CDF Measurement of the Top Quark Mass in the Lepton+Jets Channel Using the Multivariate Template Method

John Freeman
Lawrence Berkeley National Laboratory

Outline

- Basic description of the Multivariate Template Method (MTM) analysis
- Distinguishing features of the analysis
- Top mass reconstruction technique
- How we compute the likelihood
- Choices we made for our data measurement
- The top mass result
- Discussion of systematics

Overview of the Analysis

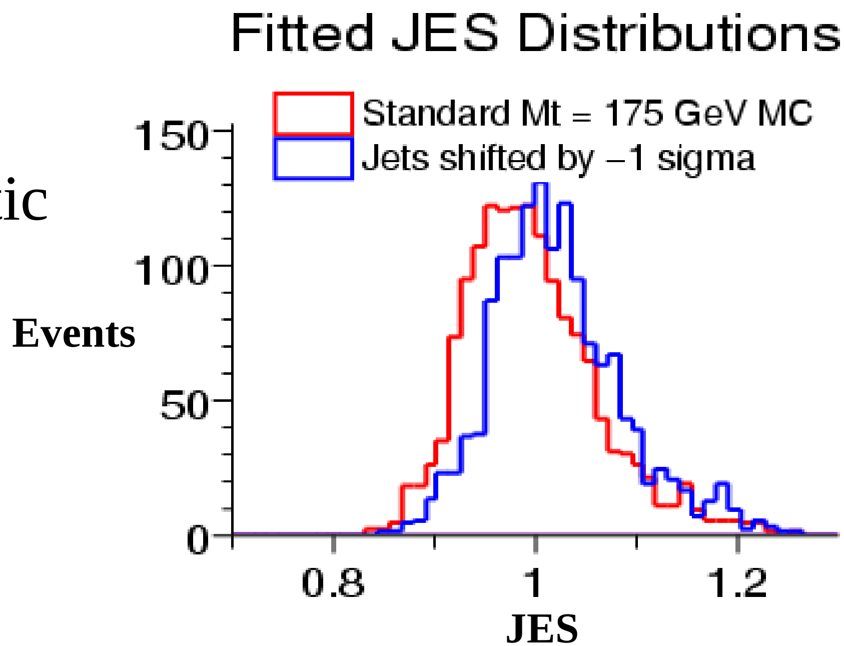
- We measure the top mass in the lepton + jets channel ($t\bar{t} \rightarrow bq\bar{q}\bar{b}l\nu$) using 1 or 2 b-tagged jets
- Analysis works as follows:
  - Extract top mass and other observables from each event
  - Construct density functions of these observables using events from both background and signal MC samples
  - Mass measurement is taken from the joint likelihood of our events calculated with these density functions

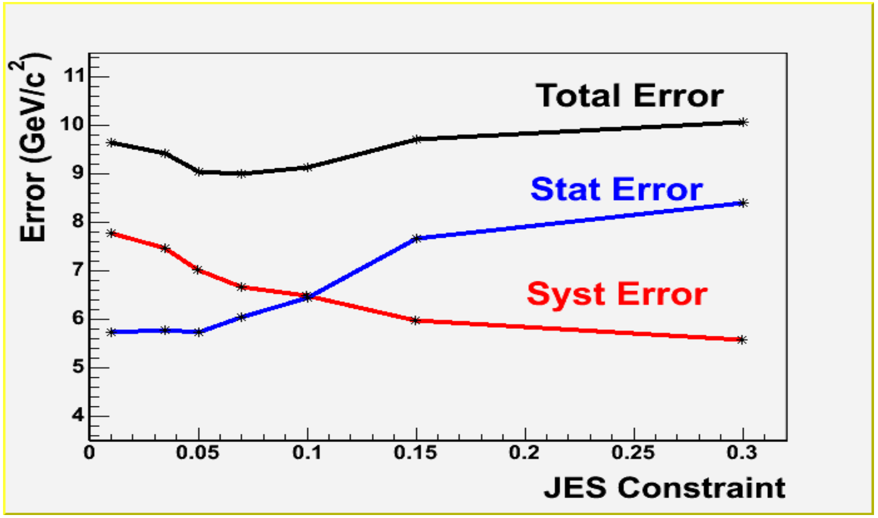
Analysis Features

- A number of features distinguishes MTM from the other methods of top mass measurement
  - With the goal of reducing systematic error, a kinematic fit to the hadronic W mass includes the jet energy scale (JES) as a variable
  - Statistical error is reduced by estimating the probability that the correct jet-parton assignment in an event was selected
  - Using event variables besides the reconstructed mass in the likelihood function should provide more info + improve signal/background discrimination
  - We employ KDE, a non-parametric method of density estimation, in our likelihood calculation

Mass Reconstruction

- In each event, for every jet-parton permutation we use the JES as a constrained parameter in a kinematic W-mass fit. We keep the permutation with the best χ^2.
- This serves to improve the systematic error, but also increase the statistical error
- The upper plot indicates that for correctly chosen permutations, the systematic is compensated by the JES shift
- The lower plot shows that by altering the JES constraint, we can alter the resulting tradeoff in errors





Overview of the Likelihood

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

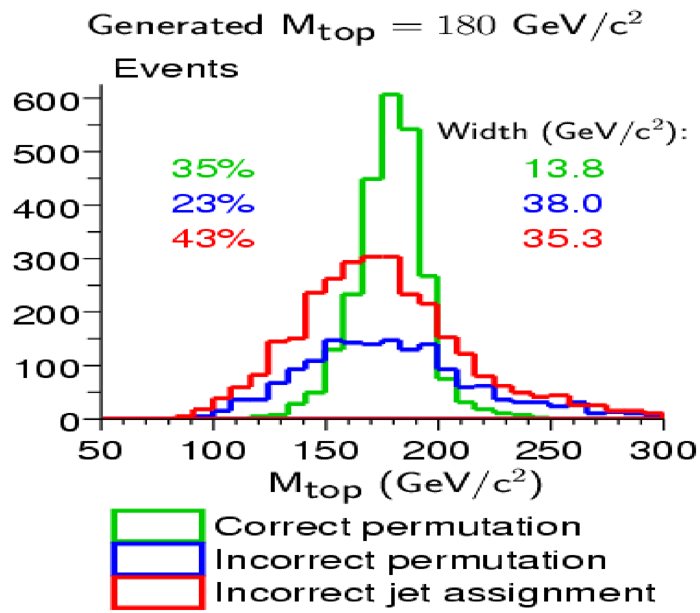
Background: $P_b(m, x) = \sum_{\text{bg types}} a_j B_j(m, x), \qquad \sum_{\text{bg types}} a_j = 1$

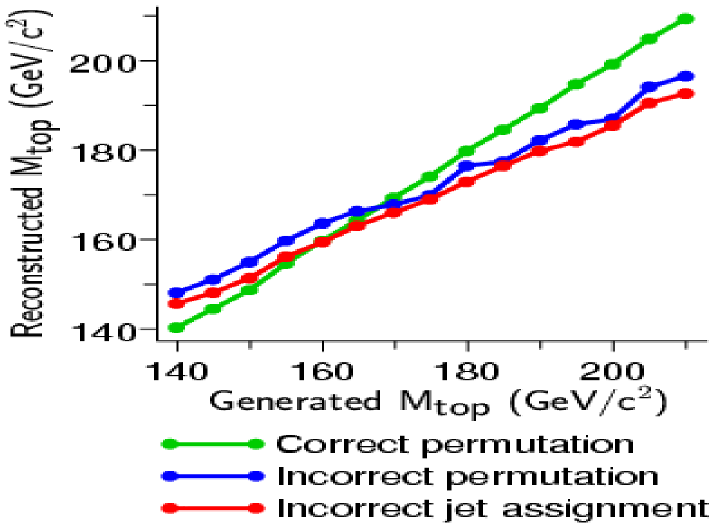
Signal: $P_s(m, x, m_t) = \sum_{\text{sig types}} c_j S_{j,m_t}(m, x), \qquad \sum_{\text{sig types}} c_j = 1$

- Density function templates are assigned to classes of background and signal (B's and S's, respectively)
- These are functions of the reconstructed mass (m) and any other observables we add (x)
- The likelihood of one event is the weighted sum of its probability in each template

Separate Signal Templates

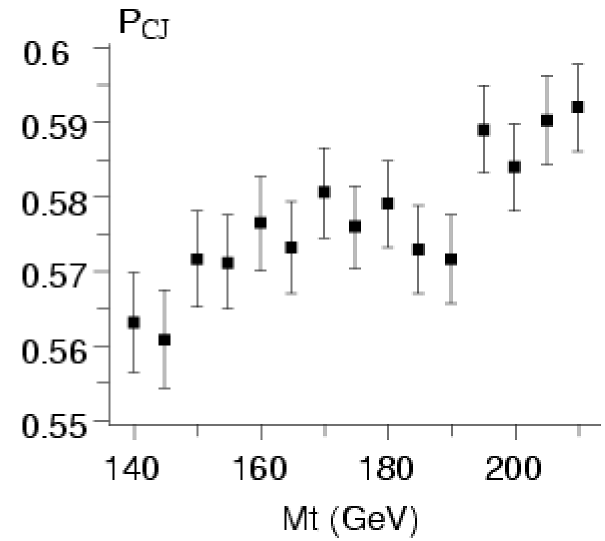
- The signal is divided into three templates: one for good permutation (GP) events, one for correct jets, bad permutation (BP), and one for incorrect jets (IJ) events. GP + BP = CJ, "Correct Jets"
- In an ideal situation (no background and perfect assignment of signal events to their corresponding templates) there could be an improvement in top mass resolution of ~ 1.7





Estimating an Event's Signal Template Probability

- We take the probability that an event belongs to CJ simply from the ratio of CJ events to all events in our MC files at a given top mass



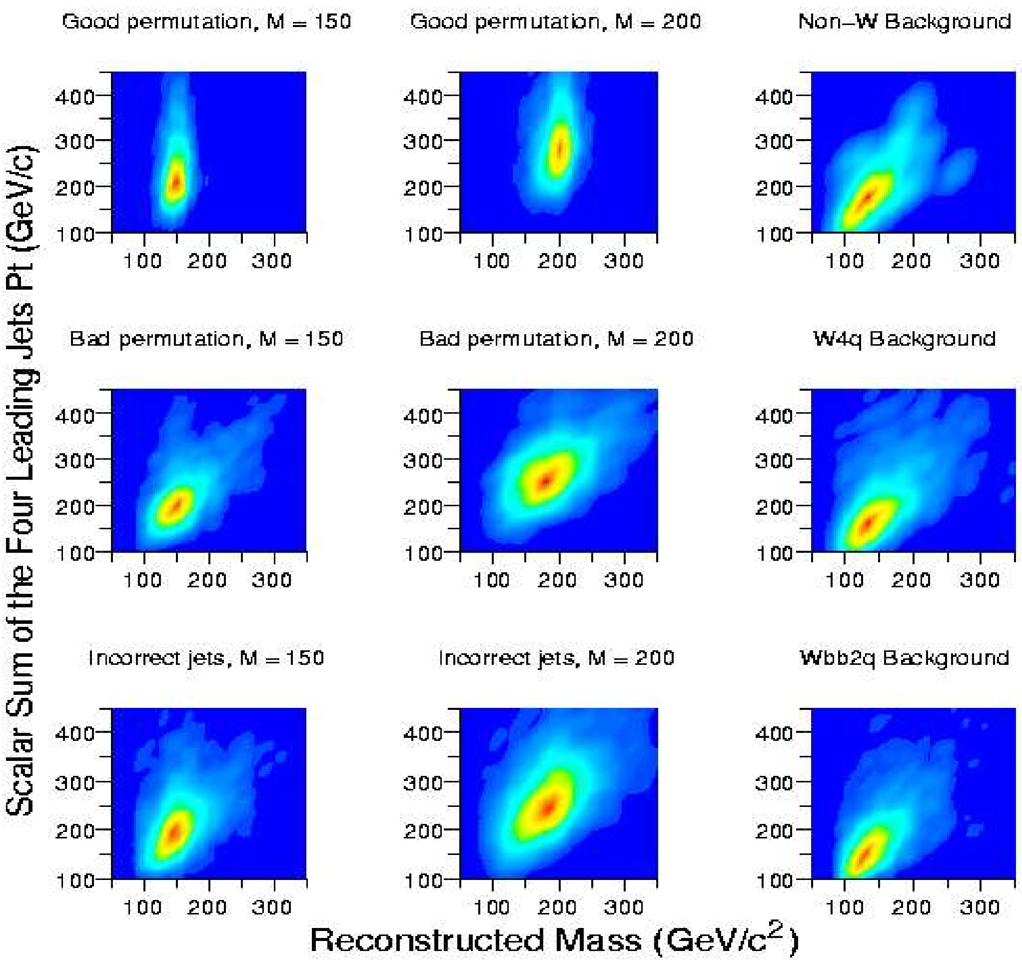
- Then, we split the remaining probability into GP and BP probabilities by employing a formula which uses the difference between the best χ^2 and the others

$$p_{\mathrm{GP}} = \frac{a_b}{\sum_{i=1}^{12} a_i \exp\left(-\frac{\chi_i^2-\chi_b^2}{w_e(\chi_i^2+\chi_b^2)}\right)}$$

- Finally, a Bayesian update of the GP probability is performed using event angle variables related to the leptonic W polarization and the $t\bar{t}$ spin correlations

Kernel Density Estimation

- In our likelihood, rather than use a fitted function to represent a continuous density function using a discrete # of MC points, we use KDE
- This technique performs a weighted sum of the surrounding “training points” to estimate the density
- Its advantage is that it frees us from making assumptions as to the form of the density – especially useful in higher dimensions!

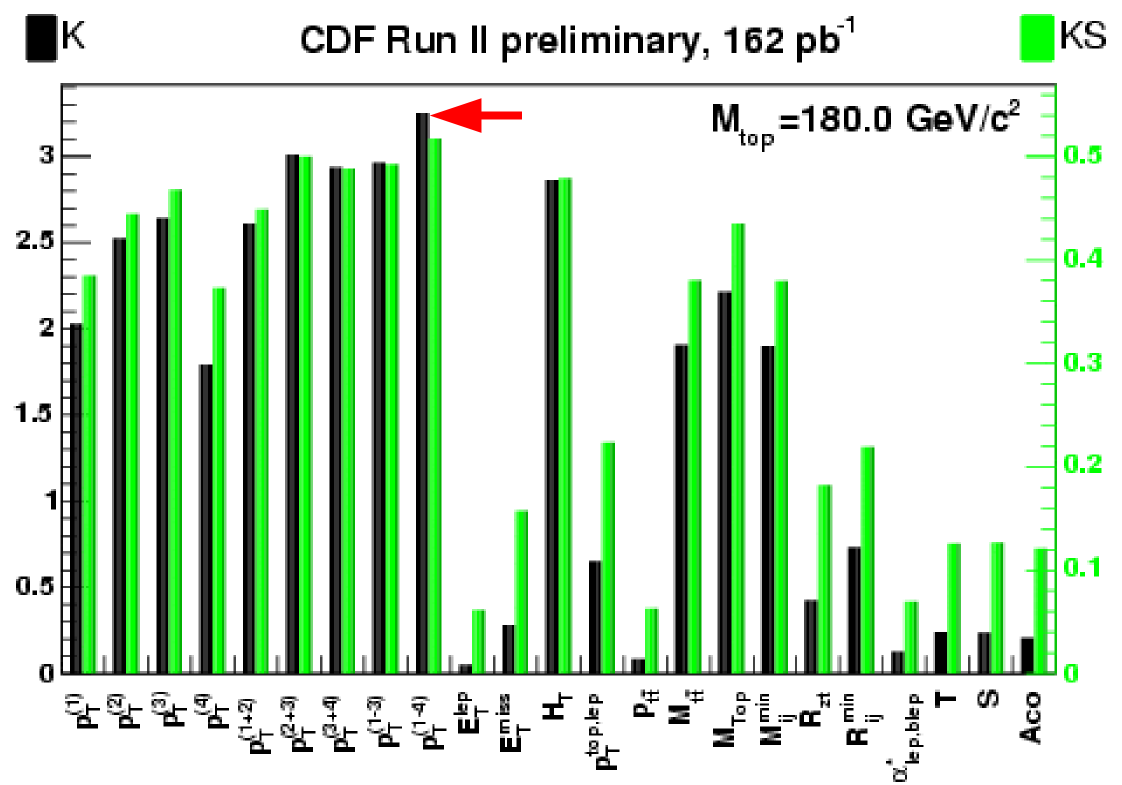


Choosing a Variable Set

- Using the Kolmogorov-Smirnov distance and K-divergence tests, we attempt to determine which variables do the best job at distinguishing signal from background

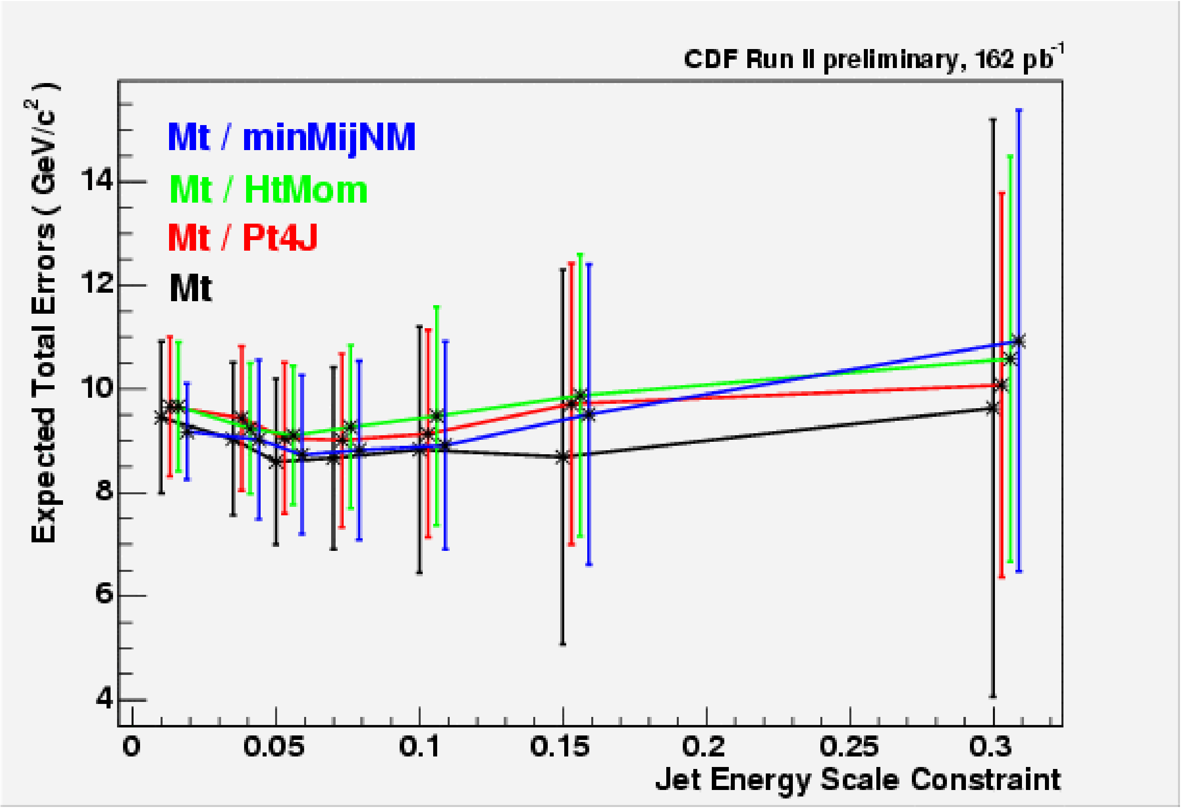
- We find that the scalar sum of the transverse momenta of the leading four jets is the best choice

- For our likelihood, then, we use 2-d templates of reconstructed mass and this momentum variable



Choosing a JES constraint

- Using candidate variable sets suggested by divergence tests, we run thousands of PE's to determine what value to tune our JES constraint to.
- JES constraint yields lowest error at 0.07
- Relative to the expected error of a single top mass measurement, there appears to be no preferred choice of variables



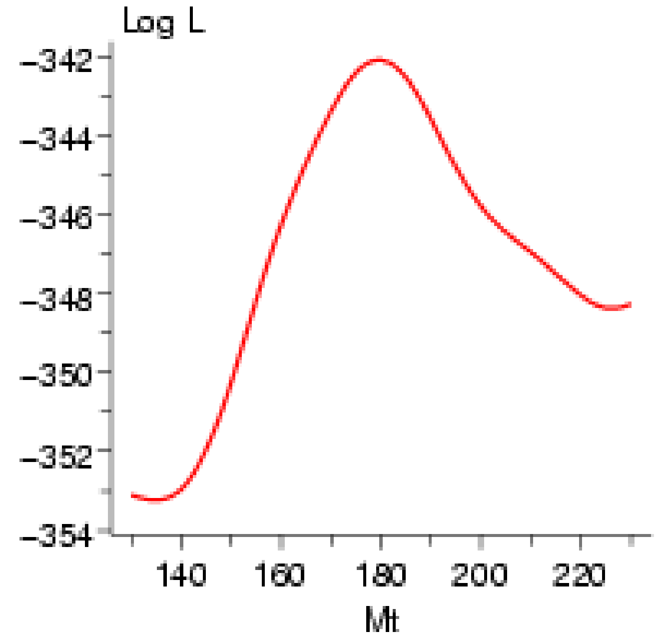
Our Measurement

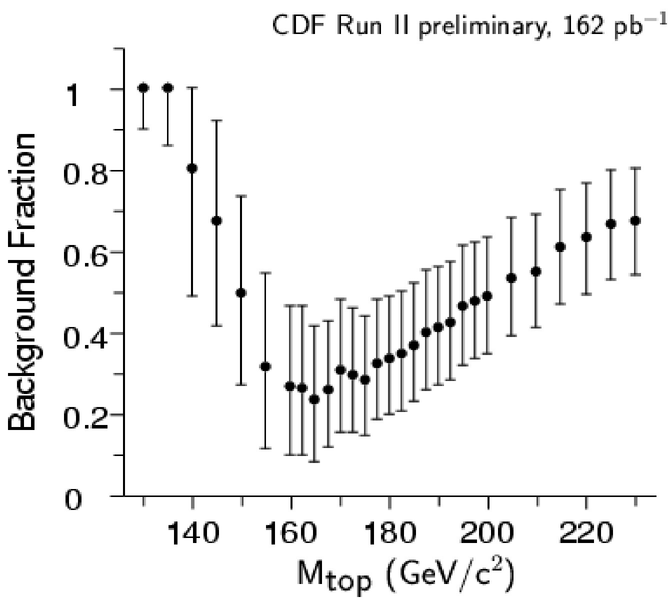
Data consists of 33 events from 162 inverse pb

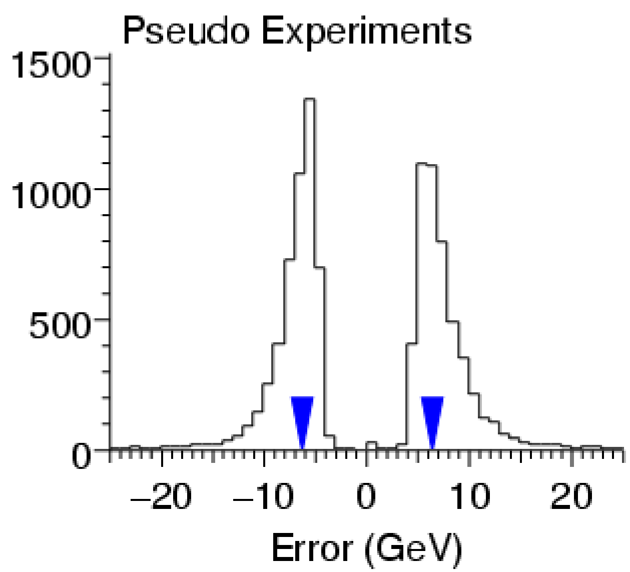
Top Mass: $m_t = 179.6^{+6.4}_{-6.3}$ (stat error)

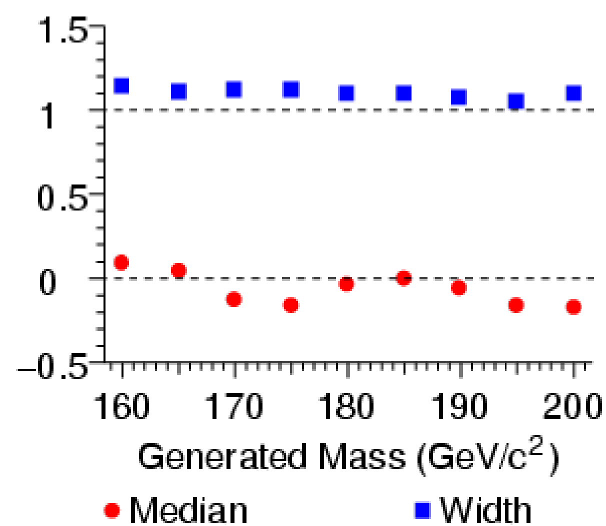
Estimated background fraction : $f_b = 0.34 \pm 0.14$

Average pull width : 1.10

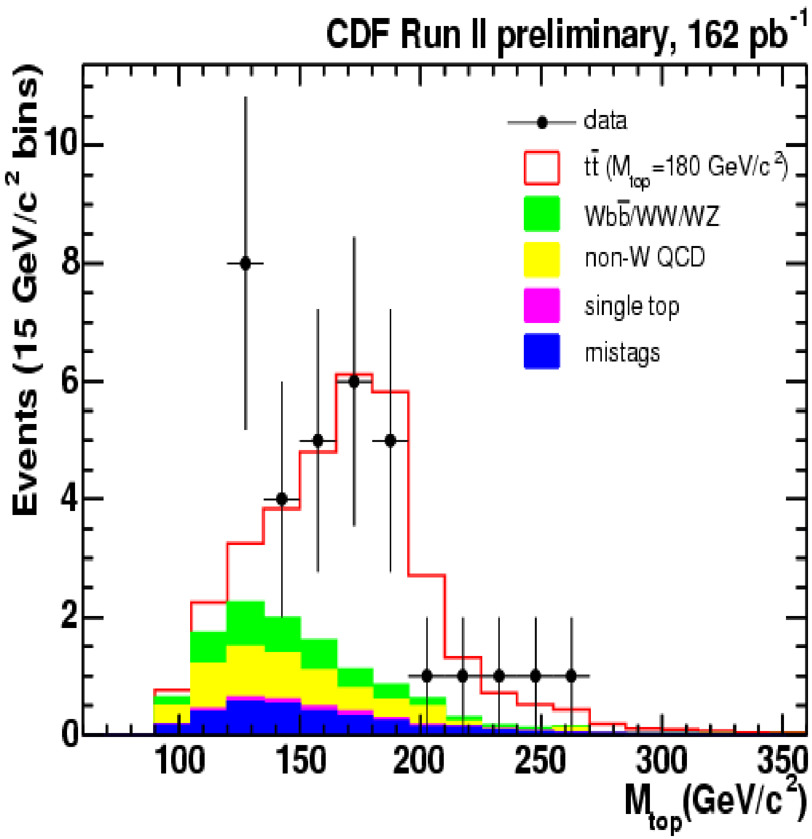


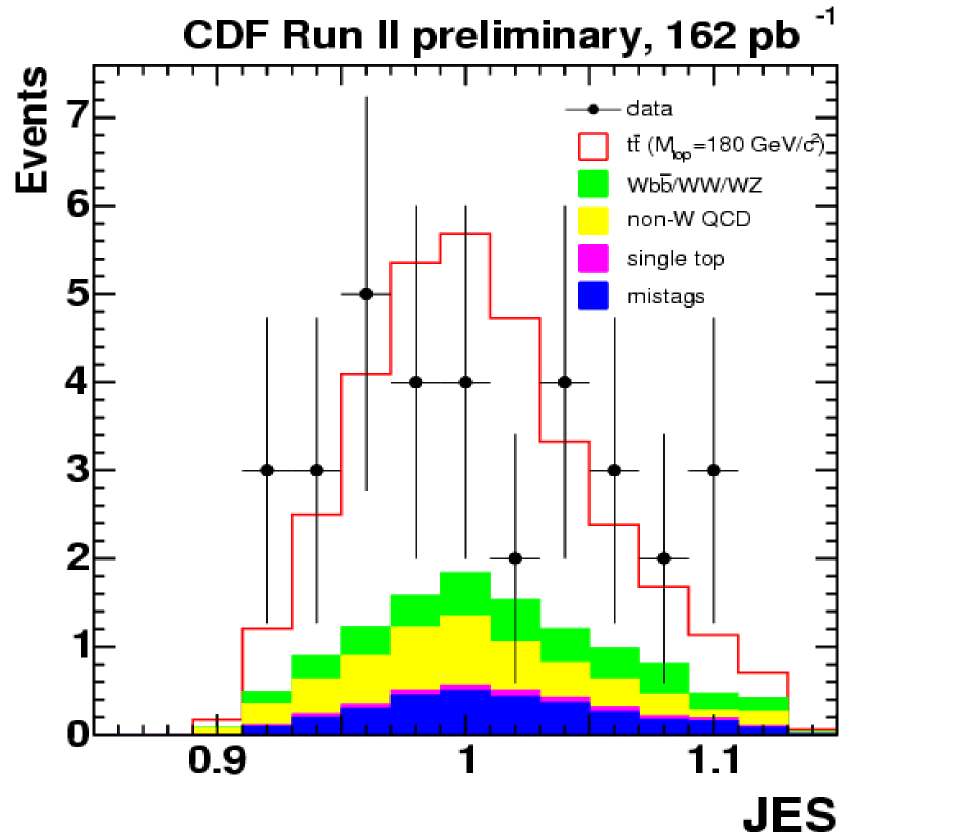






Comparison of Data w/ MC





- Here we see a comparison of our data sample with a normalized combination of signal and background MC. Results are in reasonable agreement.

Systematics

Systematic	ΔM_{top} (GeV/c^2)
Jet Energy	6.7
Generators	0.2
ISR	0.2
FSR	0.6
PDF	0.6
Background Shape	0.4
b Tagging	0.3
Fitting Procedure	0.7
Total	6.8

As you can see from this table, the greatest contribution to the systematic error overwhelmingly comes from the jet energy measurement!

The jet energy systematic error is being worked on by many at CDF and should improve soon

Conclusions

- MTM techniques allow us to find an optimal systematic-statistical error tradeoff
- Our likelihood calculation is improved by estimating probabilities that a given event belongs to a given template
- We look forward to:
  - Re-tuning the JES constraint and re-examining our choice of variables as statistics improve
  - Adding new features to our analysis
  - And improving on our current blessed measurement of

$$179.6^{+6.4}_{-6.3}(\mathsf{stat.}) \pm 6.8(\mathsf{syst.})\ \mathsf{GeV}/c^2$$

BACKUP SLIDES

Bayesian Update

- In order to increase the amount of info our GP/BP probabilities are based on, we apply Bayes theorem to X, our info variable of interest, in the standard way:

$$P(GP|CJ,X) = \frac{\kappa P(GP|CJ)}{\kappa P(GP|CJ) + (1 - P(GP|CJ))}$$

where

$$\kappa = \frac{P(X|GP)}{P(X|BP)}$$

Alpha-Skew Tests

- The general alpha-skew formula is as follows:

$$
\begin{aligned}
d(p_1, p_2) &= \int \mathrm{d}x \; p_1(x) \times \log\left(p_1(x) / [\alpha \cdot p_2(x) + (1-\alpha) \cdot p_1(x)]\right) \\
d_s(p_1, p_2) &= d(p_1, p_2) + d(p_2, p_1)
\end{aligned}
$$

- We use a value for alpha of 0.5. This is known as a “K-divergence”.

The Data Set

- As we use the cross-section group's background fraction calculations in our analysis, we attempt to use their data set
- Criteria include:
  - 3.5 jets (4^{th} jet > 8 GeV)
  - Wrong beam line runs removed, but wrong luminosity measurement run accepted
  - Trident electrons removed
  - All jet permutations used have to agree with the SVX tag information
  - Phoenix electrons non reconstructed and therefore not used in the dilepton veto

At a JES constraint of 0.07, we end up with 33 events

Smoothing the Likelihood

- Since we only have a discrete number of top masses in our signal MC's, it's necessary to interpolate smoothly between the likelihood values of an event

- We do this by employing local polynomial regression

- For several events, each event is interpolated separately and then summed