MEASUREMENT OF THE TOP QUARK MASS AT DØ

SERBAN PROTOPOPESCU

Brookhaven National Laboratory, Upton, NY 11973, USA

For the DØ Collaboration

The mass of the top quark is measured using a sample of 93 lepton + 4 or more jets events collected with the DØ detector at the FNAL Tevatron collider. We find the top quark mass is $169 \pm 8(stat.) \pm 8(syst.)$ GeV/c^2.

Since the discovery of the top quark[1], DØ has more than doubled its data sample and improved its methods for measuring the top quark mass. Preliminary results are presented for final states with a single isolated lepton and four or more jets for an integrated luminosity close to 115 pb^{-1}. The analysis assumes that top quarks are produced as $t\bar{t}$ pairs that decay to W bosons and b quarks. Our final states result when one W decays to $e\nu$ or $\mu\nu$ and the other W to $q\bar{q}$. More than four jets may be present because of final and initial state radiation.

The DØ detector and particle identification techniques are described in detail elsewhere[2]. Events are selected by requiring an isolated e or μ, with $E_T^l > 20$ GeV, and $|\eta_e| < 2.0$ or $|\eta_\mu| < 1.7$, missing E_T ($\not{E}_T$)> 20 GeV, at least 4 jets (cone $R \equiv \sqrt{\Delta\phi^2 + \Delta\eta^2} = 0.5$) with $E_T > 15$ GeV and $|\eta(jets)| < 2.0$, $\not{E}_T^{cal} > 25$ GeV for e+jets, and $\not{E}_T^{cal} > 20$ for μ+jets (where $\not{E}_T^{cal}$ is calculated by summing E_x and E_y of all cells in the calorimeter). The events are separated into two classes: 8 with a μ-tag and 85 without. A μ-tagged event has a μ within $R = 0.5$ of a jet with $p_T > 4$ GeV and $|\eta| < 1.7$. Events without a μ-tag must satisfy two additional requirements that significantly reduce the background from events without W bosons: $E_T^W (\equiv \not{E}_T + E_T^l) > 60$ GeV and $|\eta_W| < 2.0$.

For every event, a mass ($m_{\rm fit}$) is calculated assuming that the event has a $t\bar{t}$ pair of unknown mass and that the 4 jets with highest E_T are either from $W \rightarrow q\bar{q}$ or the b jets. If one assumes $\not{E}_T$ is due to the ν from the W decaying leptonically, there is only one unmeasured quantity, p_z^ν, with three constraints: two from requiring two W bosons in the event and one from requiring the masses of t and $\bar{t}$ to be equal. There are 12 possible ways of assigning jets and 2 solutions for $p_z(\nu)$ for a total of 24 solutions (12 for μ-tag events).

Jet energies are corrected to more closely approximate the 4-momenta of the original partons. These corrections are derived using $t\bar{t}$ HERWIG Monte Carlo events, and the procedure was checked using Z+jets data as well as Monte Carlo[3]. From this study the uncertainty on the overall hadronic energy scale is determined to be $\pm$ (4%+1 GeV/c^2) and limited by the statistics of Z+jets events.

Of the 93 events, 73 have at least one mass fit solution with $\chi^2 < 7$. If more than one solution satisfies this requirement, $m_{\rm fit}$ of the solution with smallest χ^2 is chosen. The $m_{\rm fit}$ distribution for Monte Carlo $t\bar{t}$ events peaks at the correct m_t but has a width that is dominated by jet combinatorics and not by the intrinsic detector resolution.

Most of the events without a μ-tag are not $t\bar{t}$ events and we therefore had to devise ways to distinguish signal from background. We have identified four kinematic variables weakly correlated with the mass of the top quark (m_t), that have very different distributions for signal and background[4]:

$$v_1 \equiv \not{E}_T$$

$$v_2 \equiv \mathcal{A} \equiv \tfrac{3}{2} \times \text{least eigenvalue of } \mathcal{P}$$

$$v_3 \equiv \frac{H_{T2} \equiv H_T - E_T^{j1}}{H_\parallel}$$

$$v_4 \equiv \frac{K_{T\min} \equiv (\text{min of 6 } \Delta\mathcal{R}_{jj}) \cdot E_T^{\text{lesser } j}}{E_T^W} .$$

where $\Delta\mathcal{R}_{jj}$ is the distance in η, ϕ between any two jets, $\mathcal{A}$ is the aplanarity and $\mathcal{P}$ is the normalized momentum tensor of the jets and the W in the laboratory frame, H_T is the scalar sum of E_T of the jets, and $H_\parallel$ is the scalar sum of $|p_z|$ of the jets, charged lepton, and neutrino. The variables $v_5 \equiv H_{T2}$ and $v_6 \equiv K_{T\min}$ are more strongly correlated with m_t.

To exploit the differences between signal and background we use two multivariate techniques:
1) Construct a top likelihood discriminant ($\mathcal{D}$) using normalized signal to background ratios of v_1-v_4.
2) Use neural network (NN) techniques[5] to construct a top probability discriminant ($top_{\rm prob}$) using variables v_1-v_6.

In the first method we define a normalized ratio of distributions

$$\ln \text{top/bkgnd} \equiv \ln \mathcal{L}_i(v_i)$$

and form a "likelihood" $\mathcal{L}$ for each event

$$\ln \mathcal{L} \equiv \sum_i \omega_i \ln \mathcal{L}_i$$

where ω_i is a constant weight. Ordinarily, all the ω_i would be unity. We fix the ω_i to nullify any correlation between $\mathcal{L}$ and m_t. Because the variables v_1-v_4 are only weakly correlated with m_t the ω_i's are close to 1. Finally we construct

$$\mathcal{D} \equiv \frac{\mathcal{L}}{1+\mathcal{L}} .$$

In the second method we construct a discriminant $top_{\rm prob}$ defined by

$$top_{\rm prob} = \frac{s(x)}{s(x)+b(x)},$$

where s and b are the signal and background distributions of x, with $x = n2 \times n4$. Here $n2$ is the output of a two-variable NN using v_5 and v_6 and $n4$ is that of a four-variable NN using v_1-v_4.

The neural networks were trained using a Monte Carlo sample of HERWIG generated $t\bar{t}$ events with m_t =180 GeV/c^2 as signal and as background a Monte Carlo sample of VECBOS[6] generated W+4 or more jets mixed with 20% non-W background from data. Figure 1 shows the expected distributions of $top_{\rm prob}$ as function of $m_{\rm fit}$ for expected top signal (for m_t = 170 GeV/c^2), expected background and actual data. Note that the signal and background have very distinct distributions and the data show contributions from both. We found all variables to be well reproduced by the Monte Carlo; Fig. 2 shows examples for observed and expected distributions v_3 and v_5 for events with only 3 jets (10% $t\bar{t}$), and $\geq$ 4 jets (35% $t\bar{t}$).

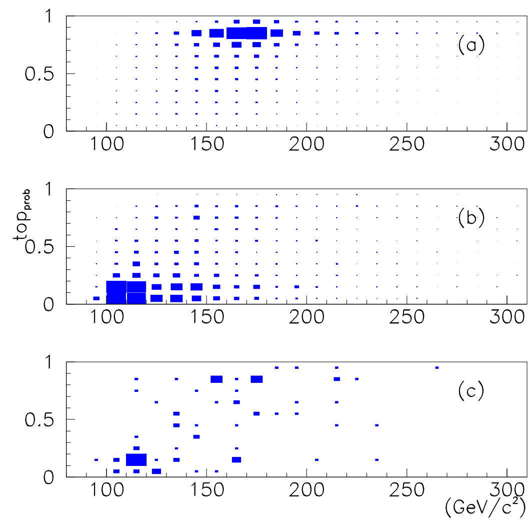

Figure 1: Scatterplots of $top_{\rm prob}$ vs $m_{\rm fit}$ for a) $t\bar{t}$ Monte Carlo events with m_t=170 GeV/c^2, b) background, c) data.

To determine the signal to background ratio and the most likely value of m_t, we binned the data and Monte Carlo in 10 GeV bins of $m_{\rm fit}$ and six regions of top "quality":
1) All μ-tag events.
2) $H_{T2} \geq 90\,$GeV, $\mathcal{D} \geq 0.59$.
3) $H_{T2} \geq 90\,$GeV, $0.43 \leq \mathcal{D} < 0.59$.
4) $H_{T2} \geq 90\,$GeV, $0.27 \leq \mathcal{D} < 0.43$.
5) $H_{T2} \geq 90\,$GeV, $\mathcal{D} < 0.27$.
6) $H_{T2} < 90\,$GeV.
The entire sample of events (all 6 regions) is refered to as PR (precut), and contains an estimated signal-to-background ratio (S/B) of $\approx$1/2. The first 3 regions are $t\bar{t}$ enriched and have S/B$\approx$ 2/1, we refer to that subsample of 32 events as LB (low bias). Figure 3 shows the expected distributions in $m_{\rm fit}$ for $t\bar{t}$ events with m_t = 150 and m_t = 170 GeV/c^2, W+jets background and non-W multijet background for the PR and LB samples. Note that the LB samples have nearly the same peak values of $m_{\rm fit}$ for $t\bar{t}$ Monte Carlo events and have large acceptance ($\approx$ 80%) relative to PR, while the backgrounds are reduced by a factor of 5 and display no significant peaking in $m_{\rm fit}$.

To extract m_t we maximize the following Poisson-statistics likelihood function using dis-

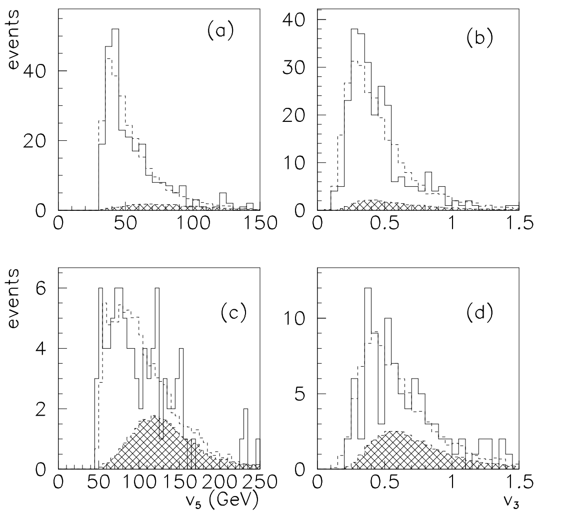

Figure 2: Observed and predicted distributions for variables v_5 and v_3 (see text), solid histogram is observed, dashed is predicted sum of signal and background, shaded is predicted signal (m_t=170 GeV/c^2): a) v_5 for lepton + 3 jets (exclusive), b) v_3 for lepton + 3 jets (exclusive), c) v_5 for lepton + 4 jets (inclusive) d) v_3 for lepton + 4 jets (inclusive)

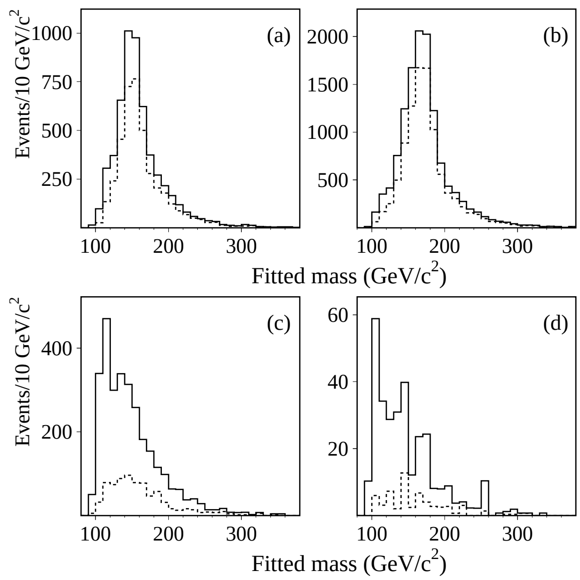

Figure 3: Distributions of $m_{\rm fit}$ for $t\bar{t}$ Monte Carlo events a) m_t= 150 GeV/c^2 b) m_t= 170 GeV/c^2, and background c) W+jets d) non-W multi-jet. Solid (dashed) histograms are for events satisyfing PR (LB) criteria (see text).

crete values of m_t

$$L(m_t, p_s, p_b) = \prod_j q(N_j, p_s a_j^s + p_b a_j^b) \, q(A_j^s, a_j^s) \, q(A_j^b, a_j^b), \qquad (1)$$

where N_j is the number of observed events, and $A_j^s(a_j^s)$ and A_j^b (a_j^b) are the number of generated (true) signal(s) and background(b) events in any bin j; $q(N, a)$ is the Poisson probability $e^{-a}a^N/N!$; $p_s = n_s/n_o$ and $p_b = n_b/n_o$; n_s (n_b) is the expected number of signal (background) events, and n_o is the total number of observed events. The total number of bins is the number of bins in $m_{\rm fit}$ multiplied by the number of regions of top "quality".

Monte Carlo samples of 100,000 events were generated with HERWIG 5.7 as function of m_t every 5 GeV/c^2. Data samples with poor electron fits or non-isolated muons were used for the non-W background, along with $W+\geq 4$ jets Monte Carlo background. We searched for the values of n_s and n_b that maximized the likelihood for the PR sample as a function of m_t. The 5 points closest to the minimum of $-\ln L$ were fitted to a quadratic form to find the minimum value, and the closest 9 points to a cubic to determine the uncertainty. The same procedure was repeated for the LB sample, but in this case n_s was constrained within errors to the amount expected from the fit to the PR sample. The distribution in $m_{\rm fit}$ for the LB sample and the fits to $-\ln L$ are shown in Fig. 4.

The PR and LB samples were also fitted using 10 bins of $top_{\rm prob}$, instead of top "quality", to calculate a Bayesian posterior probability:

$$P(n_s, n_b, m_t) = \frac{L \cdot p(n_s, n_b, m_t)}{\int L \cdot p(n_s, n_b, m_t) dn_s dn_b dm_t},$$

where $p(n_s, n_b, m_t)$ is the prior probability , chosen here to be flat. The probability as a function of any of the variables (n_s, n_b or m_t) can be obtained by integrating the posterior probability over the other variables. The mean and width of the probability distributions provide the most likely value and its uncertainty. Figure 4(c) shows the m_t probability distribution for the LB sample.

The estimated values and statistical errors from both methods from the PR and LB samples are given in Table 1. The estimates for the mass and its error were checked by generating a large number of ensembles varying the expected number of signal events. For an input m_t of 170 GeV/c^2 the ensemble fits give an average mass of

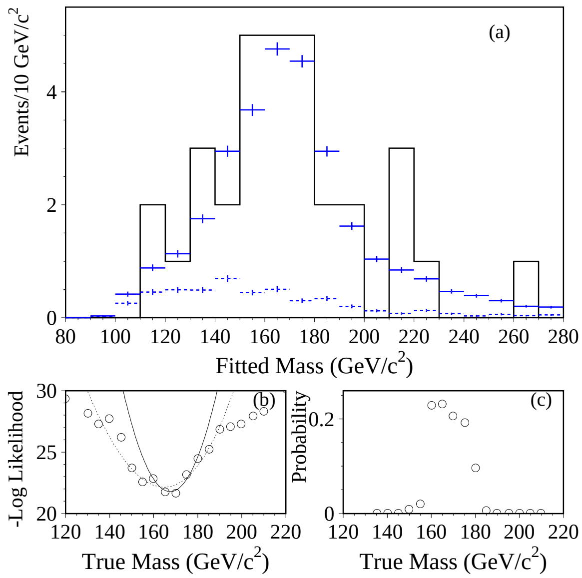

Figure 4: a) Distribution of $m_{\rm fit}$ for the LB data sample, where the solid (dashed) crosses are predictions from fitted values for the sum of signal and background (just background). (b) Shows the likelihood as function of m_t and the curves used to determine mass and error for method 1. (c) Shows the probability as function of m_t for method 2.

169 GeV/c^2 and an error of 7 GeV/c^2. The expected error is close to that obtained in Table 1, while the mass has a 1 GeV/c^2 shift, for which we correct.

The systematic error on the mass estimate is dominated by uncertainties in corrections of the energy scale that amount to 6.9 GeV/c^2. The next largest uncertainty comes from Monte Carlo modeling. This is estimated by comparing HERWIG and ISAJET Monte Carlo for signal, and by generating background W + jets events with different scales in VECBOS. The total uncertainty from modeling is estimated as 3.3 GeV/c^2. Other sources of systematic errors were found to be small compared to the two mentioned above. The total systematic error is 8 GeV/c^2. Our best estimate of the top quark mass is therefore $169 \pm 8(stat.) \pm 8(syst.)$ GeV/c^2.

In conclusion, we analyzed a sample of events with a single isolated lepton and four or more jets obtained with a 115 pb^{-1} exposure of the DØ detector at the Tevatron collider to determine the top quark mass. Multivariate techniques were used and very similar results were obtained with two different methods. The measured top quark

Table 1: Results of fits to PR and LB samples, for the two methods as indicated in the first column.

		m_t (GeV/c^2)	n_s	n_b
1	PR	168 ±10	27.5±7.0	45.5±10.0
	LB	168±8	$24.5^{+3.7}_{-5.0}$	$4.9^{+7.7}_{-2.2}$
2	PR	169 ±10	26.4±7.6	39.5±7.6
	LB	168 ±7	26.6±5.5	2.4±2.0

mass is $169 \pm 8(stat.) \pm 8(syst.)$ GeV/c^2.

We thank the staffs at Fermilab and the collaborating institutions for their contributions to the success of this work, and acknowledge support from the Department of Energy and National Science Foundation (U.S.A.), Commissariat à L'Energie Atomique (France), Ministries for Atomic Energy and Science and Technology Policy (Russia), CNPq (Brazil), Departments of Atomic Energy and Science and Education (India), Colciencias (Colombia), CONACyT (Mexico), Ministry of Education and KOSEF (Korea), CONICET and UBACyT (Argentina), and the A.P. Sloan Foundation.

References

1. CDF collaboration, F. Abe *et al.* *Phys. Rev. Lett.* **74**, 2626 (1995). DØ collaboration, S. Abachi *et al.* *Phys. Rev. Lett.* **74**, 2632 (1995).
2. DØ collaboration, S. Abachi *et al.* *Phys. Rev.* D **52**, 4877 (1995), *Nucl. Instrum. Methods* **A338**, 185 (1994).
3. G. Marchesini *et al.*, Comp. Phys. Commun. 67, 465 (1991)
4. DØ collaboration, M. Strovink to be published in Proceedings 11th Topical Workshop on $p\bar{p}$ Collider Physics, Padova, Italy (1996), Fermilab-Conf-96/336-E.
5. E.K. Blum and L.K. Li, Neural Networks 4,511(1991), D.W. Ruck et al., IEEE Trans. Neural Networks 1,296 (1990). DØ collaboration, P.C. Bhat in Proceedings 10th Topical Workshop on $p\bar{p}$ Collider Physics, FNAL, p.308 (1995), Fermilab-Conf-95/211-E.
6. F.A. Berends *et al.*, *Nucl. Phys.* B **357**, 32 (1991); ISAJET was used for fragmentation, F. Paige and S. Protopopescu, BNL report No. BNL38034, 1986 (unpublished).