Measurement of the Top Quark Pair Production Cross Section in the All-Jets Decay Channel


(DO Collaboration)
We present a measurement of \( t \bar{t} \) production in \( pp \) collisions at \( \sqrt{s} = 1.8 \) TeV from 110 pb\(^{-1}\) of data collected in the all-jets decay channel with the DØ detector at Fermilab. A neural network analysis yields a cross section of 7.1 ± 2.8 (stat.) ± 1.5 (syst.) pb, at a top quark mass (\( m_t \)) of 172.1 GeV/c\(^2\). Using previous DØ measurements from dilepton and single lepton channels, the combined DØ result for the \( t \bar{t} \) production cross section is 5.9 ± 1.2 (stat.) ± 1.1 (syst.) pb for \( m_t = 172.1 \) GeV/c\(^2\).
The standard model predicts that, at Tevatron energies, top quarks are produced primarily in \( t\bar{t} \) pairs, and that each top quark decays into a b quark and a W boson. 44% of these events are expected to have both W bosons decay into quarks. These pure hadronic, or "all-jets", \( t\bar{t} \) events are among the rare collider events with several quarks in the final state. With no final state energetic neutrinos, the all-jets mode is the most kinematically constrained of the top quark decay channels, but is also the most challenging to measure due to the large QCD multijet background. This compelled us to use unique tools such as quark/gluon jet differences, and to make extensive use of neural networks, to separate the \( t\bar{t} \) final states from the QCD background \[1\]. The comparison of \( t\bar{t} \) cross sections from the all-jets and lepton + jets channels allows a search for new phenomena in top decays; for example, top decay via a charged Higgs boson could be observed as a deficit, relative to the all-jets final states with energetic leptons.

The signal for these all-jets \( t\bar{t} \) events is at least six reconstructed jets. The main background is from QCD multijet events that arise from a 2\( \rightarrow \)2 parton process producing two energetic ("hard") leading jets and less energetic ("soft") radiated gluon jets.

The DØ detector is described in Ref. 2. We used the same reconstruction algorithms for jets, muons, and electrons as those used in previous top quark analyses 3. The muons in this analysis are used to identify b jets, and are restricted to the pseudorapidity range \( |\eta| \leq 1.0 \), where \( \eta = \tanh^{-1}(\cos \theta) \), and \( \theta \) is the polar angle relative to the beam axis.

The multijet data sample was selected using a hardware trigger and an online filter requiring five jets of cone size \( R = 0.5 \), pseudorapidity \( |\eta| < 2.5 \) and transverse energy \( E_T > 10.0 \text{ GeV} \). Here, \( R = \sqrt{\Delta \phi^2 + \Delta \eta^2} \), where \( \phi \) is the azimuthal angle around the beam axis. Additionally, we required the total transverse energy of the event (\( H_T \)) to be > 115 or 120 GeV (depending on run conditions). The data sample after the initial cuts has \( 600,000 \) events. With about 200 expected top events in this channel, the background overwhelms the signal by a factor of \( 3000 \). As discrimination from many variables was needed to separate signal from background, most of which are significantly correlated, we used neural networks (NN) as an integral part of this analysis.

The offline analysis proceeded by excluding events with an isolated muon or electron to maintain a data sample independent of the other \( t\bar{t} \) samples. We required events to have at least six \( R=0.3 \) cone jets and less than nine \( R=0.5 \) cone jets, with jet \( E_T > 8.0 \) GeV. We generally used \( R=0.3 \) cone jets because of their greater reconstruction efficiency, but used \( R=0.5 \) cone jets to calculate mass-related variables. We required that at least one jet have an associated muon which satisfied muon quality criteria and which was kinematically consistent with a \( b \rightarrow \mu X \) decay within the jet. As about 20% of \( t\bar{t} \) all-jets events have such a \( \mu \)-tagged jet in the acceptance region for \( t\bar{t} \) signal, compared to approximately 3% of the QCD multijet background in that region, the tagging requirement reduces the background-to-signal ratio by about an order of magnitude. Of the total 280,000 events surviving the offline cuts, 3853 have at least one \( \mu \)-tagged jet. These tagged events comprise the data sample used to extract the cross section.

Compared with the QCD multijet background, \( t\bar{t} \) events typically have more energetic jets, have the total energy more uniformly distributed among the jets, are more isotropic, and have their jets distributed at smaller \( \eta \). To discriminate \( t\bar{t} \) signal from QCD background, we defined at least two variables describing each of these qualities (total energy, jet energy distribution, event shape, and rapidity distribution) 4:

1. \( H_T \): The sum of the transverse energies of jets.
2. \( \sqrt{s} \): The invariant mass of the jets in the final state.
3. \( E_T / H_T \): \( E_T \) is the transverse energy of the leading jet.
4. \( H_T^\text{H} \): \( H_T \) without the transverse energy of the two leading jets.
5. \( N^A_{\text{jets}} \): The number of jets averaged over a range of \( E_T \) thresholds (15 to 55 GeV), and weighted by the \( E_T \) threshold. This parameterizes the number of jets taking their hardness into account.
6. \( E_T_{\text{s, o}} \): The square root of the product of the transverse energies of the fifth and sixth jets.
7. \( \mathcal{A} \): The aplanarity, calculated from the normalized momentum tensor.
8. \( S \): The sphericity, calculated from the normalized momentum tensor.
9. \( C \): The centrality, \( C = H_T / H_E \), where \( H_E \) is the sum of all the jet total energies. This characterizes the transverse energy flow.
10. \( <\eta^2> \): The \( E_T \)-weighted mean square of the \( \eta \) distribution of jets in an event.

These ten variables are the inputs to the first neural network (NN1), whose output is used as an input variable for the second neural network (NN2). The three other inputs to NN2 are:

11. \( p_{\text{T}}^\mu \): The transverse momentum of the tagging muon.

The \( p_{\text{T}}^\mu \) distribution is harder for tagged jets in \( t\bar{t} \) events than for tagged jets in QCD multijet events.
came from a solutions. We did not assume that the muon tagged jet simple two and three jet combinations using DØ jet res-

erminized when there are two invariant masses consistent average, narrower than gluon jets \[5,6\]. The Fisher dis-

FIG. 1. Comparison of the absolute number of data-based multijet background (histogram) with the observed 3853 muon-tagged events in data (a) for (a) \(H_T\) (in GeV), (b) \(E_{T1}/H_T\), (c) sphericity, and (d) jet-width Fisher discriminant. Shown also are the distributions in these variables for HERWIG \(t\bar{t}\) events (•).

12. \(M\): The mass-likelihood variable. This variable is defined as \(M = (M_{W1} - M_{W})^2/\sigma_{mW}^2 + (M_{W2} - M_{W})^2/\sigma_{mW}^2 + (m_t - m_{\bar{t}})^2/\sigma_{mT}^2\), with the parameters \(M_{W}, \sigma_{mW}, \) and \(\sigma_{mT}\) set to 80, 16 and 62 GeV/c\(^2\), respectively. \(M_{W}\) and \(m_t\) refer to the jet combinations that best define the \(W\) boson and top quark masses in an event.

The mass likelihood variable \(M\) is a \(\chi^2\)-like quantity, minimized when there are two invariant masses consistent with the \(W\) mass, and two candidate top quark masses that are identical. \(\sigma_{mW}^2\) and \(\sigma_{mT}^2\) were determined from simple two and three jet combinations using DØ jet res-

olutions. We did not assume that the muon tagged jet came from a b quark.

13. \(F\): The jet-width Fisher discriminant. This is defined as \(F_{\text{jet}} = (\sigma_{\text{jet}} - \sigma_{\text{quark}}(E_T))^2/\sigma_{\text{quark}}^2(E_T) - (\sigma_{\text{jet}} - \sigma_{\text{gluon}}(E_T))^2/\sigma_{\text{gluon}}^2(E_T)\), where \(\sigma_{\text{quark}}(E_T)\) and \(\sigma_{\text{gluon}}^2(E_T)\) are mean square jet widths calculated from HERWIG \([4]\) Monte Carlo, for quarks and gluons respectively, as functions of jet \(E_T\).

It has been demonstrated that quark jets are, on average, narrower than gluon jets \([3,4]\). The Fisher discriminant, based on the \(\eta-\phi\) RMS jet widths, is calculated for the four narrowest jets in the event, and indicates whether the jets were most probably “quark-like” (\(t\bar{t}\)) or “gluon-like” (QCD multijet).

Figure \([1]\) shows a comparison of distributions from the modeled background discussed below, the data, and HERWIG \(t\bar{t}\) events for four of the above variables.

The top quark production cross section is calculated from the output of NN2. Both networks were trained to force their output near 1 for \(t\bar{t}\) events, and near 0 for QCD multijet events, using the back-propagation learning algorithm in JETNET \([6]\).

The very large background-to-signal ratio in the untagged data indicates an almost pure background sample. With a correction for the very small \(t\bar{t}\) component expected, and with a method of assigning a muon tag to the untagged event, the background estimate can be determined directly from the data. Separate sets of un-
tagged data with added muon tags were used for network background response training and background modeling. HERWIG \(t\bar{t}\) events were used for the \(t\bar{t}\) network signal re-

The correct assignement of muon tags to the untagged data was critical to our background model. We derived a “tag rate function” from the entire multijet data set, defined as the probability for any individual jet to have a tagging muon. We chose a function that factorized into two pieces: \(\epsilon\), the detector efficiency dependent on \(\eta\) of the jet and the run number of the event (to account for chamber aging), and \(f(E_T)\), the probability that a jet of transverse energy \(E_T\) has a tagging muon. We studied two parametrizations of \(f(E_T)\), and used the difference to estimate the systematic error from this source. Finally, a small dependence of the tag rate on \(\sqrt{s}\) of the event was found, which was incorporated into \(f(E_T)\). A detailed discussion of the tag rate function is given in Ref. \([1]\).

We established that the \(p_T\) of the tagging muon and the \(E_T\) of the tagged jet (uncorrected for the muon and neutrino energy) are uncorrelated. Therefore, the muon \(p_T\) factors out of the tag rate function, and can be generated independently. By applying the tag rate function to each jet in the untagged data sample, and generating a muon \(p_T\) for those jets determined as tagged, we produced the background model sample.

The NN2 output distributions for data, modeled back-

ground and HERWIG \(t\bar{t}\) signal are plotted in Fig. \([2]\). We excluded events in the region of NN2 output < 0.02. Jets
in that region tend to have low $E_T$, where the tag rate is not well determined due to the low tagging probability (low statistics), and consequently, the background modeling may be less accurate. The cross section is obtained from a simultaneous fit of the data to the background and HERWIG $t\bar{t}$ shapes, with the background normalization ($A_{bkg}$) and the $t\bar{t}$ cross section ($\sigma_{t\bar{t}}$) as free parameters. The result of this fit is also shown in Fig. 2.

The stability of our result can be checked by successively eliminating data points at the lowest values of the NN2 output. Figure 3 shows the values of the background normalization and $t\bar{t}$ cross section as the data points are removed and the remaining points are refitted. The refitted cross sections are independent of NN2 output region, confirming that the initial NN2 output cut at 0.02, and choosing the region NN2 output > 0.1 for our final cross section calculation, does not bias the result. Because of the preponderance of background at the low end of NN2 and the stability of our fits, we use the region NN2 > 0.1 for our quoted cross section results.

Values of the cross section and background normalization are obtained from similar fits with HERWIG $t\bar{t}$ events generated at different top quark masses. The results are shown in Table I. Interpolating to the top quark mass as measured by DO (mt = 172.1 GeV/c^2), we obtain $\sigma_{t\bar{t}} = 7.1 \pm 2.8$ (stat.) $\pm 1.5$ (syst.) pb, consistent with a previous measurement in this channel [2], and the most precise value for this channel to date. Table II summarizes the contributions to the systematic error on the cross section. These were determined by varying each source by its uncertainty, and calculating the difference in the cross section.

As a check, we calculated the cross section from the excess events over expected background, using the efficiency of the criteria for $t\bar{t}$ selection (calculated using HERWIG), along with the branching ratio and the measured luminosity. For NN2 > 0.85 (chosen to minimize the error on the cross section) we observed 41 events with 24.8 $\pm$ 2.4 expected background events for an excess of 16.2 events. The excess corresponds to a $t\bar{t}$ cross section of 7.3 $\pm$ 3.3 $\pm$ 1.6 pb at mt = 172.1 GeV/c^2, consistent with our result above.

The significance of the excess is characterized by the probability $P$ of the observed number of events being due to fluctuation. For an NN2 output threshold of $\approx$ 0.94, where Monte Carlo studies predict maximal expected significance, we observe 18 events where 6.9 $\pm$ 0.9 background events are expected, for which $P = 0.0006$, corresponding to a 3.2 standard deviation effect. This is sufficient to establish the existence of a $t\bar{t}$ signal in the all-jets final state.

To further check the validity of the tag rate function and hence the background model, we looked at events with more than one tagged jet. The modeled background here consists of those untagged events that had two jets tagged by application of the tag rate function. We assumed that the fraction of the double-tagged events from correlated sources, such as $b\bar{b}$ production, is constant over the NN2 output, and refitted the background normalization for a possible overall correlation. A total of 32 double-tagged events are observed for NN2 output > 0.02 where 28.7 $\pm$ 8.2 events are expected from background. Two events are observed for NN2 output > 0.85 with 0.7 $\pm$ 0.1 expected background events, and 1.2 top events expected from Monte Carlo. The small excess in the double-tagged sample is consistent with our conclusion that the more significant excess in the singly-tagged sample is from $t\bar{t}$ production.

Previous DO measurements of $t\bar{t}$ production in the dilepton and single lepton channels [11] give an average cross section of 5.6 $\pm$ 1.4 (stat.) $\pm$ 1.2 (syst.) pb at mt = 172.1 GeV/c^2, in very good agreement with that from the all-jets channel. We combine the all-jets cross section with these results, assuming the statistical errors are uncorrelated, and that the systematic errors have the appropriate correlation coefficients. The combined DO result for the $t\bar{t}$ production cross section is 5.9 $\pm$

---

**TABLE I.** Results of the fits to neural network output.

<table>
<thead>
<tr>
<th>Top Quark</th>
<th>$A_{bkg}$</th>
<th>$\sigma_{t\bar{t}}$</th>
<th>$\chi^2$ / d.o.f.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass (GeV/c^2)</td>
<td></td>
<td>(pb)</td>
<td></td>
</tr>
<tr>
<td>140</td>
<td>1.05 $\pm$ 0.03</td>
<td>18.4 $\pm$ 7.8</td>
<td>17.6 / 17</td>
</tr>
<tr>
<td>160</td>
<td>1.06 $\pm$ 0.03</td>
<td>9.3 $\pm$ 3.8</td>
<td>17.2 / 17</td>
</tr>
<tr>
<td>170</td>
<td>1.07 $\pm$ 0.02</td>
<td>7.2 $\pm$ 3.0</td>
<td>17.1 / 17</td>
</tr>
<tr>
<td>180</td>
<td>1.07 $\pm$ 0.03</td>
<td>6.3 $\pm$ 2.5</td>
<td>16.9 / 17</td>
</tr>
<tr>
<td>200</td>
<td>1.07 $\pm$ 0.03</td>
<td>5.1 $\pm$ 2.0</td>
<td>16.8 / 17</td>
</tr>
<tr>
<td>220</td>
<td>1.07 $\pm$ 0.03</td>
<td>4.4 $\pm$ 1.7</td>
<td>16.7 / 17</td>
</tr>
</tbody>
</table>
TABLE II. Summary of statistical and systematic uncertainties for the cross section.

<table>
<thead>
<tr>
<th>Background Source</th>
<th>Size of Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistical error</td>
<td>4 %</td>
</tr>
<tr>
<td>Functional Form of the Muon-Tag Rate</td>
<td>7 %</td>
</tr>
<tr>
<td>Background Correction for $t\bar{t}$ Signal</td>
<td>6 %</td>
</tr>
<tr>
<td>Background $E_T$ scale</td>
<td>9 %</td>
</tr>
<tr>
<td>Signal Source</td>
<td></td>
</tr>
<tr>
<td>Statistical Error</td>
<td>3 %</td>
</tr>
<tr>
<td>Trigger Turn-on</td>
<td>5 %</td>
</tr>
<tr>
<td>Luminosity Error</td>
<td>5 %</td>
</tr>
<tr>
<td>Jet Energy Scale</td>
<td>6 %</td>
</tr>
<tr>
<td>$t\bar{t}$ Tag Rate</td>
<td>7 %</td>
</tr>
<tr>
<td>Model Dependence</td>
<td>6 %</td>
</tr>
<tr>
<td>$b\rightarrow\mu$ Branching Fraction</td>
<td>6 %</td>
</tr>
<tr>
<td>muon $p_T$ Dependence</td>
<td>7 %</td>
</tr>
<tr>
<td>$F$ Dependence</td>
<td>2 %</td>
</tr>
</tbody>
</table>

1.2 (stat.) $\pm$ 1.1 (syst.) pb for $m_t=172.1$ GeV/$c^2$.

We thank the Fermilab and collaborating institution staffs for contributions to this work and acknowledge support from the Department of Energy and National Science Foundation (USA), Commissariat à L’Energie Atomique (France), Ministry for Science and Technology and Ministry for Atomic Energy (Russia), CAPES and CNPq (Brazil), Departments of Atomic Energy and Science and Education (India), Colciencias (Colombia), CONACyT (Mexico), Ministry of Education and KOSEF (Korea), and CONICET and UBACyT (Argentina).

* Visitor from Universidad San Francisco de Quito, Quito, Ecuador.
† Visitor from IHEP, Beijing, China.