ID Efficiency SFs</td>
<td>10.7</td>
</tr>
<tr>
<td>Fake Matrix</td>
<td>7.8</td>
</tr>
<tr>
<td>Calorimeter Modeling</td>
<td>7.7</td>
</tr>
<tr>
<td>Tag Matrix</td>
<td>6.8</td>
</tr>
<tr>
<td>Jet Environment Correction</td>
<td>5.4</td>
</tr>
<tr>
<td>Total Tagger Uncertainty</td>
<td>17.6</td>
</tr>
<tr>
<td>Total</td>
<td>20.6</td>
</tr>
</tbody>
</table>

TABLE VIII: Summary of systematic uncertainties.

VIII. CONCLUSIONS

We have performed the first measurement of the $t\bar{t}$ production cross section with SLT$_e$ tags in Run II of the Tevatron. This measurement, $\sigma_{\bar{t}t} = 7.8 \pm 2.4$ (stat) $\pm 1.6$ (syst) $\pm 0.5$ (lumi) pb, is consistent with the theoretical value $\sigma_{\bar{t}t} = 6.7 \pm 0.8$ pb ($m_t = 175$ GeV/$c^2$), as well as the current CDF average $\sigma_{\bar{t}t} = 7.02 \pm 0.63$ pb. While statistically limited, this measurement demonstrates the consistency of the top quark production cross section in the lepton+jets final state with soft electron $b$-tagging. This measurement also provides an experimental basis for investigating other high-$p_T$ physics measurements with the soft electron tagging technique.

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 di Fisica 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 World Class University Program, the National Research Foundation of Korea; the Science and Technology Facilities Council and the Royal Society, UK; the Institut National de Physique Nucléaire et Physique des Particules/CNRS; the Russian Foundation for Basic Research; the Ministerio de Ciencia e Innovación, and Programa Consolider-Ingenio 2010, Spain; the Slovak R&D Agency; and the Academy of Finland.


   A. Abulencia et al. (CDF Collaboration), Phys. Rev. D 74, 072006 (2006);
   N. Kidonakis and R. Vogt, Phys. Rev. D 78, 074005 (2008);
[19] We characterize the transverse energy of soft electrons by the track $p_T$, rather than the more typical calorimetric $E_T$, because of the presence of the jet.
[21] Although L00 is used for track reconstruction, it is only one-sided and excluded from the the conversion-finding algorithm. The last layer of the ISL is also too far forward to be fiducial for SLT$e$ tracks. Therefore, a total of at most six layers may be missing for any given SLT$e$ track.
[25] The dilepton channel — where both W bosons decay leptonically — contributes 2–3% of the total event selection as one lepton may escape identification.