Measurement of the $t\bar{t}$ production cross section in the all-hadronic channel
Conference Note

The CDF Collaboration
(Dated: August 1, 2005)

Abstract

We present here the measurement of the $t\bar{t}$ production cross section in the all-hadronic channel, where both $W's$ decay hadronically. The analysis is performed using 311 pb^{-1} of $p\bar{p}$ collisions collected with a multijet trigger at $\sqrt{s}$ = 1.96 GeV with the Collider Detector at Fermilab. After the application of an optimized kinematical selection we observe an excess of events with 6 or more jets, including one or more b jets, relative to background expectations. Based on this excess we measure the production cross section $\sigma_{t\bar{t}} = 7.5^{+3.7}_{-2.8}$ pb.

1 Introduction

In this analysis we search for events in which both W bosons decay into a $q\bar{q}'$ pair, leading to an all-hadronic final state characterized by a nominal 6-jet topology. This final state has the advantage of a large branching ratio ($BR \approx 4/9$) and of being fully reconstructed. The major drawback is that it competes against a very large QCD multijet background which dominates by three order of magnitude even after the application of a specific multijet trigger.

To improve the signal-to-background ratio, S/B, we apply a set of kinematical requirements which favors $t\bar{t}$ candidates leading to $S/B \approx 1/25$ with a good efficiency (ϵ_k = 6.7% for a top mass of 178 GeV/c^2). This is clearly not enough and we need to require the presence of displaced secondary vertices (*tags*) in the event to increase the S/B and to extract a $t\bar{t}$ signal. The analysis presented here is one of the various strategies used by CDF in Run I for the observation of $t\bar{t}$ production and the measurement of the top mass in this final state [1],[2].

2 Monte Carlo samples

In this analysis, Monte Carlo simulations are used only for the generation of $t\bar{t}$ samples intended for the optimization of the kinematical selection and the evaluation of its efficiency and systematics. Samples generated with PYTHIA (6.203) and HERWIG (6.4) under different conditions of ISR/FSR modeling and using different p.d.f. sets have been produced for the evaluation of efficiency and its systematic uncertainties. All samples have been generated with a top mass of 178 GeV/c^2 and processed with a realistic simulation of the CDF detector response in particular in order to reproduce the performances of the Silicon detector.

3 Data sample and event selection

We summarize here the procedure used to isolate all-hadronic $t\bar{t}$ candidates in the $p\bar{p}$ collisions collected by CDF in the period 2002/2004 and corresponding to 311 pb^{-1}.

3.1 Multijet trigger

A special multijet trigger has been devised specifically for this analysis. This trigger requires at least four calorimeter clusters with $E_T \geq 15$ GeV (with respect to $z = 0$) and a total transverse energy exceeding 125 GeV. Such a trigger has a cross section of about 14 nb and an efficiency of about 85% on all-hadronic $t\bar{t}$ events. The signal-to-background ratio at this level amounts to about 1/3500 (assuming $\sigma_{t\bar{t}} = 6.1$ pb, as expected for a top mass of 178 GeV/c^2).

3.2 Pre-requisites

We define some pre-requisites in order to have a sample of events as well defined and as clean as possible:

1. at least one *good quality* vertex with $|z_{vert}| < 60$ cm;
2. $\not{E}_T/\sqrt{\sum E_T} < 3$ to avoid overlap with the $\not{E}_T$ + jets analysis;
3. removal of events with a *good, high-P_T central electron or muon* to avoid overlap with other leptonic analyses.

After such pre-requisites we remain with about 3 846 000 events.

3.3 Kinematical selection

We seek an optimal kinematical selection based on the best $S/\sqrt{B+S}$ achievable. As signal we consider $t\bar{t}$ Monte Carlo events, while as background we consider the multijet events themselves (this is because the S/B is very small). We consider jets with a transverse energy $E_T \geq 15$ GeV (corrected for detector and reconstruction effects), and pseudo-rapidity $|\eta| \leq 2.0$. The quantities used for the selection are, as suggested by [1], number of jets (N_{jet}), total transverse energy ($\sum E_T$), centrality (C), aplanarity (A) and the sub-leading energy ($\sum_3 E_T \equiv \sum E_T - E_T^1 - E_T^2$) obtained by removing the two jets with the highest E_T. The comparison of the data (background dominated) and signal MC for the chosen kinematical variables is shown in Fig.1.

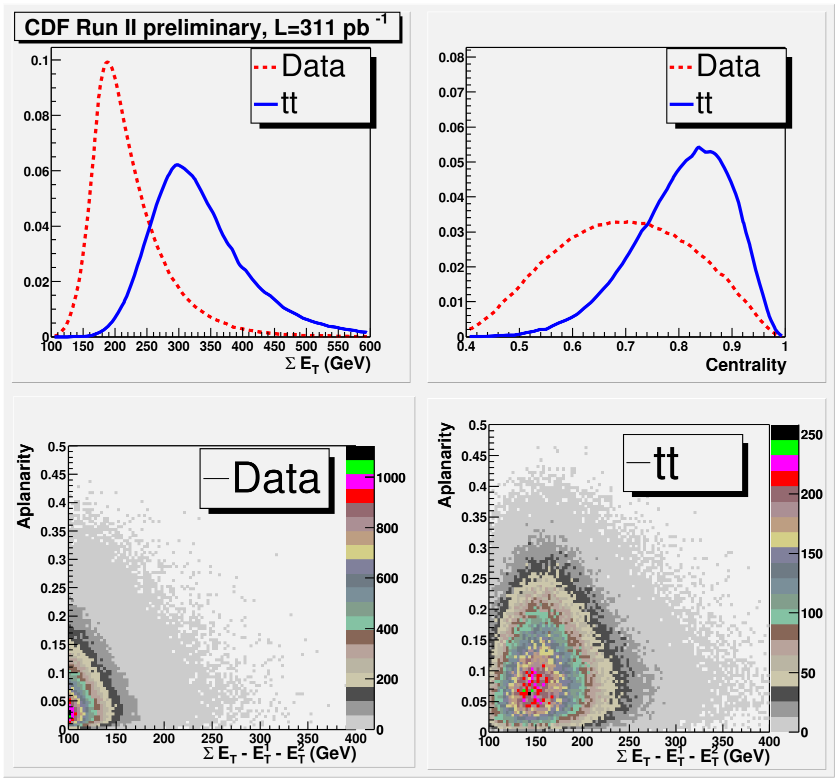

Figure 1: *Kinematical variables distributions for $t\bar{t}$ signal and multijet data for events with* $6 \leq N(jet) \leq 8$.

The optimal selection is the following (in cascade):

- all pre-requisites
- $6 \leq N_{jet} \leq 8$
- $A + 0.005 \sum_3 E_T \geq 0.96$
- $C \geq 0.78$
- $\sum E_T \geq 280$ GeV

Such a selection isolates about 3300 candidates with an efficiency of $6.7 \pm 1.4\%$ and $S/B \approx 1/25$. The relative systematic uncertainty on the efficiency (20%) is dominated by the uncertainties on jet energy scale (see Table 1).

Table 1: *Relative systematic uncertainties.*

Source	$\delta\epsilon/\epsilon$ (%)
Generation/fragmentation	1.7
ISR/FSR modeling	3.9
p.d.f.	2.6
Jet energy scale	19.4
Total	20.0

4 Displaced vertex tagging

In order to improve the S/B we request the presence of secondary vertices reconstructed with at least two good quality tracks with hits in our Silicon vertex detector. To be called *tags* the jets with secondary vertices need to have a signed displacement from the primary vertex with a significance larger than 3.0. The cross section measurement is performed using the total number of tagged jets (not events) in order to avoid the explicit calculation of the background for double tagged events (this is the subject of another ongoing analysis).

The average number of tags in a $t\bar{t}$ event passing the kinematical selection is $n_{tag}^{ave} = 0.846 \pm 0.073$, as derived from simulations and comparison to b-enriched events.

4.1 Background estimate

The method we use to estimate the background is based on the verified assumption that even if the relative contribution from all QCD processes changes as a function of jet multiplicity, the probability for a *taggable* jet (that is a jet inside the vertex detector acceptance) to be tagged remains flat increasing the jet multiplicity. This allows to use the probability extracted from a control sample (events with $N_{jet} = 4$) on jets in events with higher jet multiplicity in order to estimate the number of tags expected in the signal sample. Such a probability (*tag rate*) is defined as:

$$\mathcal{P} = \frac{\#\ of\ tagged\ jets}{\#\ of\ taggable\ jets}$$

To account for the different conditions of the control and the signal samples, the tag rate is parametrized in terms of jet-E_T, number of tracks reconstructed in the vertex detector, N_{trk} and number of primary vertices, N_{vert}.

Before the kinematical selection the multijet sample is composed, essentially, by background events only. The goodness of the parametrization and of the resulting estimates across different jet multiplicities is shown in Table 2 where we see that the expectations agree with the observed number of tags within a few %, also in terms of average number of tags per event, see Fig. 2. The small differences are accounted for as a systematic uncertainty on the jet multiplicity.

Table 2: *Observed and Expected Tags. Statistical uncertainties only.*

N_{jet}	4	5	6	7	8
Events	1542092	855290	282894	68449	13182
Taggable jets	4163173	2859146	1127964	317045	69653
Observed Tags	133367	86414	32589	8925	1904
Expected Tags	133367 ± 356	87372 ± 239	33244 ± 95	9148 ± 27	1993 ± 6
$\frac{N_{obs}-N_{exp}}{N_{exp}}$ (%)	0 ± 0.4	-1.0 ± 0.4	-2.0 ± 0.6	-2.4 ± 1.0	-4.5 ± 2.2

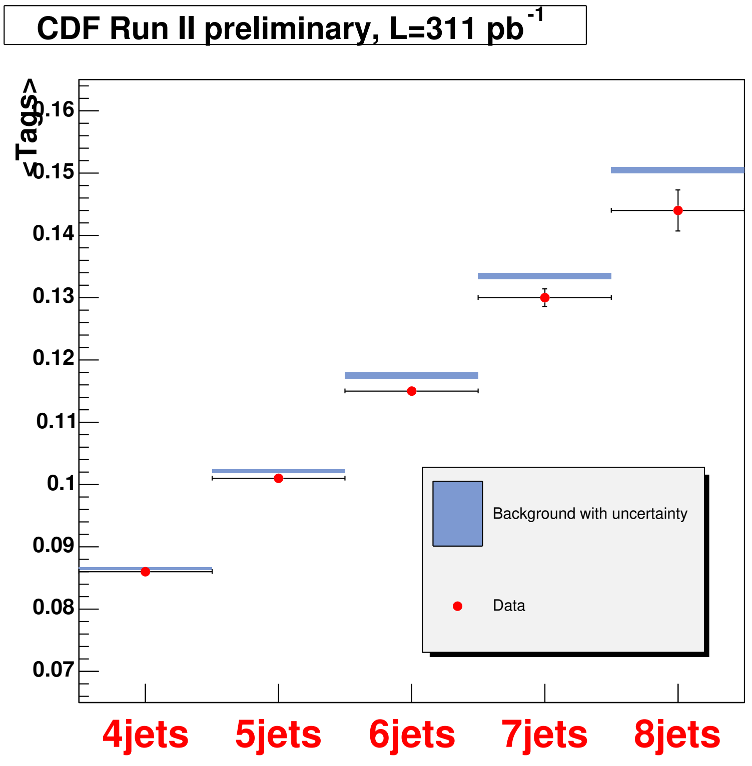

Figure 2: *Average number of tags (observed and expected) per event before kinematical selection as a function of the jet multiplicity.*

4.2 Systematics uncertainties on the background estimate

The kinematical selection changes the event characteristics with respect to those found in the control sample (events with $N_{jet} = 4$) where the parametrization has been derived. Possible biases due to the selection are treated as systematic uncertainties on the background prediction and are evaluated with the help of suitable control samples which are essentially composed of multijet events as highly populated as possible but with the smallest possible presence of $t\bar{t}$ events. We define several samples where we exclude $t\bar{t}$ events by "negating,, the all-hadronic kinematical selection. With these samples we evaluate a relative systematic uncertainty of 4.1% on the background prediction after the kinematical selection, see Table 3. We also checked for possible dependecies on instantaneous luminosity and detector effects and found those negligible. The so-called "common" uncertainty refers to the differences found between observed and expected tags in the various control samples, including the difference found due to the jet multiplicity.

Table 3: *Systematic uncertainties on the kinematical requirements.*

Variable	relative uncertainty (%)
$\sum E_T \geq 280$ GeV	0.4
Centrality ≥ 0.78	0.1
Aplanarity $+0.005 \sum E_T^3 \geq 0.96$	2.6
common	3.2
Total	4.1

4.3 Background estimate after kinematical selection

We now estimate the background expected after the kinematical selection. The background is corrected in each bin for the presence of $t\bar{t}$ events expected in the pretag sample.

Table 4: *Observed tags and expected background after the kinematical selection*

N_{jet}	4	5	6	7	8
Events	101	695	1396	1300	646
Background	17.6 ± 0.3	133.4 ± 7.8	255.5 ± 15.3	267.1 ± 16.3	148.8 ± 6.1
$t\bar{t}$ ($\sigma = 6.1$ pb)	0.4 ± 0.1	14.9 ± 3.0	62.4 ± 12.5	49.4 ± 9.9	17.7 ± 3.5
All SM sources	18.0 ± 0.3	148.3 ± 8.4	317.9 ± 19.8	316.5 ± 19.1	166.5 ± 7.0
Observed Tags	20	154	346	322	148
$Data - Bkgd$	2.4 ± 4.5	20.6 ± 14.7	90.5 ± 24.1	54.9 ± 24.2	-0.8 ± 13.6

The presence of an excess in the signal region ($6 \leq N_{jet} \leq 8$) is visible in Fig. 3, also when we look at the average number of tags per event as a function of the jet multiplicity (see Fig. 4).

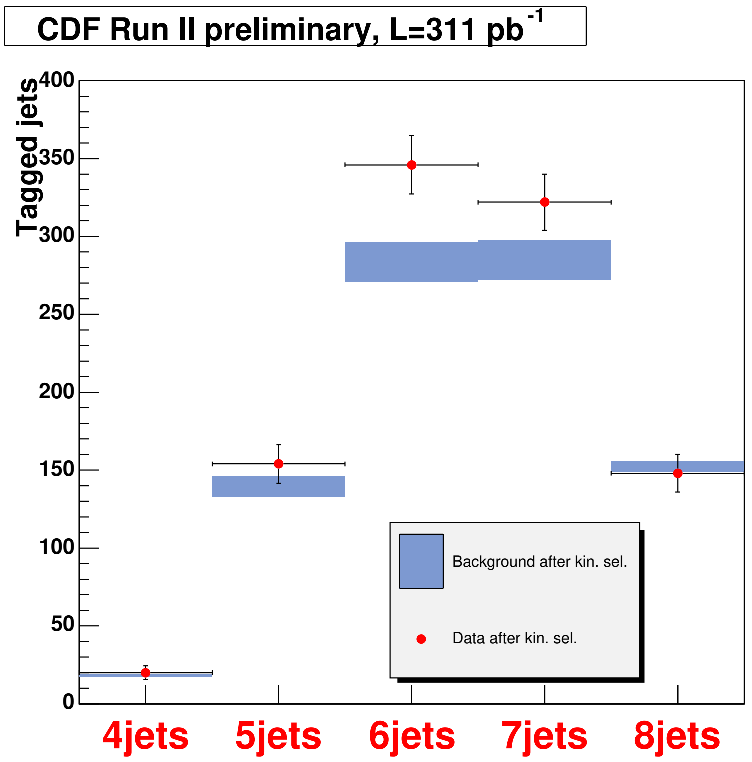

Figure 3: *Comparison between observed and expected tags in each multiplicity bin.*

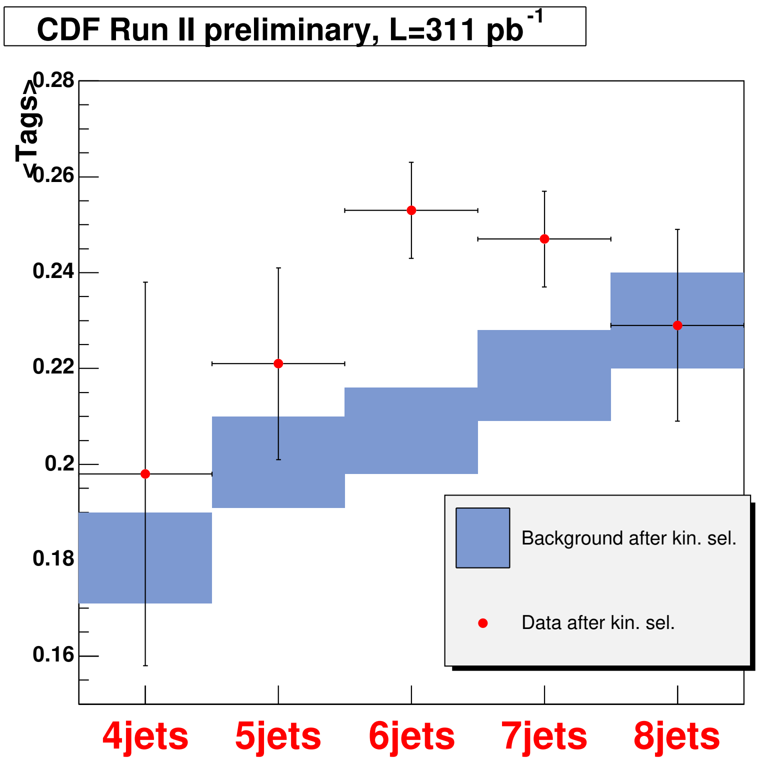

Figure 4: *Average number of tags per event as a function of the jet multiplicity.*

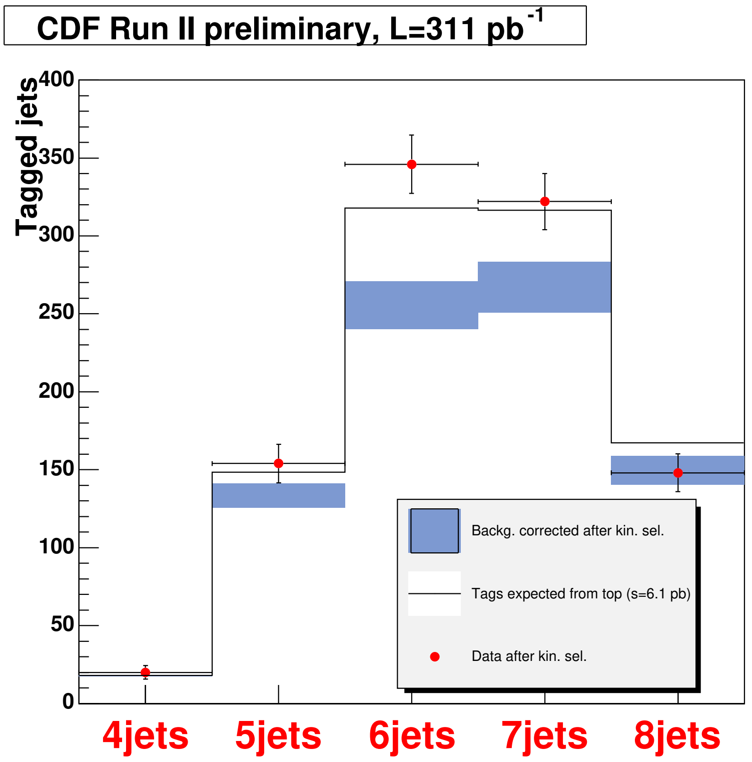

Figure 5: Number of tagged jets *vs* jet multiplicity. Data (points), expected background and $t\bar{t}$ expectation ($\sigma_{t\bar{t}} = 6.1$ pb).

5 Cross section measurement

After the application of the kinematical selection to a multijet sample of 311 pb^{-1}, we are left with $N_{evt} = 3315$ events, out of which 695 events have at least one tag, for a total of $n_{obs} = 816$ candidate tags. The total background expectation in the signal region, provided by the tag rate parametrization, amounts to $n_{exp} = 683.2 \pm 37.5$ tags after correcting[1] the background for the presence of $t\bar{t}$ events in the pretag sample.

The excess of tags $Data - Background = 133 \pm 47$ in the signal region is ascribed to $t\bar{t}$ production and from this excess we measure the production cross section. In Figure 5 we show the number of tagged jets as a function of the jet multiplicity and compare it to background and signal (assuming $\sigma_{t\bar{t}} = 6.1$ pb) expectations.

The ingredients we use for the cross section measurement are shown in Table 5 and with these

[1] Here the correction is global not bin-by-bin as done for Table 4.

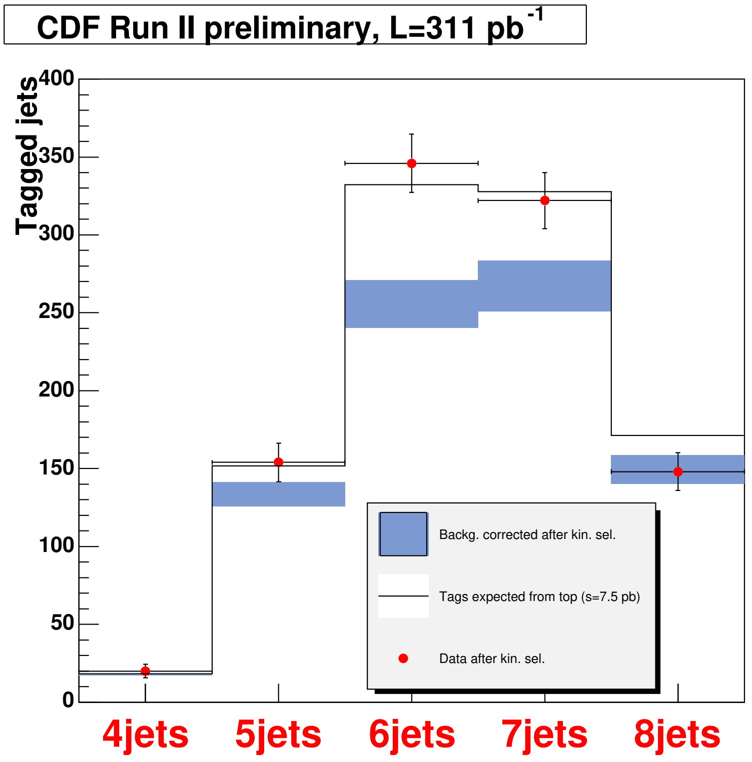

Figure 6: Number of tagged jets *vs* jet multiplicity. Data (points), expected background and $t\bar{t}$ expectation (assuming $\sigma_{t\bar{t}} = 7.5$ pb).

Table 5: *Input values for the Likelihood maximization*

Variable	Symbol	Input value
Integrated luminosity (pb^{-1})	L	311 ± 18
Candidate tags	$n \equiv n_{obs}$	816
Expected background	$b \equiv n_{exp}$	683.2 ± 37.5
Kinematical efficiency (%)	ϵ_k	6.7 ± 1.4
Average number of tags per $t\bar{t}$	n_{tag}^{ave}	0.846 ± 0.073

ingredients we build a likelihood function:

$$\mathcal{L} = e^{-(L-\overline{L})^2/2\sigma_L^2} \cdot e^{-(\epsilon_k-\overline{\epsilon_k})^2/2\sigma_{\epsilon_k}^2} \cdot e^{-(n_{tag}^{ave}-\overline{n_{tag}^{ave}})^2/2\sigma^2_{n_{tag}^{ave}}} \cdot e^{-(b-\overline{b})^2/2\sigma_b^2} \cdot$$
$$\cdot \frac{(\sigma_{t\bar{t}} \cdot \epsilon_k \cdot n_{tag}^{ave} \cdot L + b))^n}{n!} \cdot e^{-(\sigma_{t\bar{t}} \cdot \epsilon_k \cdot n_{tag}^{ave} \cdot L + b)}$$

The central value of the cross section is given by the maximum of $\log \mathcal{L}$, that is

$$\sigma_{t\bar{t}} = \frac{n-b}{\epsilon_k n_{tag}^{ave} L}$$

while the uncertainties are calculated as $\Delta \log \mathcal{L} = \frac{1}{2}$ points of the likelihood function.

The value of the $t\bar{t}$ production cross section, measured in the all-hadronic channel is then:

$$\sigma_{t\bar{t}} = 7.5 \pm 1.7(stat)\ ^{+3.3}_{-2.2}(syst) = 7.5^{+3.7}_{-2.8}\ pb$$

Figure 6 shows the number of tagged jets as a function of the jet multiplicity and its comparison to the sum of background and signal (assuming our measured value $\sigma_{t\bar{t}} = 7.5$ pb) expectations.

The cross section has been evaluated as a function of the top mass in Figure 7.

6 Conclusions

We have discussed here the measurement of the $t\bar{t}$ production cross section in the all-hadronic channel using a sample of 311 pb^{-1} events collected by CDF with a multijet trigger in the period 2002/2004.

Starting from a very small signal-to-background ratio ($S/B \approx 1/3500$), we improve it to $S/B \approx 1/5$ by applying a specific kinematical selection ($\epsilon = 6.7 \pm 1.4\%$) and the request of presence of displaced secondary vertices ($n_{tag}^{ave} = 0.846 \pm 0.073$). The background estimate is completely data driven based on a parametrization of the tag rate on a multijet sample that is signal depleted but kinematically similar to our signal region.

The counting of tags shows an excess of about 133 tags with respect to a background of 683 ± 38 tags. Abscribing such an excess to $t\bar{t}$ production, we measure the cross section to be

$$\sigma_{t\bar{t}} = 7.5 \pm 1.7(stat)\ ^{+3.3}_{-2.2}(syst) = 7.5^{+3.7}_{-2.8}\ pb$$

References

[1] F. Abe *et al.*, *First Observation of the All-Hadronic Decays of $t\bar{t}$ Pairs*, Phys. Rev. Lett. 79, 1992 (1997).

[2] A. Sidoti *et al*, *Investigation of $t\bar{t}$ in the full hadronic final state at CDF with a Neural Network approach*, Nucl.Instrum.Meth.A466:538-550 (2001)

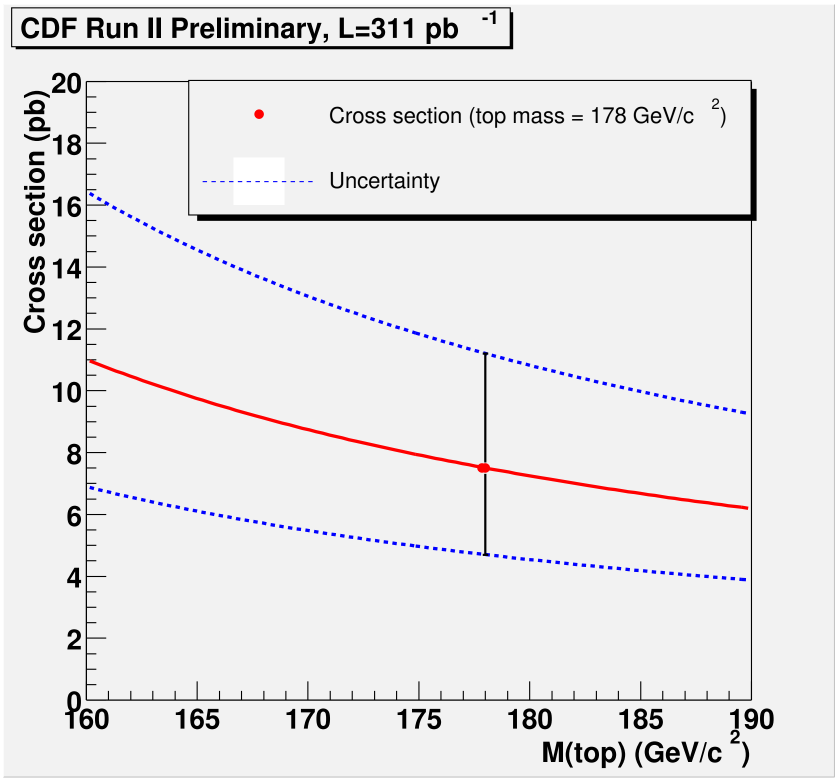

Figure 7: *Cross section as a function of the top mass: central value (red curve) and $\pm 1\sigma$ uncertainty (blue-dotted curve).*