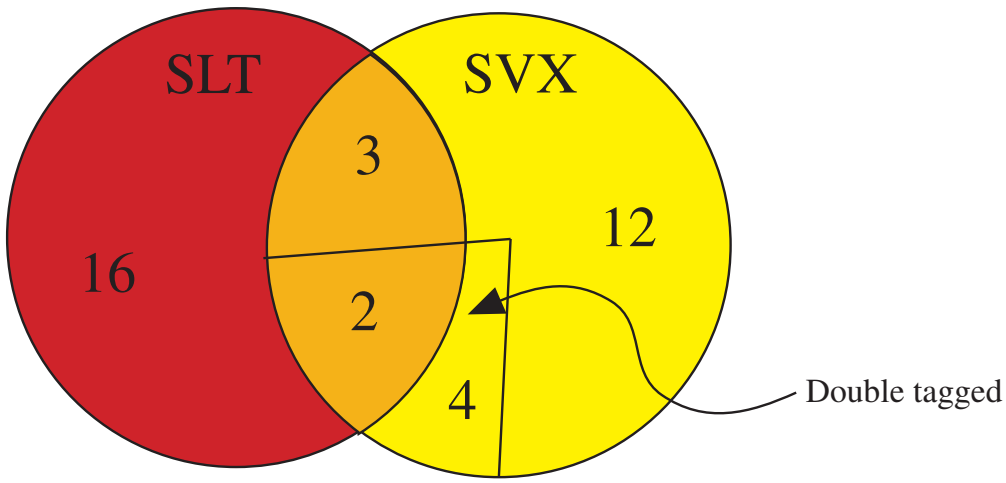
b Quark Tagging in W+Jet Events

Applied b tagging techniques to W+jet events

- The $W + 1$ and $W + 2$ jet samples are ultimate control
  - Don't expect significant top contribution
  - Theory predicts low b hadron content

Sample	SVX bkg	SVX tags	SLT bkg	SLT tags
W+1 jet	50 ± 12	40	159 ± 25	163
W+2 jet	21 ± 7	34	46 ± 7	55
W+$\geq$ 3 jet	6.7 ± 2.1	27	15.4 ± 2.3	23



DØ observes 6 tagged events

- Expect 1.2 ± 0.2 background events
- Expect about 3 events if $M_{top}\sim$ 160 GeV/c^2

Summary of CDF Counting Experiment

CDF sees significant excesses in all three channels

Channel	Background	Tags	Events	Prob
SLT	15.4 ± 2.0	23	22	6×10^{-2}
SVX	6.7 ± 2.1	27	21	2×10^{-5}
DIL	1.3 ± 0.3		6	3×10^{-3}

Probability of background $= 1 \times 10^{-6}$

Probability calculation performed by assuming no $t\bar{t}$

- "simulate" large number of experiments
- Count how many experiments would "see" more tags
- Properly takes into account
  - Uncertainties in background estimates
  - Correlations in uncertainties
  - Poisson fluctuations in total rate of events

Summary of DØ Counting Experiment

DØ combines dilepton, kinematic and μ-tagged events

- Observe 6 μ-tagged events on background of 1.2 ± 0.2
- Observe 8 events on background of 1.9 ± 0.5 from $\mathcal{A} - H_T$
- 3 dilepton events (expect 0.7 ± 0.2 background)

Gives a total of 17 events on background of 3.8 ± 0.6

- Probability of background fluctuation is 2×10^{-6}

Expectations from top are

M_{top} (GeV/c^2)	Number of Events (DØ)
140	18.0 ± 2.0
160	11.8 ± 1.7
180	6.8 ± 1.0
200	4.0 ± 0.6

So What About Multijets?

$t\bar{t} \rightarrow 6j$ faces huge QCD multijet background

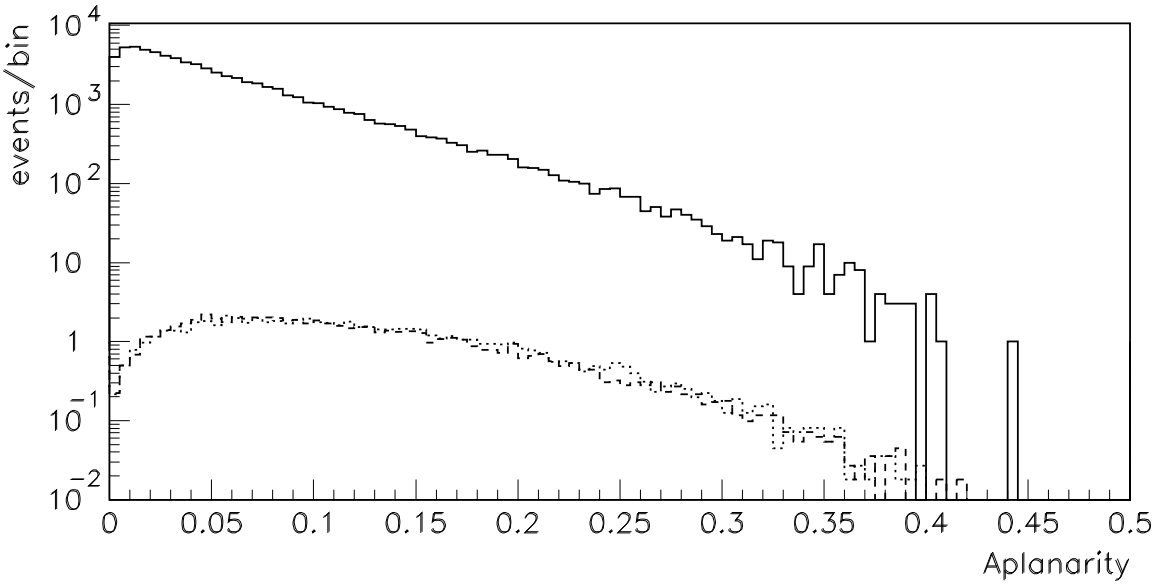
- QCD cross section is over 4 orders of magnitude larger
- Both CDF and DØ attempting to pull out signal
  - Can use kinematic cuts to get of order 10^3 rejection
    - acoplanarity > 0.15
    - $H_T^3 \equiv \sum E_T - E_T^1 - E_T^2 >$ 175 GeV
  - Use b quark tagging techniques to get $\times 10$

Expect to hear more about these in next few weeks

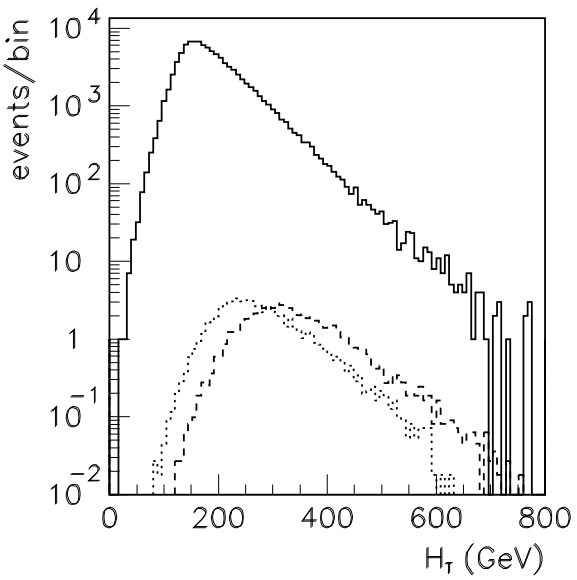
- Expect to get S/N $\sim 1/1$
- Important handle on branching fractions

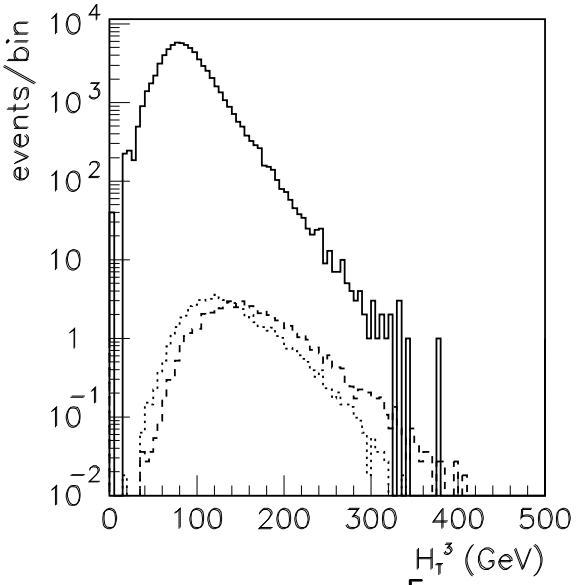
DØ Multijet Kinematic Cuts

Aplanarity



H_T^3





Cross Section for Top Production

Can use counting experiments to estimate $\sigma_{t\bar{t}}$

- Subtract the expected backgrounds
- Perform a maximum likelihood fit to estimate $\sigma_{t\bar{t}}$
- A number of issues have to be addressed
  - Make consistent background calculation
    - CDF backgrounds assume all pre-tagged events are background
    - Bootstrap to get correct estimate
  - Does one count b-tags or events?
    - relatively low tagging efficiency
    - have two b quarks for signal events
    - slightly more information in tags
  - Mass dependence of acceptance
  - Correlations between channels

Uncertainties are dominated by Poisson fluctuations

- Not a Gaussian statistic
- Systematic uncertainties have to be included
- Correlations are potentially important

Methodology

The expected number of events observed in each channel given by

$$\begin{aligned} N^i &= B^i + \int \mathcal{L}\,\mathrm{dt}\,\epsilon^i\,\sigma_{t\bar{t}} \\ \Rightarrow \sigma_{t\bar{t}} &= \frac{N^i - B^i}{\int \mathcal{L}\,\mathrm{dt}\,\epsilon^i} \end{aligned}$$

where

- B^i is expected background in ith channel
- ϵ^i is efficiency in ith channel

Maximum likelihood solution for n channels is

$$\sigma_{t\bar{t}} = \frac{\sum_i \left(N^i - B^i\right)}{\int \mathcal{L}\,\mathrm{dt} \sum_i \epsilon^i}$$

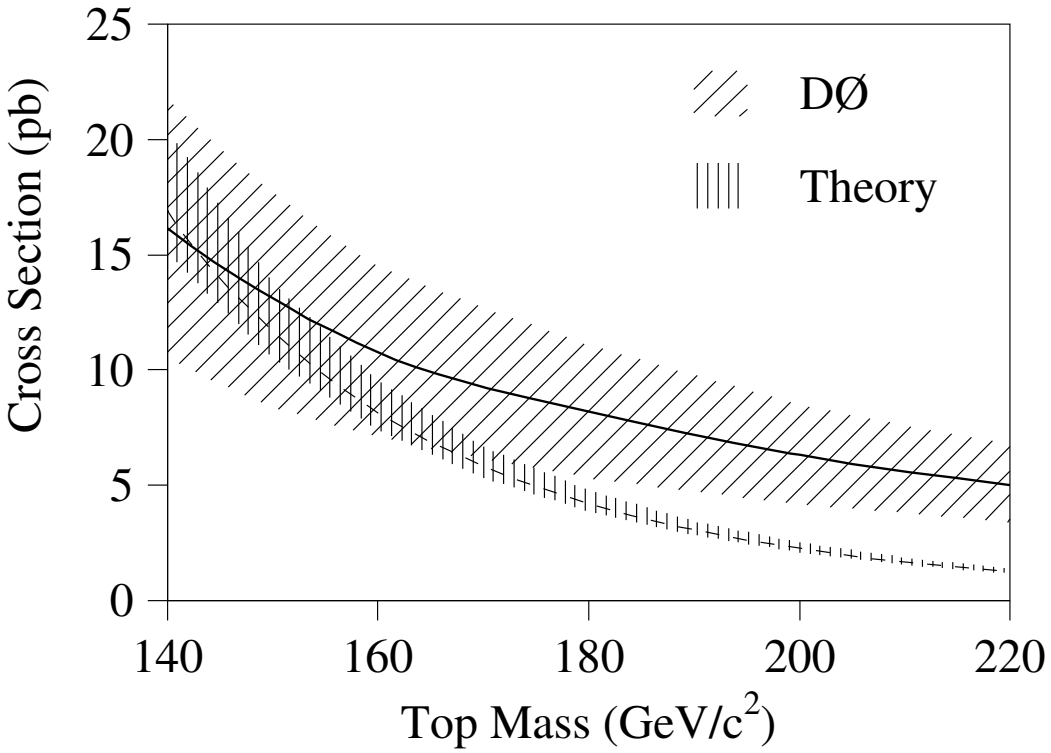
In practice, have to account for uncertainties

Source	Uncertainty % (CDF)
Integrated luminosity	3.6
Lepton detection	5
Jet energy scale (depends on M_{top})	$10 \rightarrow 3$
Initial state radiation	7
B tagging efficiency	12
Total	16%

DØ Cross Section

DØ cross section (quoted at M_{top}= 200 GeV/c^2)

$$\sigma_{t\bar{t}} = 6.3 \pm 2.2 \text{ pb}$$



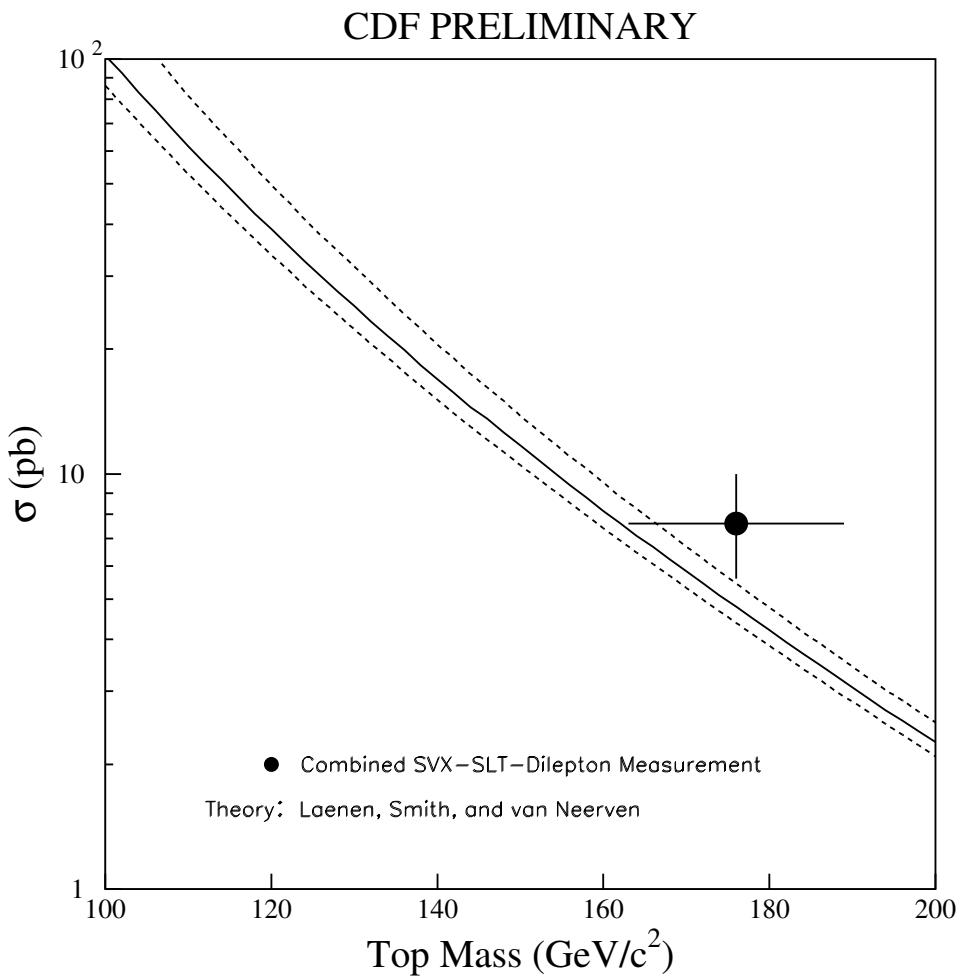
CDF Cross Section Result

CDF uses the results from the three channels

- SVX tags have the most information
- SLT and DIL channels have $\times 2$ larger uncertainty

Result of fit (quoted at $M_{\text{top}} = 176$ GeV/c^2)

$$\sigma_{t\bar{t}} = 7.6^{+2.4}_{-2.0} \text{ pb}$$



Top Quark Event Topologies

The expected $t\bar{t}$ kinematics and topologies are unusual

- Six partons in final state
- Two "real" $W^{\pm}$ bosons
- A b and $\bar{b}$ quark jet

Reluctance to take advantage of this in searches

- analysis dependent on MC model for background
  - W+jets physics background not well understood
  - Hard to constrain non-W backgrounds

- $t\bar{t}$ signal characteristics are not well established
  - Initial and final state radiation effects
  - Top quark fragmentation
    - First quark to decay weakly **before** "dressing" up
  - Production mechanisms are based on perturbative QCD

- Concern about subtle detector effects
- Kinematics-based searches have been avoided

Are forced to consider $t\bar{t}$ kinematics

- Need additional background rejection (DØ)
- Use it to verify counting experiments
- Need kinematics model to measure mass

Employ a number of different MC calculations

1. W+jets physical background described by VECBOS
   - Only tree-level calculations
     - Have to "graft" fragmentation model
     - CDF uses HERWIG, DØ uses ISAJET
   - Only up to $W + 4$ jets
   - $Wb\bar{b}$ and $Wc\bar{c}$ estimates difficult
   - Sensitive to Q^2 scale and singularities
     - Favourite choices are $Q^2 =< P_T^2 >$ and $Q^2 = M_W^2$
2. $t\bar{t}$ signal modeled by several calculations
   - DØ uses ISAJET and PYTHIA
   - CDF uses PYTHIA and HERWIG primarily
   - Initial and final state radiation differs alot

Kinematics in Counting Experiments

DØ used several kinematic features to separate top from QCD background

- $t\bar{t}$ events are more spherical (aplanarity)

$$M_{ij} \equiv \sum_{\text{jets}} \frac{p_i p_j}{p^2}$$

$\mathcal{A} \equiv \frac{3}{2} Q_1$ where Q_1 is smallest eigenvalue of M

- $t\bar{t}$ events have more energetic partons (H_T)

$$H_T \equiv \sum_{\text{jets}} E_T$$

Applied these techniques to W+$\geq$ 4 jets

- Require $\mathcal{A} > 0.05$ and $H_T > 200$ GeV
- Estimate W+jet background by scaling jet multiplicities
- Observe 8 events and expect 1.9 ± 0.5 from background

Likelihood Approach

Number of studies have been made to determine those kinematic variables that are most effective

- Focus has been on separation of W+jets and $t\bar{t}$
- Depends on MC predictions, so have to compare several
- Phase space is very large ($\geq$ 6 partons!)

Two basic methods have been discussed

- Consider 1-3 variables that "intuitively" make sense
  - CDF has looked at E_T^2 vs E_T^3
  - DØ has used H_T and $\mathcal{A}$
- Use "blender" approach
  - Employ Fischer discriminants
  - Correlated, multi-variate likelihood
  - Neural networks using kinematic variables as input

I will focus on just a few examples

- Worry more about mass measurement

CDF Likelihood Analysis

Look at spectrum of 2nd and 3rd leading jet

- Use expected distribution of jets
  - Jet 2 and 3 provide good signal-background separation
  - Form relative likelihood

$$\log(L_{\text{rel}}) \equiv \log\left(\frac{L_{\text{top}}(E_T^2, E_T^3)}{L_{\text{QCD}}(E_T^2, E_T^3)}\right)$$

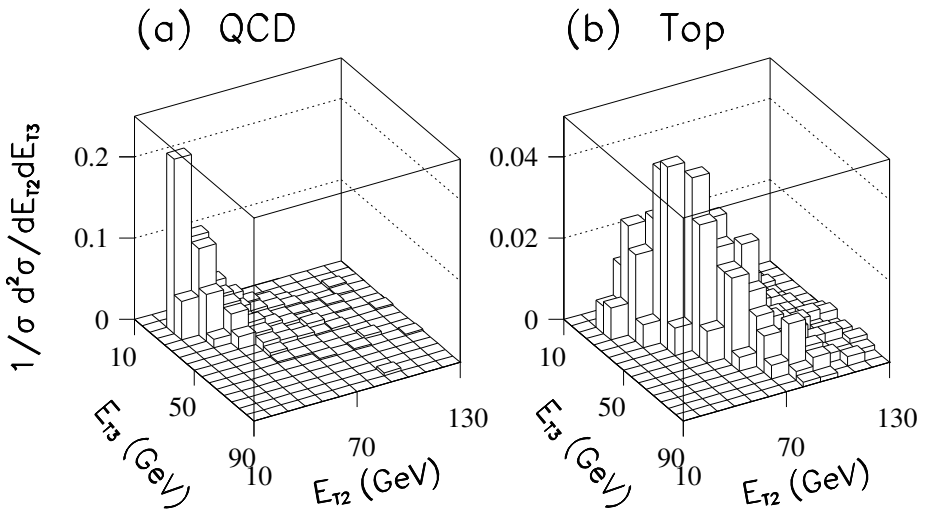
- Compute using VECBOS (QCD) and Pythia ($t\bar{t}$) MC

This technique is MC dependent

- Have to study effect of systematics on L_{rel}
- Calibrate technique on control samples
  - "background-enriched" W+jet sample
  - W+2 jet sample where little top is expected

CDF Event Structure (cont.)

Clear difference in spectra between background and signal

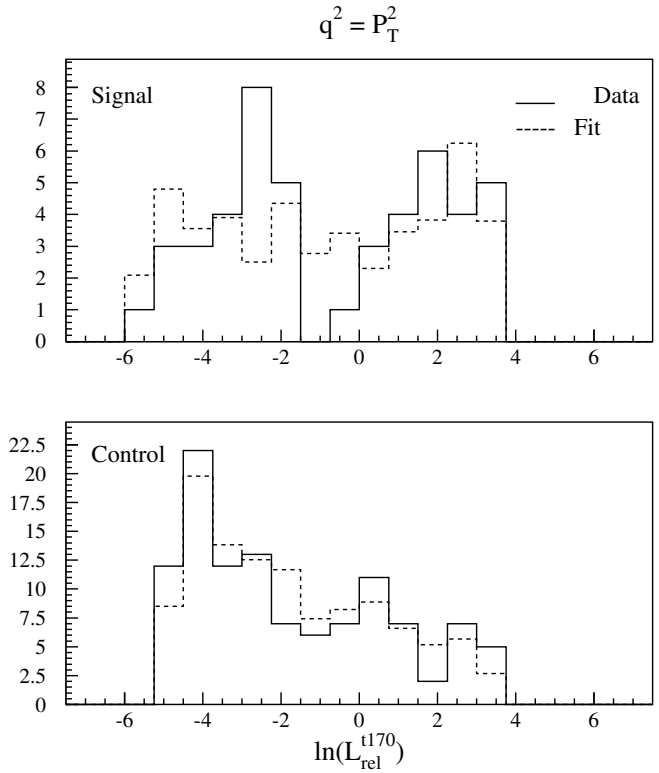


Weight each event with likelihoods and form ratio L_{rel}

- Effective way to separate background from $t\bar{t}$
- Can compare observed shape with that predicted by background and signal
- Use control samples
  - Events with jet at large η more likely background

Results of Event Structure Analysis

Analysis indicates that W + ≥ 3 jet data is mixture of top + background

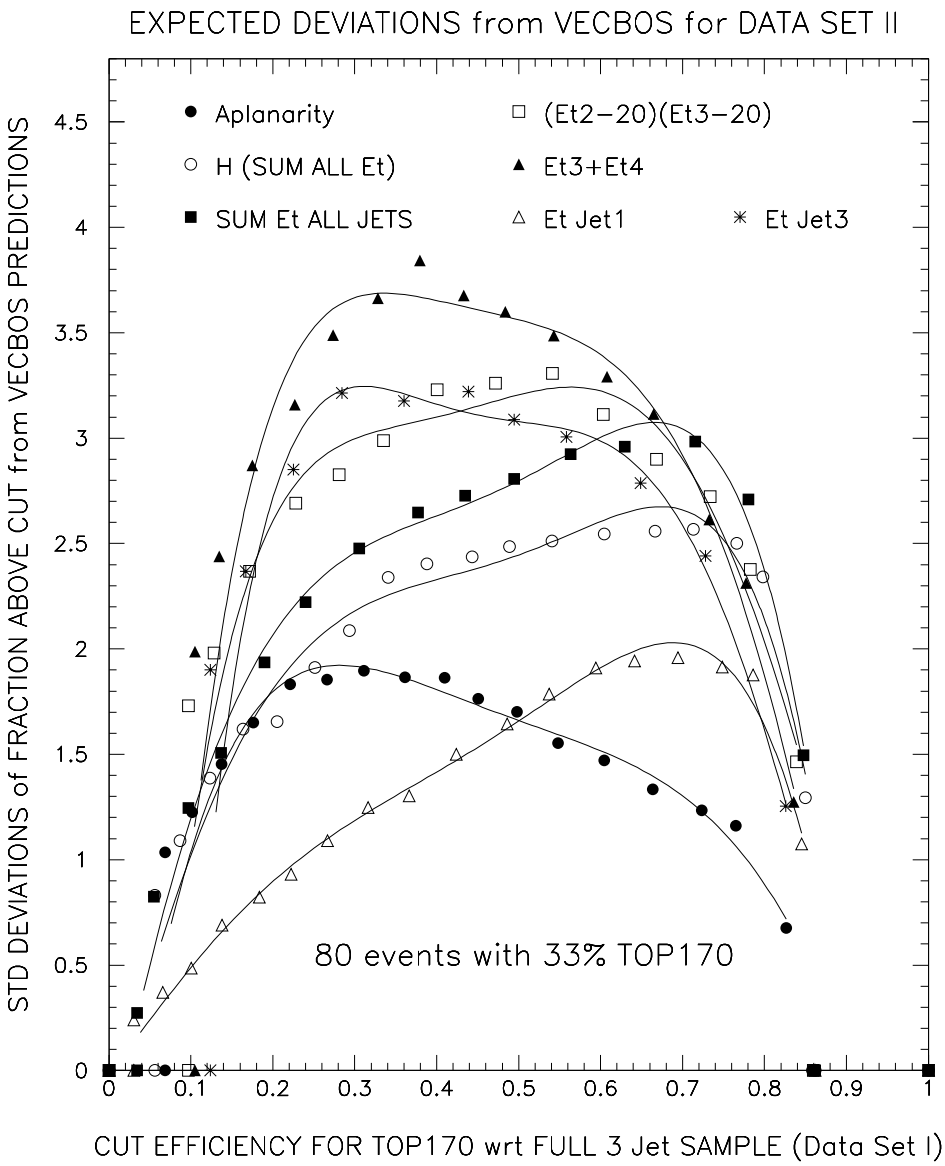


- Can readily ask how likely this is if no $t\bar{t}$
  - Ask how often background would
    - give at least 22 events with $\log(L_{\text{rel}}) > 0$
    - when observe 47 total
- Result is 0.26%

Other Kinematic Variables

Many other kinematic variables have been studied

- Can characterize by "significance"
  - Expected # of standard deviations above background
  - Function of cut efficiency



The Top Quark Mass

"Best" M_{top} measure comes from $t\bar{t}$ reconstruction
Alternatives are:

- b jet spectrum
- correlation with H_T
  - Recent CDF result gets to $\sim \pm 20$ GeV/c^2
- Can use cross section
  - CDF $\sigma_{t\bar{t}} \Rightarrow \pm 12$ GeV/c^2
  - But assumes standard model, BR, etc.
- Direct measurement preferable

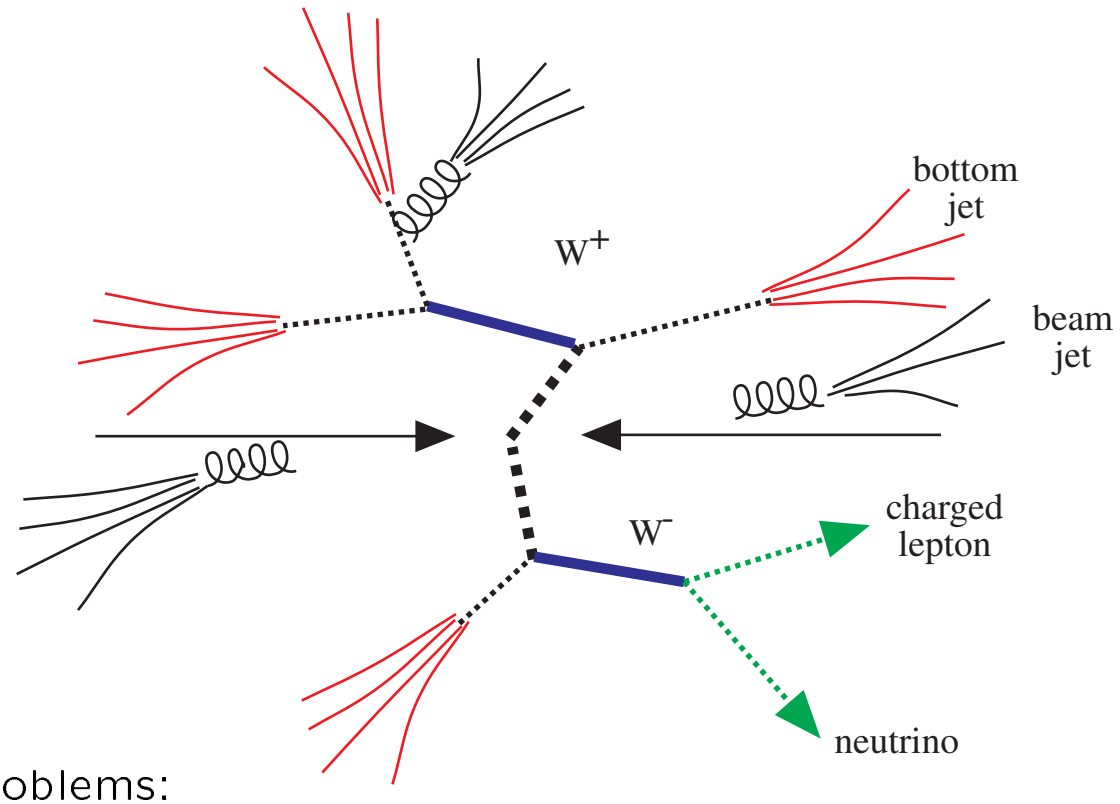
However, has lots of problems

- ≥ 6 partons need to be identified
  - 2 ν's in dilepton mode problematic
  - Have fewest constraints and fewest events
- Focus has been on lepton + jets channels

Lepton + Jets Reconstruction

In this mode, have measured almost all $t\bar{t}$ daughters

- Lepton 4-momentum solid and ν_l is missing only p_z
- $W \to q\bar{q}'$ observed as 2 jets
- bottom quarks observed as jets (perhaps tagged)



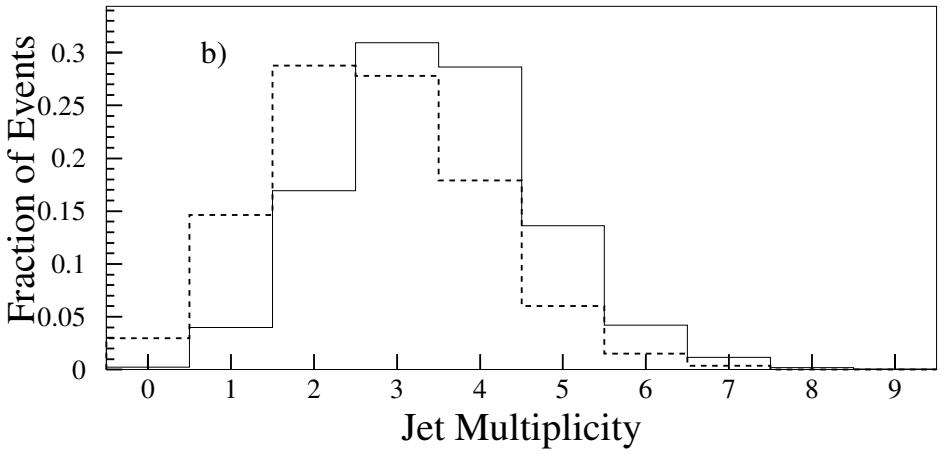
Problems:

- QCD radiation generates additional jets
- Combinatorics in identifying jets $\leftrightarrow$ partons
- Background events

Biggest Problem is Jet Assignments

Jets in real $t\bar{t}$ events arise from

- $W \to q\bar{q}'$ and b quarks
- Initial state radiation and "beam" jets
- Final state radiation



Has to be overcome by looking at likely combinations

- Fortunately highest p_T jets are most likely from $t\bar{t}$
- Can use tagging information...
- Take into account biases due to mis-identification
  - Yet another source of MC dependency

CDF Measurement of Top Mass

Select a subsample of events with $\geq$ 4 jets

- Place loose cuts on 4rth jet
  - Require $E_T >$ 8 and $|\eta| <$ 2.4
  - High efficiency ($\sim$ 85% for $t\bar{t}$)

- Combine all jet pairs to form $W \rightarrow q\bar{q}$
  - 12 jet-parton assignments
  - Correct jets depending on assignments
  - 2 unique solutions for neutrino p_z
- Fit each assignment:
  1. W daughters invariant mass constrained to M_W
  2. t and $\bar{t}$ have same mass
  3. Allow for $p\bar{p} \rightarrow t\bar{t}X$
  4. Gives 2C fit
  - Choose one with best χ^2
  - Require $\chi^2 < 10$
- Use b-tag information to identify b quark jet

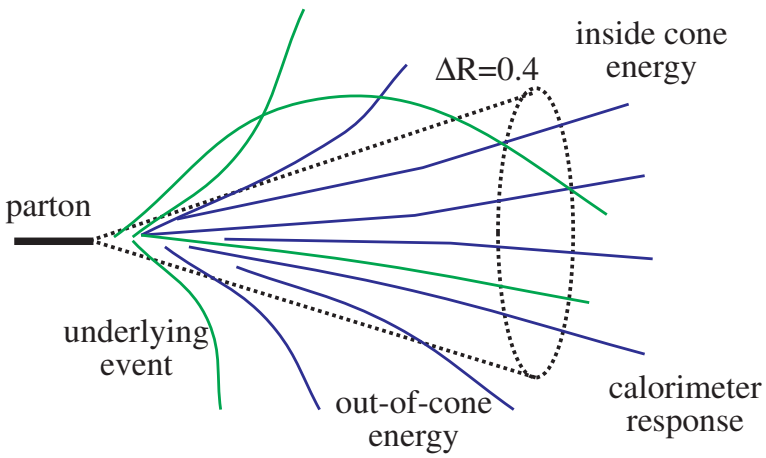
The fourth jet cut leaves 99 events out of 203

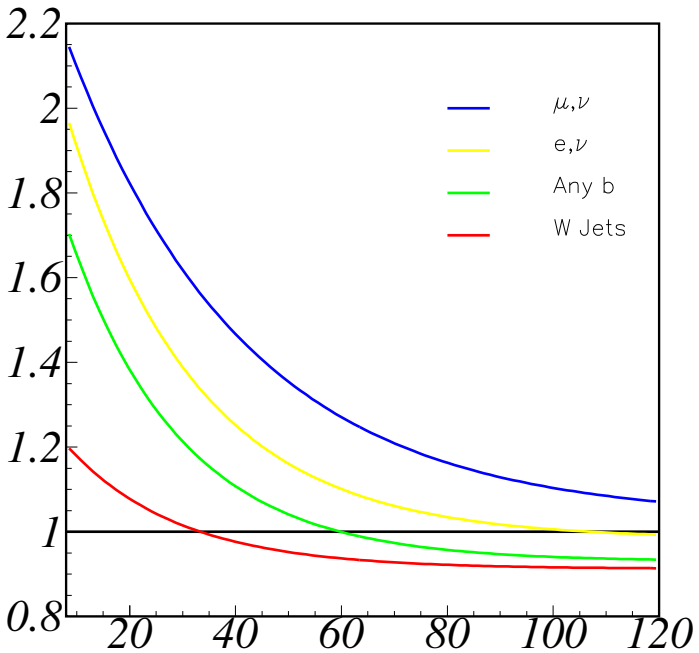
- 88 have at least one good fit
- 19 events have at least one b tag

Jet-Parton Corrections

Have to first correct observed E_T to parton E_T

- Corrections determined by MC and testbeam studies
- Make additional corrections for $t\bar{t}$ topology





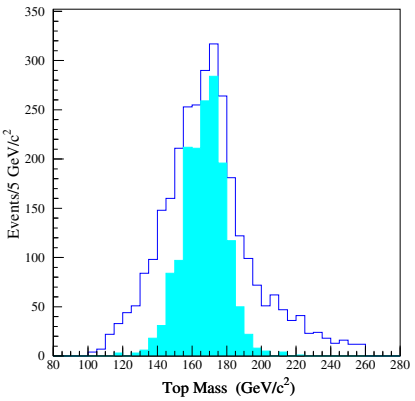
Tests of CDF Reconstruction

Can test how well procedure works with MC events

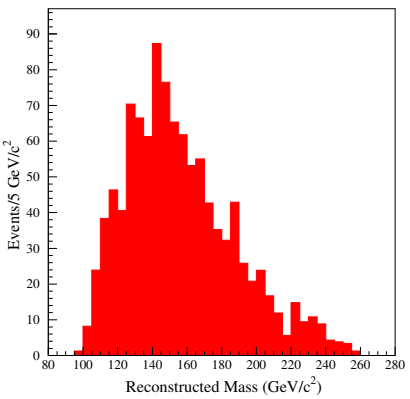
- Look at correct and "best" assignments
- Also look at how background would reconstruct

Mass Reconstruction for Top and W+jets Monte Carlo

Reconstructed mass for HERWIG MonteCarlo, $M_{top} = 170$ GeV



Reconstructed mass for the VECBOS+HERPRT W+3 jets events.

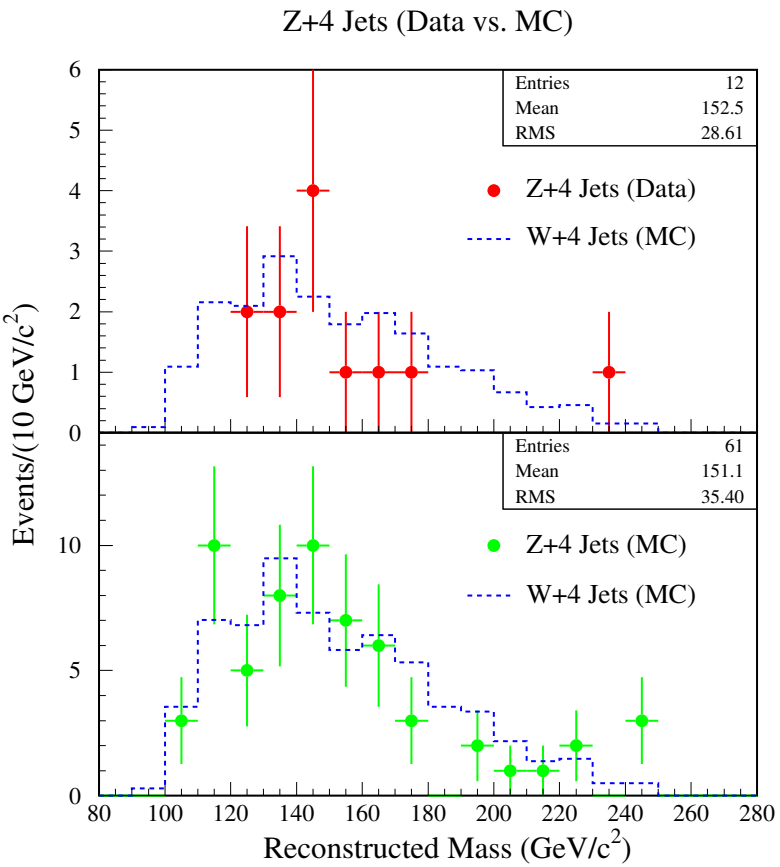


Background Effects

Primary backgrounds are from $W b\bar{b}$ and fake leptons

- Use background rates from counting experiments
- Can check $W b\bar{b}$ shape by looking at $Z+$ jets
- Use "non-isolated" lepton + jet sample for fakes
- Look identical to VECBOS $W+$ jets prediction

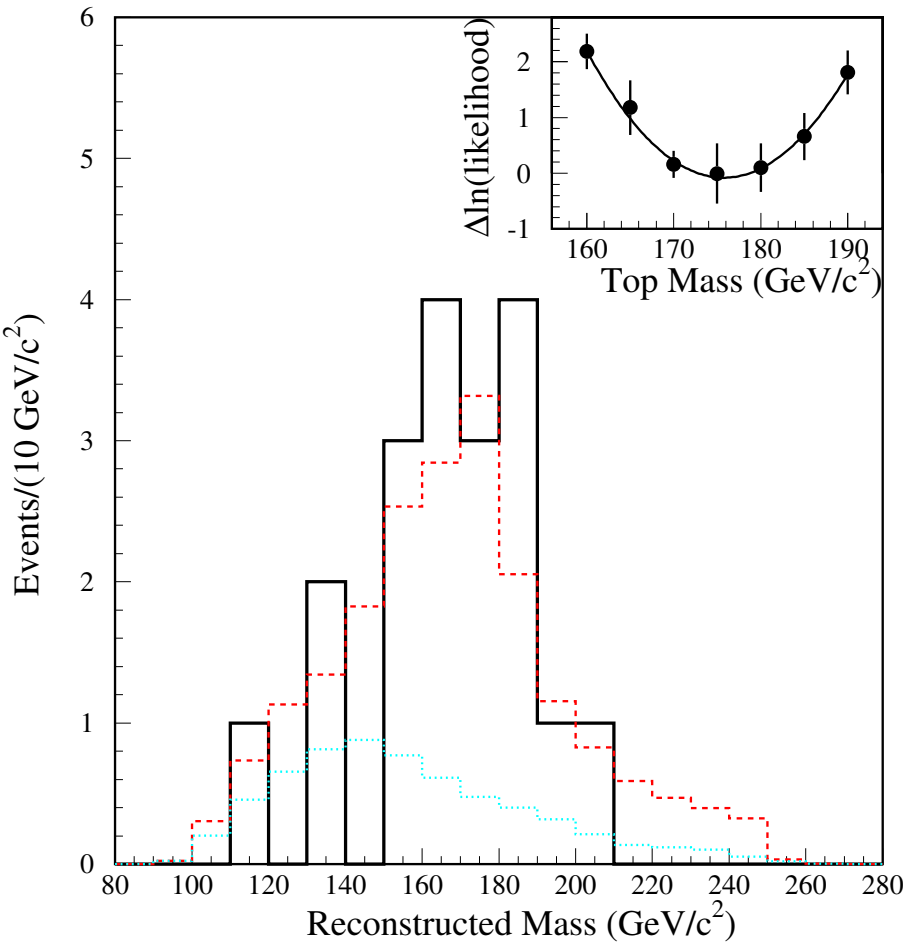
  $\Rightarrow$ Use VECBOS $W+$ jets MC for bkg shape



CDF Mass Results

Estimate background based on W+jet sample

- Use known fake tag rates and $Wb\bar{b}$
- Background shape in M_{top} comes from MC calculation
- Estimate background to be $6.9^{+3.4}_{-2.4}$ events
  - Primarily QCD W+jet



$M_{\text{top}} = 176 \pm 8 \pm 10 \quad \text{GeV/c}^2$

Systematic Uncertainties

Systematic error is dominated by energy scale

Systematic	Value (GeV/c^2)	Percentage
Jet E_T scale (ISR)	7.7	4.4
MC Statistics	3.1	1.8
Fit procedures	2.5	1.4
SLT Tagging bias	2.1	1.2
Background shapes	1.6	0.9
Minimisation procedure	1.6	0.9
Fitting likelihood	1.2	0.7
SVX Tagging bias	1.1	0.6
Total	10.0	5.7

Jet enery scale uncertainty comes largely from our lack of knowledge about the physics

- Should come down as we accumulate larger samples
- Currently focus of effort in reducing uncertainty

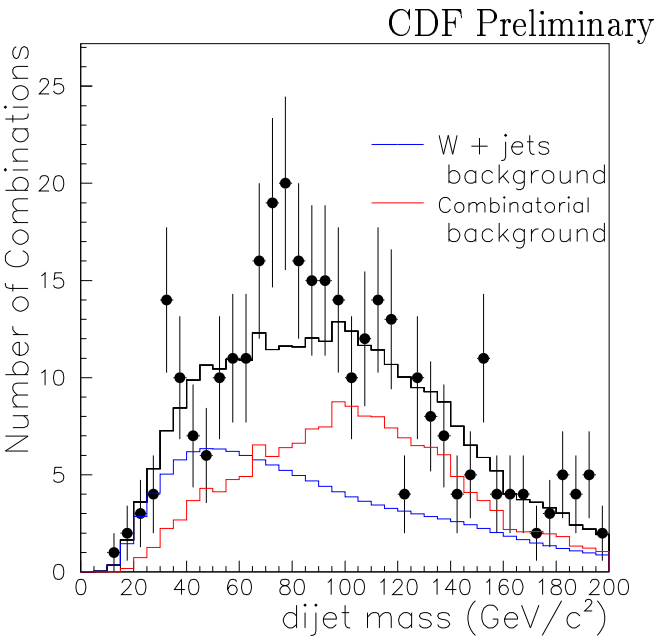
Evidence for W → qq̄

Observation of $W \to q\bar{q}$ is important

- Strong evidence of signal in $W^+W^-b\bar{b}$ channel
- Can calibrate jet energy scale

Have to refit data relaxing W mass constraint

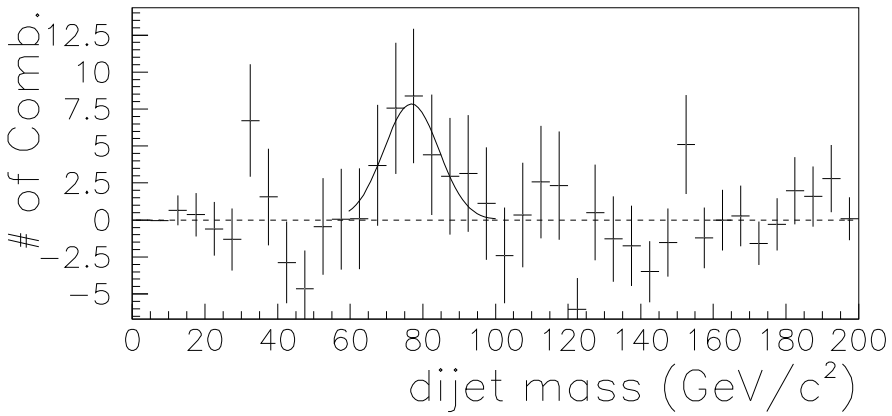
- Look for W signal plotting all mass combinations
  - Makes it tough since 6 combos/event
- Reduce backgrounds by requiring $H > 300$ GeV



Hadronic W Observation (cont.)

Can perform a background subtraction

- Observe signal of 27 ± 12 events
- Expect about 28 events from $t\bar{t}$ production



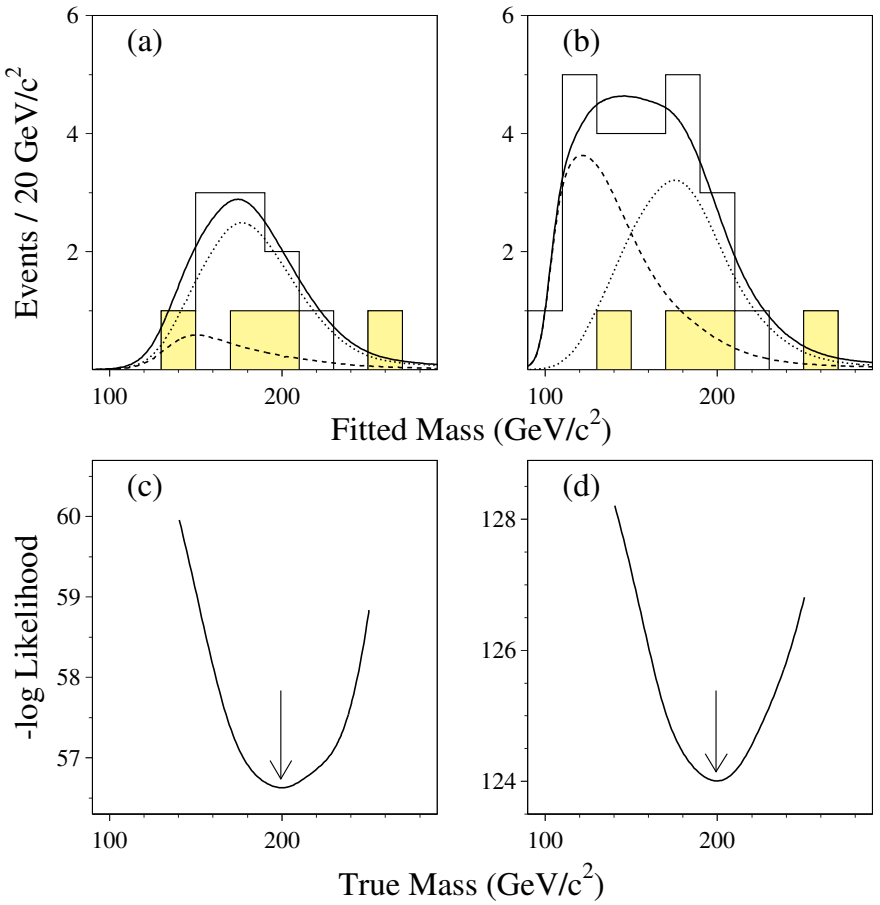
Width of signal is consistent with that expected in $t\bar{t}$ decays

- Can eventually calibrate jet energy scale to $1-2$ GeV
- Can also extract signal in other ways
  - Relax mass contraint and choose best fit
  - Similar result is obtained

DØ Mass Analysis

DØ has performed a similar lepton+jets analysis

- Have a smaller signal sample
- Use 4 highest energy jets
- Perform jet corrections somewhat differently
- Find 11 events that successfully fit $t\bar{t}$ hypothesis
- Don't constrain the background rate

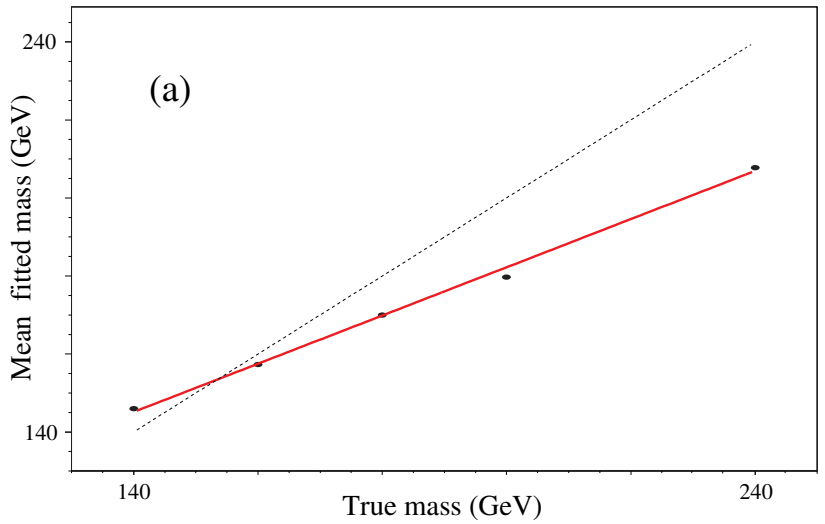


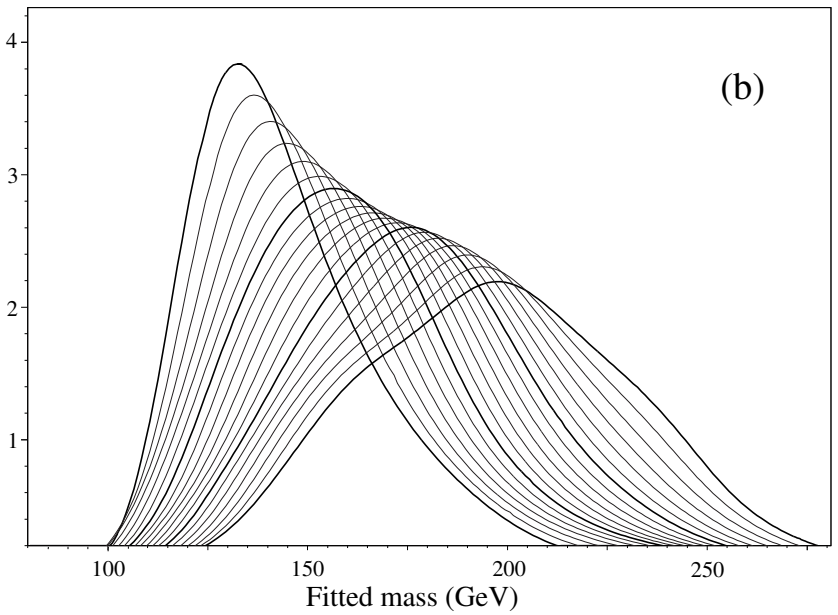
DØ Mass Systematics

Have similar sources of systematic uncertainty

- Larger jet energy scale uncertainty ($^{+12}_{-20}$ GeV/c^2)

$$M_{\text{top}} = 199^{+19}_{-21} \pm 22 \quad \text{GeV/c}^2$$





Top Quark Branching Fractions

Indication of new physics would be if $BR(t \to W^{\pm} b) \neq 1$

- SM expects V_{tb} to be unity
  - Unitarity constraints
    - V_{ud} large from nuclear β and μ decay
    - V_{us} constrained by K_{e3} decays
    - V_{cd} comes from $\nu n \to cX$
    - V_{cs} constrained by D_{e3} decays
    - V_{cb} small because large τ_b
    - V_{ub} even smaller ($b \to u$)
  - Expect V_{ts} and V_{td} much smaller than V_{tb}
- $V_{tb} \lesssim 0.9985$ violates (or bends) SM

Experimentally, we can measure

$$\mathcal{R} \equiv \frac{BR(t \to Wb)}{BR(t \to Wq)} = \frac{|V_{tb}|^2}{|V_{tb}|^2 + |V_{ts}|^2 + |V_{td}|^2}$$

Measure the rate of zero, single and double b-tagged events

- Need to be able to b tag
- Can only use zero tagged events with dileptons where backgrounds small without b tagging

BR Technique

Fraction of zero, single and double-tagged events is

$$\begin{aligned} f_0 &= (1-\mathcal{R}\epsilon)^2 \\ f_1 &= 2\mathcal{R}\epsilon(1-\mathcal{R}\epsilon) \\ f_2 &= (\mathcal{R}\epsilon)^2 \end{aligned}$$

- ϵ is tagging efficiency

Rate of single vs double tags gives

$$\mathcal{R} = \frac{2}{\epsilon(f_1/f_2 + 2)}$$

Rate of zero vs single tags gives

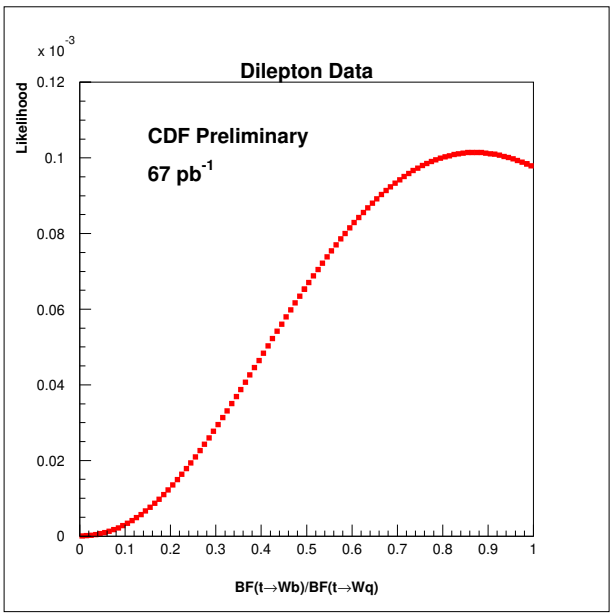
$$\mathcal{R} = \frac{1}{\epsilon(2f_0/f_1 + 1)}$$

"Best" way of estimating $\mathcal{R}$ is to use maximum likelihood technique

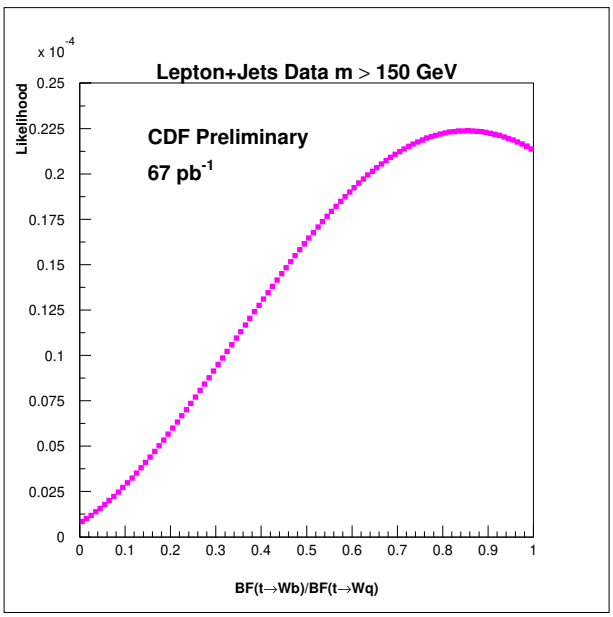
- For a given number of observed events,
  - calculate probability of observing 0, 1 and 2 tags
  - maximize probability by varying $\mathcal{R}$
- Can use probability to constraint $\mathcal{R}$

CDF BR Results

Result from dilepton data is $\mathcal{R} = 0.875 \pm 0.4$



Result from lepton + jets data is $\mathcal{R} = 0.87 \pm 0.3$



BR Conclusions

Largest systematic uncertainty comes from b-tagging efficiency

- introduces $0.08 - 0.10$ variation in R
- Overall systematic uncertainty is $^{+0.13}_{-0.11}$

$$\mathcal{R} = 0.87^{+0.13}_{-0.30}(\text{stat})^{+0.13}_{-0.11}(\text{syst})$$

Can convert that into a measure of $|V_{td}|$ by assuming

- $0.004 \leq |V_{td}| \leq 0.015$ and $0.030 \leq |V_{ts}| \leq 0.048$
- Take average values of these to obtain

$$|V_{tb}| = 0.11^{+0.89}_{-0.05}$$

This is in good agreement with the SM (note errors)!

- Really saying that $|V_{tb}| \gg |V_{ts}|$ or $|V_{td}|$
- Alternatively, 95% CL is $V_{tb} > 0.016$

New Physics Searches Using Top

Large mass of top has spurred speculation

- Top color model (Hill & Parke)
- Multiscale technicolor (Eichten & Lane)
- Lots of other variations on models

Predict structure in $t\bar{t}$ system

- Can look for "resonant' production

$$p\bar{p} \rightarrow V\,X \rightarrow t\,\bar{t}\,X$$

- Should see peak in $M_{t\bar{t}}$ distribution
- Stiffer p_T distribution for top
- Possibly larger $\sigma_{t\bar{t}}$

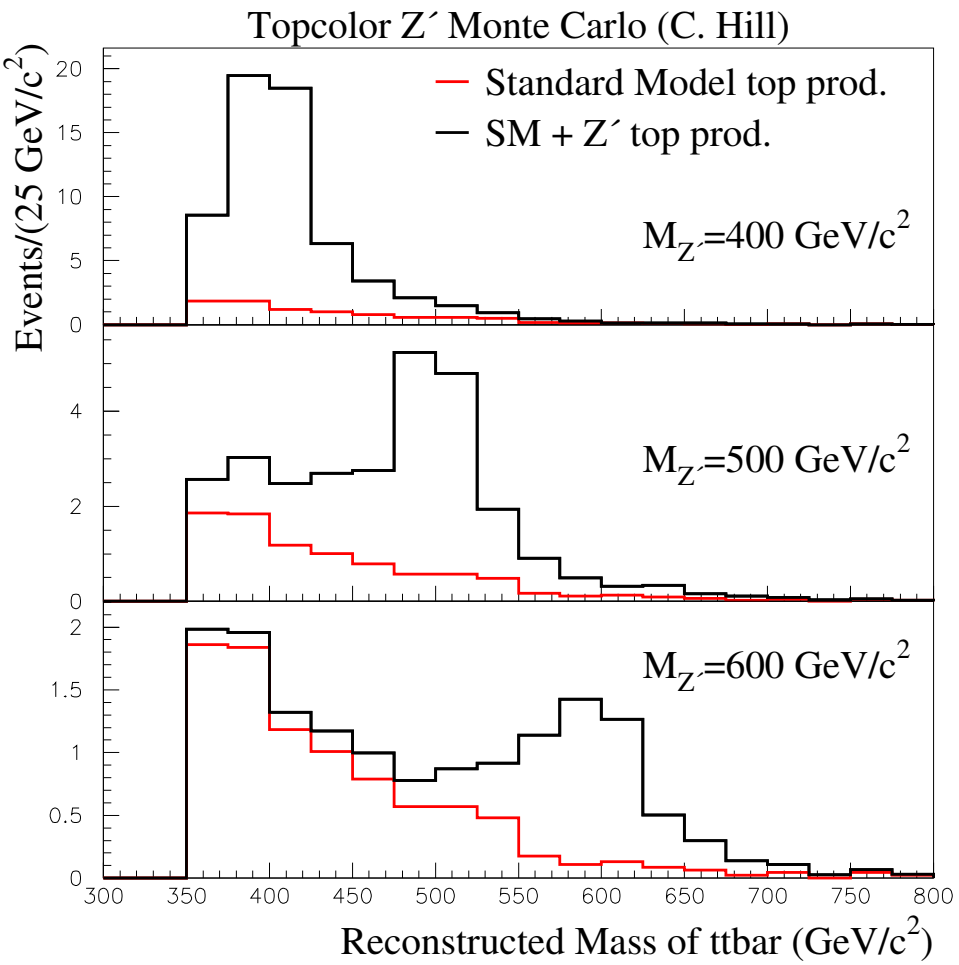
Can look at fully reconstructed $W + 4$ jet events

- Background is small
- Have sorted out (for the most part) jet confusion

Predicted signal for Top Color

Calculations by C.Hill et al. predict clear signature

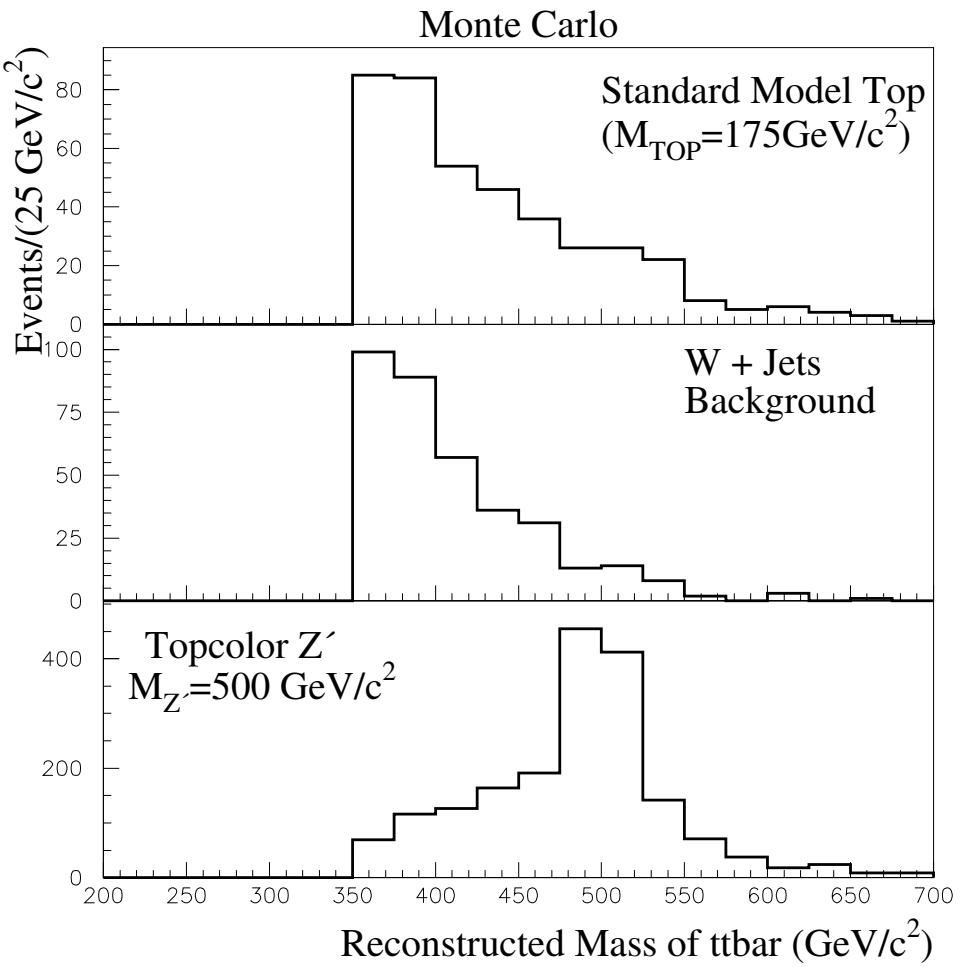
- Coupling to light quarks strong enough to yield appreciable cross section
- Invariant $t\bar{t}$ mass best place to look for effect



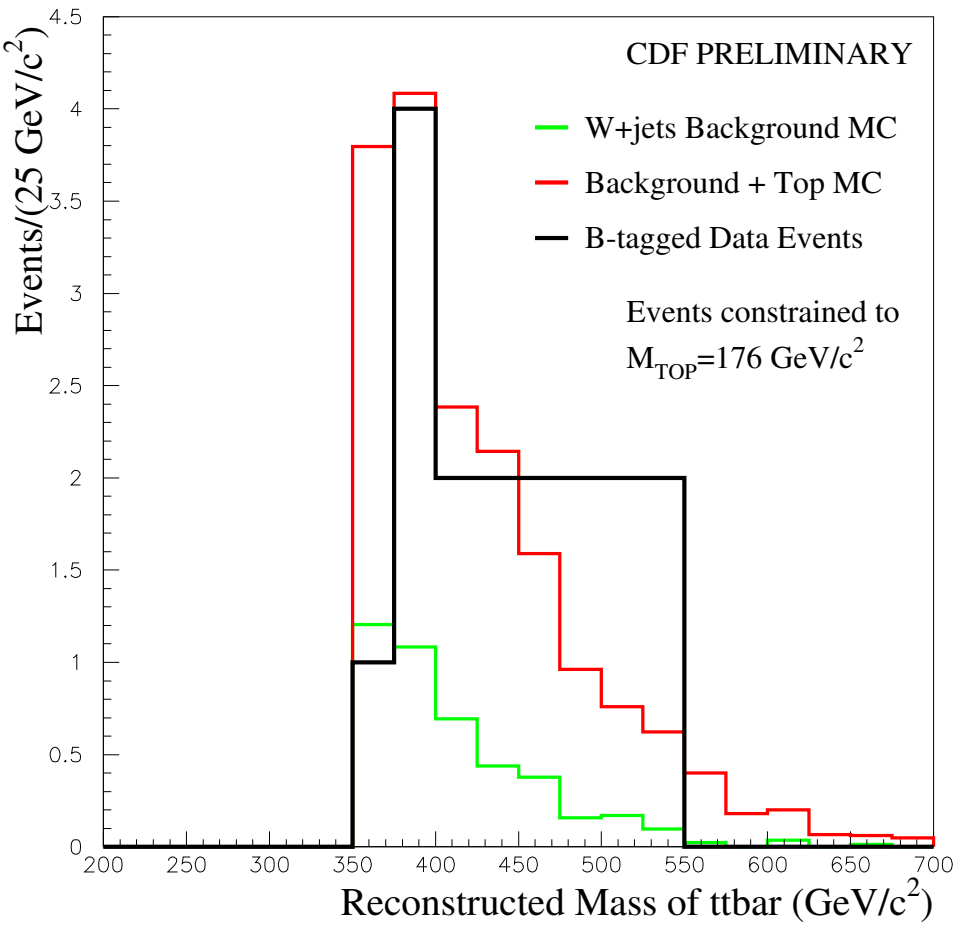
Backgrounds to Top Color

Standard model $t\bar{t}$ production most significant background

- Depends on Z' mass
- Can compare shape of backgrounds from $W+$ jets



$t\bar{t}$ Invariant Mass

Data is statistics limited

- There is no apparent enhancement
- An upper limit on Top-Color Z' mass and cross section coming soon