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

Fermilab

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

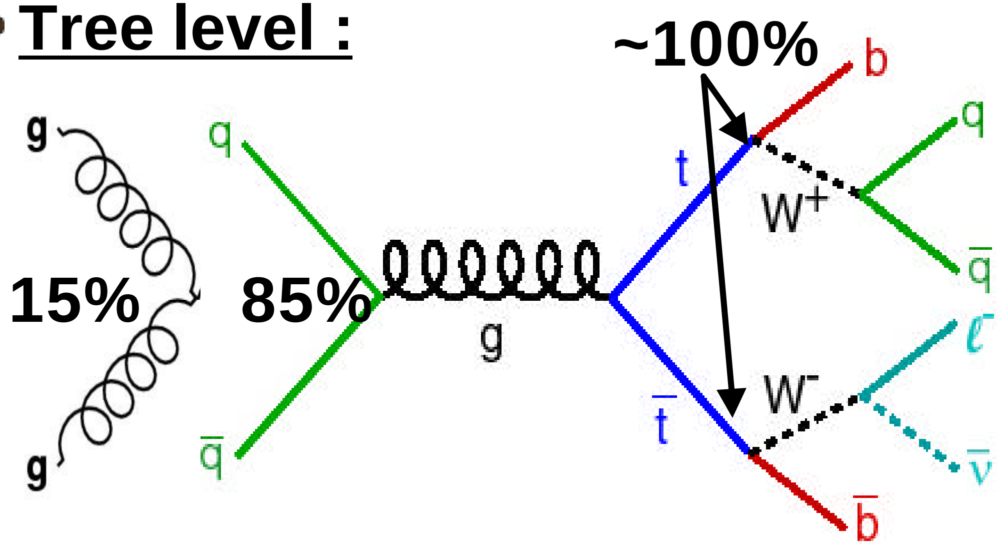
   & Some Comparisons
9. Conclusions

Ms. Particle

Q: How many particles in me?

Top quark production and decay

- **Tree level :**



- **The Tevatron is the only place which can produce top quark until LHC runs!**
  ($q\bar{q}$, gg fractions reversed at LHC)
- **s = 5.8 ~7.4 pb** at 1.96TeV **(Cacciari et al.)**

- **Final states :**

• Dilepton (e,m)	BR=5%
• Lepton (e, m) +jets	BR=30%
• All jets	BR=44%
• t + X	BR=21%

- **Use this channel with at least 1 bjet.**
  *** Higher statistics**
  *** Lower background**

Why is Top Mass interesting?

(1) Fundamental Standard Model parameter.

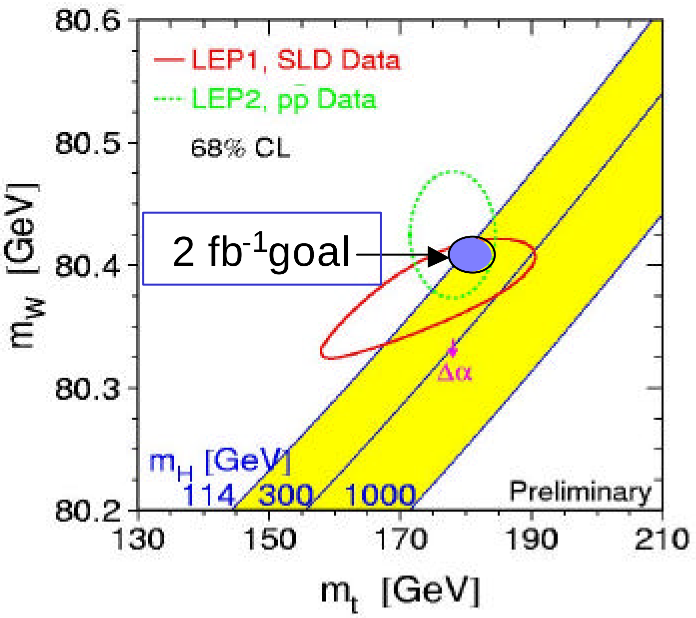
(2) Top quark is heavy (~ 180 GeV)
Yukawa coupling ~ 1.

$$y_t = \frac{\sqrt{2} m_t}{v} \approx 1$$

*** 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.

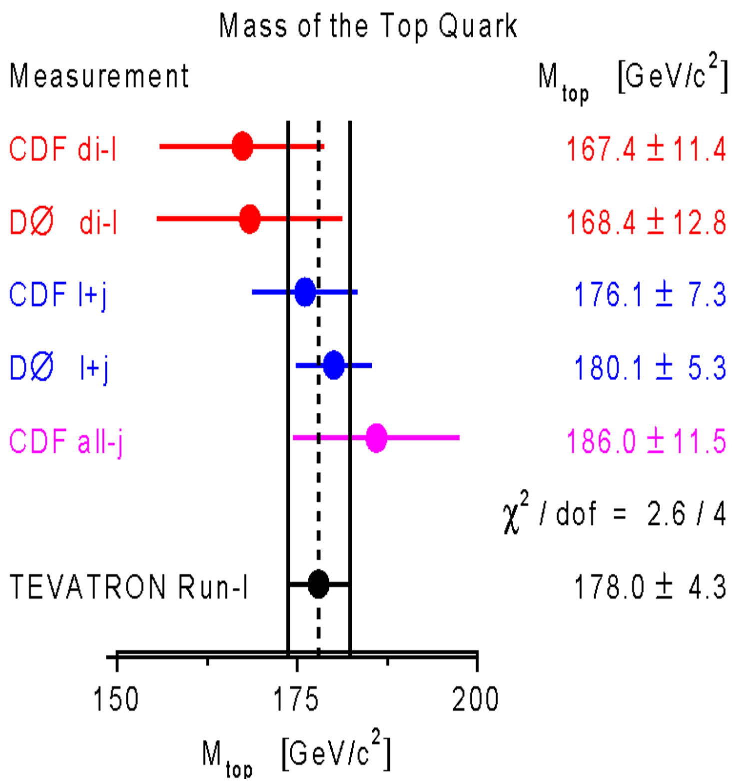


$$dM_W \propto (\, M_{top}^2,\ \ln(M_H)\,)$$

(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 ± 5.1 GeV (±3.2±4.0)

  (Fermilab-TM-2084)

  New DØ l+j measurement, (Nature 429, 638-642 (2004))

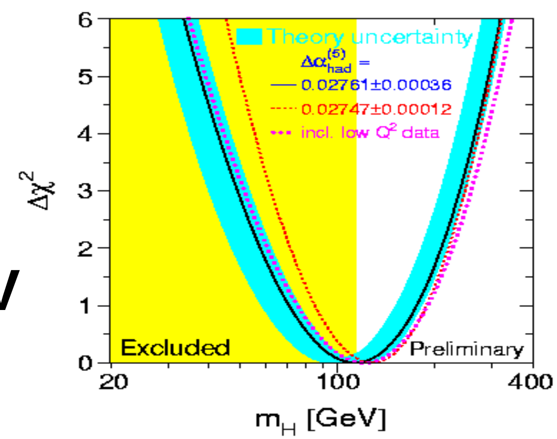
- **New World Average (2004)**

  M_t = 178.0 ± 4.3 GeV (±2.7±3.3)

  hep-ex/0404010

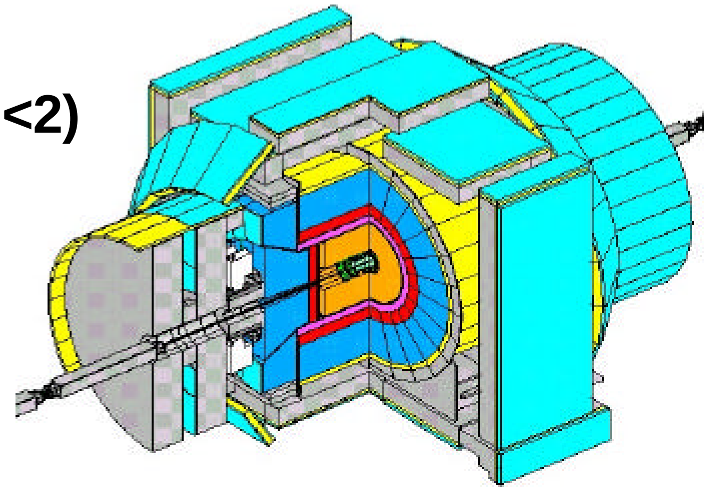
- **Standard Model Higgs Mass:**

Most probable: 96 GeV ® 113 $\pm^{62}_{42}$ GeV
Upper limit(95% CL):219 GeV ® 237 GeV



RunII Data at CDF

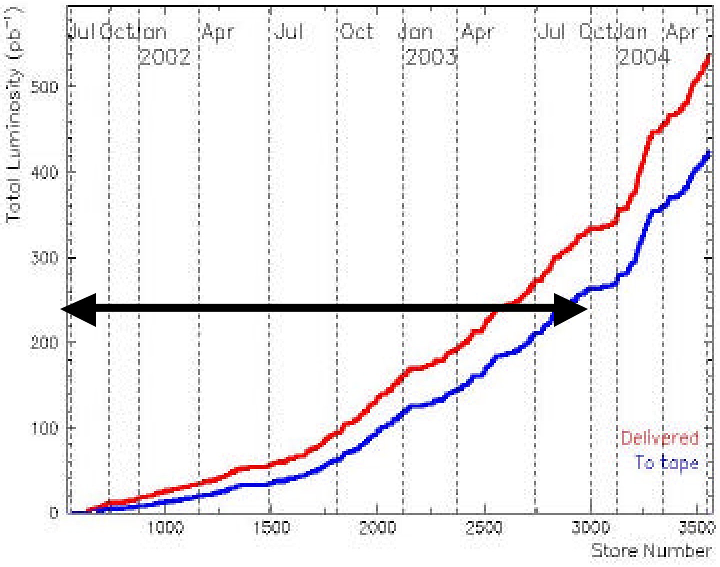
- **RunII Detector Upgrades**
  - New silicon tracker (7-8 layers) (|h|<2)
  - New central drift chamber
  - New time of flight detector
  - Extended muon coverage (|h|<1.5)
  - New DAQ



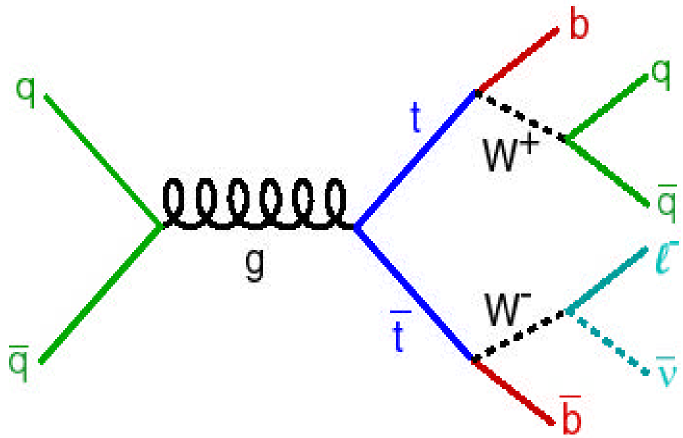
- **Record initial luminosity**
  = 8.2 ´ 10^{31} sec^{-1} cm^{-2}

- **Data taking efficiency ~80-90%**

- **In this analysis,**
  We use 162 pb^{-1} collected until September 2003.
  Cf) RunI ~110 pb^{-1}



Event Selection



- **Kinematical cuts for "lepton+jets"**
  1) One lepton : central electron / muon
     Et(Pt) > 20 GeV, |eta| < ~1.0
  2) Met > 20 GeV
  3) 4 tight jets : Et > 15 GeV, |eta| < 2.0
  4) At least one SVX b-tagged jet

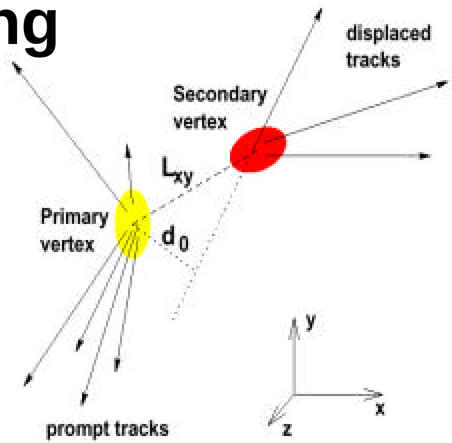
- **Exactly 4 jets**

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

- **SVX b jet tagging**

  B hadrons are long-lived.

  Identify by Vertex of displaced tracks



- **Observed events :**

  Total 22 events ; electrons 12, muons 10

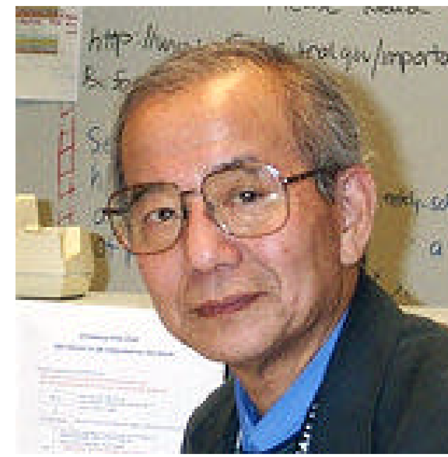
DLM Introduction

- **The Method :**

- Basic idea is to use matrix elements convoluted likelihood.
- Originally proposed in 1988 by K. Kondo.(J.Phys. Soc. 57, 4126)

The year Leon M. Lederman won the Nobel Prize !

For the neutrino beam method and the demonstration of the doublet structure of the leptons though the discovery of the muon neutrino.



- 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,***

$$L^i(M_{top}) = \sum_{I_t}\sum_{I_s}\int \frac{2\pi^4}{Flux} F(z_a, z_b, p_T)\,|M|^2\, d(s_w - (\ell + \nu)^2)\, w(I_t, \mathbf{x}\,|\,\mathbf{y}; M_{top})\, d\mathbf{x}\, ds_w$$

F : Parton distribution function for (z_a,z_b) and Pt of tt system

M : Matrix element of tt process, $|M|^2 = |M_{prod}|^2 \prod(s_\tau)\,|M_{dec}|^2$

w : Transfer function, x ; partons ⟷ y ; observables

$S_w=(l+\nu)^2$

- ***Two Summations and one integration :***

I_t : Possible combination(Jets to partons) , I_s : Two ν_z solutions,

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

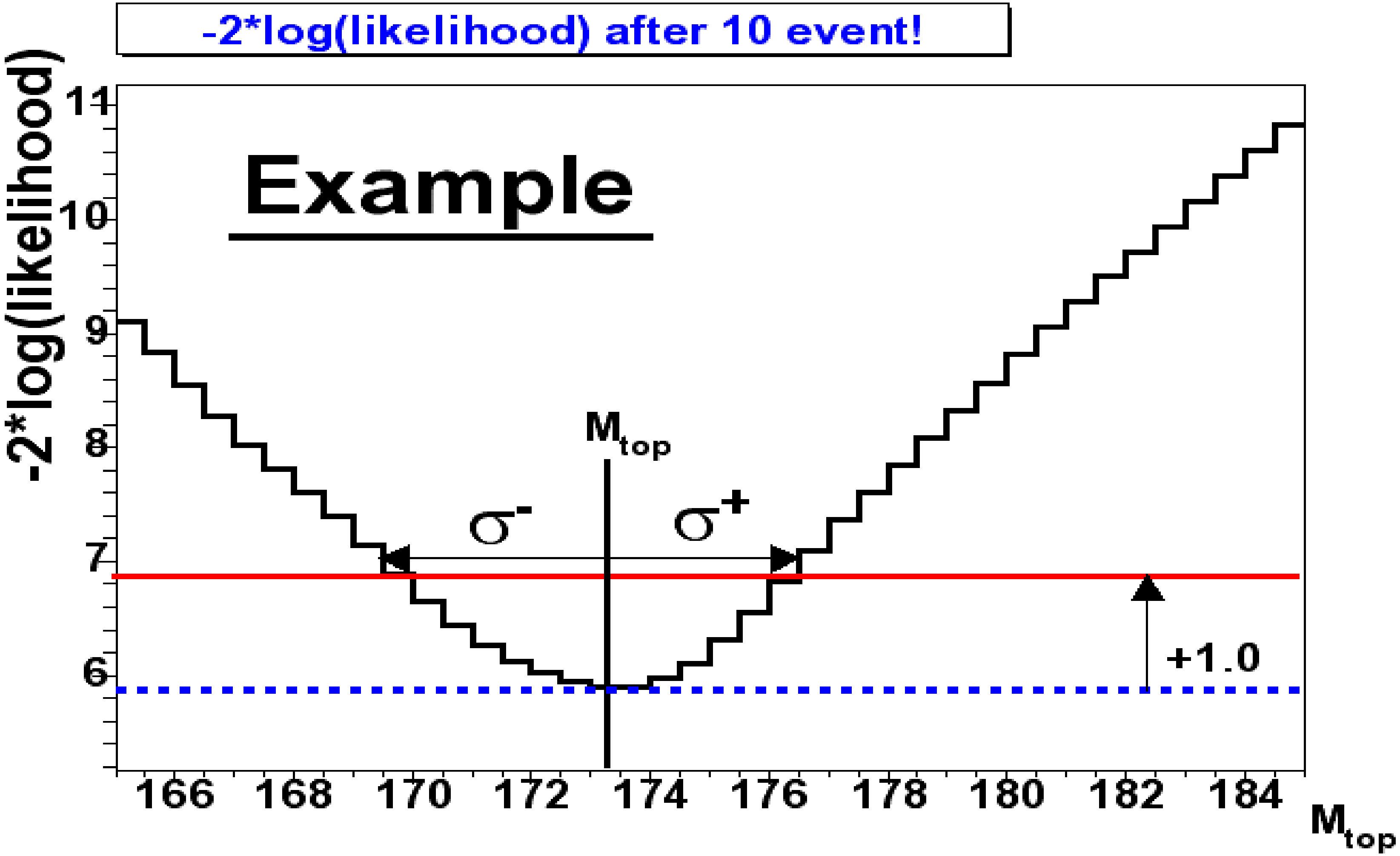
- ***Extract top mass by maximum likelihood method,***

$$\Lambda(M_{top}) = -2ln\left(\prod_{event} L^i(M_{top})\right) \longrightarrow \overline{M_{top}} = M_{top}\ min.\ \Lambda(M_{top})$$

Performance : $\Lambda(M_{top})$

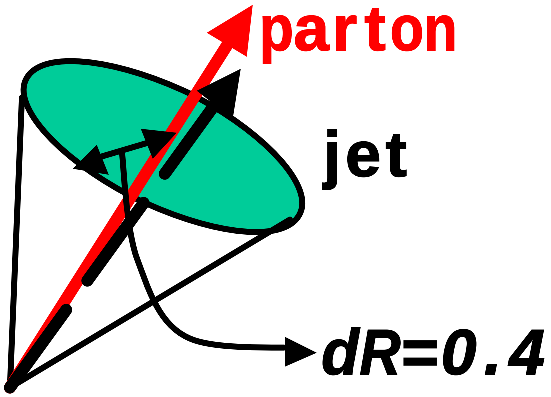
<u>**Demonstration!!!**</u>

<u>**For ppt user!**</u>



Transfer Functions ; *w*
~ *Jets to Partons* ~

- **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 = $\dfrac{E(Parton) - E(Jet)}{E(Parton)}$
  - At present, we ignore the difference of directions between parton & jets.

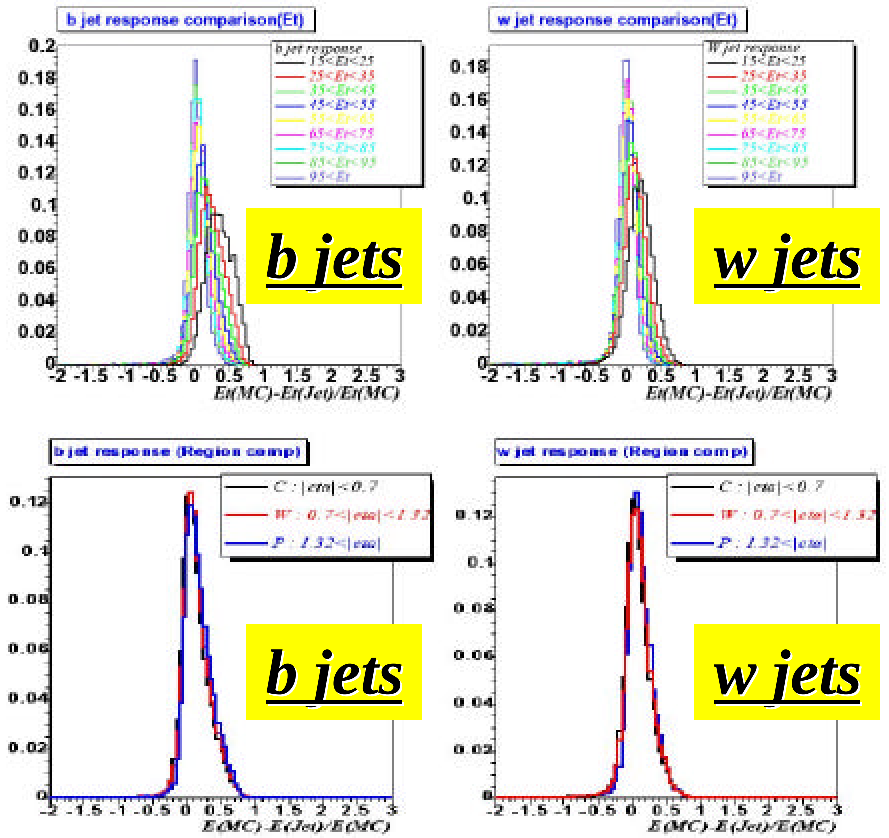
Transfer Functions ; *w*
~ *Jets to Partons* ~

- $\frac{E(Parton) - E(Jet)}{E(Parton)}$ is asymmetric and depends on Et & η 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

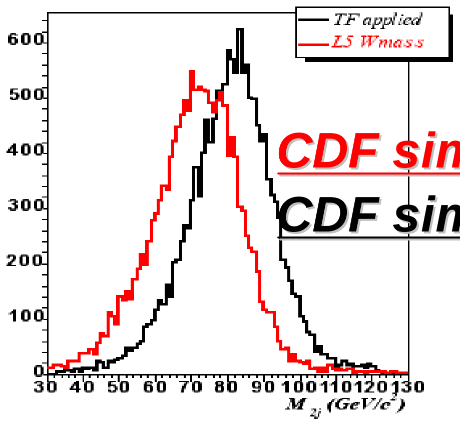


- 30 histograms for each b/w jet. We do not fit them, but random generation along the shape to get parton momenta.

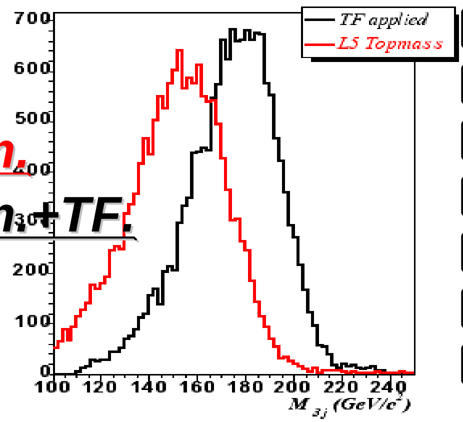
Transfer Function performance

- **Comparisons between (CDF sim.) and (CDF sim.+Transfer function)**

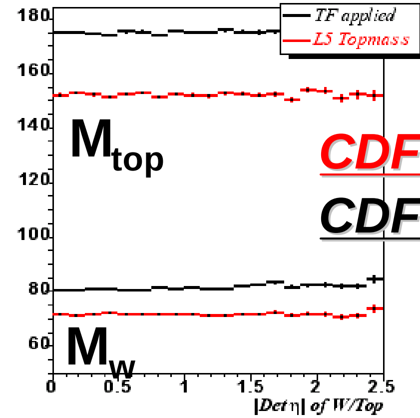
2jets W mass



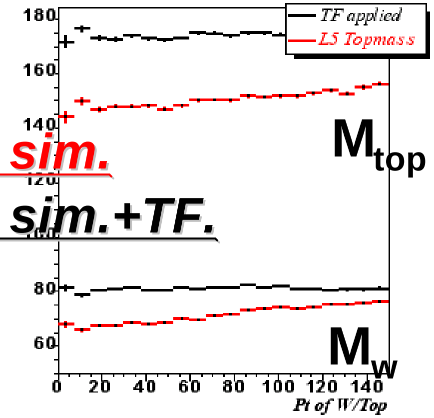
3jets Top mass



η dependence



Pt dependence



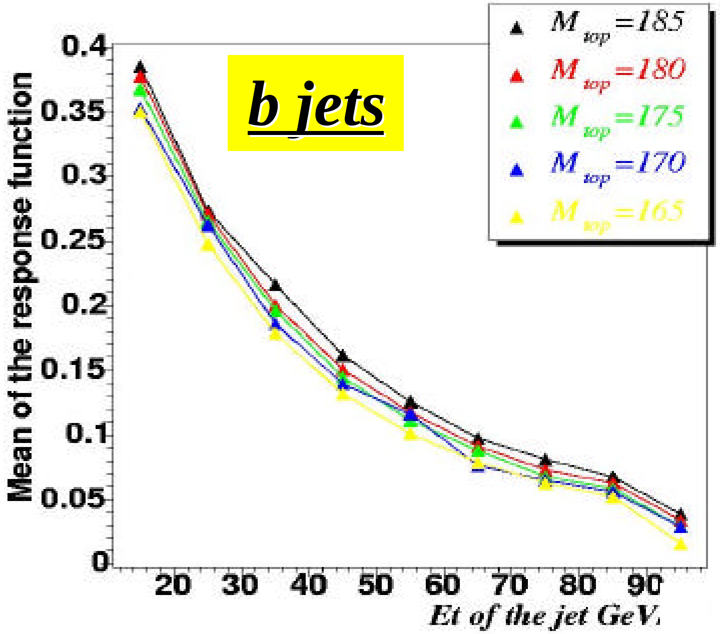
	W mass <RMS> (GeV)	Top mass <RMS> (GeV)
Before TF	71.4 <12.8>	152.1 <20.4>
After TF	80.9 <12.2>	174.8 <19.1>

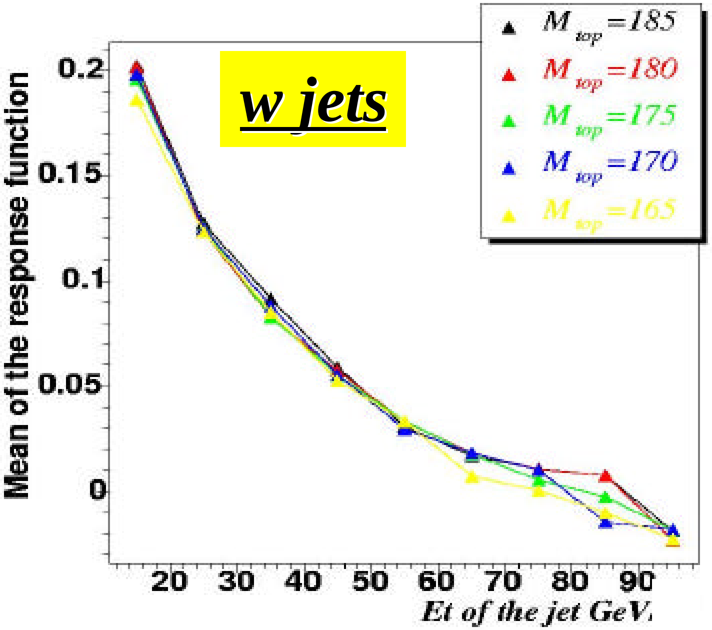
Input W : 80.4 GeV **Input top : 175 GeV**

- **Means are back to inputs and RMS gets better after TF.**

Transfer function
Top Mass dependence

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

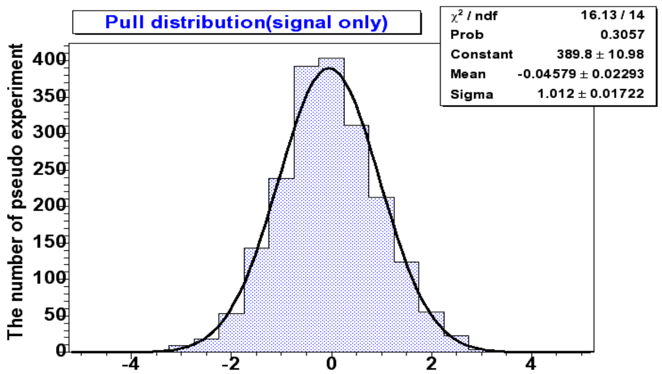




- **bjet has the top mass dependence.**
- **We use the transfer function obtained from M_{top}=175 GeV .**

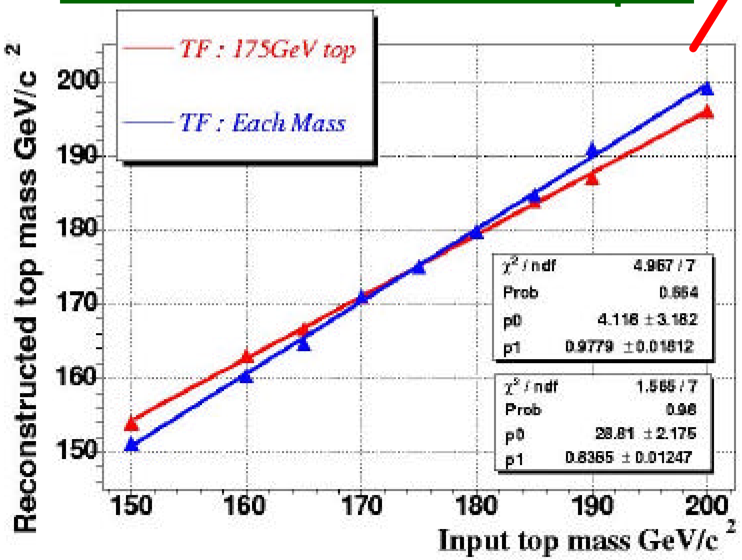
Monte Carlo signal studies

- <u>22 events Pseudo expts using M_{top} = 175 GeV sample</u>



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

- <u>Different Mass sample</u>



<u>Default Slope : 0.83</u>

- 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?

<u>A: Yes, checked,</u>

- Use Transfer function at each mass.

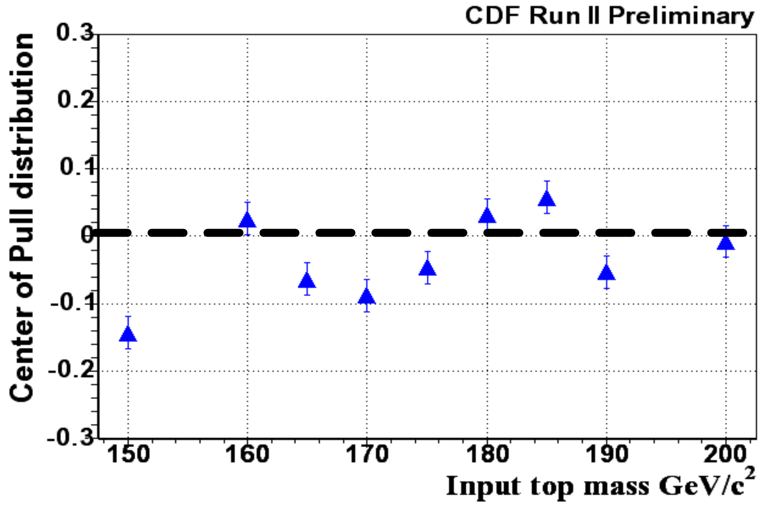
ex) 160 GeV sample ← 160 GeV TF.

→ Slope : 0.98

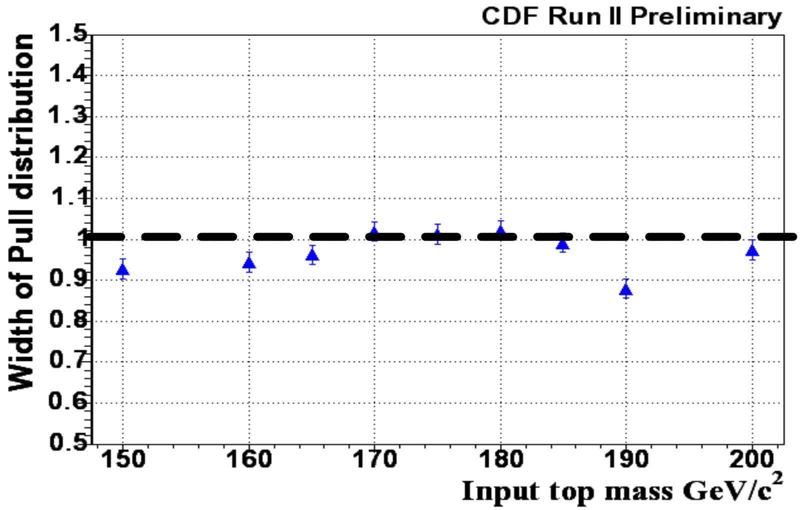
Pull distributions(signal only)

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

- **Center of Pull**



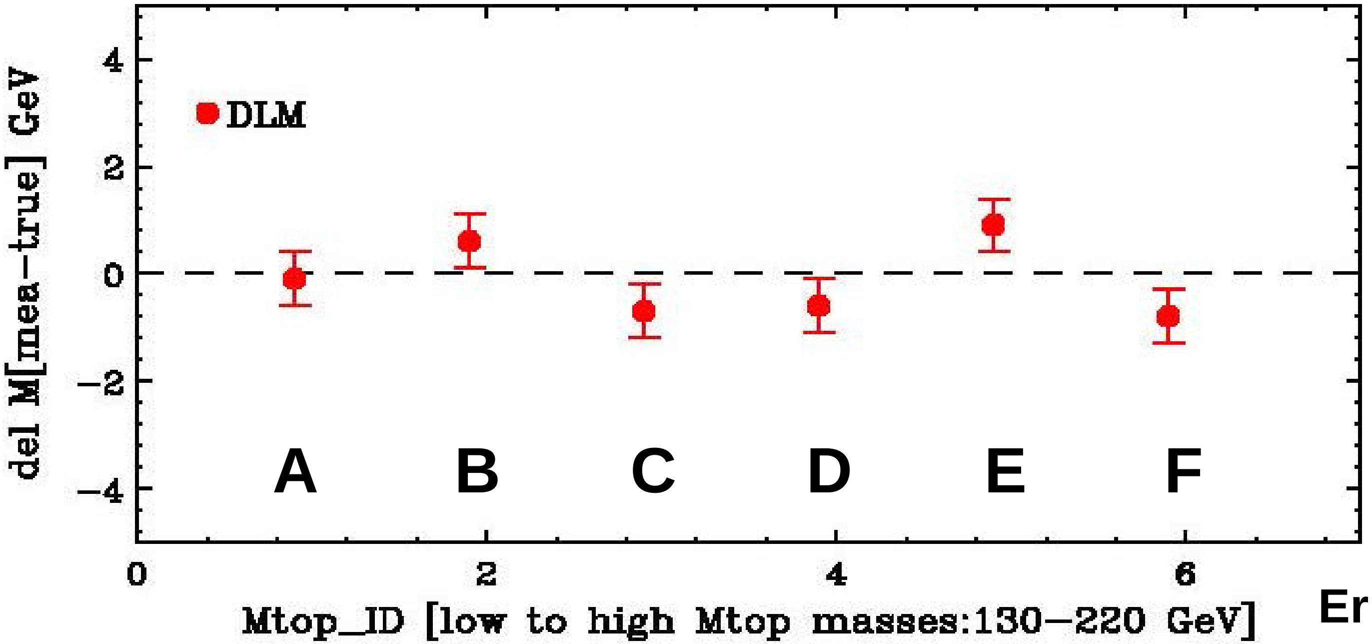
- **Width of Pull**



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)



Error stat only

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

Very Short Summary on the way…

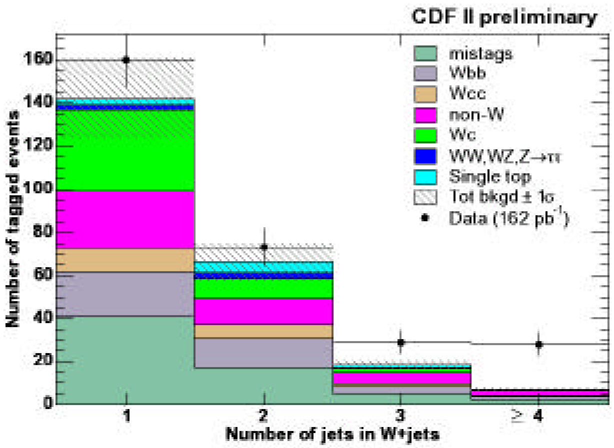
- So far, We have studied Signal Monte Carlo and got reasonable results without bias by correcting transfer function mass dependence.

- Let’s move to a treatment of the backgrounds.
  We need to understand the background effects on signal likelihood distribution.

Background Estimate

- In the method, first, all events are assumed to be signal. Need correct background effect.

Counting Method



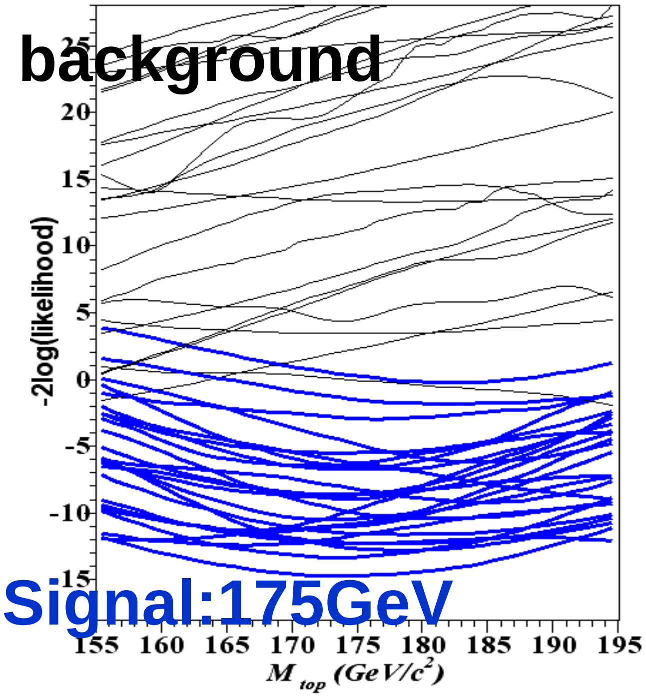
- Publication of this result coming soon.

The background fraction is estimated to be 19 ± 5 %

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

Understanding of background
~ Likelihood Distributions ~

Parton Level



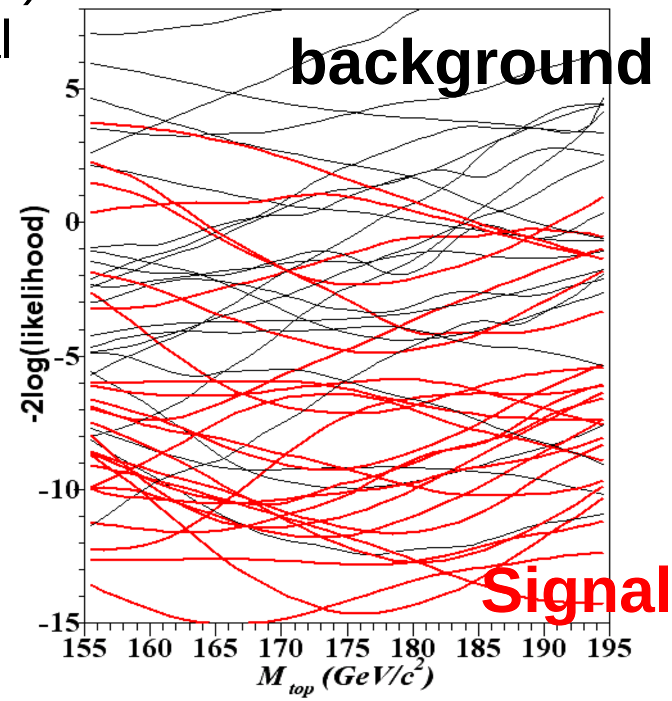
Typical –2ln(likelihood) distributions for signal and backgrounds.

Peak :

Backgrounds : lower mass
Signal : input mass

Absolute value of likelihood in signal is much larger than that in background

CDF simulation



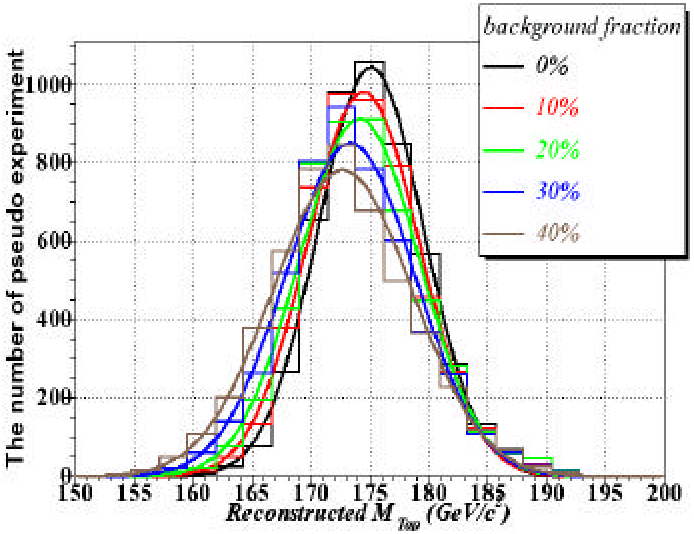
We expect the background makes likelihood peak down when it is multiplied to signal events.

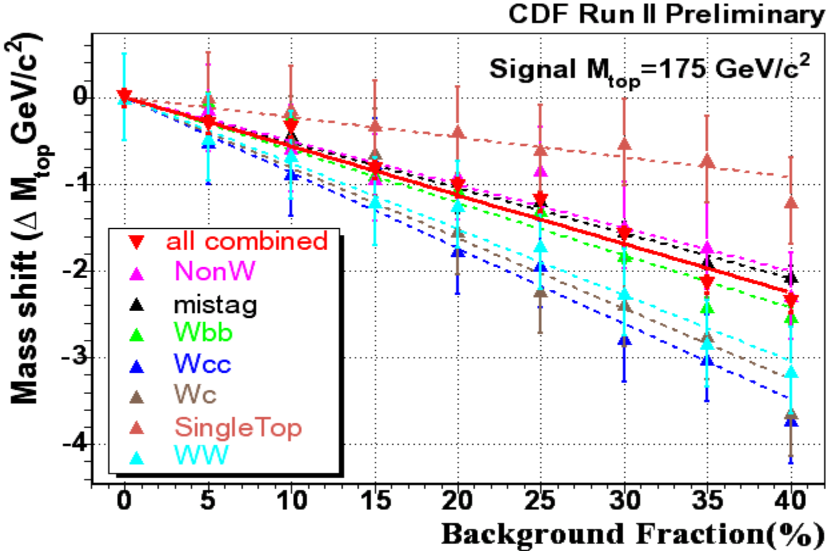
Background Effects Details

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

Reconstructed Mass Shift due to each background source.

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



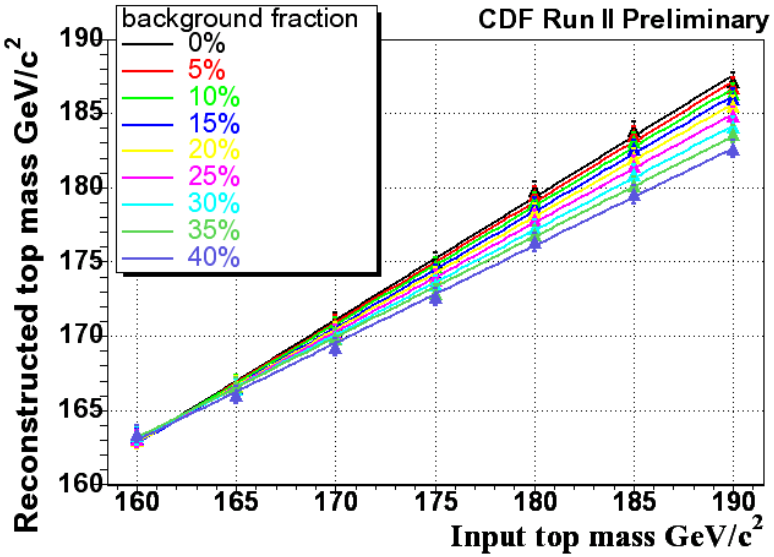


Mass is shifted lower by background increase. Resolution also gets worse.

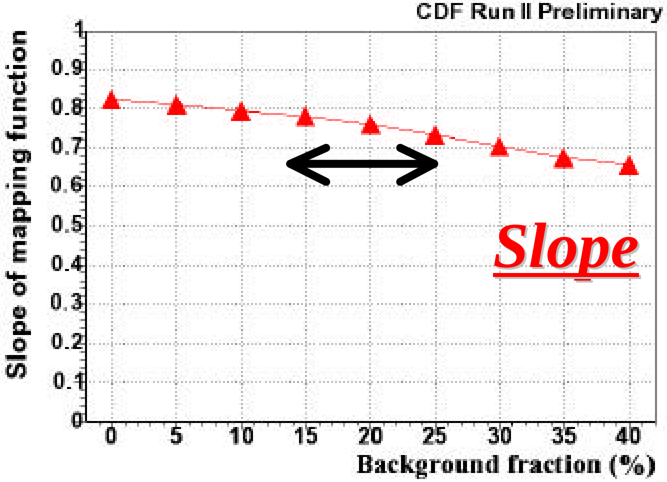
The size of shift is different in each source. But W+LF(mistag), QCD and Wbb(>80%) have similar behavior.

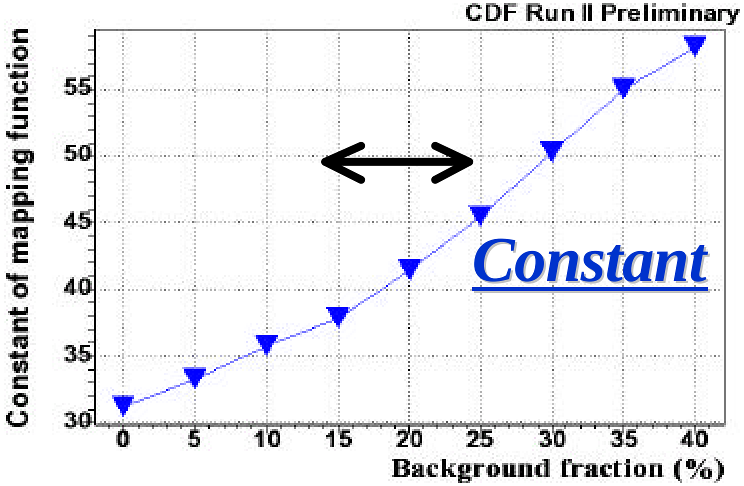
The Mapping Function

- 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

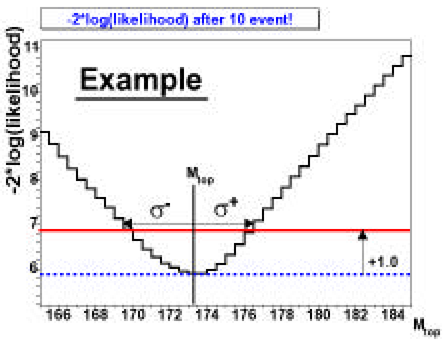


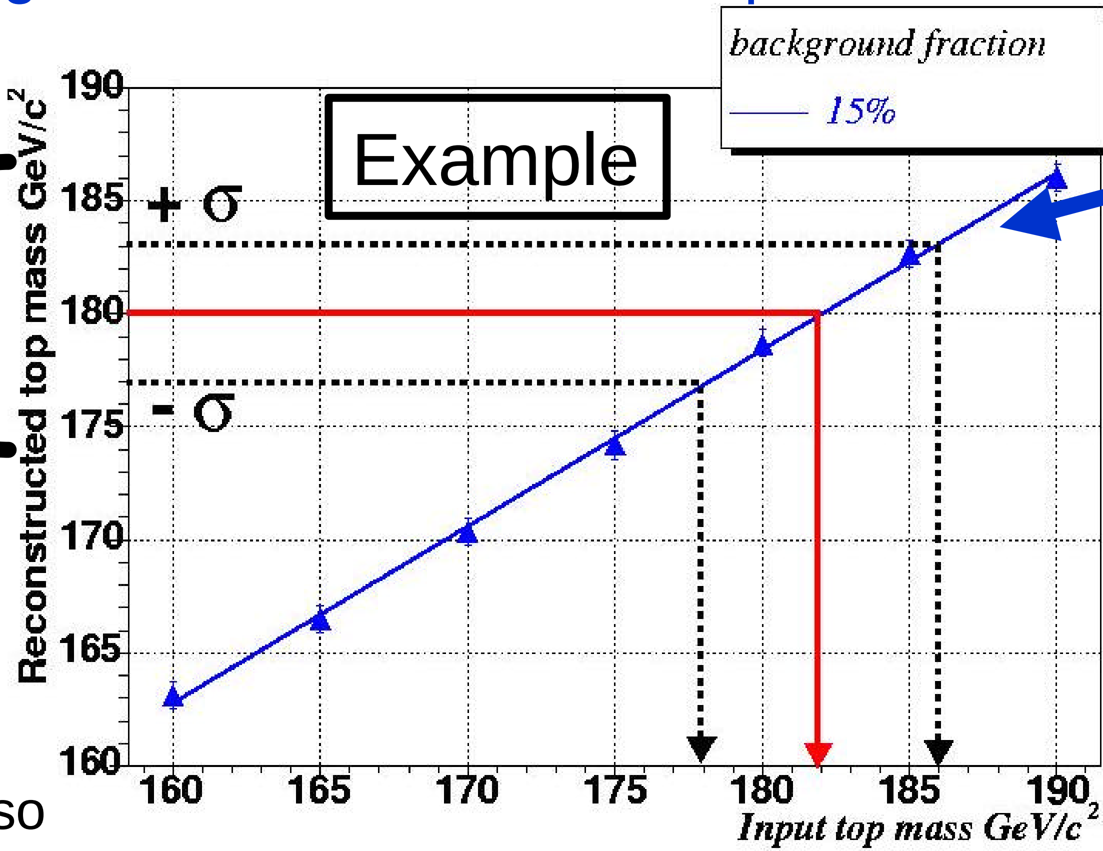


How to apply the mapping function?

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

(1)





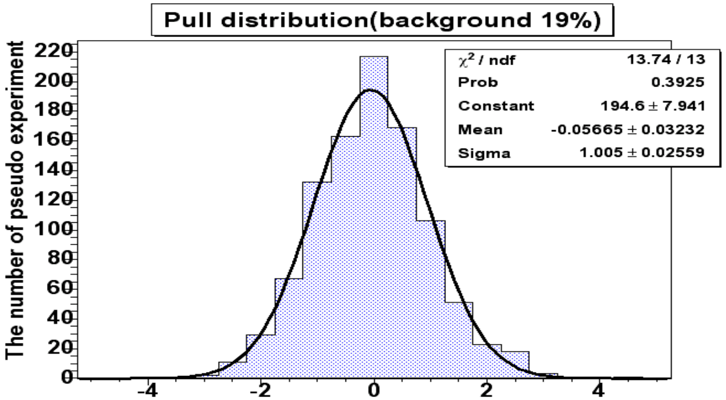
(2)

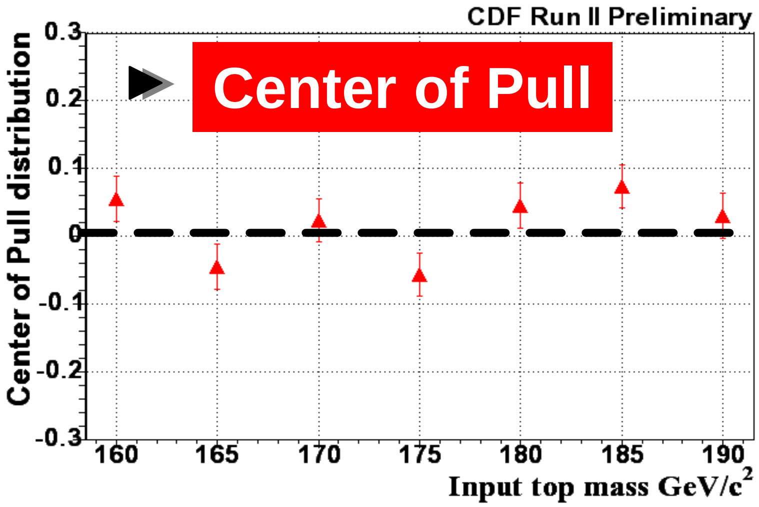
* Statistical errors are also scaled by the slope properly.

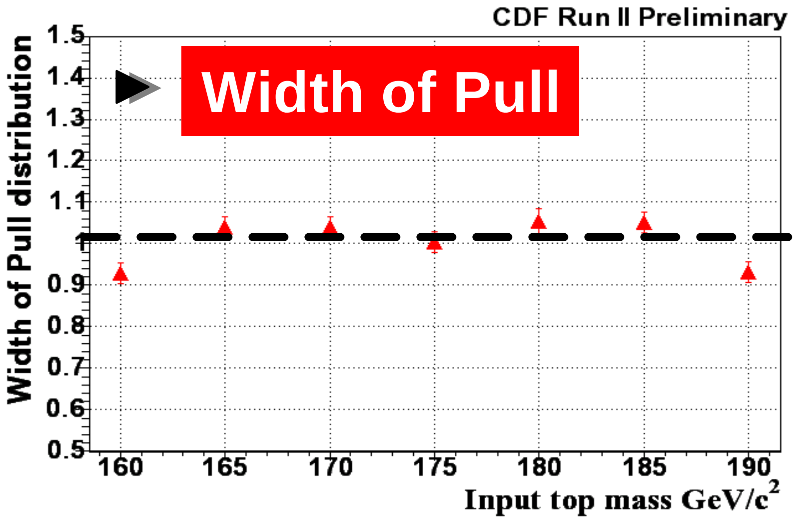
Extract the top mass

Pull distribution of S:17.8ev + B:4.2ev = 22 events

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







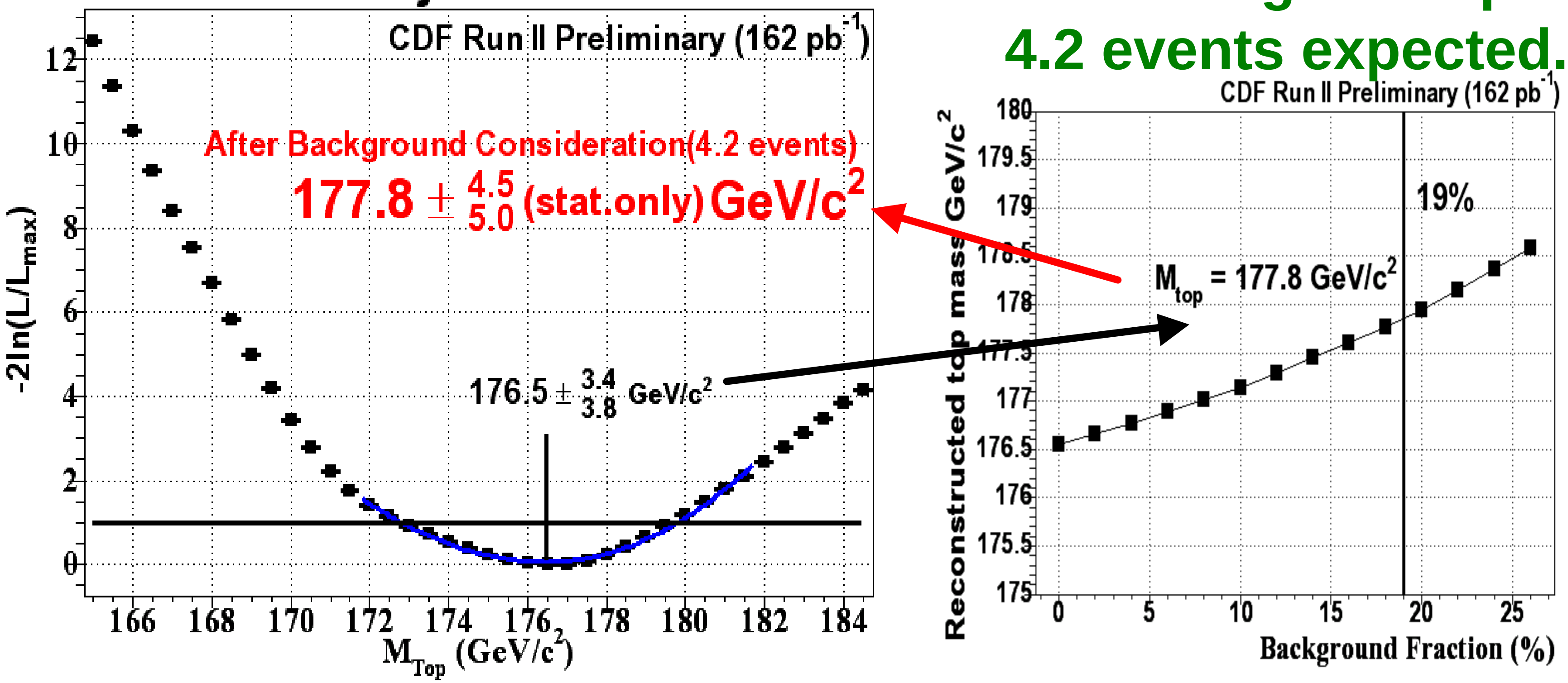
No bias ! Even if the sample includes backgrounds.

Extracted top mass from RunII

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

22 events joint likelihood

Correct background-pulling
4.2 events expected.



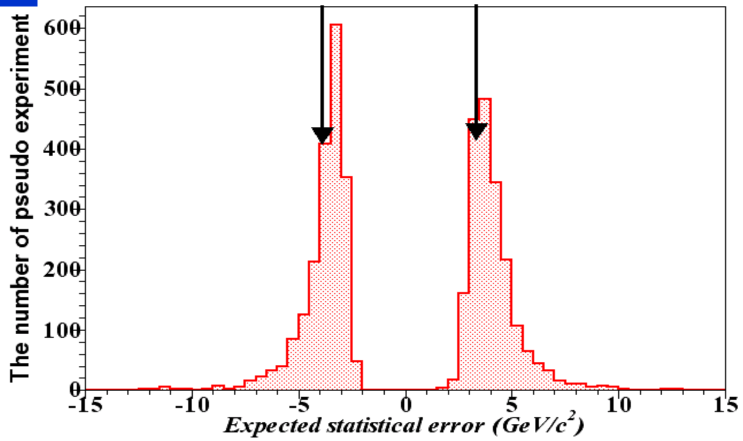
Fit : Two 2nd order polynomials for positive/negative errors.

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

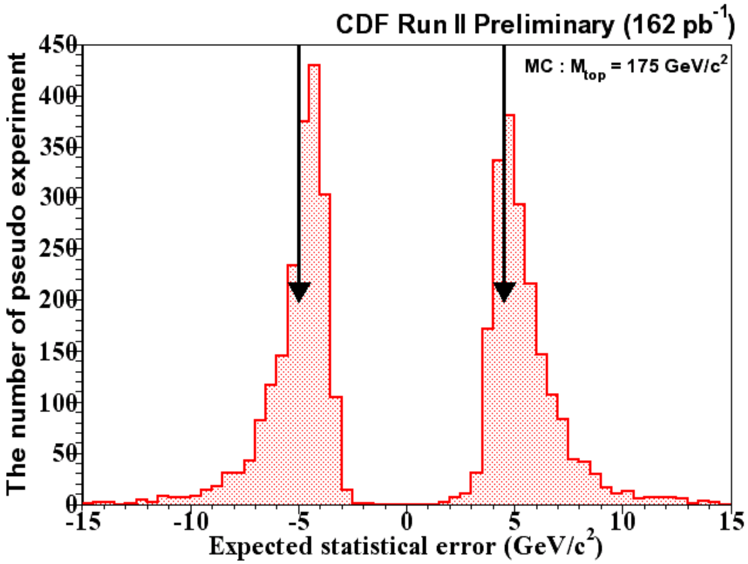


After Mapping function applied

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.

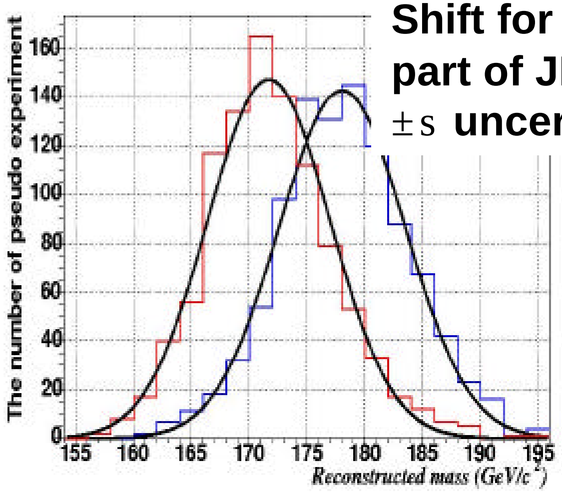


Systematic uncertainty

Jet Energy uncertainty

- **Jet Energy Scale**

Description	ΔM_{top} GeV
Calorimeter non-uniformity	2.6 ± 0.1
Central Calorimeter Response	4.2 ± 0.2
Corrections to Hadrons	2.0 ± 0.1
Total	5.3 ± 0.3



Shift for one of part of JES. ±σ uncertainty

- **Transfer Function** **"Jets ⟷ 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) bjet-gamma balance e.t.c

For now, we take 2.0 GeV error to be conservative.

ISR/FSR uncertainty

- **We checked several ways such as On/Off, diff α_s, RunI like.**

ISR (Initial State Radiation)

We have now the following two samples set the parameters range by comparing with Drell-Yan data.

Description	$\Delta M_{top} = M - M^{default}$
ISR less	-0.2 ± 0.2 GeV
ISR more	$+0.3 \pm 0.2$ GeV

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

ISR systematic : 0.5 GeV

FSR (Final State Radiation)

Description	$\Delta M_{top} = M - M^{default}$
FSR Less	-0.4 ± 0.2 GeV
FSR More	$+0.5 \pm 0.2$ GeV

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 α_s with MRST
(3) CTEQ5L vs MRST

Sources	ΔM_{top} GeV
(1)	1.4 ± 0.4
(2)	0.6 ± 0.2
(3)	1.1 ± 0.3
Total	1.9 ± 0.5

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

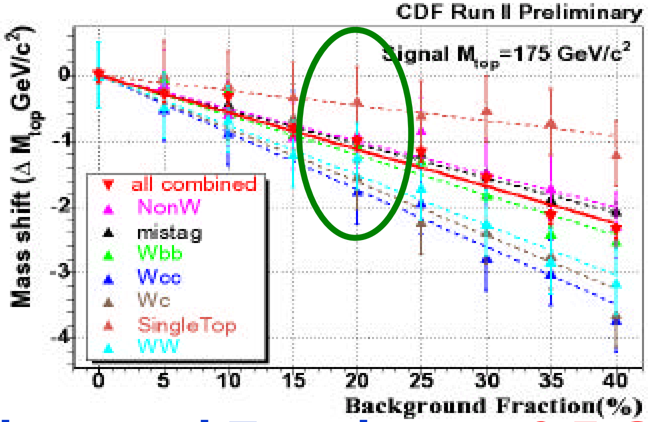
- **Additional jet smearing**

  Resolution underestimated in our MC Additional 15% gives 0.6 ± 0.2 GeV

- **Unclustered Energy** **0.1 ± 0.2 GeV**

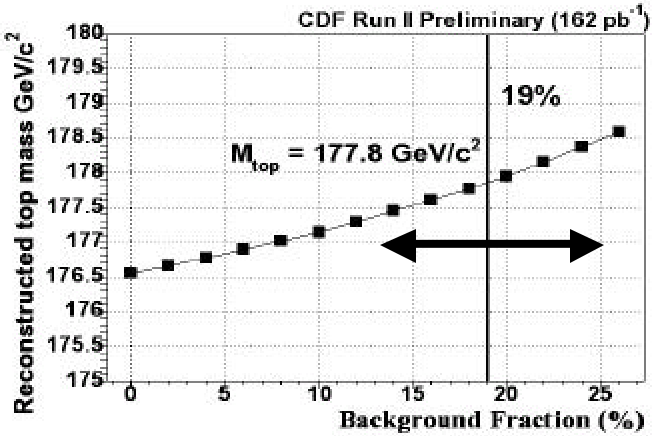
  The effect of difference for Unclustered calorimeter energy

- **Background Modeling** **0.5 GeV**



- **Background Fraction** **0.5 GeV**

  Error due to 19±5%



Systematics Summary

- 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)

Sources	ΔM_{top}(GeV/c^2)
Jet Energy Scale	5.3
ISR	0.5
FSR	0.5
PDF	2.0
Generator	0.6
Spin correlation	0.4
NLO effect	0.4
Bkg fraction(±5%)	0.5
Bkg Modeling	0.5
MC Modeling(jet,UE)	0.5
Transfer function	2.0
Total	6.2

CDF RunII Preliminary Results Summary

	Lepton+jets (btag)			Dilepton
	Template Method	Multivariate Method	Dynamical Likelihood Method	Template Method ($125pb^{-1}$)
Expected Statistical Uncertainty	7.3	6.4	+ 5.4 (4.5) – 5.0 (4.1)	13.9
Results	$174.9^{+7.1}_{-7.7} \pm 6.5$	$179.6^{+6.4}_{-6.3} \pm 6.8$	$177.8^{+4.5}_{-5.0} \pm 6.2$	$175^{+17.4}_{-16.9} \pm 8.4$
Total Measured Uncertainty	+ 9.6 – 10.1	+ 9.3 – 9.3	+ 7.7 – 8.0	+ 19.2 – 18.9

Yellow: Monte Carlo

Red: Data

- Currently DLM is the most precise measurement in RunII.

More checks of MC vs Data (1)

- Event likelihood

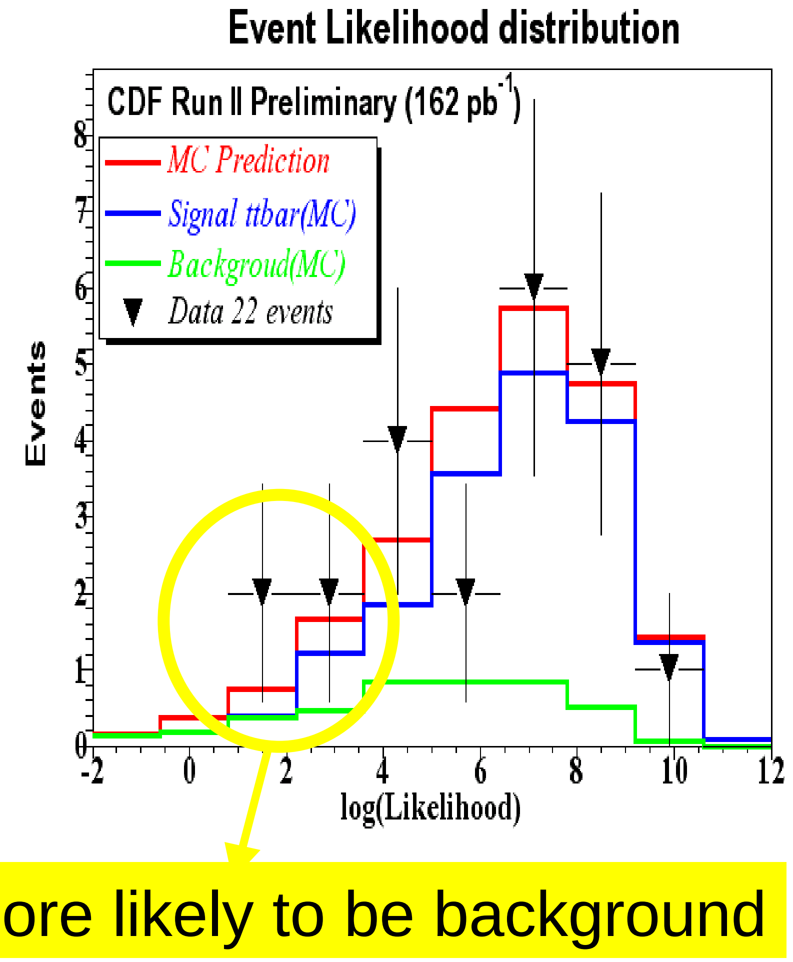
  For i-th event,

$$L_{ev}^{i} = \int L^{i}(M)dM$$

An event has one likelihood.

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

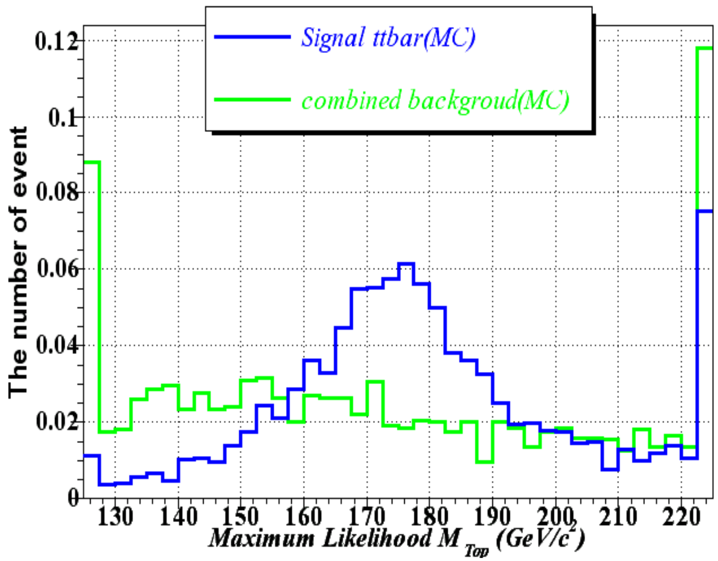
- Agreement is quite good !



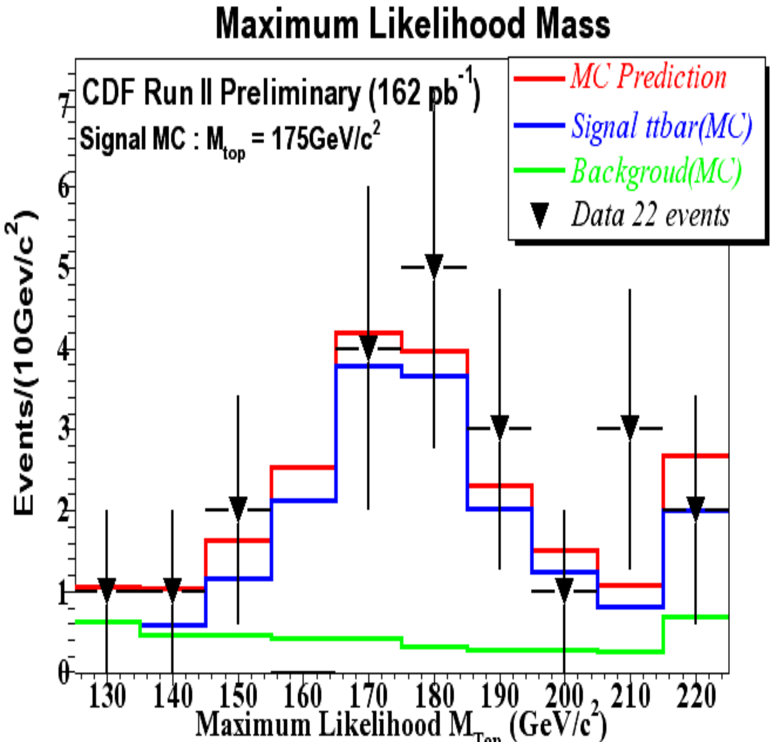
More checks of MC vs Data (2)

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

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,**

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

  - **Currently, this is the most precise measurement using RunII data. Total: $^{+7.7}_{-8.0}$ GeV/c^2**

Cf) world average : 178±4.3 (±2.7±3.3) 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, $$L^i(M_{top}) = N\left[\frac{ds}{d\Phi}\right]_c w(I_t, \mathbf{x} \mid \mathbf{y}; M_{top}) \quad (1)$$

1. Normalization,

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

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

2. Phase space,

dF : assume that each final state parton occupies a unit phase volume in each event.

(1) ⟶ $$L^i(M_{top}) = \sum_{I_t}\sum_{I_s}\int \frac{2p^4}{Flux} F(z_a, z_b, p_T)\,|M|^2\, d(s_w - (\ell + n)^2)\, w(I_t, \mathbf{x} \mid \mathbf{y}; M_{top})\, d\mathbf{x}\, ds$$

Comparisons between DLM and DØM(RunI)

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

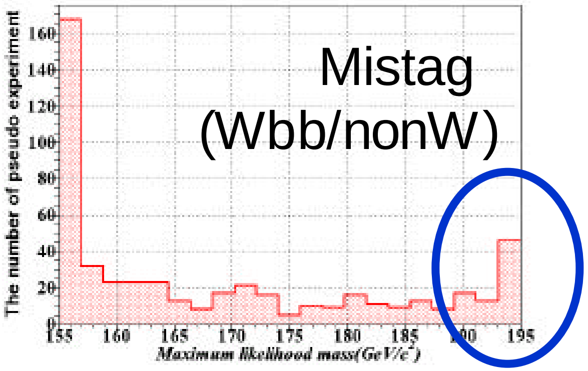
DLM : $$L^i(M_{top}) = \sum_{I_t}\sum_{I_s}\int \frac{2p^4}{Flux} F(z_a, z_b, p_T)\,|M|^2\, d(s_w - (\ell + n)^2)\, w(I_t, \mathbf{x}\,|\,\mathbf{y}; M_{top})\, d\mathbf{x}\, ds_w$$

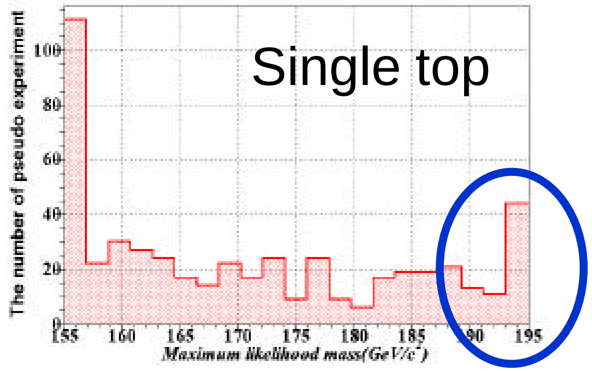
DØM : $$\overline{P}(x; a) = \frac{1}{s}\int d^n s\,(y; a)\, dq_1\, dq_2\, f(q_1)\, f(q_2)\, W(x, y)$$

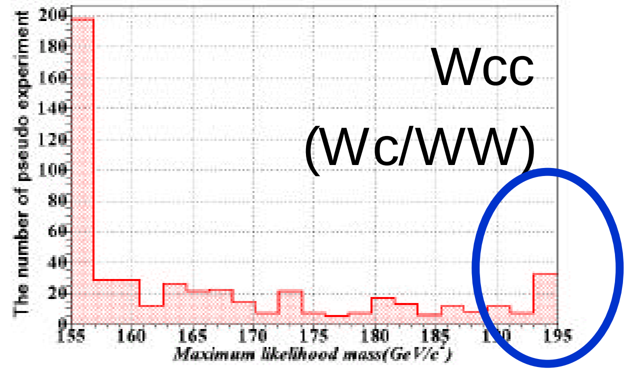
	DLM(RunII)	DØM(RunI)
Event	**b-tagged ,only 4jets (22)** **No more..**	**Non-btag, only 4jets (71)** **Background likelihood cut (22)**
Background	**4.2 events (~20%)** **(QCD Fake 40%)**	**10 events (~50%)** **(W+jets 85%)**
Treatment	**Mapping**	**In the likelihood**
Likelihood	**Signal likelihood only**	**Signal and background likelihood**
Others	**-**	**Normalization & acc correction**

More on backgrounds

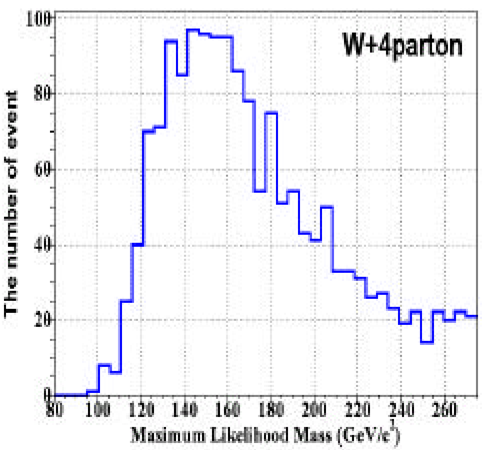
- **5 events pseudo experiment using only each background source**



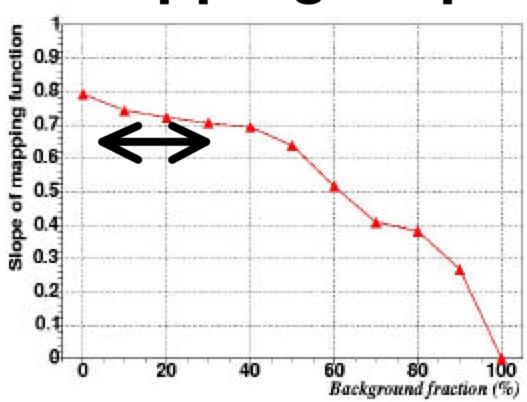




- **Max. likelihood mass in each event**



- **Mapping slope**

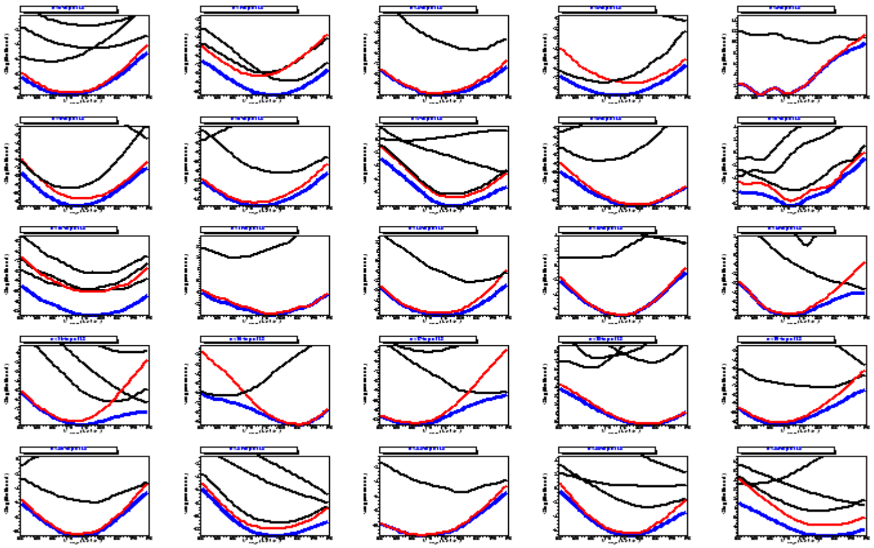


source	Joint likelihood
Mistag	~162 GeV
Wbb	~162 GeV
Wcc	~157 GeV
Wc	~158 GeV
nonW	~161 GeV
Single top	~170 GeV
WW	~160 GeV

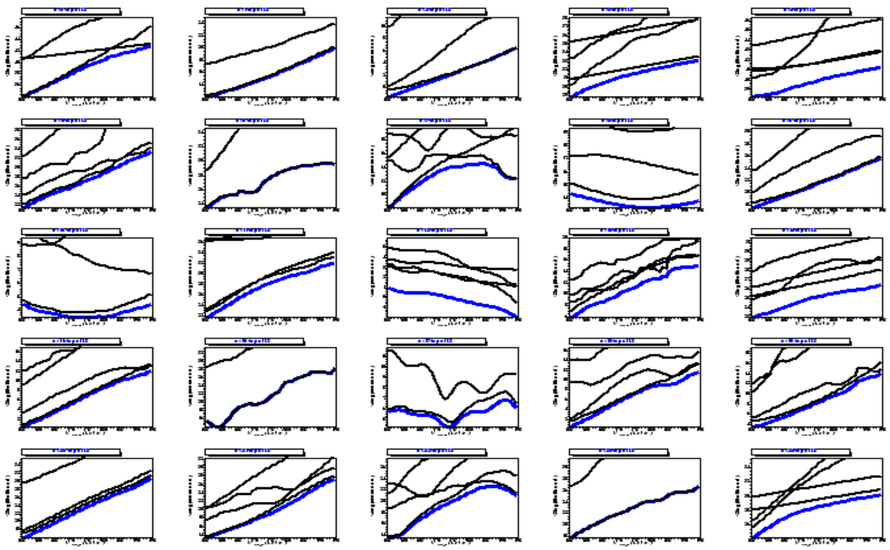
How Likelihood looks like?

25 events examples using generator level input

Signal example: - log(likelihood)



Bkg example: - log(likelihood)



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

Peak around 175 GeV

Range[155-195]GeV

Blue : all added up
Black : each comb.

Likelihood tends to be higher in lower mass region.