Top Quark Mass Measurement In Dilepton Channel Using m_{T2} at CDF

Jian Tang

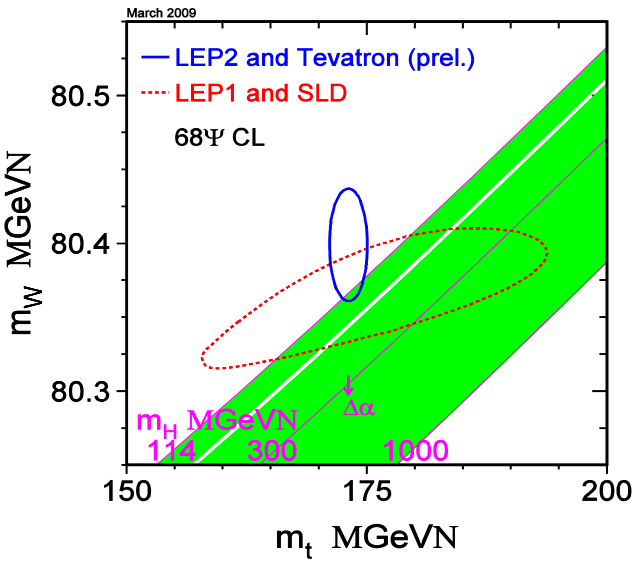
The University of Chicago

On Behalf of CDF Collaboration

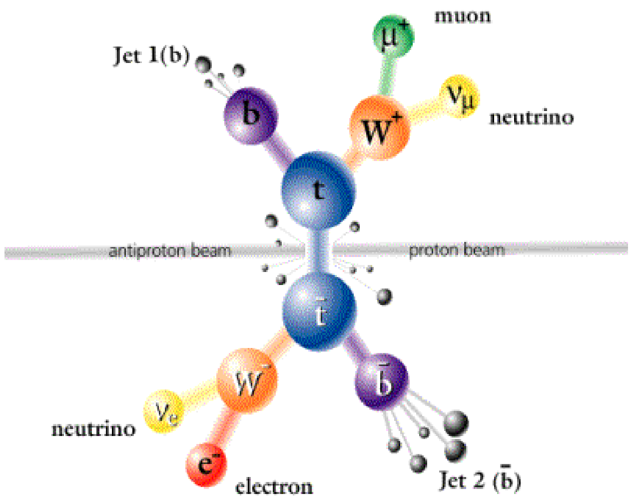
Motivation

Why Top Mass

- Test of SM
- Possible calibration for future experiments
- Constraint on Higgs Mass



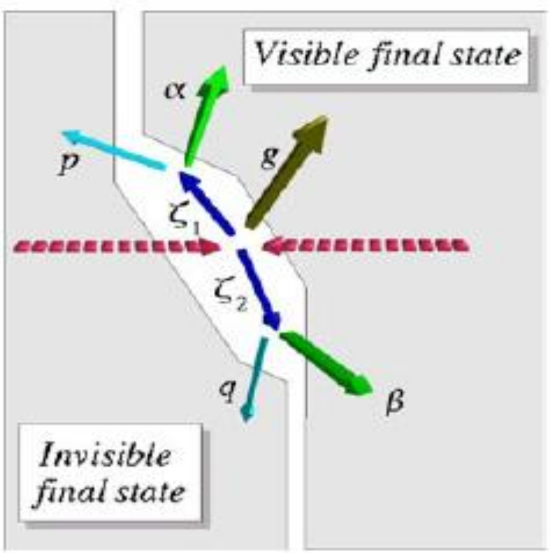
Why Dilepton Channel & m_{T2}



- 2 leptons, 2 neutrinos, 2 b-quarks, BR~5%, little background
- m_{T2} – created for pair production w/ 2 undetected particles, we want to test how well it performs

Introduction to m_{T2}

Pair production w/ 2 undetected particles



Model-independent information about mass of ζ : m_{T2}

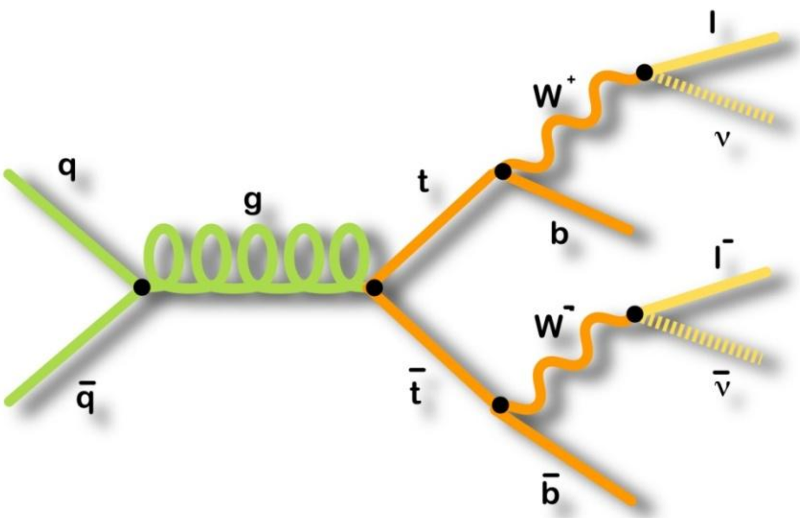
$$m_{T2} = \min[\max(m_{T(1)}, m_{T(2)})]$$
$$q_T + p_T = \text{missing } p_T$$

Alan Barr, Christopher Lester & Phil Stephens, J.Phys.G: Nucl.Part.Phys. 29(2003) 2343-2363

- Particle(ζ) decays into visible(α,β) and invisible(p,q)
- Masses of invisibles are unknown
- Masses of "parents"(ζ) of invisibles are also unknown

m_{T2} in ttbar Dilepton Channel

$$qq \rightarrow t\bar{t} \rightarrow bl_1\nu_1 + \bar{b}l_2\nu_2$$



$$p_T^{\nu 1} + p_T^{\nu 2} = p_T^{missing}$$

$$(p_T^{\nu 1}, p_T^{\nu 2})$$

$$m_T^2 = m_{bl}^2 + m_\nu^2 + 2(E_T^{bl} \cdot E_T^\nu - P_T^{bl} \cdot P_T^\nu)$$

$$m_T^{(1)}, m_T^{(2)}$$

$$(p_T^{\nu 1}, p_T^{\nu 2})_1 \rightarrow \max\{m_T^{(1)}, m_T^{(2)}\}_1$$

$$(p_T^{\nu 1}, p_T^{\nu 2})_2 \rightarrow \max\{m_T^{(1)}, m_T^{(2)}\}_2$$

$$(p_T^{\nu 1}, p_T^{\nu 2})_3 \rightarrow \max\{m_T^{(1)}, m_T^{(2)}\}_3$$

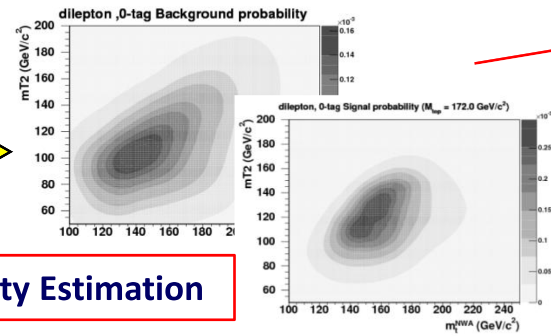
$$\vdots$$

$$m_{T2} = \min_{p_T^{\nu 1} + p_T^{\nu 2} = p_T^{\text{missing}}} \{\max[m_T^{(1)}, m_T^{(2)}]\}$$

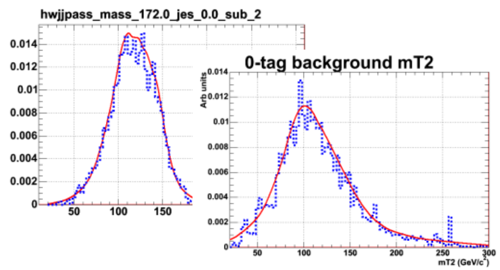
Method We Use– Template Method

MC
- tt
- bkgd

Event reconst.



Kernel Density Estimation



2d templates

m_t^{NWA} & m_{T2}

m_t^{NWA} & H_T

Old method

1d templates(for comparison of variables)

m_{T2}

m_t^{NWA}

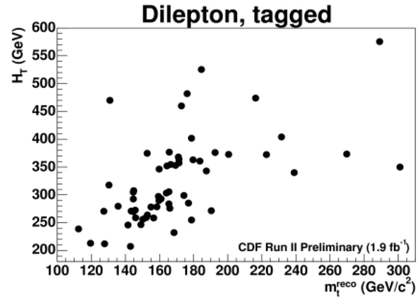
H_T

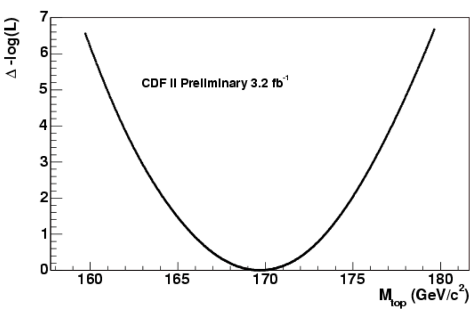
Estimation → P_{sig}, P_{bkgd}

Likelihood Fit

DATA

Event reconst

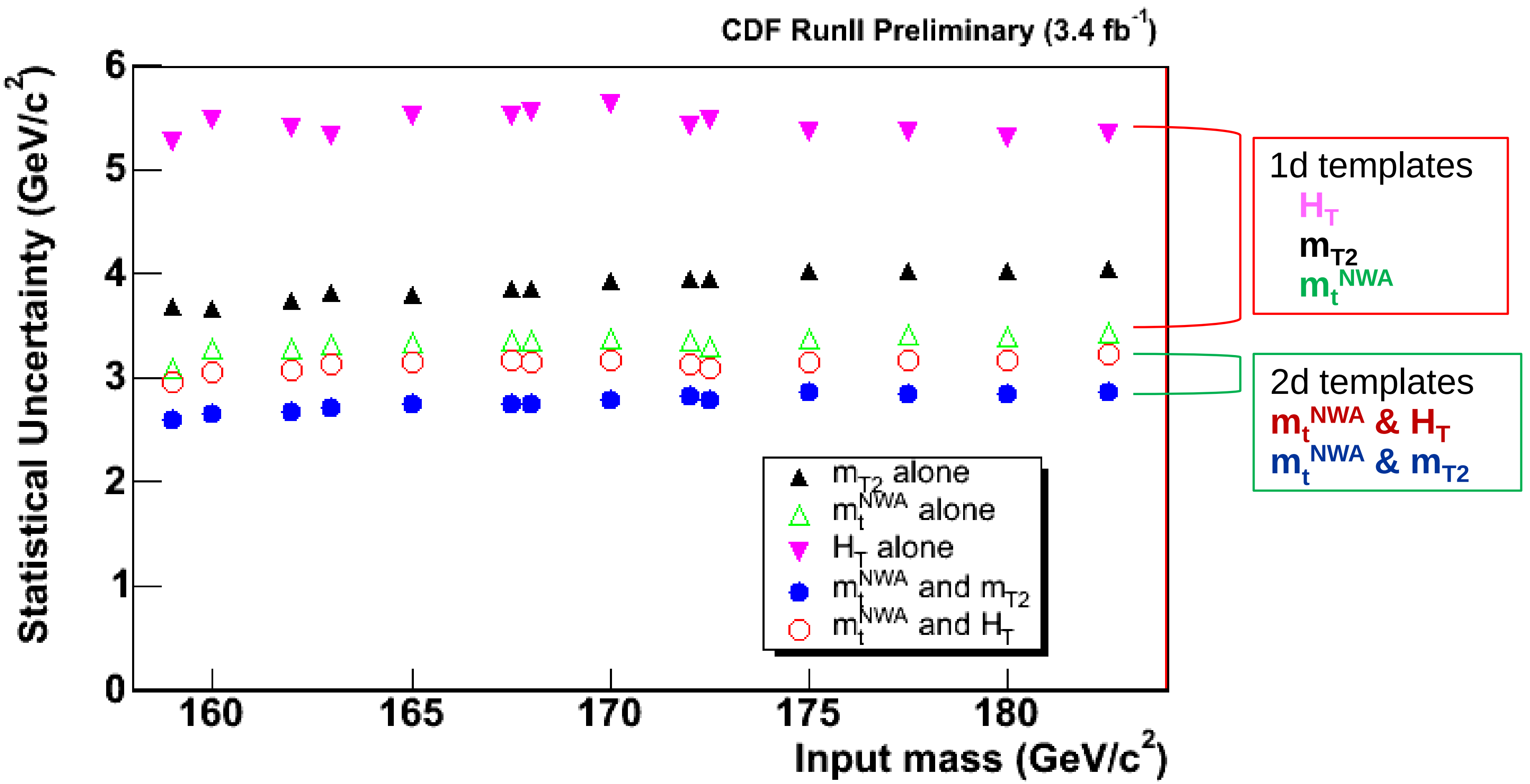




M_t^{NWA}: reconstructed mass of top(**N**eutrino **W**eighting **Al**gorithm)

H_T: scalar sum of momenta of particles in the event

Statistical Uncertainties Comparison



- Statistical Power: $H_T < m_{T2} < m_t^{NWA}$
- Measurement using m_t^{NWA} & m_{T2} is better that using m_t^{NWA} & H_T

Systematic & Total Uncertainties

Assuming 175GeV/c^2 top mass

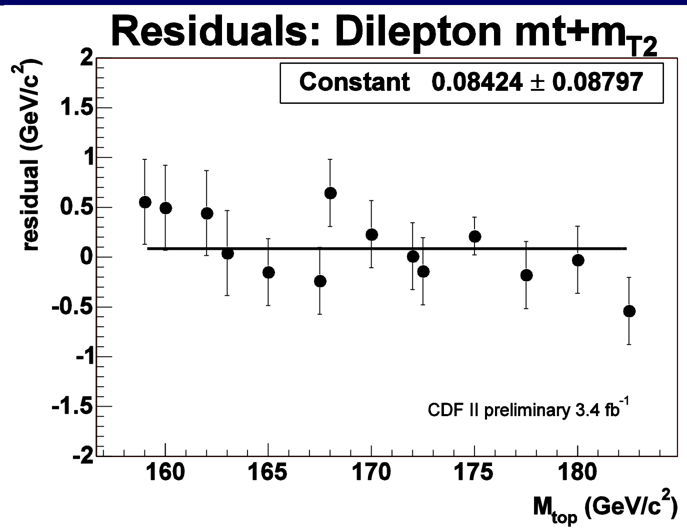
GeV/c^2		m_{T2}	m_t^{NWA}	H_T	$m_t^{NWA}+m_{T2}$	$m_t^{NWA}+H_T$
Statistical		4.0	3.4	5.4	2.9	3.2
Systematic	Jet Energy Scale	2.6	3.5	3.7	3.0	3.4
	Generator	0.3	1.0	2.6	0.5	1.3
	Parton distribution functions	0.5	0.6	1.8	0.5	0.8
	b-JES	0.2	0.3	0.2	0.2	0.3
	Background shape	0.4	0.3	0.7	0.1	0.3
	Gluon fusion fraction	0.3	0.1	0.3	<0.1	0.1
	Initial and final state radiation	0.6	0.2	0.6	0.3	0.2
	MC statistics	0.3	0.3	0.5	0.3	0.3
	Lepton energy	0.6	0.2	0.7	0.3	0.2
	Multiple Hadron Interaction	0.2	0.3	0.3	0.3	0.3
	Color Reconnection	0.7	0.6	2.5	0.6	0.6
	Total Systematic	2.9	3.8	5.7	3.2	3.8
Total		5.0	5.1	7.8	4.3	5.0

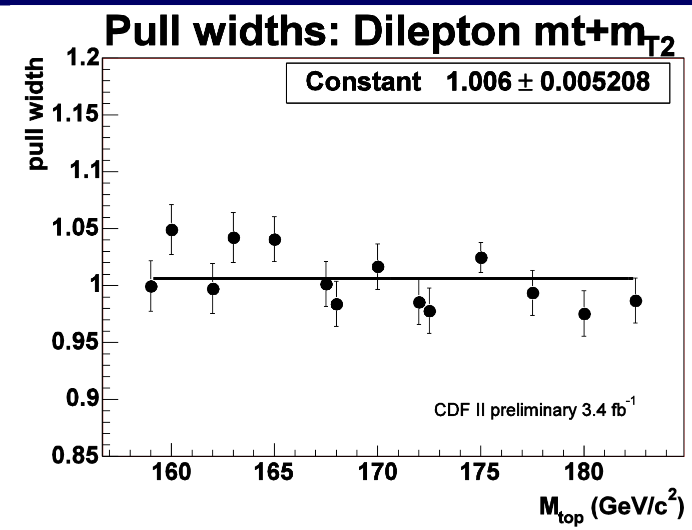
m_{T2} performs the best among all 3 variables !!!

Performance Evaluation: Residual & Pull width

2-Dim:

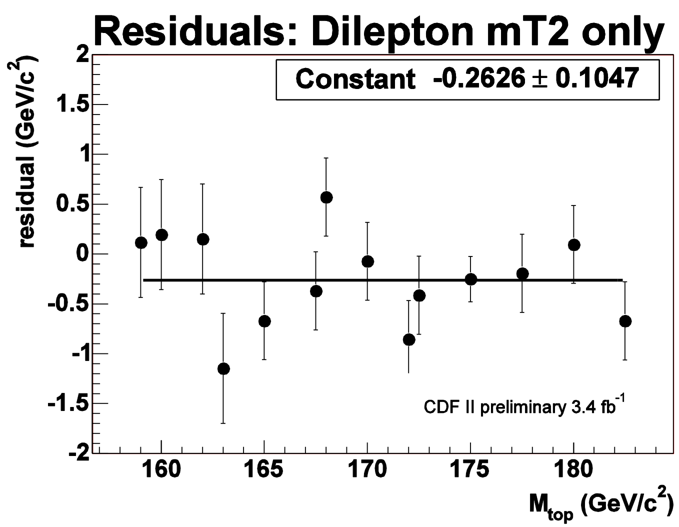
$m_{T2} + m_t^{NWA}$

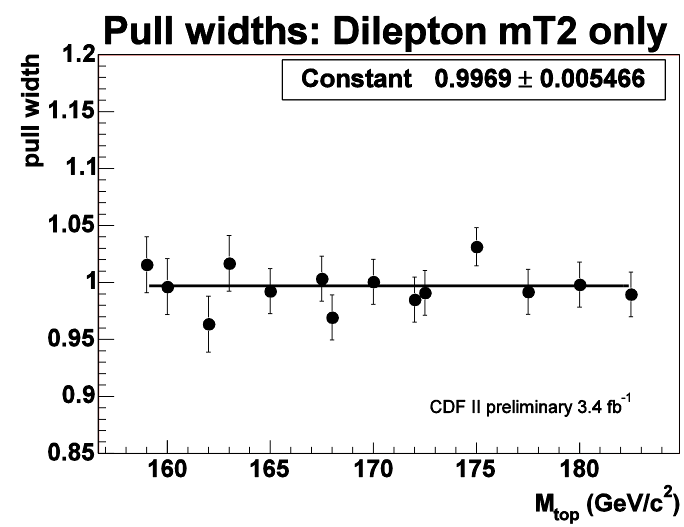




1-Dim:

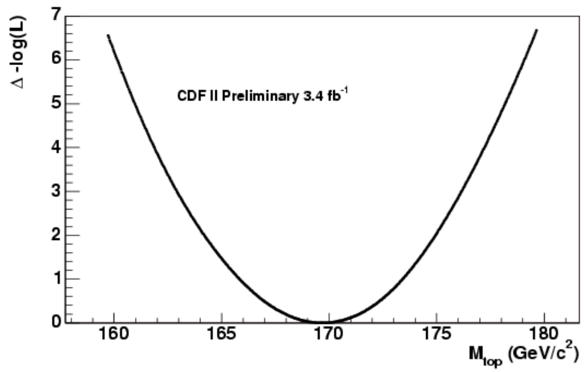
m_{T2} alone





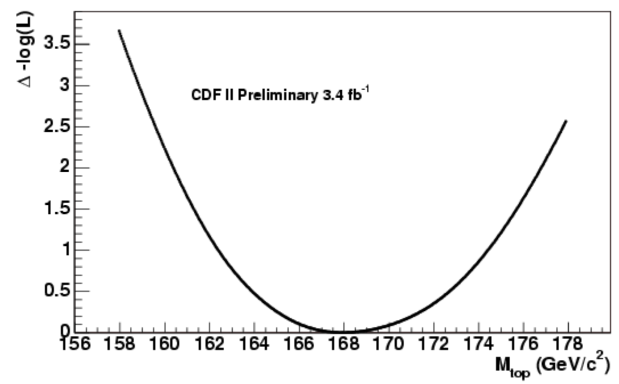
Data Fit Results

$m_{T2} + m_t^{NWA}$



169.3 ±2.7 (stat.) ± 3.2 GeV/c^2(syst.)

m_{T2} alone

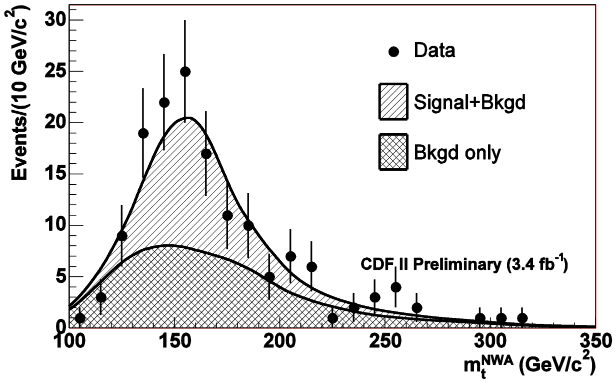


168.0 $^{+4.8}_{-4.0}$(stat.) ± 2.9 GeV/c^2(syst.)

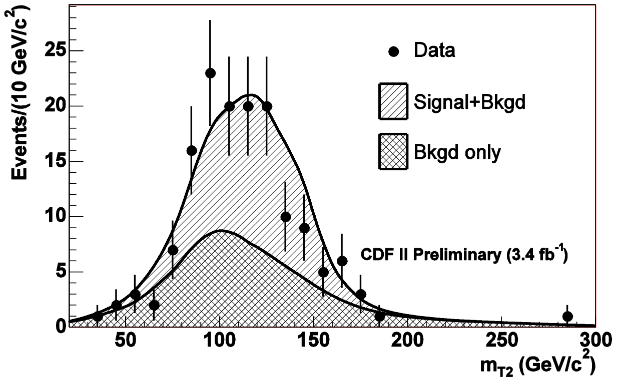
Observables	result (GeV/c^2)	result (GeV/c^2)
m_{T2}	$168.0\ ^{+4.8}_{-4.0}$ (stat.) $\pm$ 2.9 (syst.)	$168.0\ ^{+5.6}_{-5.0}$
m_t^{NWA}	$169.4\ ^{+3.3}_{-3.2}$ (stat.) $\pm$ 3.8 (syst.)	$169.4\ ^{+5.0}_{-5.0}$
H_T	$168.8\ ^{+5.1}_{-6.6}$ (stat.) $\pm$ 5.7 (syst.)	$168.8\ ^{+7.6}_{-8.7}$
m_t^{NWA} and m_{T2}	$169.3\ ^{+2.7}_{-2.7}$ (stat.) $\pm$ 3.2 (syst.)	$169.3\ ^{+4.2}_{-4.2}$
m_t^{NWA} and H_T	$169.6\ ^{+2.8}_{-2.9}$ (stat.) $\pm$ 3.8 (syst.)	$169.6\ ^{+4.7}_{-4.8}$

Data & MC(signal 169GeV/c^2 + background) Distributions

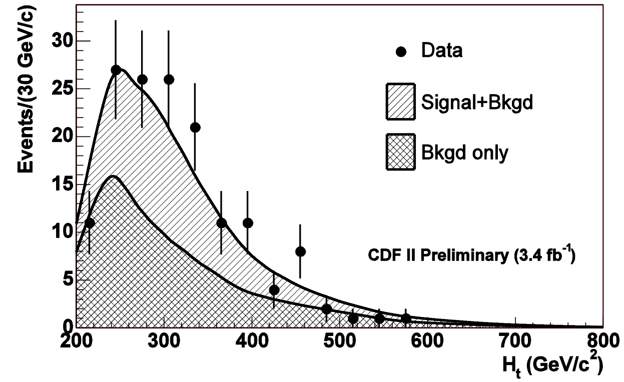
Non-tagged m_t^{NWA}



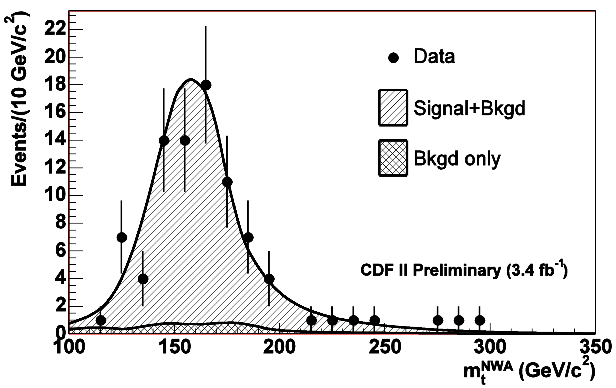
Non-tagged m_{T2}



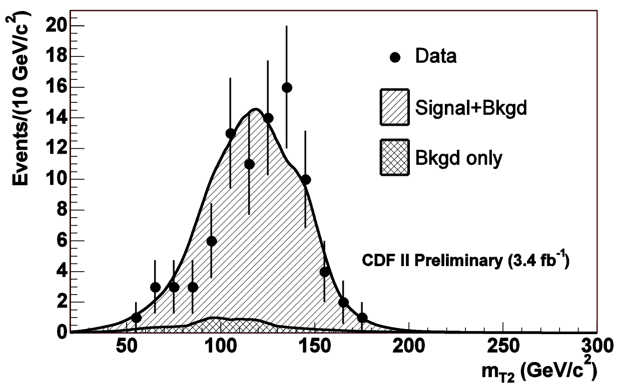
Non-tagged H_T



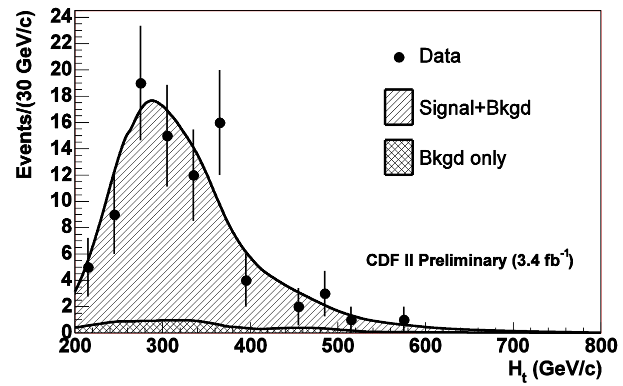
tagged m_t^{NWA}



tagged m_{T2}



tagged H_T



Combined Result with L+Jets Channel

Dilepton Channel

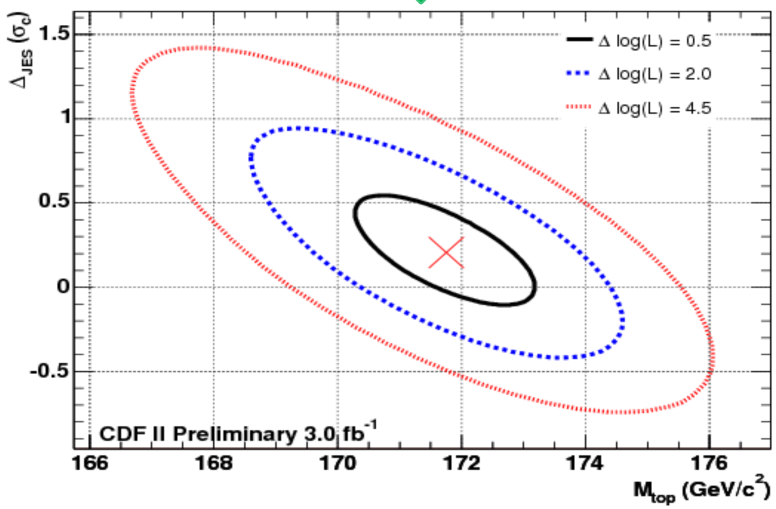
$$M_{top} = 169.3 \pm 2.7(\text{stat.}) \pm 3.2(\text{syst.}) GeV / c^2$$
$$= 169.3 \pm 4.2 GeV / c^2 (\text{Dilepton})$$

Lepton + Jets Channel

$$M_{top} = 172.2^{+1.5}_{-1.6}(\text{stat.}) \pm 1.5(\text{syst.}) GeV / c^2$$
$$= 172.2 \pm 1.9 GeV / c^2 (\text{L+Jets})$$

- 1 lepton, 1 neutrino, 4 jets
- 2D template method
- Different variables from Dilepton channel

Combined Fit Result : by simultaneously fitting both Lepton+Jets & Dilepton Channels



$$M_{top} = 171.7^{+1.4}_{-1.5}(\text{stat.}) \pm 1.1(\text{syst.}) GeV / c^2$$
$$= 171.7^{+1.8}_{-1.9} GeV / c^2 (\text{combined})$$

Conclusion

- We analyze the performance of m_{T2}
  - Very good candidate variable in pair productions w/ 2 undetected particles
- We measure the top quark mass in dilepton channel using m_{T2} with 3.4 fb^{-1}
  - 169.3 ± 2.7 (stat.) $\pm$ 3.2 GeV/c^2(syst.) with $m_t^{NWA} + m_{T2}$
  - $168.0^{+4.8}_{-4.0}$ (stat.) $\pm$ 2.9 GeV/c^2(syst.) with m_{T2} alone
  - First measurement using m_{T2}
- This method can be useful for mass determination of new particles

Back up slides

Neutrino Weighting Algorithm

1. Assume the value of top mass
2. Assume neutrino pseudorapidities
3. Using masses of W, b quark & leptons to solve px and py for each neutrino
4. Calculate a weight

$$w(m_t, \eta_1, \eta_2) = \exp\left(-\frac{E_{Tx} - p_x^{\nu} - p_x^{\bar{\nu}}}{2\sigma_{E_T}^2}\right) \bullet \exp\left(-\frac{E_{Ty} - p_y^{\nu} - p_y^{\bar{\nu}}}{2\sigma_{E_T}^2}\right)$$

5. Integrate over two pseudorapidities and obtain the weight for the assumed top mass

$$W(m_t) = \sum \iint d\eta_1 d\eta_2 g(\eta_1) g(\eta_2) \sum w(m_t, \eta_1, \eta_2)$$

6. Find the top mass with the highest weight--m_t^{NWA}

Kernel Density Estimation

➢

- Histograms are a useful but limited way to estimate the true, underlying density of observed data with an unknown distribution.
- Histograms are essentially discontinuous step functions.

➢

- A kernel histogram is a generalization of the usual histogram. It associates to each data point a function (called a kernel function). The kernel histogram is the (properly renormalized) sum of these functions.
- Kernel functions typically depend on a parameter, usually called the bandwidth, that significantly affects the roughness or smoothness of the kernel histogram that is ultimately generated.