Fermi National Accelerator Laboratory

Top Mass Measurement at CDF

Simona Rolli

For the CDF Collaboration

Fermi National Accelerator Laboratory
P.O. Box 500, Batavia, Illinois 60510

INFN-Pavia, Italy and
Lawerence Berkeley Laboratory, Berkeley, USA

June 1996

Disclaimer

Top mass measurement at CDF

Simona Rolli

INFN-Pavia, Italy and Lawrence Berkeley Laboratory, Berkeley, USA
E-mail: rolli@fnal.gov

Abstract

We present the measurement of the top quark mass using $\mathcal{L} = 110 \ pb^{-1}$ data sample of $p\bar{p}$ collisions at $\sqrt{s} = 1.8 \ TeV$ collected with the Collider Detector at Fermilab (CDF).

We show the results for the different channels and discuss with some emphasis the determination of the systematic uncertainties.

Top mass measurement at CDF

Simona Rolli

*INFN-Pavia, Italy and Lawrence Berkeley Laboratory, Berkeley, USA*
E-mail: rolli@fnal.gov

We present the measurement of the top quark mass using $\mathcal{L} = 110 \ pb^{-1}$ data sample of $p\bar{p}$ collisions at $\sqrt{s} = 1.8 \ TeV$ collected with the Collider Detector at Fermilab (CDF). We show the results for the different channels and discuss with some emphasis the determination of the systematic uncertainties.

1   Introduction

The first direct evidence for the *top* quark was presented by CDF in 1994[1] and has been confirmed by CDF and D0 in 1995[2,3] using a larger data sample. In these references the determination of the mass of the top was also presented. Here we report the results on top mass measurement obtained on a data sample of $\mathcal{L} = 109 \pm 7.2 \ pb^{-1}$, corresponding to the full statistics collected between 1991 and 1995.

At the Tevatron collider, that provides $p\bar{p}$ collisions at $\sqrt{s} = 1.8 \ TeV$, the dominant process for top production is $q\bar{q} \rightarrow t\bar{t}$ (90%). In the Standard Model the top decays almost 100% of the time (V_{tb}) in $t \rightarrow Wb$ and the classification of top events is based upon the decay mode of the W boson.
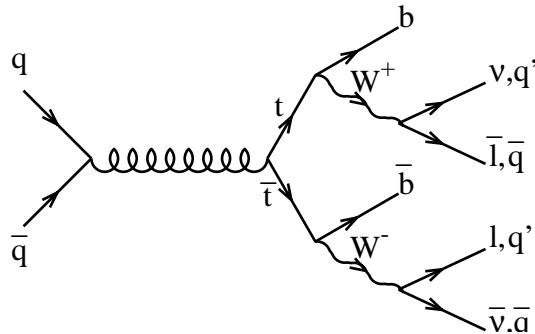
We will report the results for the following channels analysis:

- single lepton plus jets, one W decays leptonically (e or μ) the other goes to $q\bar{q}'$;

- dilepton, both W's decay leptonically (e or μ);

- all hadronic, both W's decay into quarks.

2   Mass analysis for letpon + jets

We select the events in the following way: we start from requesting a central lepton + 3 jets with $E_t > 15$ GeV and $|\eta| < 2.4$. We require an additional fourth jet with $E_t > 8$ GeV and $|\eta| < 2.4$. Loosening the E_t cut on the fourth jet increase the acceptance from 60% to about 85%, which is a factor 1.4.
We found 153 events, 37 of them with one b-tagged jet.

We fit the tagged events to the $t\bar{t}$ hypothesis with a constrained kinematic fitting method:

1) $p\bar{p} \to t_1 + t_2 + X$

2) $t_1 \to b_1 + W_1$

3) $t_2 \to b_2 + W_2$

4) $W_1 \to l + \nu$

5) $W_2 \to j_1 + j_2$

The input quantities are the 4-momentum of the lepton, the jets (*corrected quantities*) and the X system recoiling against the $t\bar{t}$ system (residual missing E_t). We impose the following conditions:

6) $M_{t1} = M_{t2}$; $M_W = 80.2$ GeV; , $_W = 2.1$ GeV

We end up with 20 equations with 18 unknowns, that means a 2 constraints fit. In the fit we use the four highest E_t jets, the b-tagged jet is requested to be one of the b-jets and both the possible solutions for P_z^ν are tried. Out of the 12 possible combinations, the solution with the lowest χ^2 is chosen ($\chi^2 < 10$).

After we determine the mass for each of the 34 fully reconstructed events we apply a likelihood method to determine the top mass in the presence of background events. We fit the mass distribution from data to top MC + background and for each top mass we calculate the likelihood that the data agrees with the MonteCarlo. Top mass corresponds to the minimum value of $-\ln \mathcal{L}$. The likelihood function is the following:

$$F(\alpha, M_{top}) = [\alpha f_b(M_{top}) + (1 - \alpha) f_t(M_{top})] \tag{1}$$

2

where $f_b(M_{top})$ and $f_t(M_{top})$ are the normalized top mass distributions obained using the fitter on HERWIG[4] and VECBOS[5] samples; α is the expected fraction of background (n_b/N).

Backgrounds in lepton + jets analysis is mainly due to W + multijets production:
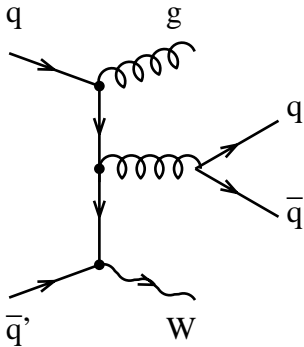
Rejection is obtained with b-tag (see also the talk on top production in this same proceeding[6]). The background to the remaining b-tagged events is calculated from the observed tag excess and measured b-tag efficiency. The major sources of background are:

- non W-events, WW, WZ, Drell-Yan, $Z \to \tau\tau$. They are derived from combination of data and MC.

- mistagged events. They are parameterized from generic jet data.

- $Wb\bar{b}$, $Wc\bar{c}$, Wc. They are determined from MC and scaled to the observed number of W events in a given jet multiplicity.

From these sources the expected number of background is $n_b = 6.4^{+2.0}_{-1.5}$.
The fitted value of background is 6.3 ± 1.7 events.

In figure 1 we report the top mass distribution for the 34 events and the likelihood fit.

We would like to spend here a couple of words on the statistical error on top mass. We check by MonteCarlo that the value of $\delta M = 5.7$ GeV is consistent with the expectations, performing a set of 1000 pseudo experiments of 34 events each with a background of 6.4 events. We used top mass distribution obtained fitting HERWIG MC at 175 GeV and VECBOS+HERWIG MC for the background. In figure 2 we sho the distribution of the masses from the 1000
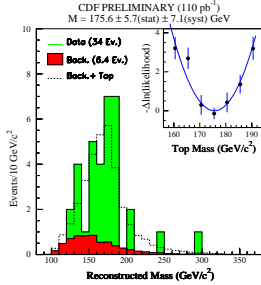
CDF PRELIMINARY (110 pb^{-1})
M = 175.6 ± 5.7(stat) ± 7.1(syst) GeV

Data (34 Ev.)
Back. (6.4 Ev.)
Back.+ Top

Events /10 GeV/c^2

Reconstructed Mass (GeV/c^2)

-Δln(likelihood)

Top Mass (GeV/c^2)

Monte Carlo Study: 1000 Samples of 34 Events

Samples/4 GeV

M = 174.8±0.2 GeV
σ = 6.6 GeV

M$_{data}$ = 175.6 GeV

Mass from Likelihood (GeV/c^2)

Samples/ GeV

σ$_{data}$ = 5.7 GeV

Mass Error for Each Sample (GeV/c^2)

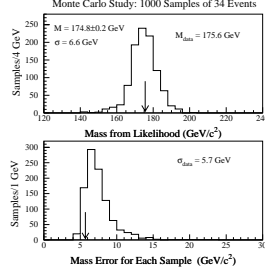Figure 1: Reconstructed top mass from
the single b-tag sample

Figure 2: Distribution of mass and un-
certainty for 1000 MC experiments of 34
events each, 6.4 of which are backgrounds.
Arrows show the real experiment.

pseudo experiments as well as the errors distribution. The arrow are pointing
to the data values.

3   Systematic uncertainties in lepton $+\geq$ 4 jets analysis

In Table I we report the systematic uncertainties in the top mass determination
for the lepton $+\geq$ 4 jets analysis.

The meaning of the different entries is the following: jet E_T scale, soft
gluon effects and hard gluon effects refer to the uncertainty we encounter re-
costructing parton from jets: we will discuss them below more extendively.
Uncertainty can also be due to the use of different MC generators (HERWIG
vs PYTHIA vs ISAJET) and fragmentation model. Fit configuration refers
to the use of different fitters: this entry has not been recently updated so the
values of the uncertainty is still large. B-taggin bias accounts for the effect of
b-tagging algorithm cuts on the shape of the signal and the background: even
in this case the estimated uncertainty is the same we quoted for the 1995 PRL.
The background spectrum can vary due to the different choices of the scale (
we obtain the background shape from a leading order MC, VECBOS, convo-
luted with a parton shower to account for jet activity), and for the presence of
contributions from production of W in association with heeavy flavours. The
uncertainty related to the likelihood method takes into account the possibility
to perform the likelihood fit leaving the background free or constraining it to
vary with gaussian shape or fixing it to the calculated value. Finally we quote
an uncertianty due to the finite size of the MC statistic.

Systematics	Value GeV	Value (%)
Jet E_T scale	3.1	1.8
Soft Gluon Effects	1.9	1.1
Different Generators	0.9	0.6
Hard Gluon Effects	3.6	2.1
Fit Configuration	2.5	1.4
b-tagging Bias	2.3	1.3
Background Spectrum	1.6	0.9
Likelihood method	2.0	1.1
Monte Carlo statistics	2.3	1.3
Total	7.1	4.0

Table I

The final result for the top mass in the lepton $+ \geq 4$ jets with one b-tagged is the following:

$$M_{top} = 175.6 \pm 5.7 \text{ (stat) } \pm 7.1 \text{ (syst) GeV/c}^2$$

Let us devote our attention to the jet energy corrections, a point of particular importance in deriving the systematic uncertainty in top mass determination.

4 Jet energy corrections

The mass fit requires going back from the jet to the parton energy. There are several sources of uncertainties in reconstructing the quark 4-momentum and for this reason we usually apply several sets of corrections, derived using MC events. The raw jet P_T (as calculated by the clustering algorithm) is corrected in the following way:

$$P_T^{corr} = P_T^{raw}(R) \times f_{rel} \times f_{abs}(R) - UE(R) - OC(R) \qquad (2)$$

where f_{rel} is the η response relative to the central region; $f_{abs}(R)$ is the absolute energy scale correction; UE(R) takes into account the underlying events effects; and OC(R) is the correction factor for the jet energy outside the cone of radius R.

The corrections are derived for a topology with single jets and are then *corrected* to the top topology with a further set of corrections (top specific or AA corrections). On each of these corrections we have a systematic uncertainty that is going to propagate to the fitted top mass.

The general philosophy to determine the systematic effects of these corrections is the following: we have a procedure we will call P:

P = (lepton and jet corr. + AA corr. + mass fitter + likelihood)

5

For each event we make our best (pre-AA) estimate of energies and angles of jets, lepton, and X. We input these quantities into P and obtain M_{out}, our best estimate of the top mass. The likelihood method uses some templates generated with the MC and that are part of the procedure. P is in fact calibrated with MC simulations. We assume that the MonteCarlo we used has the $t\bar{t}$ production and decay done correctly, and has the experimental smearing done properly. We find on average that $M_{in} = M_{out}$.
What if the energy is wrong?
Then the MC incorrectly simulates the experiment. We then rerun the calibration with an *amended* MC. For example, if the energy scale has a 10% uncertainties, then all energy scales are multiplied in turn by $f = 1.1$ or 0.9 and M'_{out} is compared to M_{out} for the different values of f. The difference is the appropriate systematic error.
The same kind of procedure and calibration is also performed with different type of generators.

4.1 "Detector effects" systematic error

The systematic error on f_{rel} is $\pm 2\%$. The effect of calorimetry stability is $\pm 1\%$ for RUN 1A and $\pm 2\%$ for RUN 1B; UE(R) is estimated to be 30% of the corrections itself (100 MeV for cone of 0.4); and finally the P_t-dependent response of the central calorimeter is given in terms of a polynomial function for cones of radius R=0.4.
The global uncertainty is 4.1% at 8 GeV to 3.7% at 150 GeV.

4.2 Soft gluon radiation uncertainties

Information on soft gluon comes from annulus studies of radiation between cones of different radius in the jets. We study the jet shape, i.e. the energy flow in the cone centered around the jet direction, defining the following quantity:
$$F = [P_T(1.0)\text{-}P_T(0.4)]/P_T(0.4)$$
where $P_T(1.0)$ $(P_T(0.4))$ is the fully corrected jet P_T using a cone of R=1.0 (R=0.4).
The difference DF between data and MC provides the systematic uncertainty on the top mass from this source. Three different data samples are used to make this comparison: W + 1 jet, Z + 1 jet and $b\bar{b}$ data. We show in figure 3 the result for W + 1 jet.
 If the MonteCarlo simulates the data correctly, we should find DF consistent with zero. That DF is not consistent with zero implies that the MonteCarlo is not a perfect simulation. We will assume that the sole reason for the discrepancy is soft gluon emission.
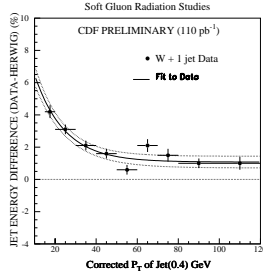
6

Figure 3: Jet Annulus study. Difference in energy in an annulus between R=0.4 and R=1.0. The difference between data and HERWIG MC is shown.

4.3 Hard gluon radiation uncertainties

We define *gluon* jet a jet with $\Delta R > 0.4$ from the closest parton from top. In HERWIG at $M_{top} = 175$ GeV 55% of the events have such *gluon* jet among the four highest E_t jets. Difference in the reconstructed top mass when this percentage in varied from 25% to 85% ($\pm 1\sigma$ on a flat distribution with mean of 55%) gives us the systematic uncertainty. The change in mass mean is about 1 GeV/c^2, while the change in mass error is about 3.5 GeV/c^2.
Added in quadrature we quote a total uncertainty of $\Delta M = 3.6$ GeV/c^2.

5 Mass analysis for the lepton $+ \geq 4$ jets sample with two b-tagged jets

This sample is a subsample of the one used for the lepton + jets analysis. We start from the lepton $+ \geq 3$jets ($E_T^{jet} > 15$ GeV GeV and $|\eta| < 2.0$) and require a fourth jet with $E_T > 8$ GeV and $|\eta| < 2.4$.
A secondary b-tagging in addition to the standard vertex algorithms (SECVTX and SLT) is performed: jets with *jet probability* less than 5% will be considered b-jets. Jet probability is defined by comparing the tracks in the jet with positive signed impact parameters to the measured SVX resolution. This distribution is flat for zero-lifetime jets, while it is sharply peaked at zero for b-jets. Physically it represents the probability that the observed set of impact parameters is consistent with resolution effects alone.

We say we have b-tag if jet probability is less than 5%. In fact charm and strange-intiated jets produce also zero peaked distributions, but less sharply
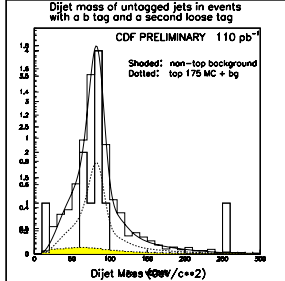
7

Figure 4: Hadronic W's by tagging two jets and plotting the dijet invariant mass of the remaining two. Superimposed are the MC expectations for background and top.
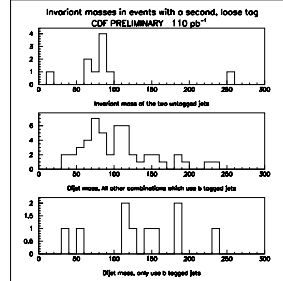
Figure 5: Comparing the previous plot with other dijet invariant masses in the same events: those combinations using one tagged jet and those using both tagged jets.

peaked than b-jets. We place an additional cuts on the reconstructed mass of the two non b-jets: $60 < M_{jj} < 100$ GeV/c^2. This cut will reduce the background by a factor 3 (from 1.3 to 0.4).

In figure 4 we show the reconstructed dijet invariant mass of the two untagged jets. Superimposed are MonteCarlo expectations for background and top. In figure 5 we compare the previous plot with other dijet invariant masses in the same events; those combinations using one tagged jet, and those combinations using both tagged jets. Finally in figure 6 we show the results for the fit on the 9 fully reconstructed events of this sample.

Systematics	Value GeV	Value (%)
Jet E_T scale	2.9	1.7
Soft Gluon Effects	1.7	1.0
Different Generators	0.9	0.5
Hard Gluon Effects	3.6	2.1
Fit Configuration	0.9	0.5
b-tagging Bias	2.0	1.1
Background Spectrum	0.1	< 0.1
Likelihood method	0.6	0.3
Monte Carlo statistics	1.0	0.6
Total	5.6	3.2

Table II

The resulting mass (statistical errors only) is $M_{top} = 174.8 \pm 7.6$ (stat) GeV/c^2. In Table II we show the systematic uncertainties in top mass measurement on 2 b-tagged W+$\geq$ 4 jet events. We refer to the previous section
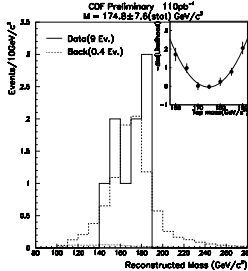
Figure 6: Result of mass fitting on double b-tagged events with constrained background events.

for the explanation of the meaning of the different entries.

The final result is: $M_{top} = 174.8 \pm 7.6$ (stat) ± 5.6 (syst) GeV/c^2

6 Mass analysis for dilepton events

The event selection is performed as follows: we chose the two highest E_t jets with $E_t > 10$ GeV and $|\eta| < 2.0$; the leptons are requested to be central and oppositely charged: they have $(P_T > 20$ GeV, $|\eta| < 1.0)$. We place a cut on H_T: $H_t = \sum E_T(l_1) + E_T(l_2) + E_T(j_1) + E_T(j_2) + \not{E}_T >170$ GeV. We fins 8 events with 1.1 background event expected.

In this sample we cannot perform a kinematic fitting as for the two other ones. In fact it is not possible to reconstruct the top mass from the measured final state because of the presence of the two neutrinos which make the fit unconstrained. The likelihood is then obtained by comparing the energy distribution of the two highest E_T jets from the data to the templates of top + background MC. There is a caveat to this kind of approach: we assume that the kinematical distributions for top production are well reproduced by top MC. This has indeed been proved to be true for the E_T of tagged jets in the lepton + jets sample and for the P_T of the $t\bar{t}$ system.

The mass is calculated from the four momentum of the 2 leptons, two leading jets and missing transverse energy. The P_z component of the two neutrinos is set equal to zero. The likelihood function has the following form:

$$\mathcal{L}_m = \Pi_i \mathcal{L}_m^i = \Pi_i \frac{n_B \times B(x_i) + n_t T(x_i)}{n_B + n_t} \quad (3)$$
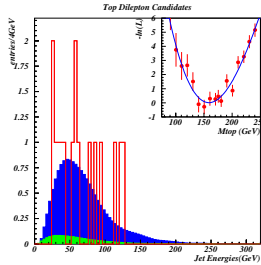
9

Figure 7: Jet energy distribution of
the two highest E_T jets for the dilepton
events (histogram). Superimposed is the
same distribution for MC events + back-
ground and background alone.

In figure 7 we show the resulting jet energy distribution and the likelihood.
The final result is $M_{top} = 159^{+24}_{-22}$ (stat) ± 17 (syst) GeV.

The calculated background for this sample of events is 2.0 ± 0.4 events
reduced to 1.1 ± 0.3 when the H_T cut is applied. In Table III we show the
different source of background both derived from data and MC.

Background type	Expected # of events	after H_T cut
Drell-Yan	0.69	0.35
$Z \to \tau\tau$	0.60	0.26
Fake leptons	0.38	0.29
WW	0.33	0.24
$b\bar{b}$	0.03	
Total	2.0 ± 0.4	1.1 ± 0.3

Table III

We note here that the H_T cut make results less sensitive to background
fluctuation, even if it of course introduces a bias. Applying or not the H_T cut in
fact shift the top mass by 5 GeV. We also take the spread in the mass values
about our nominal background point resulting from the various background
assumptions, weighted by their Poisson probability, as an additional statistical
uncertainty of 2.5%. Finally we check with other kinematical distributions for
the consistency of the results. The uncertainties in the dilepton samples are
slightly different that in the other two: in Table IV we report them. The mean
statistical uncertainty has been studied fitting the jet energy of a large number
of sets with 10, 20 , 50 and about 1000 events. This uncertainty scales with
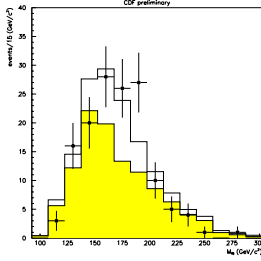
10

Figure 8: Best χ^2 top mass after kinematical selection. Bullet = b-tagged events, shaded = background normalized to its estimate, white = backgrounds + top contribution (Herwig, M_{top} =175 GeV/c^2).
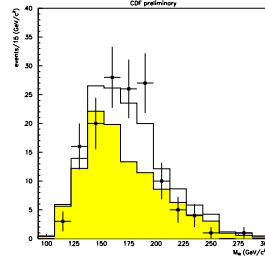

Figure 9: The same as previous plot with Herwig, M_{top} =185 GeV/c^2.

$\sqrt{n}$.

Source	Error (%)
Statistical Uncertainties	
Likelihood fit	14.5
Template Statistics	2.0
Total (statistical)	14.6
Systematic Uncertainties	
Jet energy scale	< 10
Background Shapes	5.0
Structure Functions	2.0
Different Top MC	2.0
Total (Systematics)	11

Table IV

7   Mass analysis for the multijets sample

The counting experiment for the multijets sample has been extensively described in another talk in this same proceeding[7].

For the mass analysis the event sample is kinematically selected in the following way: $N_j et \geq 6$ (with $E_t > 15$ GeV, $|\eta| < 2.0$ and $\Delta R_{min} \geq 0.5$); $\sum E_t \geq 200$ GeV and $\sum E_t/\sqrt{s} \geq 0.75$; aplanarity$+0.0025 \times \sum_3^N E_T \geq 0.54$ and at least one SVX b-tag.

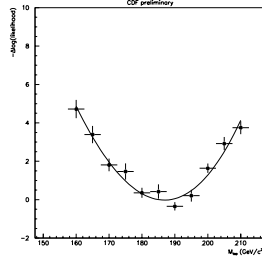The fit is a 4-vertices fit:

1) $t_1 \rightarrow j_1 + j_2 + b$

11

Figure 10: Likelihood fit of the top mass
for the multijets sample.

2) $t_2 \rightarrow j_3 + j_4 + \bar{b}$

3) $W^+ \rightarrow j_1 + j_2$

4) $W^- \rightarrow j_3 + j_4$

We have with 16 equations, 13 unknowns and then a 3 constraints fit. We chose the best χ^2 solution ($\chi^2 < 10$). We have 142 one b-tagged events and 28 double tagged events.

The fit is applied to the multijet sample distinguishing between events with or without b-tagged jets. For the latter $S/B \sim 1/30$ so we can infer the spectrum shape for backgrounds events from this *depleted* sample. In figure 8 (9) we show the best χ^2 top mass after kinemtical selection. The dotts are b-tagged data events, the shaded histogram is the background normalized to its estimate and the white is the background + top for HERWIG top175 (HERWIG top185). The likelihood is shown in figure 10. The final result is $M_{top} = 187 \pm 8$ (stat) ± 12 (syst) GeV, with $n_s = 41.2 \pm 11.0$ and $n_b = 108.8 \pm 6.6$.

The background is calculated by applying the tag rate parametrization obtained on a sample of $N_{jet} = 4$ events ot the kinematically selected sample. In Table V we report the total systematic for the multijet sample:

12

Systematics	Value (%)
Jet E_T scale	+2.9 -2.1
Soft Gluon Effects	±1.6
Different Generators	+0.3 -1.0
Hard Gluon Effects	±4.3
Fit Configuration	±2.1
b-tagging Bias	< 0.1
Background Spectrum	-1.5
Likelihood method	+0.4 -2.4
Monte Carlo statistics	±0.3
Total	+6.6 -6.3

Table V

8   Conclusions

We reported the results on top mass measurement obtained in 110 pb^{-1} of data collected with the CDF detector. We presented the results for the different channels and described in details the possible sources of systematic uncertainties. After one year from the top quark discovery, the focus of CDF is aimed to a deeper understanding of its properties, and as far as the top mass is concerned work is in progress in order to combine the results of the different channels, as it has been done for the cross section.

References

1. F. Abe *et al.* (CDF Collaboration), Phys. Rev. **D50**, 2966 (1994)
2. F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **74**, 2626 (1995)
3. S. Abachi *et al.* (D0 Collaboration), Phys. Rev. Lett. **74**, 2632 (1995)
4. G. Marchesini et al., Comp. Phys. Comm. 67(1992)465
5. F.A. Berends et al., Nucl. Phys. **B 357** 32 (1991)
6. Mark Kruse, Production and Decay of Top at CDF, this proceeding
7. Sandra Leone, New evidence for Top at CDF, this proceeding