Top cross section measurement at CDF

Gabriele Compostella* **(on behalf of the CDF Collaboration)**
CNAF and University of Padova
Department of Physics "G. Galilei"
via Marzolo 8, 35131 Padova, Italy
E-mail: compostella@pd.infn.it

This paper describes the latest measurements of the $t\bar{t}$ pair production cross section performed by the CDF Collaboration analyzing $p\bar{p}$ collisions at a center-of-mass energy of 1.96 TeV from Fermilab Tevatron, as presented at the XVIII International Workshop on Deep-Inelastic Scattering and Related Subjects. In order to test Standard Model predictions, several analysis methods are explored and all the top decay channels are considered, to better constrain the properties of the top quark and to search for possible sources of new physics affecting the pair production mechanism. Experimental results using an integrated luminosity up to 5.1 fb^{-1} are presented.

---

*Speaker.

1. Introduction

After the top quark discovery in 1995 at the Tevatron by both the CDF and D0 collaborations [1, 2], the number of top events available for experimental studies has become more than an order of magnitude larger, thanks to the increasing amount of data delivered by the accelerator, corresponding to more than 7 fb^{-1} to date. At the Tevatron $p\bar{p}$ collider, at a center-of-mass energy of 1.96 TeV, top quarks are mainly produced in pairs through quark-antiquark annihilation (85%) or gluon-gluon fusion (15%), with a Standard Model (SM) predicted cross section at next-to-leading order (NLO) of $\sigma_{t\bar{t}} = 7.45^{+0.72}_{-0.63}$ pb for a reference top quark pole mass of 172.5 GeV/c^2 [3]. Accurate measurements of the $t\bar{t}$ production cross section play an important role as tests of QCD NLO calculations and can provide probes towards new physics signals involving non-SM $t\bar{t}$ production mechanisms or decays. Since the CKM element V_{tb} is close to unity and m_t is large, the SM top quark decays almost exclusively to a W boson and a b quark, so that the top pair production experimental signatures can be classified with respect to the decay modes of the W boson. In 5% of all the $t\bar{t}$ decays, when both W's decay into electrons or muons, we have the so called dilepton channel; in 30% of the decays, when only one W decays into electrons or muons, the lepton plus jets channel. The 14% of the decays involve hadronically decaying taus that are difficult to isolate from the QCD background, while in the 44% of the cases when both Ws decay to quarks, we have the so called all hadronic channel.

Moreover, the presence of a b-quark in the final state can be exploited as a useful handle to identify the top quark production: b-quarks give rise to jets containing long lived b-hadrons, and those jets can be identified by looking for the presence of tracks in the detector compatible with a secondary decay vertex distinct from the primary interaction point. In CDF, the Silicon Vertex detector systems are crucial for precise determination of particle positions and are exploited in conjuction with a specialized algorithm (SecVtX) to detect such jets, that are usually referred to as b-tagged. The requirement of at least one b-tagged jet in the event is a powerful tool employed in many analyses to increase the signal to background ratio S/B in data.

In the following, a review of several cross section measurements performed by the CDF collaboration in the various decay channels using different methods will be given.

2. The dilepton channel

The experimental signature for this channel consists of two opposite sign, high transverse momentum ($p_T > 20$ GeV) leptons (e or μ), large missing energy due to neutrinos from W decays ($\not{E}_T > 25$ GeV) and at least two high-E_T central jets ($E_T > 15$ GeV and pseudorapidity $|\eta| < 2.5$). Physics backgrounds to the $t\bar{t}$ production are mainly due to Drell Yan +jets events and diboson production (WW, WZ, ZZ +jets) and are estimated using Monte Carlo (MC) simulations. In addition, instrumental backgrounds are present for all analyses: for example, detector resolution effects or mis-measured jets may induce artificial $\not{E}_T$, or poorly reconstructed jets can be mis-identified as leptons. In analyses using b-tagging, heavy flavor jets can cause fake tags. These backgrounds are estimated directly from data: calorimeter response for electrons and muons energies is calibrated using well known $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$ decays, for jets by balancing the transverse momentum in events with a photon and a jet, while variations in calorimeter response are determined using dijet

events. Fake leptons are estimated by evaluating the probability of mis-identification of a jet as a lepton in data samples enriched in jets, while fake b-tags are accounted for by parameterizing the probability of identifying a heavy flavour jet as a b-jet in QCD dominated data samples. The $t\bar{t}$ pair production cross section is evaluated by performing a counting experiment on the number of candidate signal events as $\sigma_{t\bar{t}} = (N_{obs} - N_{bkg})/\varepsilon_{kin}\mathscr{L}$ where N_{obs} is the number of selected signal events, N_{bkg} the number of background events, ε_{kin} the efficiency of the kinematical selection, estimated on $t\bar{t}$ signal MC, and $\mathscr{L}$ is the integrated luminosity of the data sample used in the analysis. In 5.1 fb^{-1} of CDF data in the dilepton channel 343 signal candidates are found, giving a cross section value of $\sigma_{t\bar{t}} = 7.4 \pm 0.6$ (stat.) ± 0.6 (syst.) ± 0.5 (lumi.) pb for a reference top quark mass $m_t = 172.5$ GeV/c^2. After the requirement of the presence of at least one jet in the event identified as coming from the decay of a b-quark by the SecVtX algorithm, 137 candidate events are found, for a measured cross section value of $\sigma_{t\bar{t}} = 7.3 \pm 0.7$ (stat.) ± 0.5 (syst.) ± 0.4 (lumi.) pb [4].

3. The lepton plus jets channel

Events with this signature, compared to other possible decay channels, have an advantageous mixture of small contamination from background physics processes and high branching ratio of $t\bar{t}$ events, and are widely used for top quark properties measurements. CDF has recently performed cross section measurements in this channel with two techniques [5], one based on b-tagging and the other on a topological selection based on a Neural Network (NN) exploiting the kinematical properties of the event. The basic lepton plus jets event selection is common to the two analyses and requires one central well identified high-p_T e or μ ($p_T > 20$ GeV), large amount of missing transverse energy as evidence of a neutrino from the W boson decay ($\not{E}_T > 25$ GeV) and the presence of at least three central jets with $E_T > 20$ GeV and $|\eta| < 2$. In the b-tagging measurement, to further reduce background, an additional requirement is placed on the scalar sum H_T of the transverse energy of the lepton, $\not{E}_T$, and jets ($H_T > 230$ GeV). Several physics processes can contribute as background to the selected data samples, such as W+jets, Z+jets, diboson (WW, ZZ, WZ), electroweak produced single-top, and QCD multijet: a mixture of data and MC techniques are used to estimate the contribution of each process. In the b-tagging analysis, once the pretag estimate for all processes is found, the number of events with at least one b-tagged jet for Z+jets, diboson, and single-top events is found by applying a MC-based tagging efficiency factor. As the relative fraction of jets associated with heavy flavor (HF) in the W+jets MC simulation is found to be underpredicted, a correction factor is determined using a NN trained to discriminate heavy flavor from light jets (LF), and comparing its prediction to data in the single jet sample. The number of b-tags due to W+HF is found by applying a tagging efficiency factor to the predicted MC fraction, while the W+LF contribution is determined by first parameterizing the probability of wrongly identifying a jet as HF (mistag) from independent multijet data samples and then multiplying the mistag rate by the number of expected W+jets events. QCD multijet events are modeled from data passing the analysis cuts but failing the primary lepton identification and their contribution to the selected sample in terms of b-tags is determined using a template fit in the low $\not{E}_T$ region of b-tagged data. Finally, a likelihood is formed from the data, the $t\bar{t}$ cross section, and the predicted background. Using collected data corresponding to an integrated luminosity of 4.3 fb^{-1}, the cross section is measured to be $\sigma_{t\bar{t}} = 7.22 \pm 0.35$ (stat.) ± 0.56 (syst.) ± 0.44 (lumi.) pb for $m_t = 172.5$ GeV/c^2.

The largest systematic uncertainties to this measurement are due to the measured luminosity, the correction to the W+HF background, and the b-tag modeling in the simulation.

The topological measurement uses a NN to discriminate $t\bar{t}$ events from background instead of relying on the b-tagging requirement. Because of the large mass of the top quark, $t\bar{t}$ events are more energetic, central, and isotropic compared with the dominant backgrounds such as W+jets and QCD multijet events, giving rise to differences in the kinematics. The NN is trained to discriminate simulated samples of $t\bar{t}$ signal from the dominant W+jets background exploiting different kinematical variables; then templates of the NN output distribution for W+jets, QCD events modeled as in the b-tagging analysis and $t\bar{t}$ signal are used in a binned likelihood fit to the NN output distribution of data. Using data corresponding to an integrated luminosity of 4.6 fb^{-1}, this analysis determines $\sigma_{t\bar{t}} = 7.71 \pm 0.37$ (stat.) ± 0.36 (syst.) ± 0.45 (lumi.) pb for $m_t = 172.5$ GeV/c^2. Since this measurement does not use b-tagging in the event selection, it shares only two of the largest sources of systematic uncertainty with the previous b-jet tagging based measurement, those due to the measured luminosity and the W+HF background normalization.

The largest systematic contribution in both measurements is the uncertainty on the measured luminosity, coming from the uncertainty on the inelastic $p\bar{p}$ cross section and acceptance of the detector luminosity counters. This uncertainty can be removed by measuring the $t\bar{t}$ cross section relative to the inclusive $Z/\gamma^* \rightarrow ll$ cross section and multiplying by the theoretical $Z/\gamma^* \rightarrow ll$ cross section. The uncertainties on the theoretical and measured $Z/\gamma^* \rightarrow ll$ cross sections are propagated to the final $t\bar{t}$ cross section measurement and are smaller than the luminosity uncertainty. The inclusive $Z/\gamma^* \rightarrow ll$ cross section is measured requiring events with two oppositely charged e or μ with invariant mass M_{ll} between 66 and 116 GeV/c^2, and the $t\bar{t}$ to $Z/\gamma^* \rightarrow ll$ cross sections ratio is calculated. Finally, multiplying this ratio by the theoretical $Z/\gamma^* \rightarrow ll$ cross section, CDF obtains $\sigma_{t\bar{t}} = 7.32 \pm 0.36$ (stat.) ± 0.59 (syst.) ± 0.14 (theory) pb for the b-tagging measurement and $\sigma_{t\bar{t}} = 7.82 \pm 0.38$ (stat.) ± 0.37 (syst.) ± 0.15 (theory) pb for the NN based measurement. The luminosity systematic uncertainty for both measurements has been replaced by a smaller uncertainty on the theoretical $Z/\gamma^* \rightarrow ll$ cross section. The two measurements can also be combined using a best linear unbiased estimate method, obtaining $\sigma_{t\bar{t}} = 7.70 \pm 0.52$ pb, consistent with Standard Model NLO prediction.

4. The all hadronic channel

The all hadronic channel has the highest branching ratio, but suffers from the largest background contribution due to QCD multijet events, which overwhelms the signal by three orders of magnitude. The experimental signature of this channel is characterized by the presence of six or more high E_T jets, two of which originate from b-quarks. Applying a specific topological and kinematic NN based selection together with the b-tagging requirement, S/B can be enhanced up to 1/2, enabling the clear observation of top pair signal. In a recent CDF measurement in the all hadronic channel [6] this technique is used to reconstruct the top quark mass for each event using a kinematical fit of the six leading jets in the event to a $t\bar{t}$ final state. The background contribution to the b-tagged sample is parameterized from data with four jets, dominated by QCD multijet events, estimating tagging rate probabilities, which are then propagated to the selected sample. The distribution of reconstructed top quark masses is then fit to the distributions expected from background

and $t\bar{t}$ events simulated at various values of the top quark mass, to obtain the value which best describes the data. This method allows to determine the number of $t\bar{t}$ events in the sample along with the top quark mass, and to measure the top pair production cross section. The measurement yields a top quark mass $m_t = 174.8 \pm 2.4$ (stat.) $^{+1.2}_{-1.0}$ (syst.) GeV/c^2, and a pair production cross section of $\sigma_{t\bar{t}} = 7.2 \pm 0.5$ (stat.) ± 1.0 (syst.) ± 0.4 (lumi) pb using data corresponding to an integrated luminosity of 2.9 fb^{-1}.

5. The $\not{E}_T$ plus jets channel

The signature of this hybrid channel focuses on $\not{E}_T$ due to the presence of a neutrino in the W decay rather than on lepton identification, and requires four or more energetic jets in the final state ($E_T > 15$ GeV and $|\eta| < 2$) with at least one b-tagged jet. Orthogonality with respect to lepton based and all hadronic searches is ensured respectively by explicitly vetoing well reconstructed leptons and by requiring statistically significant $\not{E}_T$ in the event. Because of this choice, this analysis carries an important weight in the combination of the results produced by the experiment. Additionally, the lepton veto enhances the τ+jets component in the data sample. A NN is trained to discriminate between MC simulated $t\bar{t}$ signal and data-driven background: since before the tagging requirement S/B is around 3.5%, the data sample after prerequisites is treated as background during the training. A cut on the NN output is used to select a $t\bar{t}$ enriched sample, and the top pair production cross section is evaluated performing a counting experiment on the number of b-tags in the selected sample. The contribution to the number of b-tagged jets due to QCD multijet and W+HF backgrounds is estimated parameterizing the tag rate probability on jet related variables in the sample with exactly three jets, where the signal contamination is below 0.1%. Using data corresponding to an integrated luminosity of 2.2 fb^{-1} CDF reports a measurement of $\sigma_{t\bar{t}} = 7.99 \pm 0.55$ (stat.) ± 0.76 (syst.) ± 0.46 (lumi) pb [7].

Acknowledgments

References

[1] F. Abe, *et al.* The CDF collaboration, *Phys. Rev. Lett.* **74** (1995) 2626.

[2] S. Abachi, *et al.* The D0 collaboration, *Phys. Rev. Lett.* **74** (1995) 2632.

[3] S. Moch, P. Uwer, *Nucl. Phys. Proc. Suppl.* **183** (2008) 75 [hep-ph/0807.2794].

[4] The CDF Collaboration, *CDF Conference Note* **10163** (2010).

[5] The CDF Collaboration, to appear in *Phys. Rev. Lett.* [hep-ex/1004.3224].

[6] The CDF Collaboration, *Phys. Rev. D* **81** (2010) 052011.

[7] The CDF Collaboration, *CDF Conference Note* **9988** (2009).