International Journal of Modern Physics A

CDF MEASUREMENT OF THE TOP QUARK MASS IN THE LEPTON+JETS CHANNEL USING THE MULTIVARIATE TEMPLATE METHOD

JOHN FREEMAN
on behalf of the CDF Collaboration

Lawrence Berkeley National Laboratory
1 Cyclotron Road, Bldg. 50B-5239
Berkeley, CA 94720, USA

Received (Day Month Year)
Revised (Day Month Year)

We measure the mass of the top quark using 162 pb^{-1} of data collected by the CDF experiment at the Fermilab Tevatron in Run II. The decay chain $t\bar{t} \to bq\bar{q}\bar{b}l\nu$ is studied using a novel technique called the Multivariate Template Method (MTM). Using this technique we obtain a result of $M_{top} = 179.6^{+6.4}_{-6.3}(\mathsf{stat.}) \pm 6.8(\mathsf{syst.})$ GeV/c^2 for the top quark.

Keywords: top; CDF; Tevatron

1. Introduction

The MTM technique was developed as an alternative to the earlier Run I method used to measure the top quark mass in the lepton+jets channel.[1] It is designed in particular to compensate the systematic error introduced by the jet energy scale (JES) in the CDF detector, which to date has dominated the total Run II systematic error. The MTM method possesses several noteworthy features. A kinematic fit to the hadronic W mass using the JES as a variable is designed to reduce the systematic error. The statistical error is reduced by an estimation of the probability that the correct jet-parton assignment in an event was selected. Furthermore, we use of event variables in addition to the reconstructed mass in the likelihood function, which should provide more information and improve signal/background discrimination. Lastly, we employ Kernel Density Estimation in our likelihood calculations, a non-parametric method for constructing a continuous density function from a discrete number of values [2]

2. Reconstructing the top mass from an event

The candidate $t\bar{t}$ events selected were essentially the same as those used in the CDF lepton+jets cross section measurement. [3] In calculating the top mass for each of those events, a fit using the JES as a variable is performed on the hadronic W mass for every possible jet permutation in an event. The mass from the permutation yielding the best fit χ^2 is taken as the event's mass. The Gaussian constraint on the JES is a tunable parameter in our analysis. In general, the looser the constraint, the more we can reduce the systematic error, but since the calculated JES now depends on fluctuations of jet fragmentation in the $q\bar{q}$ system the statistical error increases. See Section 4 for a discussion of this parameter's optimization.

3. The Likelihood Calculation

The likelihood function used for our set of candidate ttbar events is as follows

$$L(m_t) = \prod_{i=1}^{N} (f_b P_b(m_i, x_i) + (1 - f_b) P_s(m_i, x_i, m_t)) \tag{1}$$

where f_b is the background fraction, P_s and P_b are the signal and background density functions, respectively, m_t is the Monte Carlo (MC) top mass and m_i and x_i are the reconstructed mass and any other event observables used in the density functions.

The likelihood is calculated using $t\bar{t}$ MC samples at 2.5 GeV$/c^2$ intervals. Local quadratic polynomial least-squares regression is used to introduce continuity into the likelihood curve. For each m_t, the background fraction is allowed to float so as to maximize the likelihood value.

The value of an event's density in P_s or P_b is actually a weighted sum of its density in different classes of background and signal density spaces known as "templates", i.e.

$$P_s(m, x, m_t) = \sum_{\text{sig types}} c_j S_{j,m_t}(m, x), \qquad \sum_{\text{sig types}} c_j = 1 \tag{2}$$

and a similar formula for $P_b(m_i, x_i)$. The backgrounds we used were $W + 4$ light jets, $W + b\bar{b}$ jets, and non-W. The S_{j,m_t} in (2) are the different signal event templates. They consist of MC events in which the correct permutation was chosen, those in which the incorrect permutation was chosen, and those in which at least one of the leading four jets was incorrectly chosen. A motivation for this is that the width of the reconstucted mass distribution of good permutation events was approximately one-third that of the other two signal templates.

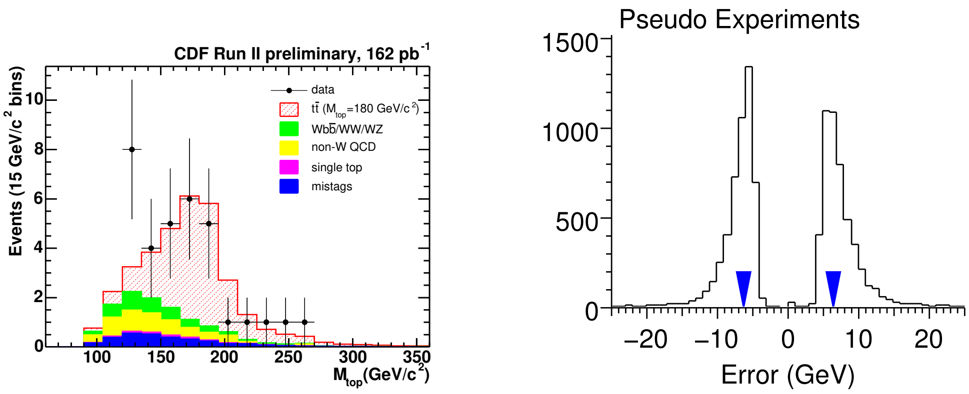
The relative background weights for every event were constant, whereas the weights for the signal templates were calculated on an event-by-event basis. These calculations were performed using a formula which employed the difference between the chisquare of the best permuation and the other permutation chisquares, as well as a Bayesian update using angular correlation variables.

4. Choices of Variables and JES Constraint

For the mass measurement, two major decisions were the choice of event observable(s) to use in addition to the reconstructed mass in the likelihood function and the choice of the JES constraint to use in the kinematic fit. Performing several thousand pseudo-experiments (PEs) with different sets of observables did not appear to indicate an ideal choice within the context of the expected total error for a single PE. As a result, we chose an observable which statistical divergence tests (Kolmogorov-Smirnov and K-divergence) indicated would be best for distinguishing signal from background, the scalar sum of the transverse momenta of the leading four jets. It was also found from the PEs that the expected total error was typically lowest at a JES constraint of 0.07 .

5. Mass Measurement

The value we obtained for the top mass is $M_{top} = 179.6^{+6.4}_{-6.3} \pm 6.8$ GeV$/c^2$. The individual reconstructed masses of our data sample are shown superimposed on MC expectation in Fig. 1a . This error is in agreement with the expected total errors shown in Fig. 2a . The value of the floating background fraction which maximized the likelihood at the measured value was 34%. In obtaining the systematic error of 6.8 GeV$/c^2$, numerous systematic sources were included, but the systematics on the jet measurement alone contributed 6.7 GeV$/c^2$.



(a) (b)

Fig. 1. Plot (a) shows the individual reconstructed masses of our data sample superimposed on the expected MC values for a mass of 180 GeV$/c^2$ and a 34% background fraction. Plot (b) indicates the expected positive and negative errors on our measurement, with the blue arrows marking our actual result

References

1. T. Affolder *et al. Phys. Rev. D* **63**, 032003 (2001).
2. D. Scott, *Multivariate Density Estimation : Theory, Practice, and Visualization* Wiley-Interscience, 1992.
3. D. Acosta *et al.* hep-ex/0410041