Top Quark Mass Measurement with Dynamical Likelihood Method at CDF RunII

Kohei Yorita
Kunitaka Kondo
Waseda University, Japan
For the CDF collaboration

~ FNAL Wine & Cheese Seminar ~
on June 11th, 2004
Outline of this talk

1. Introduction
2. Data Sample & Event Selection
3. Dynamical Likelihood Method (DLM)
4. Signal Monte Carlo Studies
5. Background Effects
6. Current Results from RunII
7. Systematic Uncertainty
8. Various Checks & Some Comparisons
9. Conclusions

Q: How many particles in me?

Ms. Particle
Top quark production and decay

Tree level:

- Dilepton (e, μ) BR=5%
- Lepton (e, μ) +jets BR=30%
- All jets BR=44%
- τ + X BR=21%

Final states:

- Higher statistics
- Lower background

The Tevatron is the only place which can produce top quark until LHC runs!

qq, gg fractions reversed at LHC

σ = 5.8 ~7.4 pb at 1.96 TeV

(Cacciari et al.)

Use this channel with at least 1 bjet.
- Higher statistics
- Lower background
Why is Top Mass interesting?

(1) Fundamental Standard Model parameter.
(2) Top quark is heavy (~ 180 GeV)  
   Yukawa coupling ~ 1.  
   * The mass is near the Electro-Weak Symmetry breaking scale.  
   * If we can measure strength of this coupling (i.e. ttH),  
     a test of the Higgs sector in the SM can be done.  
(3) Special Relation to Higgs mass, together with W boson mass.  
   \[ y_t = \frac{\sqrt{2} m_t}{v} \approx 1 \]  
   \[ \delta M_W \propto (M_{\text{top}}^2, \ln(M_H)) \]  
(4) More detailed studies of top events  
   by using “M_{top}” can be performed.  
   i.e. ttbar resonance, P/CP test,  
   W helicity, new particle search etc.

RunII Goal : 2 ~ 3 GeV!
Review of Run I Results

**Old World Average (1999)**

\[ M_t = 174.3 \pm 5.1 \text{ GeV} \ (\pm 3.2 \pm 4.0) \]

(Fermilab-TM-2084)

**New DØ l+j measurement,**

(Nature 429, 638-642 (2004))

**New World Average (2004)**

\[ M_t = 178.0 \pm 4.3 \text{ GeV} \ (\pm 2.7 \pm 3.3) \]

**Standard Model Higgs Mass:**

Most probable: \[ 96 \text{ GeV} \rightarrow 113 \pm ^{62}_{42} \text{ GeV} \]

Upper limit (95% CL): \[ 219 \text{ GeV} \rightarrow 237 \text{ GeV} \]
RunII Data at CDF

- RunII Detector Upgrades
  - New silicon tracker (7-8 layers) ($|\eta|<2$)
  - New central drift chamber
  - New time of flight detector
  - Extended muon coverage ($|\eta|<1.5$)
  - New DAQ

- Record initial luminosity
  $$= 8.2 \times 10^{31} \text{ sec}^{-1} \text{ cm}^{-2}$$

- Data taking efficiency $\sim 80$-90%

- In this analysis, We use 162 pb$^{-1}$ collected until September 2003.
  Cf) RunI $\sim 110$ pb$^{-1}$
Event Selection

Kinematical cuts for “lepton+jets”

1) One lepton: central electron / muon
   \( \text{Et(Pt)} > 20 \text{ GeV}, \ |\text{eta}| < \sim 1.0 \)
2) \( \text{Met} > 20 \text{ GeV} \)
3) 4 tight jets: \( \text{Et} > 15 \text{ GeV}, \ |\text{eta}| < 2.0 \)
4) At least one SVX b-tagged jet

Exactly 4 jets
We do not use the events with more than 4 jets, to minimize the contaminations by initial and final state radiation.

Observed events:
Total 22 events; electrons 12, muons 10

SVX b jet tagging
B hadrons are long-lived.
Identify by Vertex of displaced tracks
The Method:
- Basic idea is to use matrix elements convoluted likelihood.

The year Leon M. Lederman won the Nobel Prize!

- Waseda colleagues have worked on the method for 5 years.
- The latest formulation was submitted to JPS.

Analysis information: Please visit CDF public web page,
http://www-cdf.fnal.gov/physics/PublicResults.html
Likelihood Definition in DLM

For $i$-th event, likelihood is defined as,

$$L^i(M_{\text{top}}) = \sum_{I_t} \sum_{I_s} \int \frac{2\pi^4}{\text{Flux}} F(z_a, z_b, p_T) \left| M \right|^2 \delta(s_w - (\ell + v)^2) w(I_t, x \mid y; M_{\text{top}}) dx ds_w$$

- $F$: Parton distribution function for $(z_a, z_b)$ and Pt of tt system
- $M$: Matrix element of tt process, $\left| M \right|^2 = \left| M_{\text{prod}} \right|^2 \prod(s_x) \left| M_{\text{dec}} \right|^2$
- $w$: Transfer function, $x$; partons $\leftrightarrow y$; observables

Two Summations and one integration:

- $I_t$: Possible combination (Jets to partons)
- $I_s$: Two $\nu_z$ solutions

In practice, integration of $x, s_w$ made by Monte Carlo Method.

Extract top mass by maximum likelihood method,

$$\Lambda(M_{\text{top}}) = -2ln \left( \prod_{\text{event}} L^i(M_{\text{top}}) \right) \rightarrow \overline{M}_{\text{top}} = M_{\text{top}} \text{ min. } \Lambda(M_{\text{top}})$$
**Performance**: \( \Lambda(M_{top}) \)

**Demonstration!!!**

For ppt user!

**Example**

-2*log(likelihood) after 10 event!
Jet Measurement and Energy Scale Correction at CDF

- All jets are formed by \( dR = 0.4 \) cone cluster algorithm.
- We start with jets corrected by,
  1. Calorimeter non-uniformity
  2. Calorimeter Scale
  3. Jets to hadrons

Transfer function
- To Go back to partons from jets, it is necessary.

Transferred variable = 
\[
\frac{E(\text{Parton}) - E(\text{Jet})}{E(\text{Parton})}
\]

- At present, we ignore the difference of directions between parton & jets.
Transfer Functions; \( w \) ~ Jets to Partons ~

\[
\frac{E(\text{Parton}) - E(\text{Jet})}{E(\text{Parton})}
\]
is asymmetric and depends on \( \text{Et} \) & \( \eta \) of the jets.

- **Et bin**: 9 bins
  - 15-25-35-45-55-65-75-85-95-<

  Strong Et dependence

- **Eta bin**: 3 bins
  - Central < 0.7
  - Wall 0.7~1.32
  - Plug > 1.32

30 histograms for each b/w jet. We do not fit them, but random generation along the shape to get parton momenta.
Comparisons between (CDF sim.) and (CDF sim.+Transfer function)

<table>
<thead>
<tr>
<th></th>
<th>2jets W mass</th>
<th>3jets Top mass</th>
<th>(\eta) dependence</th>
<th>Pt dependence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CDF sim.</strong></td>
<td><img src="image1" alt="Graph" /></td>
<td><img src="image2" alt="Graph" /></td>
<td><img src="image3" alt="Graph" /></td>
<td><img src="image4" alt="Graph" /></td>
</tr>
<tr>
<td><strong>CDF sim.+TF.</strong></td>
<td><img src="image5" alt="Graph" /></td>
<td><img src="image6" alt="Graph" /></td>
<td><img src="image7" alt="Graph" /></td>
<td><img src="image8" alt="Graph" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>W mass &lt;RMS&gt; (GeV)</th>
<th>Top mass &lt;RMS&gt; (GeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Before TF</strong></td>
<td>71.4 &lt;12.8&gt;</td>
<td>152.1 &lt;20.4&gt;</td>
</tr>
<tr>
<td><strong>After TF</strong></td>
<td>80.9 &lt;12.2&gt;</td>
<td>174.8 &lt;19.1&gt;</td>
</tr>
</tbody>
</table>

Input W : 80.4 GeV  
Input top : 175 GeV

Means are back to inputs and RMS gets better after TF.
Transfer function
Top Mass dependence

Plots: Mean of \( \frac{E(\text{Parton}) - E(\text{Jet})}{E(\text{Parton})} \) as a function of Et of jets.

- **b jets** has the top mass dependence.
- **We use the transfer function obtained from** \( M_{\text{top}} = 175 \text{ GeV} \).
Monte Carlo signal studies

- 22 events Pseudo expts using $M_{top} = 175$ GeV sample

Center of pull is consistent with 0. Width is consistent with 1.0.

Default Slope : 0.83
- Transfer function from 175 GeV.
- Due to mass dependence of TF.
- Use this slope to get input mass.
  Signal only mapping slope!

Q: Is this really coming from only TF?
A: Yes, checked,
- Use Transfer function at each mass.
  ex) 160 GeV sample $\rightarrow$ 160 GeV TF.
  Slope : 0.98
Pull distributions (signal only)

Outputs from 1000 sets of 22 signal events pseudo expts. after the correction to each set.

By taking into account the slope as a mapping function, top mass can be reconstructed correctly, no input mass bias for the center and width. In fact, ……… Go next slide.
Test Using Blind Samples made by CDF top mass group

- Six top mass samples: generated with randomly selected top masses using Herwig or Pythia. (200k events each).
- For users: unknown (top mass, Herwig or Pythia) known (no backgrounds, masses are in reasonable range)

DLM promises no biases and good performance (<1.0 GeV)!

Error stat only
Very Short Summary on the way…

• So far, We have studied Signal Monte Carlo and got reasonable results without bias by correcting transfer function mass dependence.

• Let’s move to a treatment of the backgrounds. We need to understand the background effects on signal likelihood distribution.
In the method, first, all events are assumed to be signal. Need correct background effect.

### Counting Method

![CDF II preliminary]

- **Expected Number of events**
  - **W+light flavor**: $1.2 \pm 0.37$
  - **Wbb**: $0.7 \pm 0.29$
  - **Wcc**: $0.3 \pm 0.12$
  - **Wc**: $0.2 \pm 0.12$
  - **Single top**: $0.17 \pm 0.03$
  - **WW**: $0.08 \pm 0.05$
  - **QCD**: $1.6 \pm 0.38$
  - **Bkg total**: $4.2 \pm 0.71$
  - **N observed**: 22
  - **tt (6.7pb)**: 20.9

Publication of this result coming soon.

The background fraction is estimated to be **19 ± 5 %**
We expect the background makes likelihood peak down when it is multiplied to signal events.
Background Effects Details

Reconstructed mass from 22 events Pseudo expts. by varying background fraction.

Mt=175 signal sample is used and Background Numbers are Poisson fluctuated.

Mass is shifted lower by background increase. Resolution also gets worse.

The size of shift is different in each source. But W+LF(mistag), QCD and Wbb(>80%) have similar behavior.
The Mapping Function

- Mass dependent correction factor.
- The mapping function is obtained from 2000 sets of 22 events (fixed) pseudo experiments in each point, by varying the background fraction with Poisson distributed.
- The background estimate in our sample is 19%.

Parameterization of the slope and the constant of the fit

![Graph showing the relationship between slope and background fraction.]

![Graph showing the relationship between constant and background fraction.]

June 11th, 2004

FNAL W&C Seminar - Kohei Yorita
How to apply the mapping function?

Input:
1. Reconstructed Mass & Error from the sample
2. Background fraction in the sample

Example

* Statistical errors are also scaled by the slope properly.

Extract the top mass

June 11th, 2004  FNAL W&C Seminar - Kohei Yorita
Pull distribution of $S:17.8\text{ev} + B:4.2\text{ev} = 22\text{ events}$

Plots after applying the mapping function of 19% background to each output mass.

No bias! Even if the sample includes backgrounds.
Extracted top mass from RunII

- Observed: Total 22 events; electrons 12, Muons 10

22 events joint likelihood

Correct background-pulling
4.2 events expected.

Fit: Two 2nd order polynomials for positive/negative errors.
**Expected statistical uncertainties**

**Before Mapping function applied**

Black arrows: Data + 3.4, − 3.8 GeV

**Mean:** + 4.2, − 3.8 GeV  
**MPV:** + 3.5, − 3.2 GeV

**After Mapping function applied**

Black arrows: Data + 4.5, − 5.0 GeV

**Mean:** + 5.4, − 5.0 GeV  
**MPV:** + 4.5, − 4.1 GeV

Stat. error is scaled by ~30%, due to the mapping slope.
Systematic uncertainty
Jet Energy uncertainty

Jet Energy Scale

<table>
<thead>
<tr>
<th>Description</th>
<th>$\Delta M_{top}$ GeV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calorimeter non-uniformity</td>
<td>$2.6 \pm 0.1$</td>
</tr>
<tr>
<td>Central Calorimeter Response</td>
<td>$4.2 \pm 0.2$</td>
</tr>
<tr>
<td>Corrections to Hadrons</td>
<td>$2.0 \pm 0.1$</td>
</tr>
<tr>
<td>Total</td>
<td>$5.3 \pm 0.3$</td>
</tr>
</tbody>
</table>

Transfer Function “Jets $\leftrightarrow$ Parton Probability density”

- Took out-of-cone corrections uncertainty for fragmentations, $1.6$ GeV
- Hadronization model is already included in generator systematics.
- Correlated ISR/FSR, jet smearing 15%.
- It is very hard to validate whether the shape of transfer function in MC can represent that of data correctly. We will continue the works by doing,
  1. hadronic W mass,
  2. $Z \rightarrow b\bar{b}$,
  3. $b$jet-gamma balance e.t.c

For now, we take 2.0 GeV error to be conservative.
We checked several ways such as On/Off, diff $\alpha_s$, Run I like.

**ISR (Initial State Radiation)**
We have now the following two samples set the parameters range by comparing with Drell-Yan data.

<table>
<thead>
<tr>
<th>Description</th>
<th>$\Delta M_{top} = M - M_{\text{default}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISR less</td>
<td>$-0.2 \pm 0.2$ GeV</td>
</tr>
<tr>
<td>ISR more</td>
<td>$+0.3 \pm 0.2$ GeV</td>
</tr>
</tbody>
</table>

More ISR : $\Lambda_{QCD} = 384, K = 0.5$
Less ISR : $\Lambda_{QCD} = 100, K = 2.0$
Run I: no ISR: $K = \text{infinite}$

ISR systematic : 0.5 GeV

**FSR (Final State Radiation)**

<table>
<thead>
<tr>
<th>Description</th>
<th>$\Delta M_{top} = M - M_{\text{default}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FSR Less</td>
<td>$-0.4 \pm 0.2$ GeV</td>
</tr>
<tr>
<td>FSR More</td>
<td>$+0.5 \pm 0.2$ GeV</td>
</tr>
</tbody>
</table>

More FSR : $\Lambda_{QCD} = 384, K = 0.5$
Less FSR : $\Lambda_{QCD} = 100, K = 2.0$

FSR systematic : 0.5 GeV
Parton Distribution Function

**PDF General**

Add in quadrature the following,
(1) CTEQ6M 20 pairs of eigenvectors
(2) Two different $\alpha_s$ with MRST
(3) CTEQ5L vs MRST

<table>
<thead>
<tr>
<th>Sources</th>
<th>$\Delta M_{top}$ GeV</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>1.4 ± 0.4</td>
</tr>
<tr>
<td>(2)</td>
<td>0.6 ± 0.2</td>
</tr>
<tr>
<td>(3)</td>
<td>1.1 ± 0.3</td>
</tr>
<tr>
<td>Total</td>
<td>1.9 ± 0.5</td>
</tr>
</tbody>
</table>

PDF systematic : 2.0 GeV

**NLO vs LO PDF**

PDF :gg 15%(NLO), 5%(LO)
We reweight gg contribution to 15%(NLO) from 5%(LO).
This makes a difference of 0.4 GeV.

NLO PDF effect : 0.4 ± 0.2 GeV
Other Systematic uncertainties

- **Generator** 0.6 ± 0.2 GeV
  Pythia and Herwig
  (also checked Grappa, 0.6 GeV)

- **Spin correlations (Herwig)**
  We ignore spin corr. matrix element in the likelihood.
  On/Off difference is 0.7 GeV
  0.4 ± 0.2 GeV

- **Additional jet smearing**
  Resolution underestimated in our MC
  Additional 15% gives 0.6 ± 0.2 GeV

- **Unclustered Energy** 0.1 ± 0.2 GeV
  The effect of difference for Unclustered calorimeter energy

- **Background Modeling** 0.5 GeV

- **Background Fraction** 0.5 GeV
  Error due to 19±5%
Dominated by Jet Energy Scale.

Improvements by a better understanding of our simulation i.e. calorimeter response will be coming very soon.

More understanding of transfer function will reduce the error.

Avoid both under/over estimate.(correlations)

<table>
<thead>
<tr>
<th>Sources</th>
<th>$\Delta M_{top}$ (GeV/c$^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jet Energy Scale</td>
<td>5.3</td>
</tr>
<tr>
<td>ISR</td>
<td>0.5</td>
</tr>
<tr>
<td>FSR</td>
<td>0.5</td>
</tr>
<tr>
<td>PDF</td>
<td>2.0</td>
</tr>
<tr>
<td>Generator</td>
<td>0.6</td>
</tr>
<tr>
<td>Spin correlation</td>
<td>0.4</td>
</tr>
<tr>
<td>NLO effect</td>
<td>0.4</td>
</tr>
<tr>
<td>Bkg fraction($\pm$5%)</td>
<td>0.5</td>
</tr>
<tr>
<td>Bkg Modeling</td>
<td>0.5</td>
</tr>
<tr>
<td>MC Modeling(jet,UE)</td>
<td>0.5</td>
</tr>
<tr>
<td>Transfer function</td>
<td>2.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>6.2</strong></td>
</tr>
</tbody>
</table>
Currently DLM is the most precise measurement in RunII.
Event likelihood
For i-th event,
\[ L_{ev}^i = \int L^i (M) dM \]

An event has one likelihood.

Although absolute value does not have any meaning, we can compare Monte Carlo with Data directly.

Agreement is quite good!

More likely to be background
More checks of MC vs Data (2)

How about Maximum likelihood mass in each event? (gives us shape information!)

- Maximum Likelihood mass in each event from MC.

Comparison between MC and Data.

- Signal RMS: ~22 GeV
- cf) Template: ~30 GeV

Signal t\bar{b}(MC)

Combined background(MC)

CDF Run II Preliminary (162 pb⁻¹)
Signal MC: $M_{\text{top}} = 175\text{GeV/c}^2$

MC Prediction

Signal t\bar{b}(MC)

Background(MC)

Data 22 events
Conclusions and Plans

**Top Mass Results**
- From 22 events with 4.2 event background (162 pb^-1), We measured top mass by DLM to be,

\[ M_{\text{top}} = 177.8 \pm 4.5^{+5.0}_{-5.0} \text{(stat.)} \pm 6.2 \text{ (syst.) GeV/c}^2 \]

- Currently, this is the most precise measurement using RunII data. Total: \(+7.7\) GeV/c\(^2\), \(-8.0\) GeV/c\(^2\), Cf) world average : \(178\pm4.3\, (\pm2.7\pm3.3)\) GeV/c\(^2\)

**Things to do**
-- Reduce systematics (JES!)
-- Hadronic W mass measurement in this channel.
-- Get more data.

DLM is a very powerful method which can test both the Standard model and beyond.
Backup slides
More comments on DLM

In principle, \( L^i(M_{\text{top}}) = N \left( \frac{d\sigma}{d\Phi} \right)_c w(I_t, x \mid y; M_{\text{top}}) \) (1)

1. Normalization,

\[
N = N(y) = \left( \int \left[ \frac{d\sigma}{d\Phi} \right]_c w(x \mid y; M) d\Phi dM \right)^{-1}
\]

N(y) gives no effect on top mass measurement. \( \longrightarrow \) constant

2. Phase space,

\( d\Phi \) : assume that each final state parton occupies a unit phase volume in each event.

\[
(1) \longrightarrow L^i(M_{\text{top}}) = \sum_{I_t} \sum_{I_s} \int \frac{2\pi^4}{\text{Flux}} F(z_a, z_b, p_T) |M|^2 \delta(s_w - (\ell + v)^2)w(I_t, x \mid y; M_{\text{top}}) dx ds
\]
Comparisons between DLM and DØM(RunI)

~ These two Method look very similar, but Not identical! ~

DLM : \[ L_i(M_{top}) = \sum_{l_x} \sum_{l_y} \int \frac{2\pi^4}{Flux} F(z_a, z_b, p_T) |M|^2 \delta(s_w - (\ell + v)^2) w(l_x, x | y; M_{top}) dxds_w \]

DØM : \[ \bar{P}(x; \alpha) = \frac{1}{\sigma} \int d^n\sigma (y; \alpha) dq_1 dq_2 f(q_1) f(q_2) W(x, y) \]

<table>
<thead>
<tr>
<th>Event</th>
<th>DLM(RunII)</th>
<th>DØM(RunI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Event</strong></td>
<td>b-tagged, only 4jets (22) No more..</td>
<td>Non-btag, only 4jets (71) Background likelihood cut (22)</td>
</tr>
<tr>
<td><strong>Background</strong></td>
<td>4.2 events (~20%) (QCD Fake 40%)</td>
<td>10 events (~50%) (W+jets 85%)</td>
</tr>
<tr>
<td><strong>Treatment</strong></td>
<td>Mapping</td>
<td>In the likelihood</td>
</tr>
<tr>
<td><strong>Likelihood</strong></td>
<td>Signal likelihood only</td>
<td>Signal and background likelihood</td>
</tr>
<tr>
<td><strong>Others</strong></td>
<td>-</td>
<td>Normalization &amp; acc correction</td>
</tr>
</tbody>
</table>
More on backgrounds

5 events pseudo experiment using only each background source

- **Mistag (Wbb/nonW)**
- **Single top**
- **Wcc (Wc/WW)**

Max. likelihood mass in each event

Mapping slope

<table>
<thead>
<tr>
<th>source</th>
<th>Joint likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mistag</td>
<td>~162 GeV</td>
</tr>
<tr>
<td>Wbb</td>
<td>~162 GeV</td>
</tr>
<tr>
<td>Wcc</td>
<td>~157 GeV</td>
</tr>
<tr>
<td>Wc</td>
<td>~158 GeV</td>
</tr>
<tr>
<td>nonW</td>
<td>~161 GeV</td>
</tr>
<tr>
<td>Single top</td>
<td>~170 GeV</td>
</tr>
<tr>
<td>WW</td>
<td>~160 GeV</td>
</tr>
</tbody>
</table>

June 11th, 2004  
FNAL W&C Seminar - Kohei Yorita
How Likelihood looks like?

Signal example: \(- \log(\text{likelihood})\)

Blue: all added up
Red: right comb.
Black: wrong comb.

Peak around 175 GeV

Bkg example: \(- \log(\text{likelihood})\)

Blue: all added up
Black: each comb.

Range \([155-195]\text{GeV}\)

Likelihood tends to be higher in lower mass region.