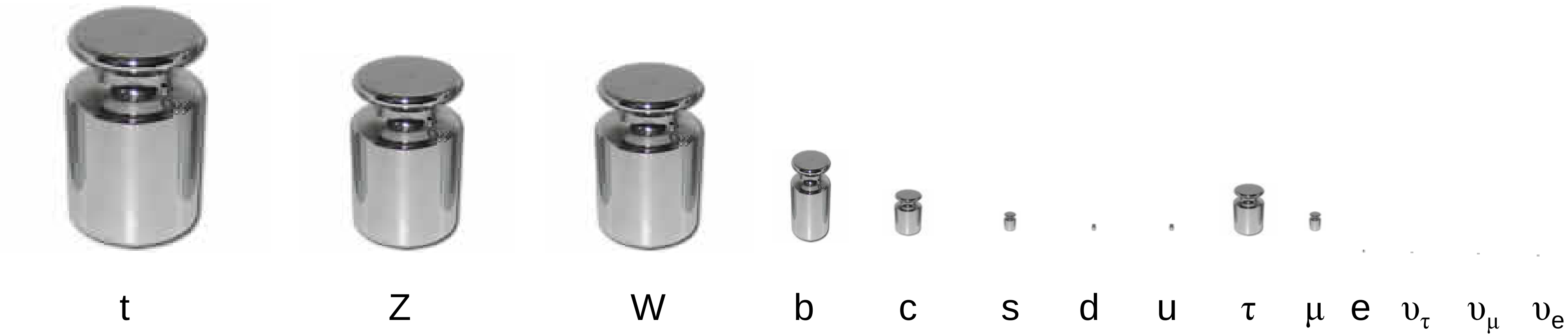
M_{top} measurement at CDF using template method

Hyunsu Lee
The University of Chicago

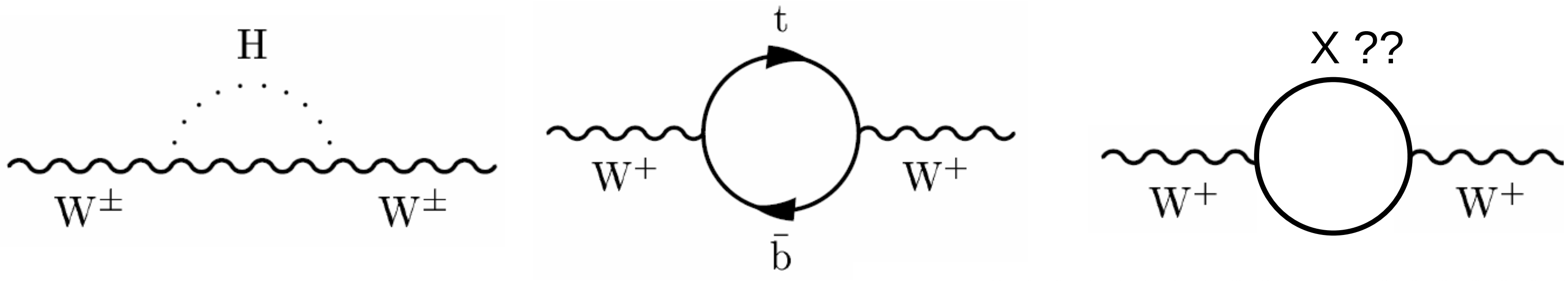
On behalf of the CDF collaboration

The top quark



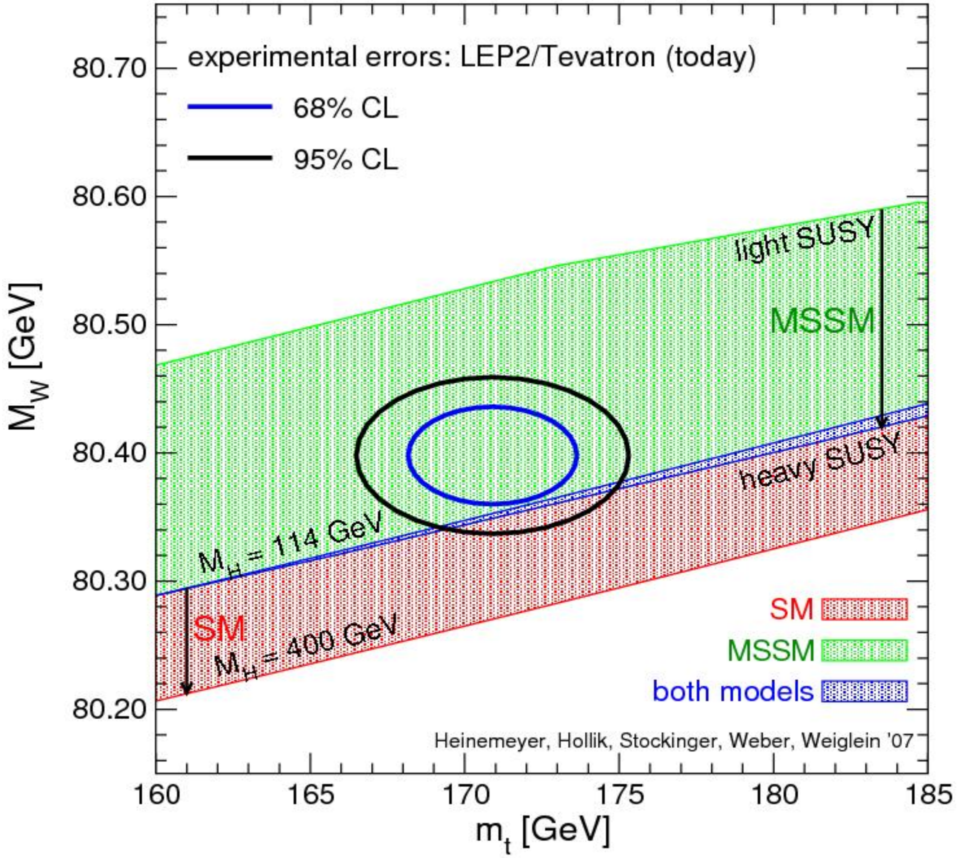
- The heaviest known fundamental particle
- Decays as a free quark – no other quark does this!
- Only Tevatron can make it until now

Why we measure top quark mass

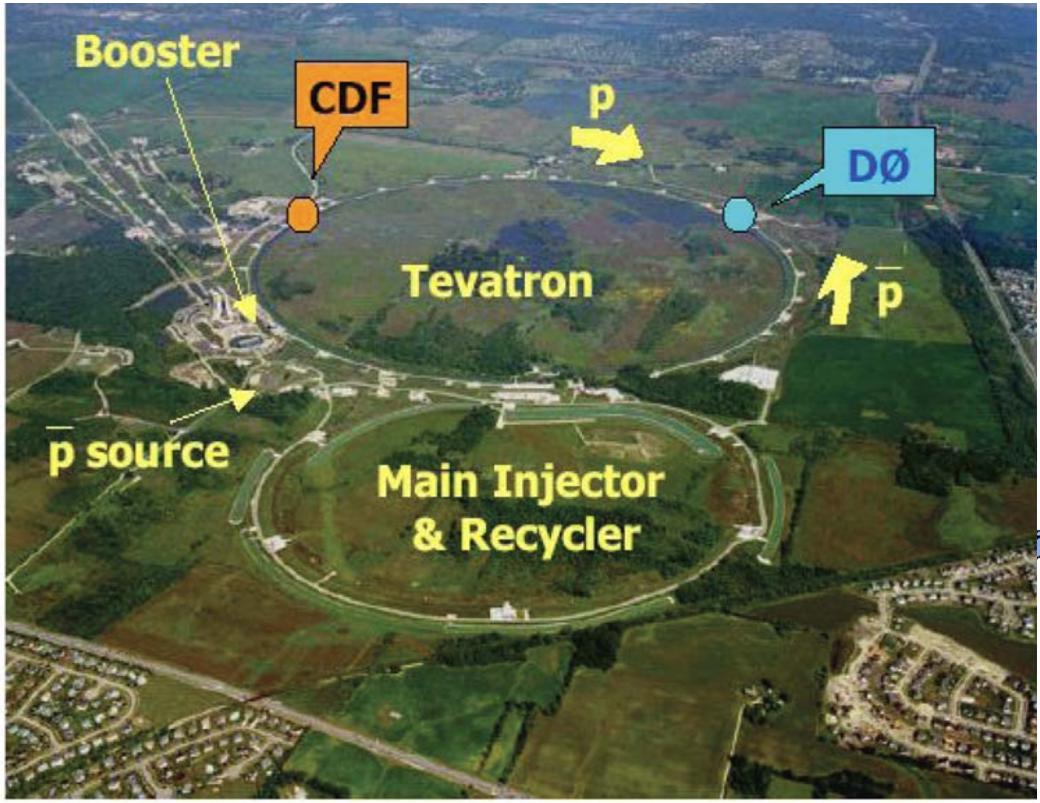


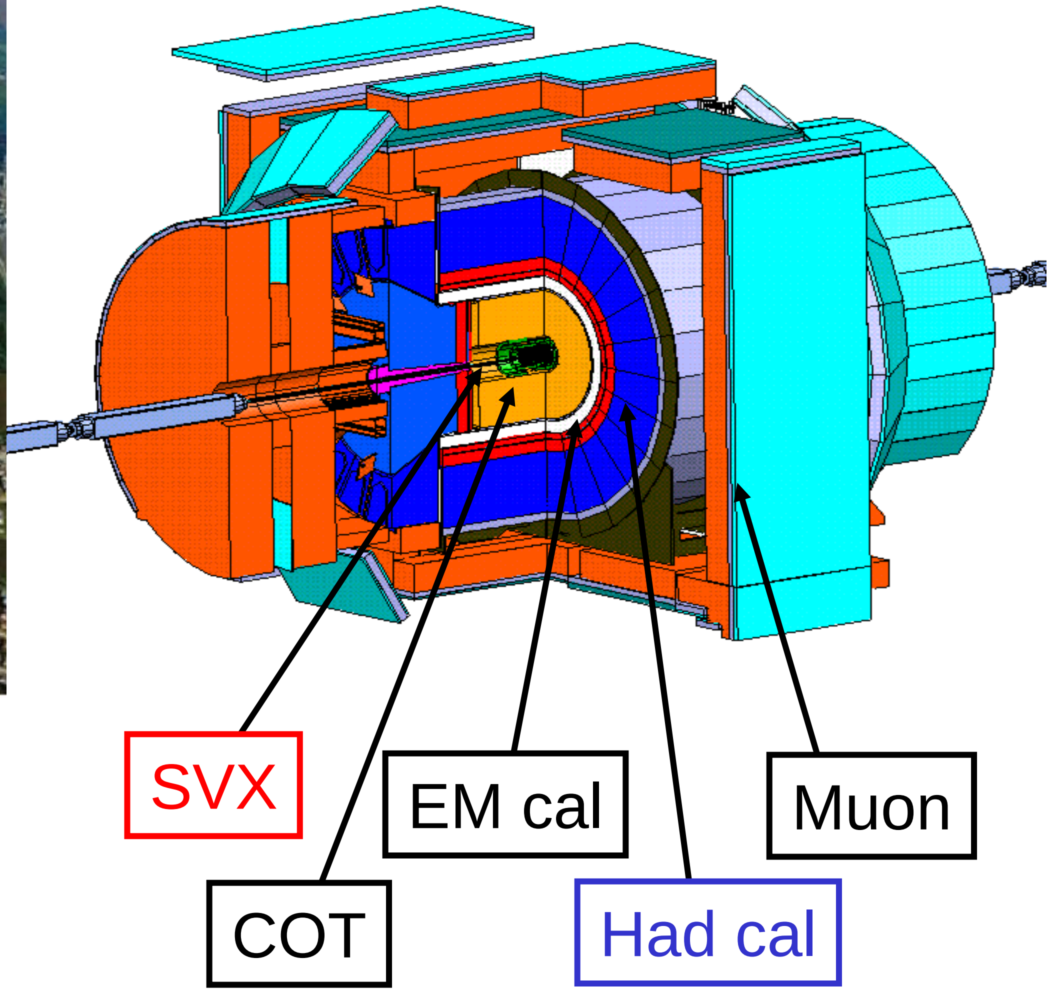
- SM Higgs Mass was constrained by M_{top} and M_W through loop correction of W mass

- Precision top quark mass measurement
  - ❖ Predict SM Higgs mass
  - ❖ Constraints for physics beyond standard model



Tevatron and CDF II detector





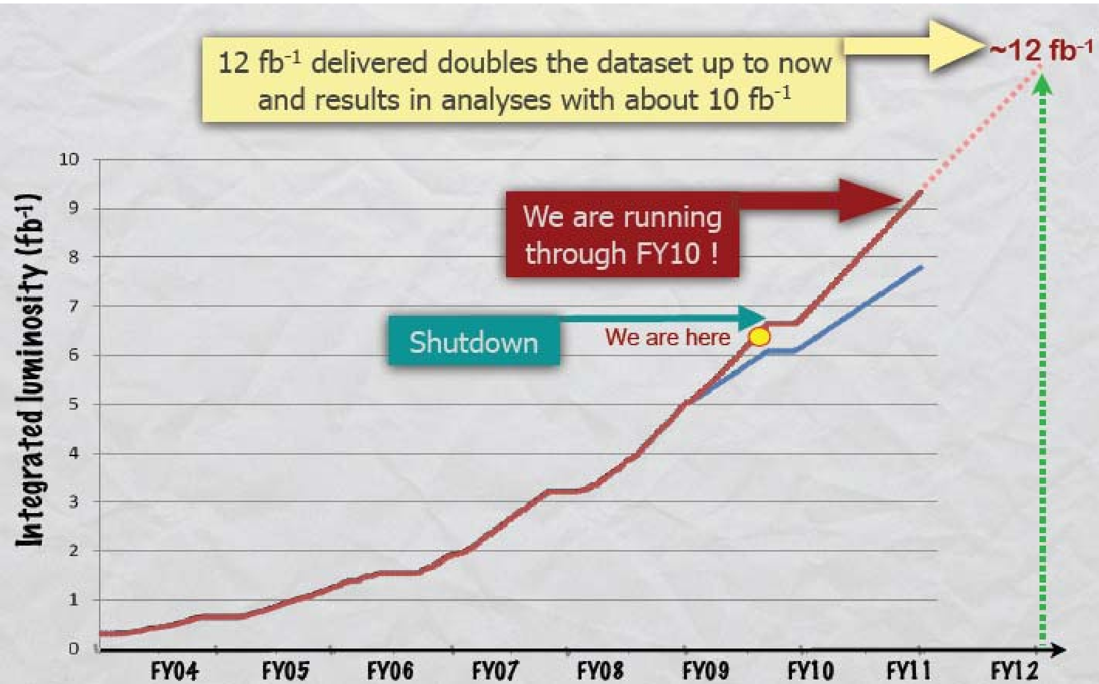
Tevatron is p pbar collider with sqrt(s)=1.96TeV

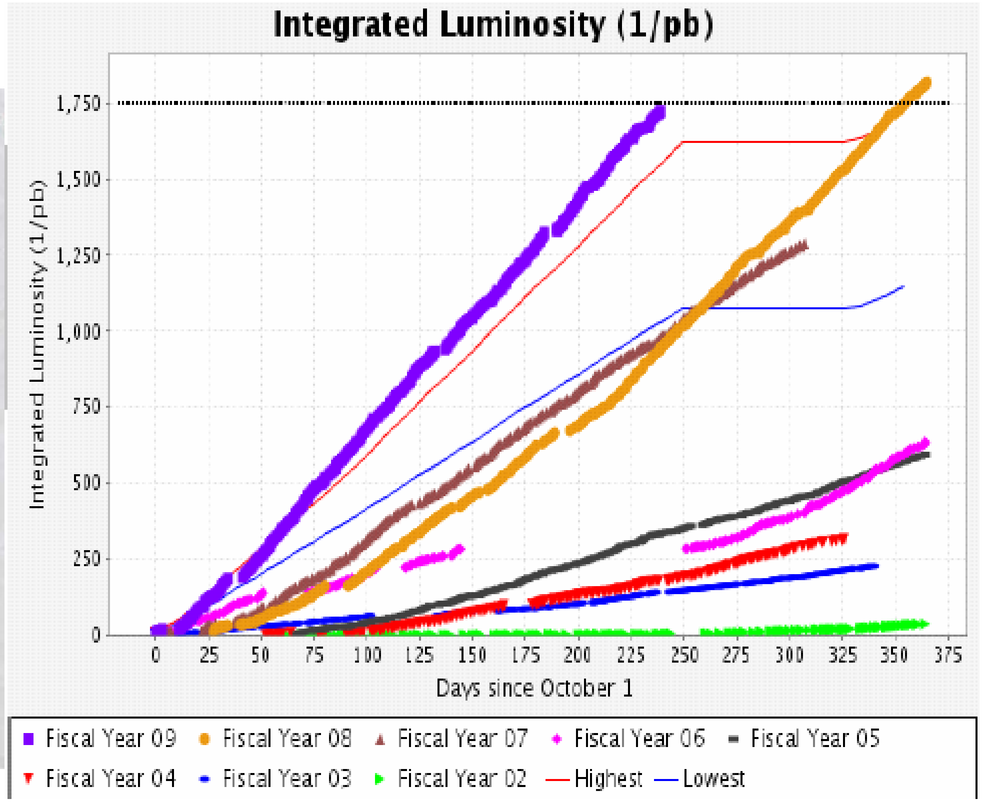
Still the highest energy in the world

Still only one machine to generate top quark

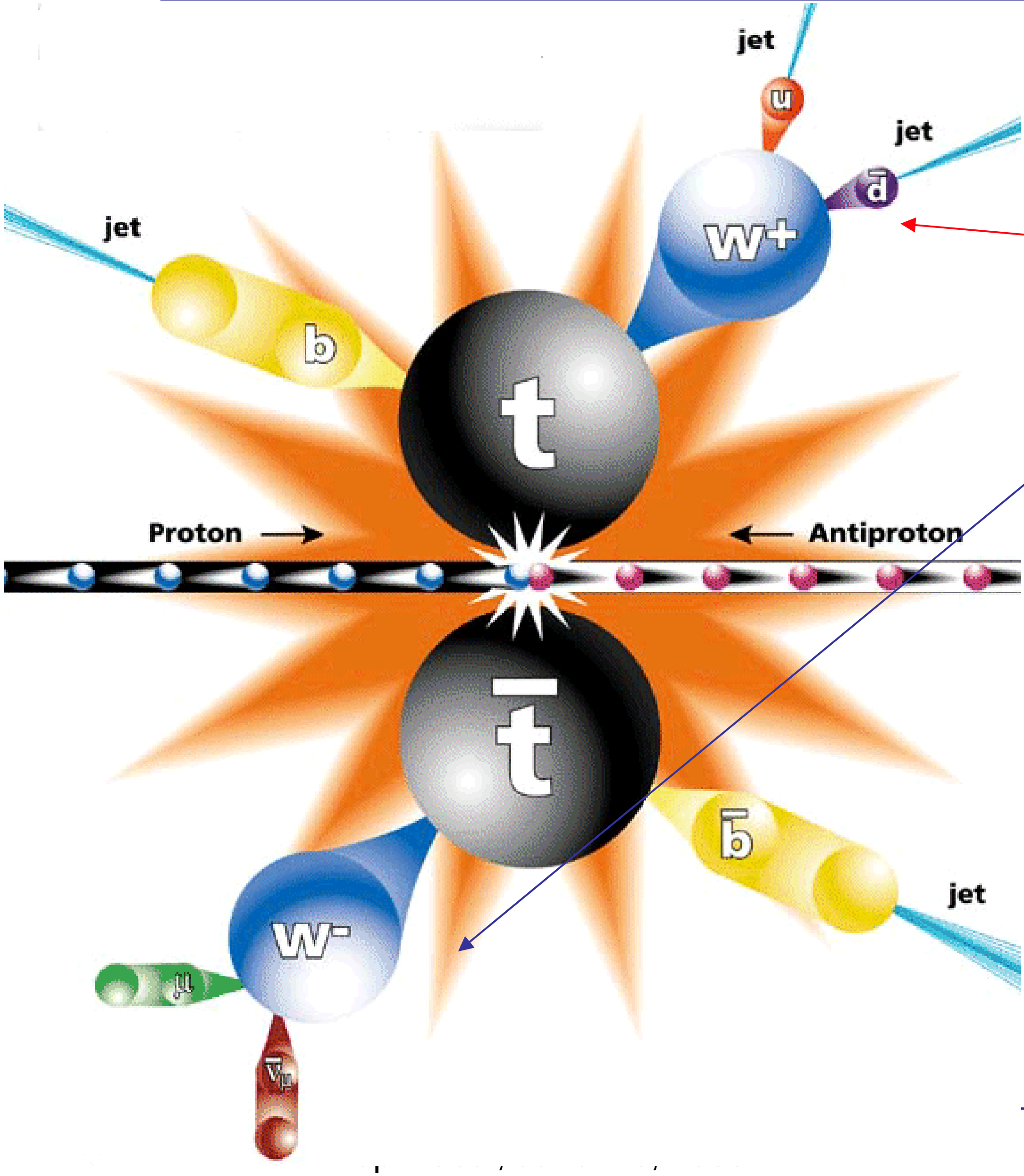
Tevatron Luminosity

- Integrated luminosity >5fb^{-1}
- Luminosity is still accelarating
- Now ~ 2fb^{-1}/year





Top quark production and decay



- Tops always decay via t->Wb
- Event topology then depends on W decays
  - Hadronic (quarks)
  - Leptonic (electron or muon + neutrino)
- This analysis uses the Lepton+Jets channel(30%)
  - One W decays to hadrons, the other to leptons
  - Signature = 4 quarks, 1 charged lepton + undetected neutrino

And Dilepton(5%)

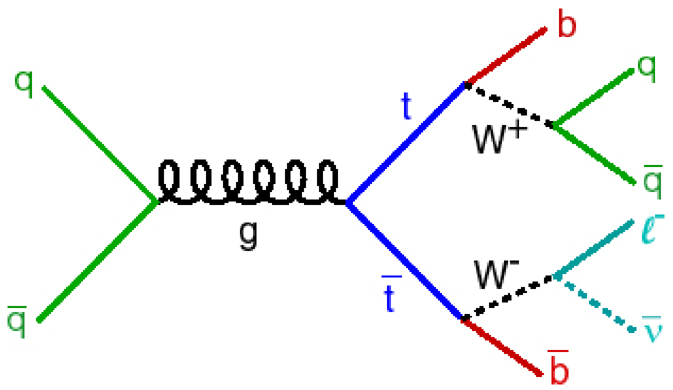
  - Both W decay to leptons
  - Signature = 2b quarks, 2 charge lepton+2 undetected

Why M_{top} is difficult

- With 4 (and only 4 jets!), there are 12 different ways of assigning jets to partons at hard scattering
- Neutrino from W decay
- Non-negligible backgrounds
- Jets are difficult

Why M_{top} is difficult

- With 4 (and only 4 jets!), there are 12 different ways of assigning jets to partons at hard scattering
- Neutrino from W decay
- Non-negligible backgrounds
- Jets are difficult

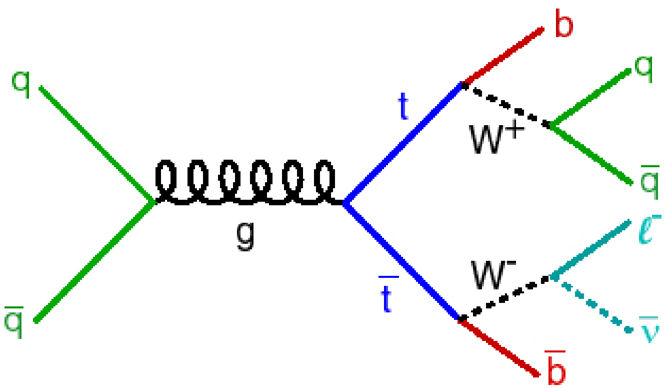


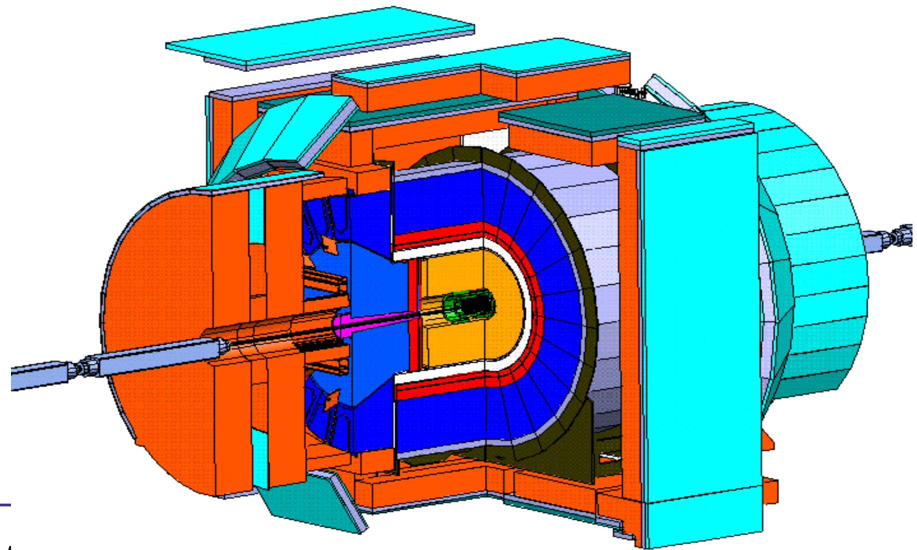
- Jet to parton assignment
- (ISR/FSR, splitting, merging)
- Use b-tagging to reduce combinatorics

Why M_{top} is difficult

- With 4 (and only 4 jets!), there are 12 different ways of assigning jets to partons at hard scattering
- Neutrino from W decay
- Non-negligible backgrounds
- Jets are difficult

- Momentum imbalance
- All of system affect – large uncertainty
- Dileton channel is even worse (under-constraint system)

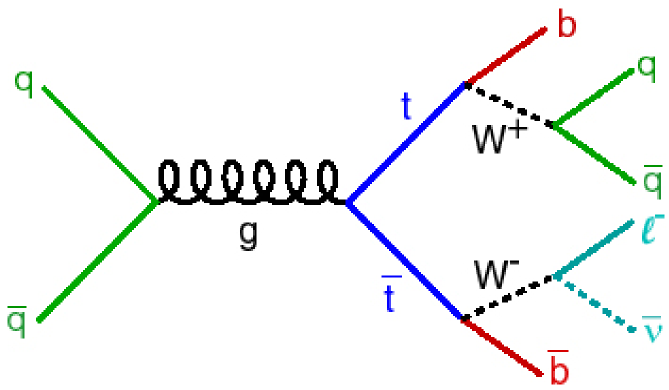




Why M_{top} is difficult

- With 4 (and only 4 jets!), there are 12 different ways of assigning jets to partons at hard scattering
- Neutrino from W decay
- Non-negligible backgrounds
- Jets are difficult

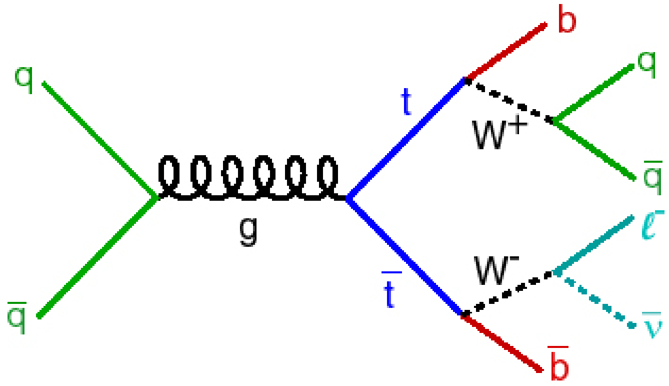
- b-tagging reduce background significantly
- check the kinematics
- Use independent estimates of background rate

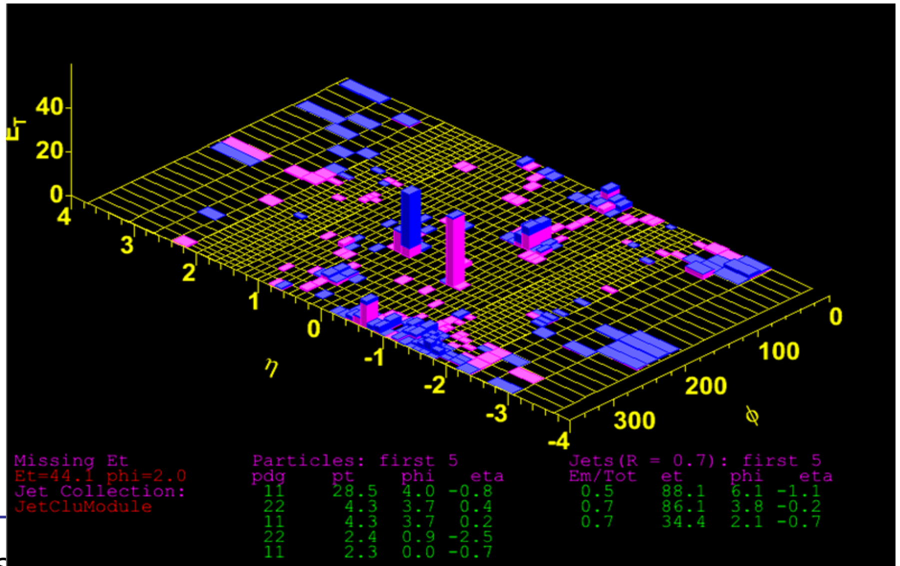


Why M_{top} is difficult

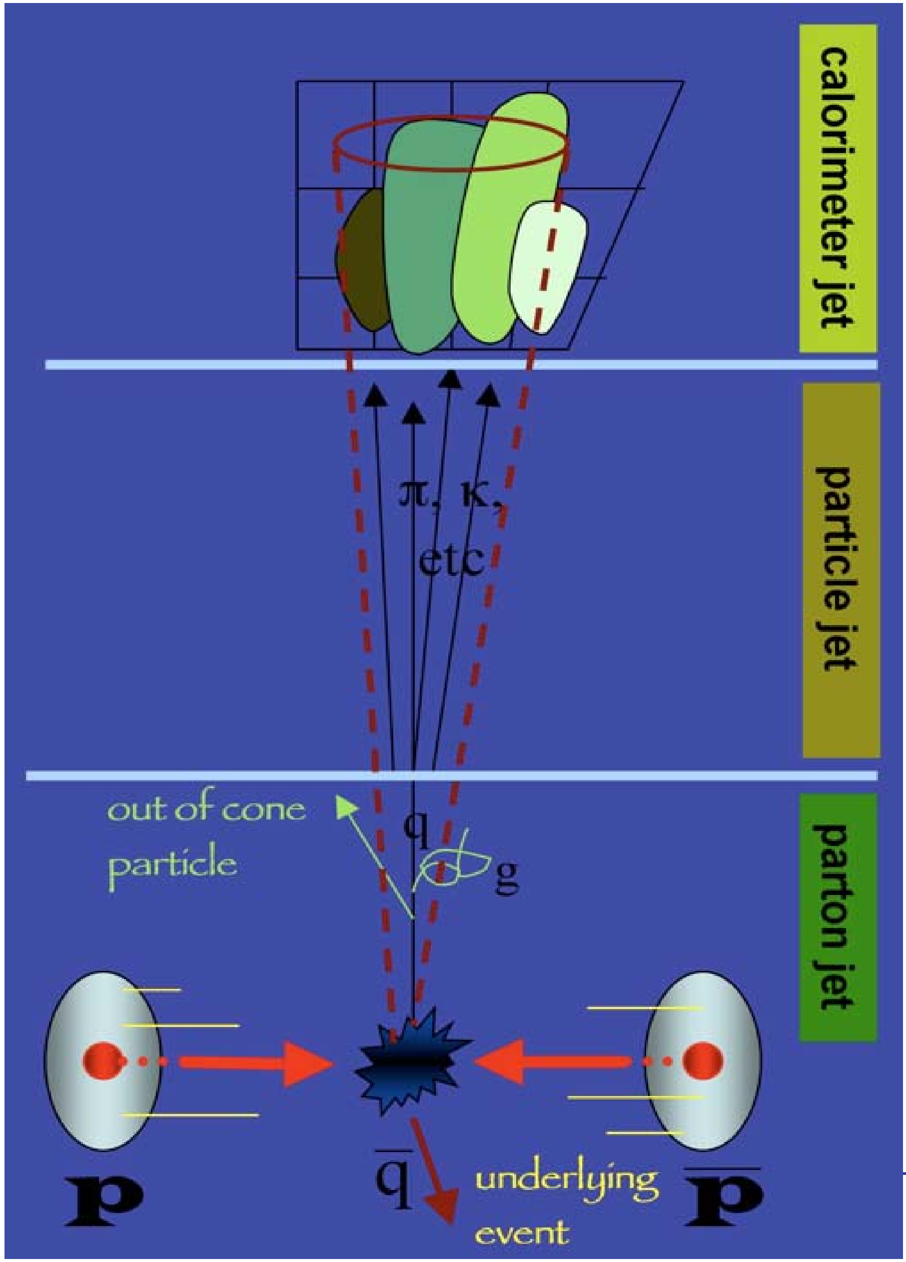
- With 4 (and only 4 jets!), there are 12 different ways of assigning jets to partons at hard scattering
- Neutrino from W decay
- Non-negligible backgrounds
- Jets are difficult

- Hard work (gamma jet balance)
- Use resonance of hardronic decay W in lepton jet

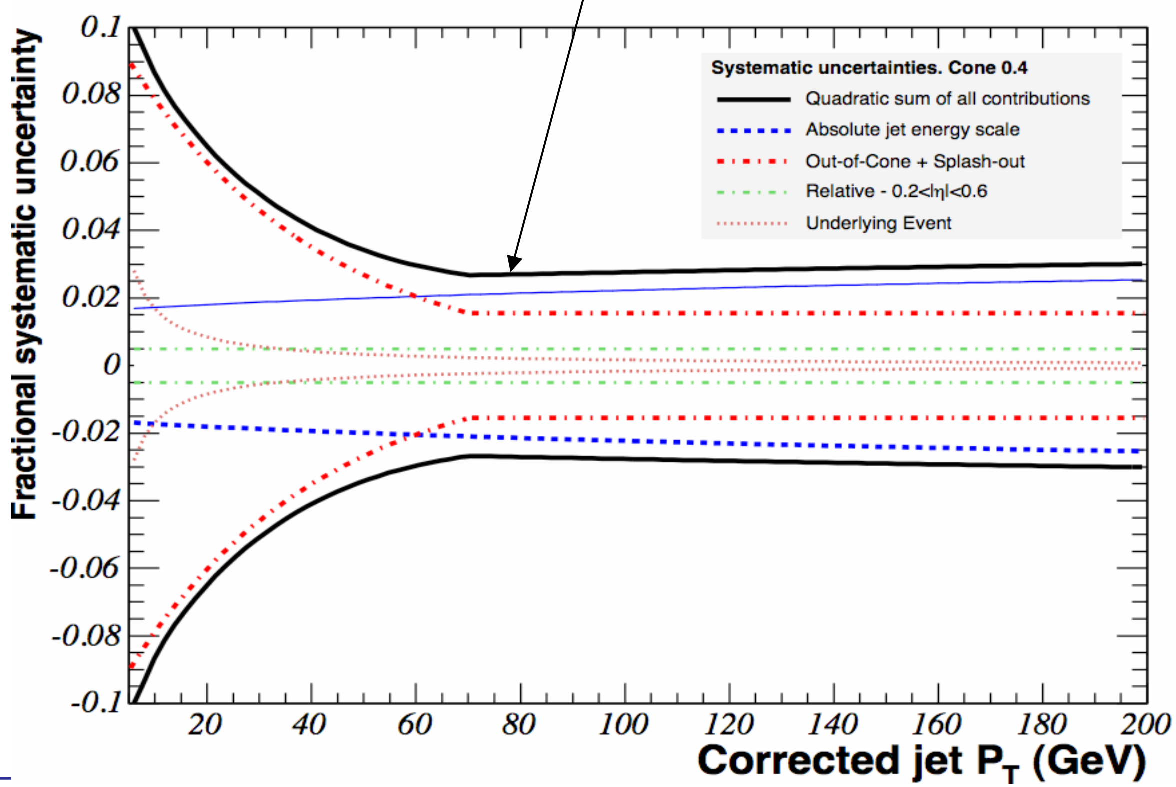




Jet Energy Scale



σ_c = unit of combined nominal CDF JES calibration uncertainty



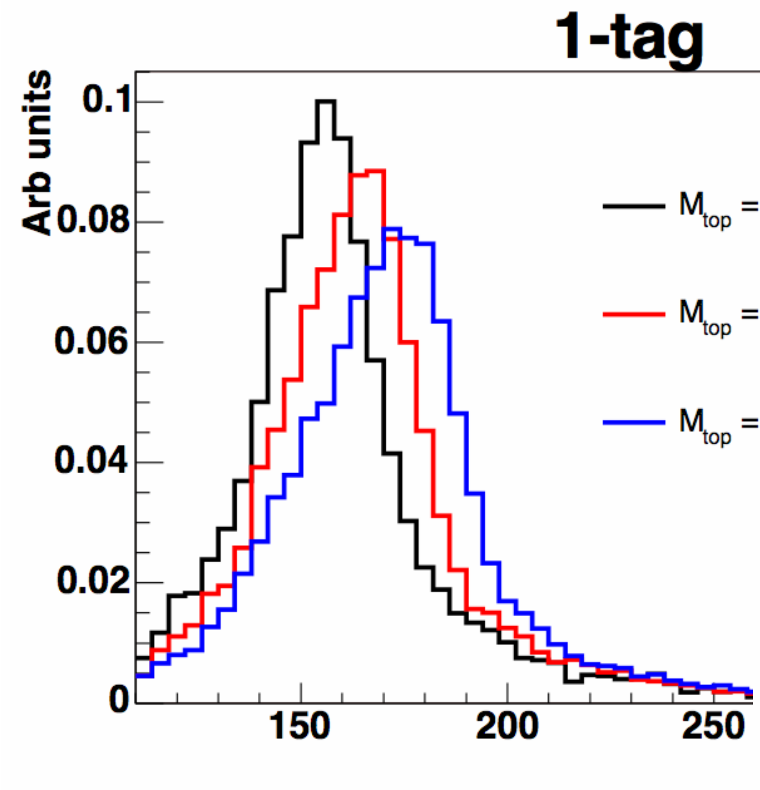
Top mass reconstruction in the lepton+jet channel

- Lepton jet channel is overconstraints system

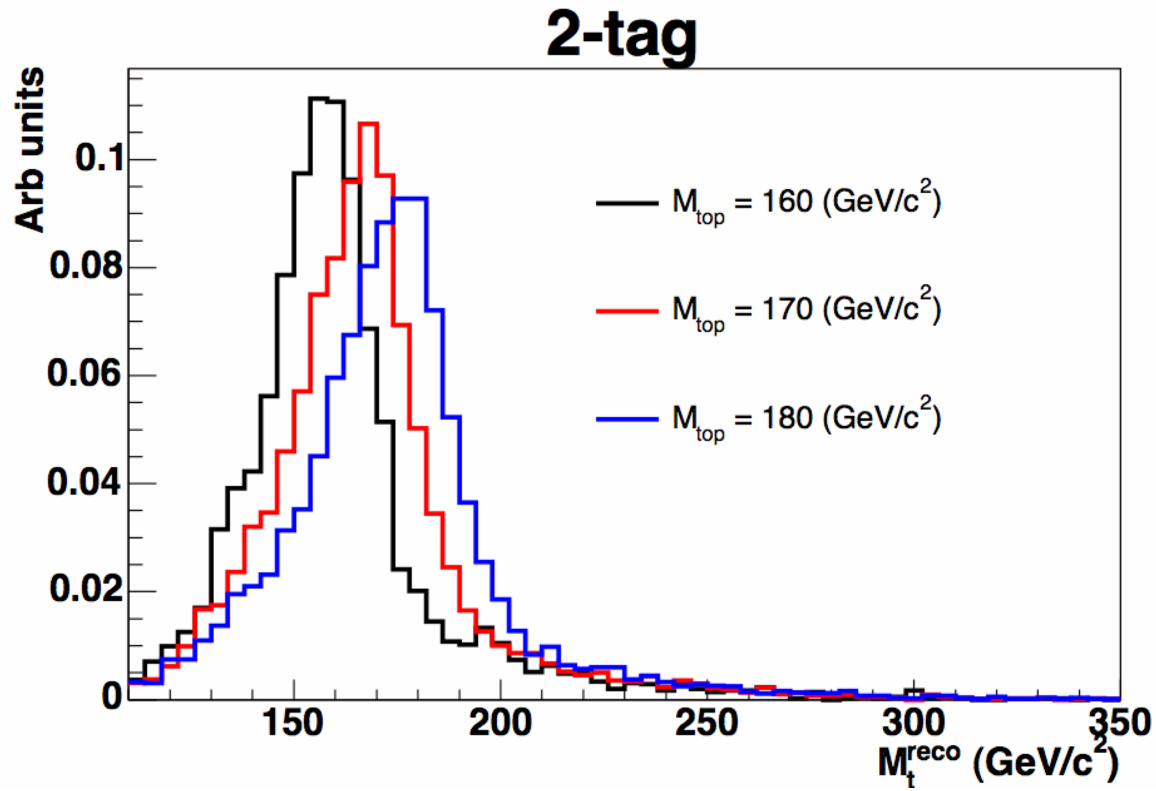
What we know	What we don't know
6 final-state particles *4 vectors = 24 needed	24 unknowns
4 jets and charged lepton 4-vectors = 4*5 = 20	4 unknowns
We know the mass of the neutrino = 1	3 unknowns
We know the W mass quite well (both of them) = 2	1 unknown
Require $m_{top} = m_{anti\text{-}top} = 1$	0 unknowns
Transverse components of p_ν from momentum conservation = 2	2 constraints

$$\chi^2 = \Sigma_{i=\ell,4jets} \frac{(p_T^{i,fit} - p_T^{i,meas})^2}{\sigma_i^2} + \Sigma_{j=x,y} \frac{(U_j^{fit} - U_j^{meas})^2}{\sigma_j^2} + \frac{(M_{jj} - M_W)^2}{\Gamma_W^2} + \frac{(M_{\ell\nu} - M_W)^2}{\Gamma_W^2} + \frac{(M_{bjj} - M_t)^2}{\Gamma_t^2} + \frac{(M_{b\ell\nu} - M_t)^2}{\Gamma_t^2}$$

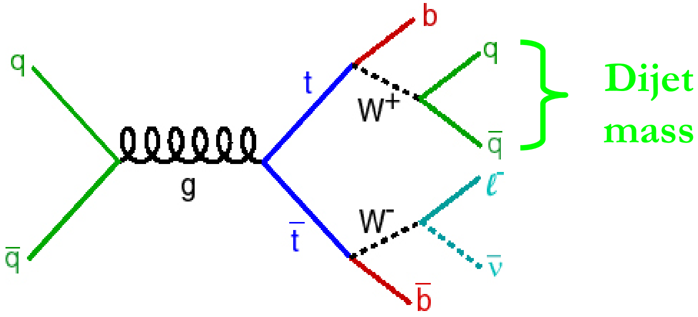
Reconstructed top mass (lepton jet)

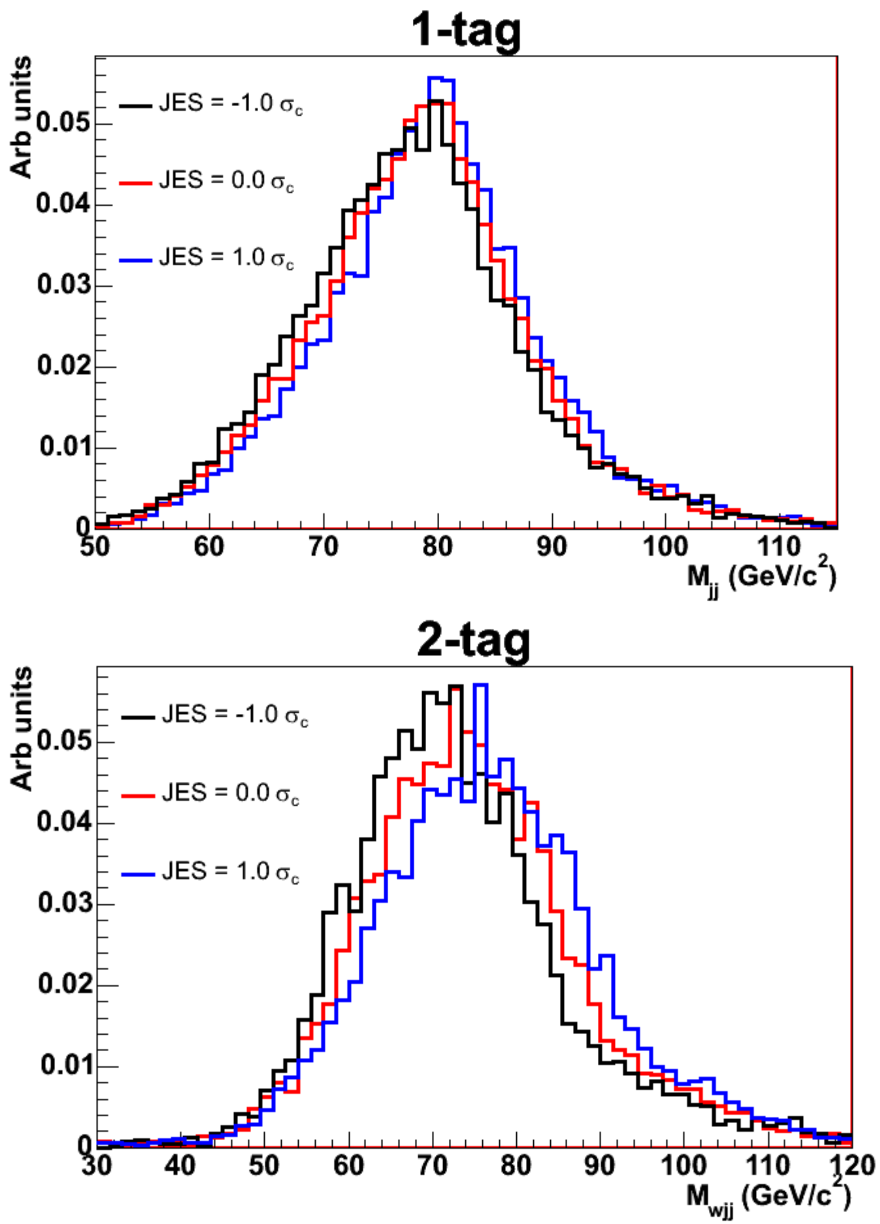


Reconstructed top quark mass is highly correlated with true top quark mass but not the same



JES (*in situ* correction)

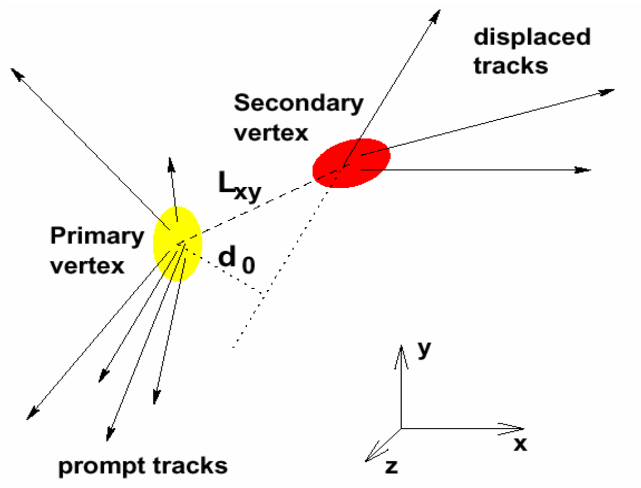




- JES is dominant systematic in the top quark mass measurement
- We can calibrate JET using the dijet mass of W decay because we very precisely know the W mass

Event Selection (lepton+jet)

- Use b-tagging in SVX to reduce combinatorics and increase S:B
- Divide events into 2 exclusive subsamples with different S:B and different reconstructed mass shapes

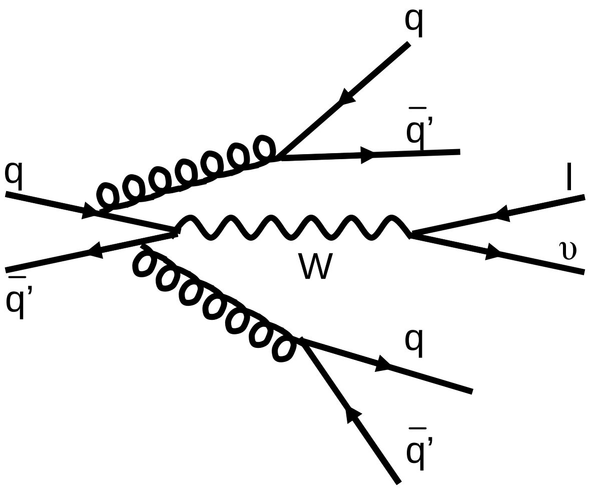


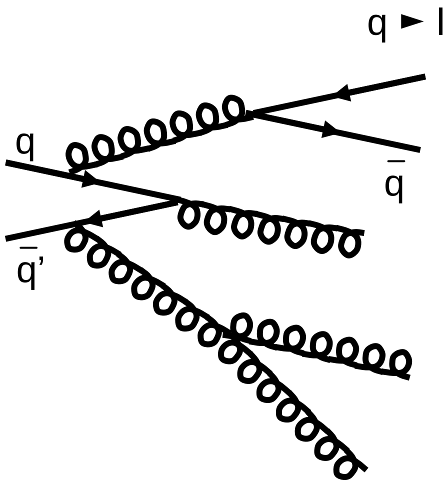
	2-tag	1-tag
B-tags	>=2	==1
Jet E_T (GeV) (jets 1-3)	> 20	> 20
Jet E_T (GeV) (jet 4)	> 12	> 20
Jet E_T (GeV) (extra jets)	any	< 20
χ^2 cut	< 9.0	< 9.0
MET (GeV)	> 20	> 20
e E_T (GeV) or μ P_T (GeV/c)	> 20	> 20

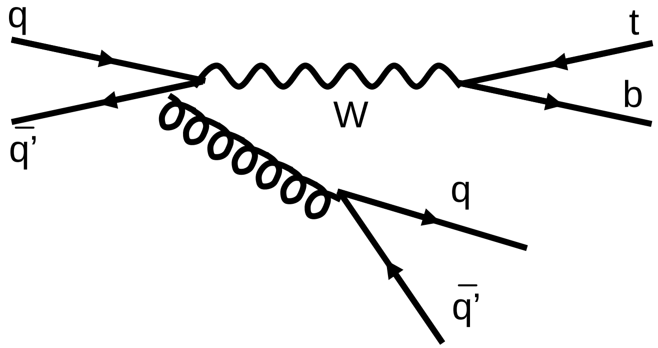
Top Event Tag Efficiency: 60%

False Tag Rate (per jet): 0.5%

Background in the lepton jet



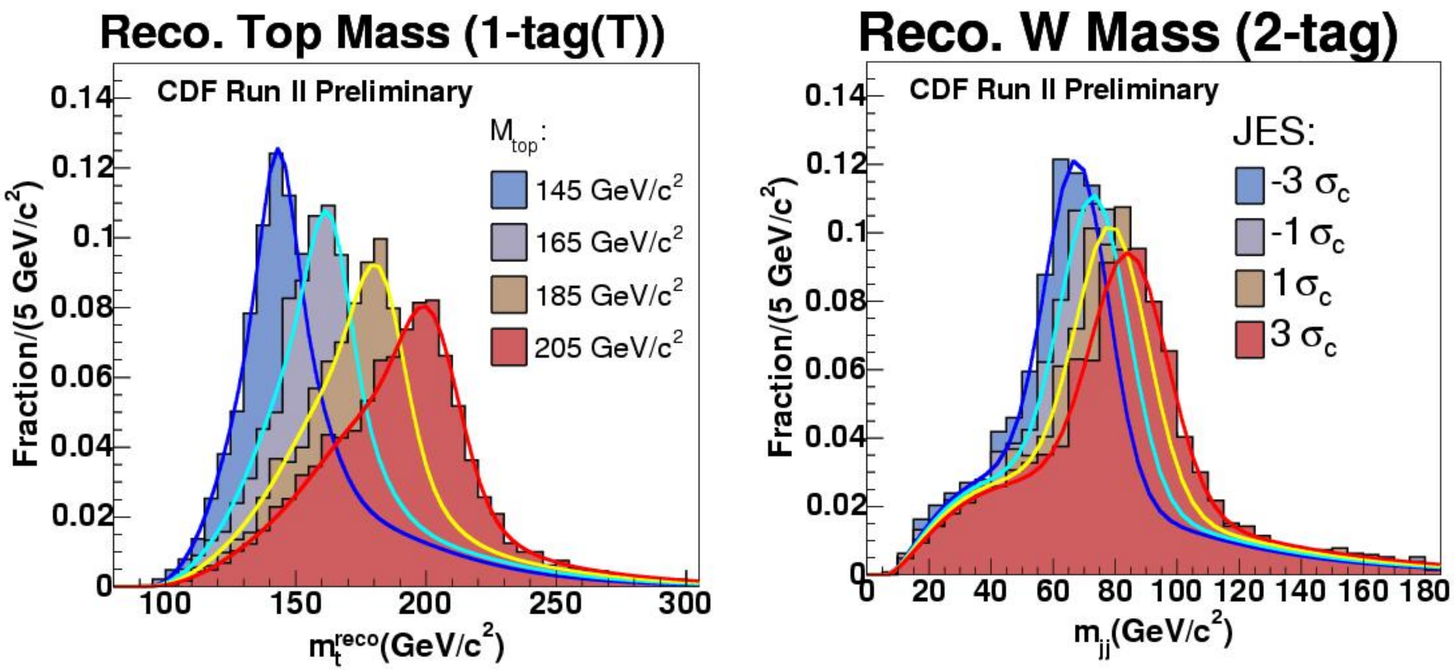




	1-tag	2-tag
Wbb	21.40 ± 8.95	5.48 ± 2.77
W$c\bar{c}$/Wc	19.95 ± 8.37	1.52 ± 0.74
W LF	14.11 ± 5.14	0.34 ± 0.16
single top	3.34 ± 0.36	1.27 ± 0.23
Diboson	4.13 ± 0.59	0.41 ± 0.12
QCD	18.32 ± 16.64	1.90 ± 2.64
Total	81.25 ± 27.96	10.91 ± 4.53
$t\bar{t}$ (σ=6.7pb M_{top}=175 GeV/c^2)	264.35 ± 37.09	131.02 ± 14.05

Building likelihood in the lepton jet

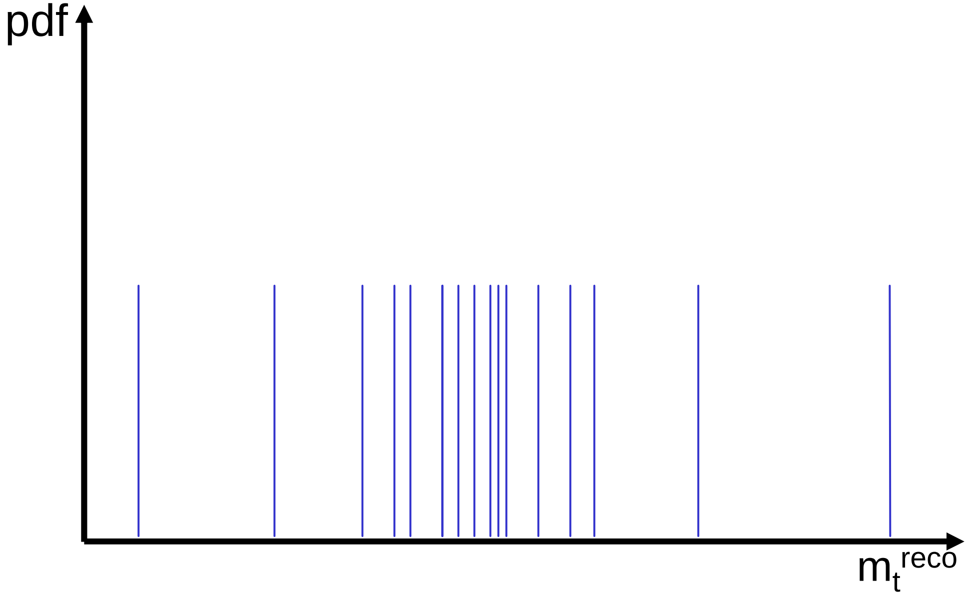
- We build probability for top quark mass using reconstructed top mass and Jet energy scale using dijet mass



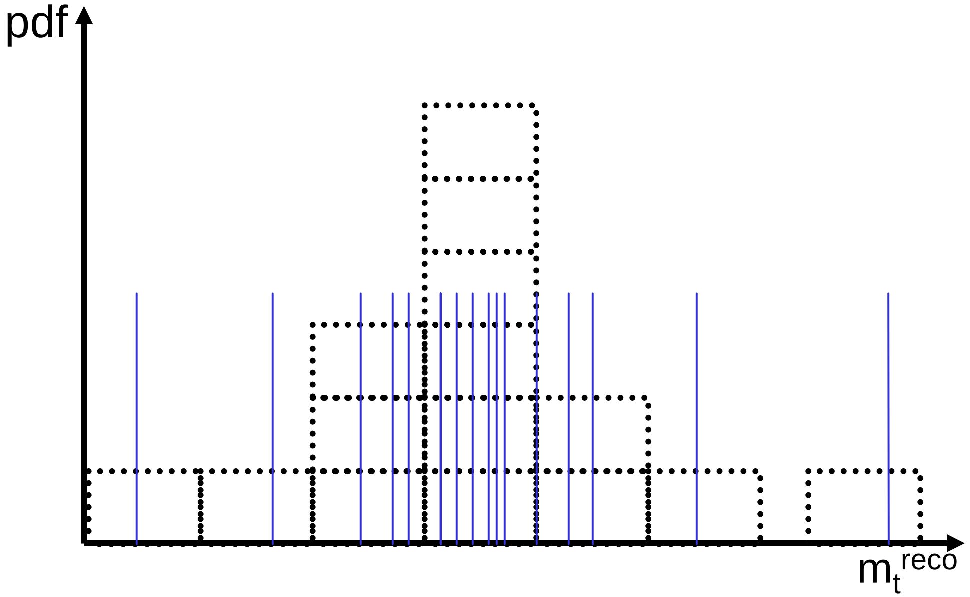
$$P(m_t^{reco}, m_{jj}; M_{top}, \Delta JES) = P(m_t^{reco}; M_{top}, \Delta JES) \times P(m_{jj}; \Delta JES)$$

- We use arbitarary function to build probability density function
- We assume no correlation between reconstructed top mass and dijet mass

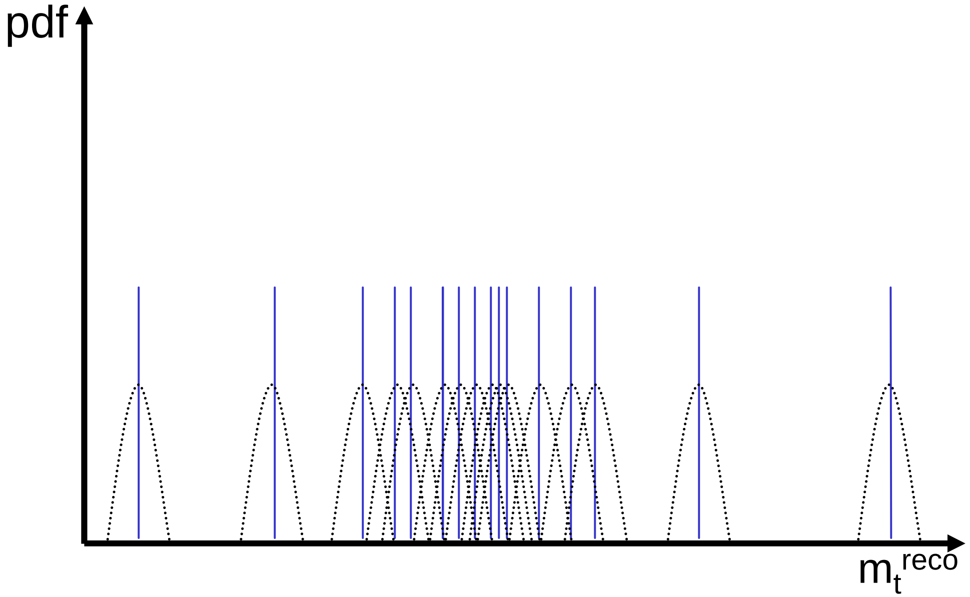
Kernel Density Estimation



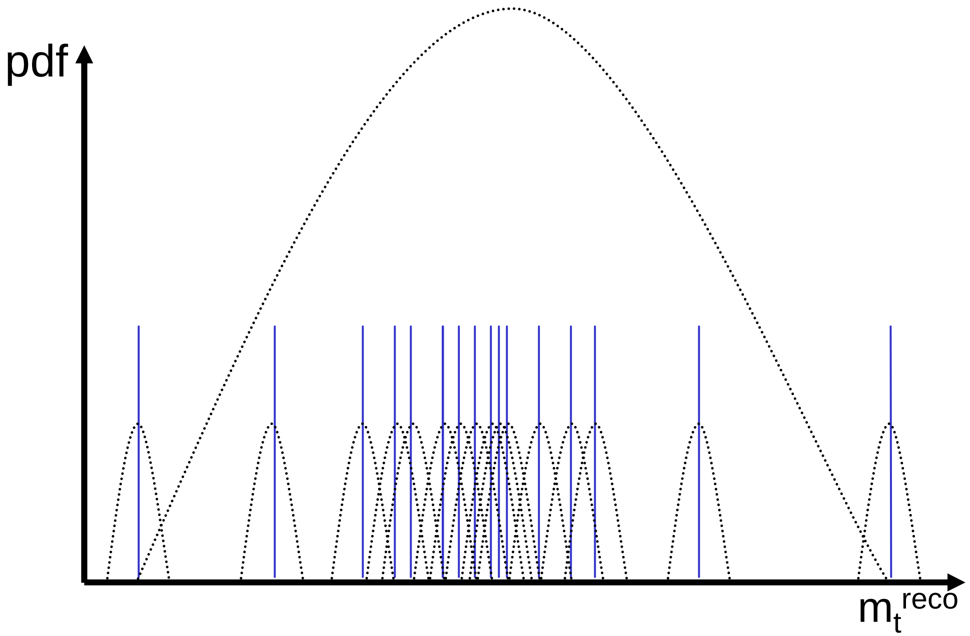
Kernel Density Estimation

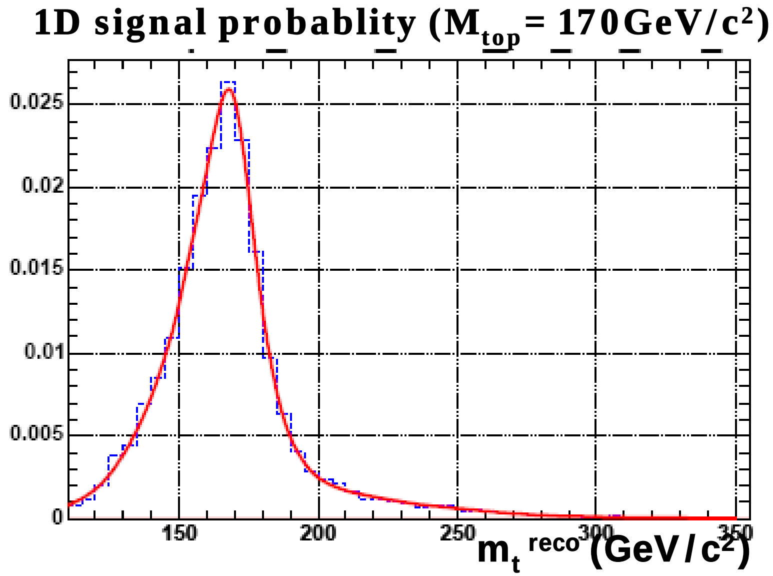


Kernel Density Estimation



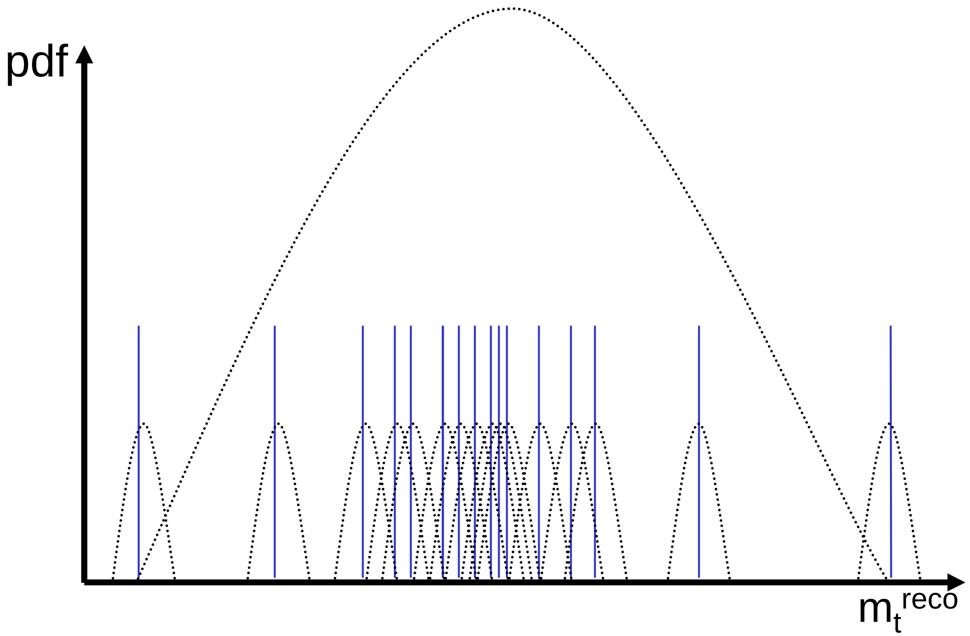
Kernel Density Estimation

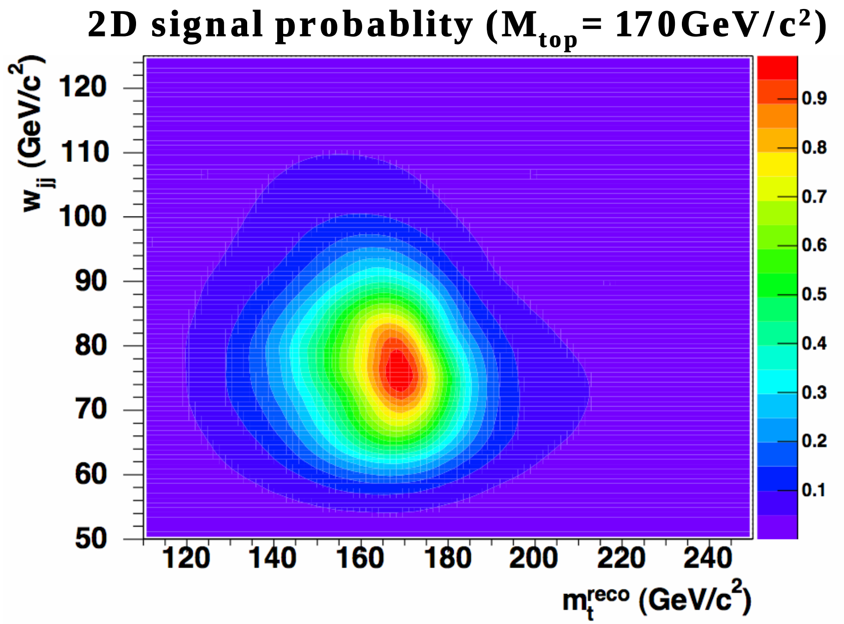




- No need to assume form of the shape
- Naturally extendible to more than 1 dimensions (correlations treated intrinsically)
- No way to interpolate between mass points

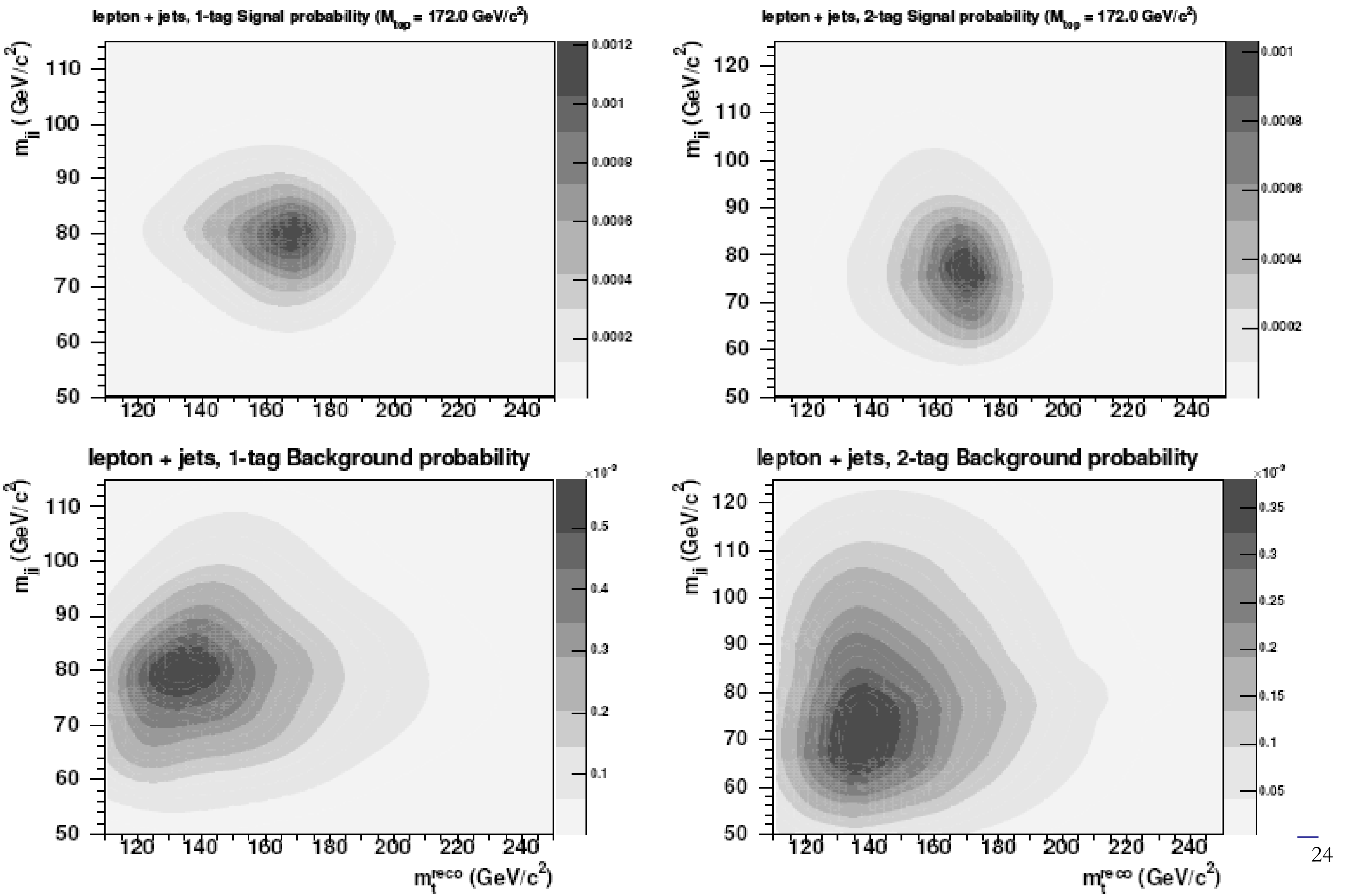
Kernel Density Estimation





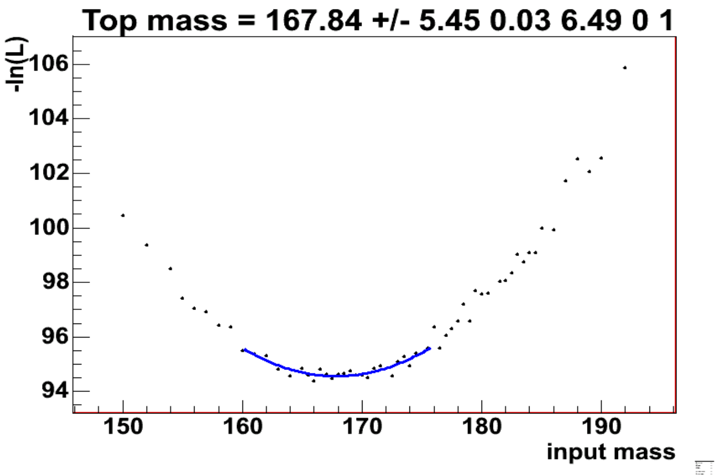
- Expand to 2D
  - ❖We can correctly account the correlation between two observables

Signal and background probability



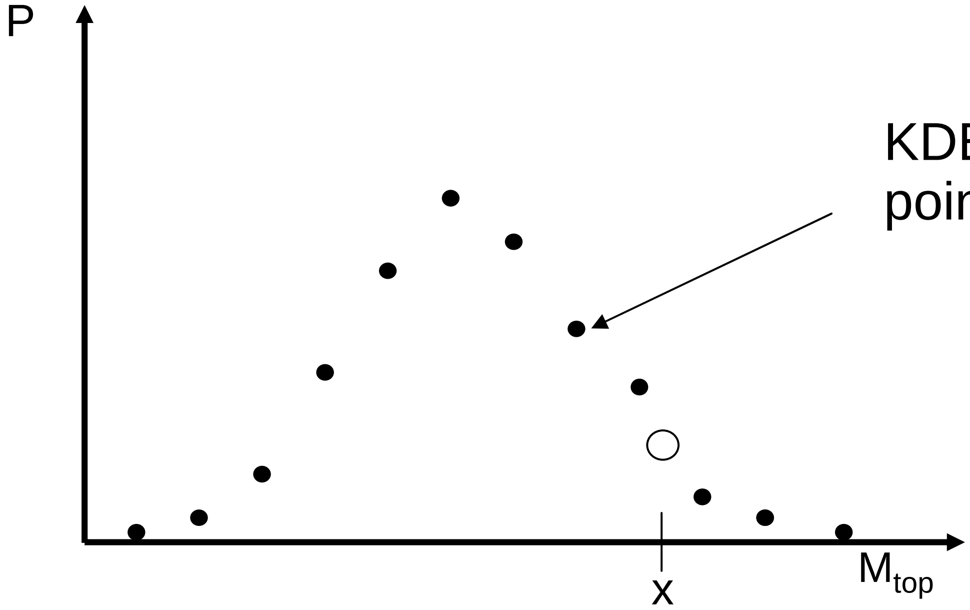
Likelihood fit and remained issue

$$\mathcal{L} = \frac{e^{-(n_s+n_b)}(n_s+n_b)^N}{N!} \times \prod_{i=1}^{N} \frac{n_s P_{sig}(m_{t,i}^{reco}, m_{jj,i}; M_{top}, \Delta_{JES}) + n_b P_{bg}(m_{t,i}^{reco}, m_{jj,i}; \Delta_{JES})}{n_s + n_b} \times e^{-\frac{(n_b - n_{b_0})^2}{2\sigma_{n_b}^2}}$$



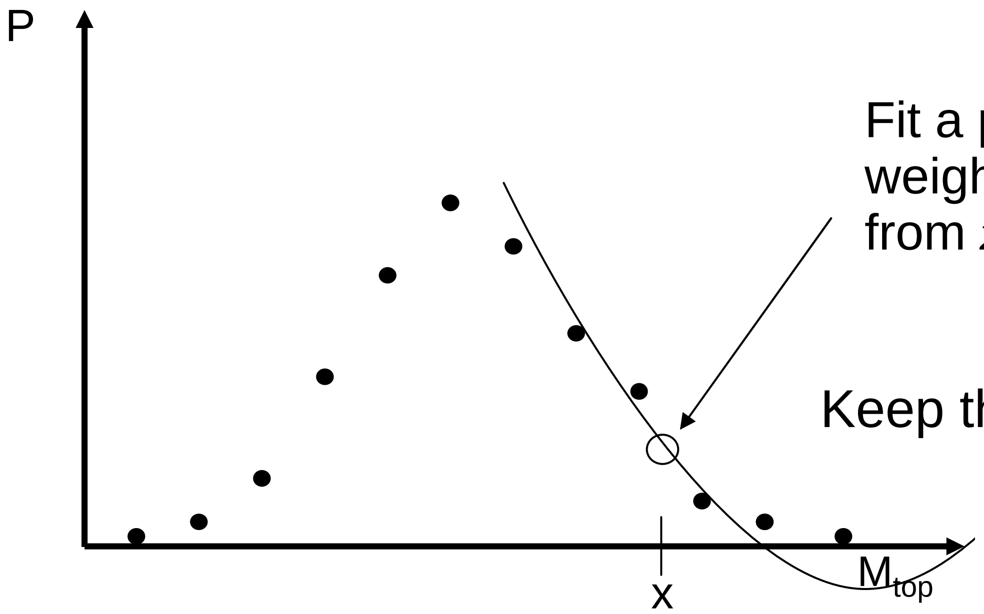
- The probability is only defined for discrete M_{top} and JES points
- We have assumption of parabola likelihood and do fit near maximum likelihood
- 2D is a bit complicate

Local polynomial smoothing



KDE gives us estimates Y_k at discrete points $\boldsymbol{x_k}$

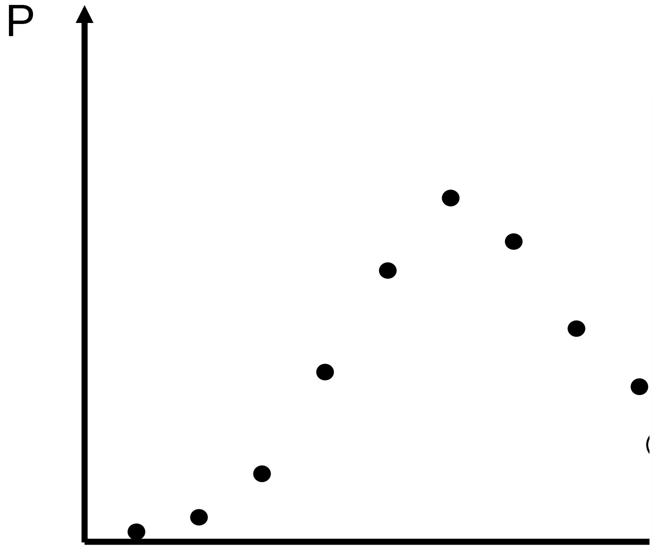
Local polynomial smoothing

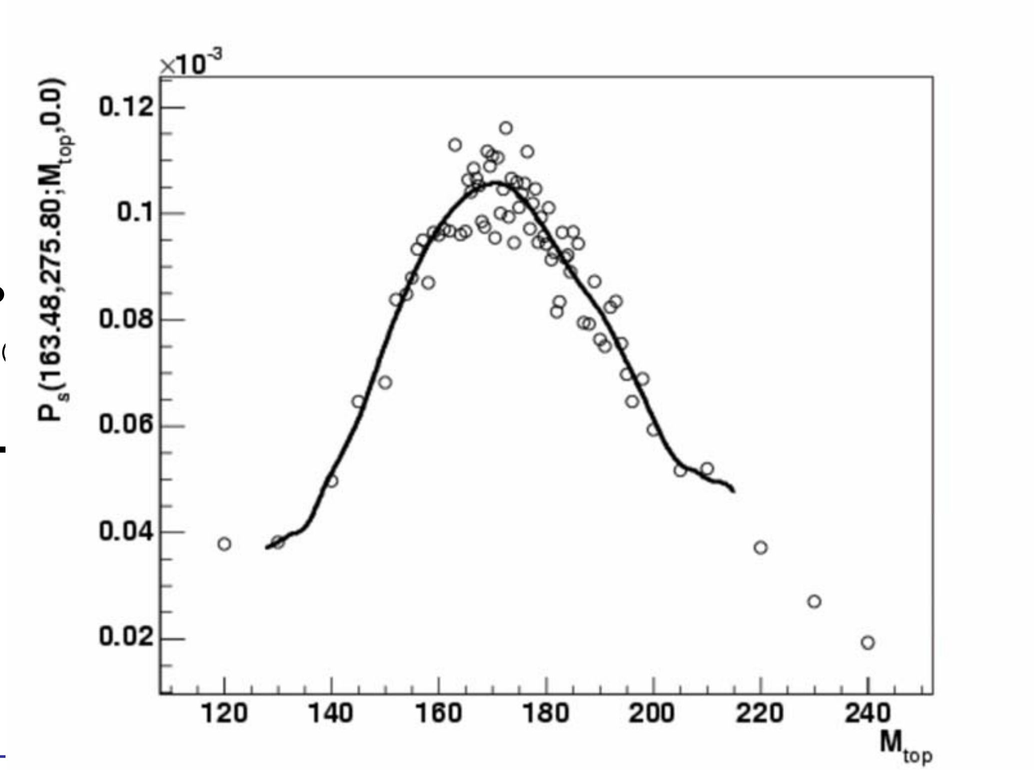


Fit a parabola- give decreasing weights to the estimates far away from ***x***

Keep the value at ***x***

Local polynomial smoothing

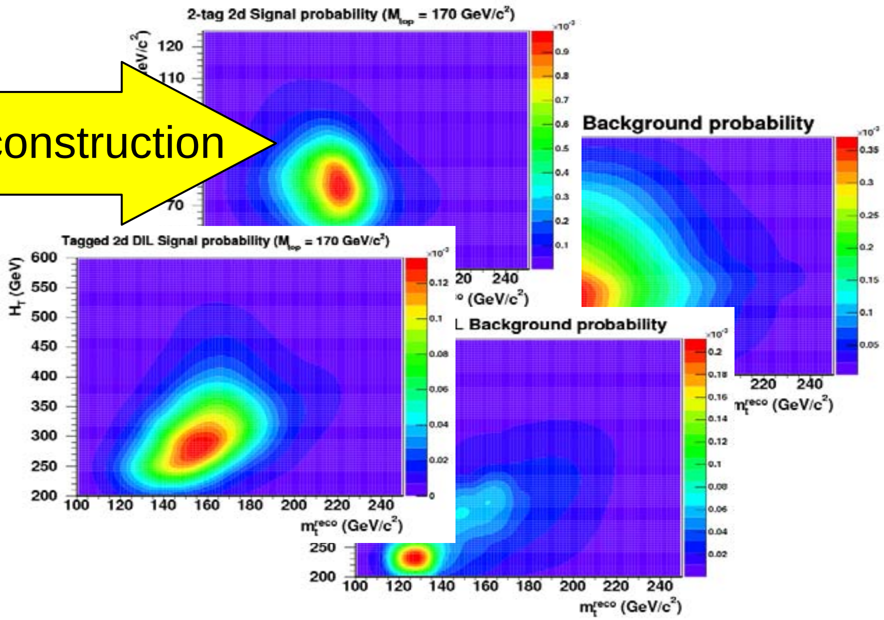




Our machinery now

MC
• $t\bar{t}$
• backgrounds

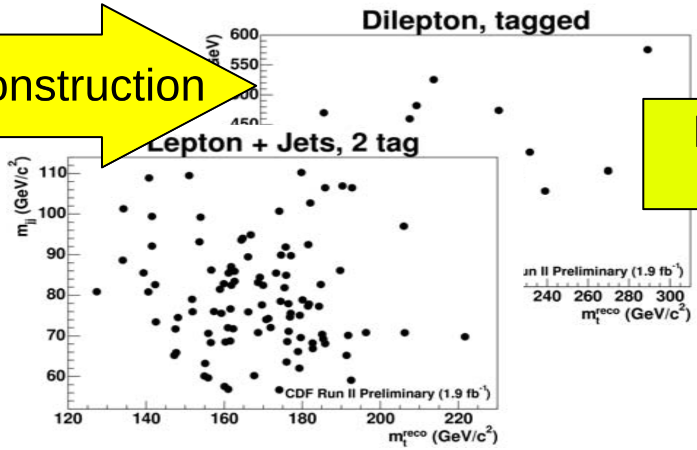
Event reconstruction



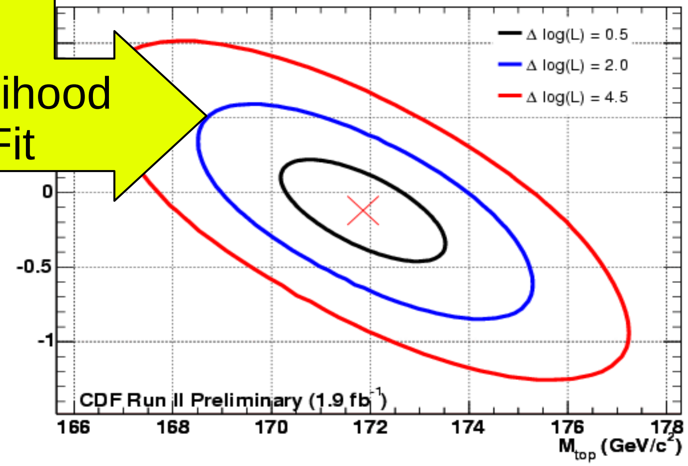
- 2d templates
- m_t^{reco}, w_{jj} in Lepton+Jets
- m_t^{reco}, m_{T2} in Dilepton
- signal pdf's depend on M_{top}, Δ_{JES}
- background pdf's depend on Δ_{JES}

DATA

Event reconstruction



Likelihood Fit



Machinery overview and dilepton channel

- Our KDE machinery allow us to use two observable without any assumption of PDF shape
- We can account correlation between two observable

- In dilepton channel
  - We can use two observable even if two observables have slightly different information
    - classical reconstructed top quark mass based on neutrino weighting algorithm (m_t^{NWA})
    - a kind of transverse mass with two missing particle case (m_{T2})
      - Similarly with W mass measurement

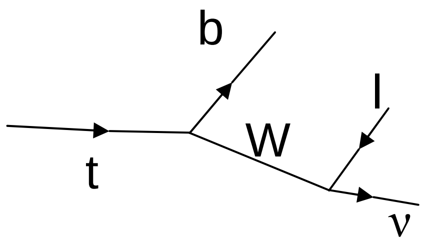
Dilepton selection

	0-tag	Tagged
b-tags	$= 0$	> 0
Leading 2 jets E_T (GeV)	> 15	> 15
Missing E_T (GeV)	> 25	> 25
H_T (GeV)	> 200	> 200
$M_t^{\rm NWA}$ boundary cut (GeV/c^2)	$100 < M_t^{\rm NWA} < 350$	$100 < M_t^{\rm NWA} < 350$
$m_{\rm T2}$ boundary cut (GeV/c^2)	$20 < m_{\rm T2} < 300$	$20 < m_{\rm T2} < 300$

- 1-lepton should be central and isolated but the other can be a forward and non-isolated
- We separate sample based on b-tagging
- Slightly lower energy of jet definition and bigger missing energy

m_t^{NWA}

- Leptonical decay of top
  - t->blv
  - We measure b and lepton but don't know neutrino
    - 4 unknown
  - Known parameter
    - W mass, neutrino mass (2 unknown)
  - If we assume the top quark mass and neutrino eta direction, we can measure neutrino x,y momentum
  - Same thing happen for the other leg
- Getting weight using measured missing transverse energy



$$
\begin{aligned}
B \equiv 2b\nu &= m_t^2 - m_W^2 - m_b^2 - 2b\ell \\
L \equiv 2\ell\nu &= m_W^2 - m_\ell^2 - m_\nu^2 \\
&= m_W^2 - m_\ell^2
\end{aligned}
$$

$$
w_i = exp\left(-\frac{(\not{E}_x - P_x^\nu - P_x^{\bar{\nu}})^2}{2\sigma_x^2}\right) \cdot exp\left(-\frac{(\not{E}_y - P_y^\nu - P_y^{\bar{\nu}})^2}{2\sigma_y^2}\right)
$$

$$
w_i = w_i(m_{top}, \eta_1^\nu, \eta_2^\nu)
$$

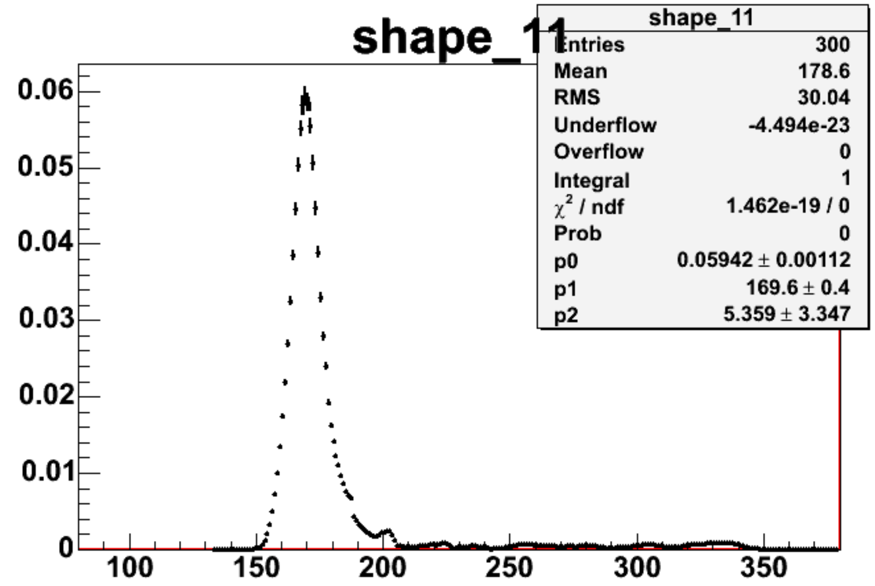
m_t^{NWA}

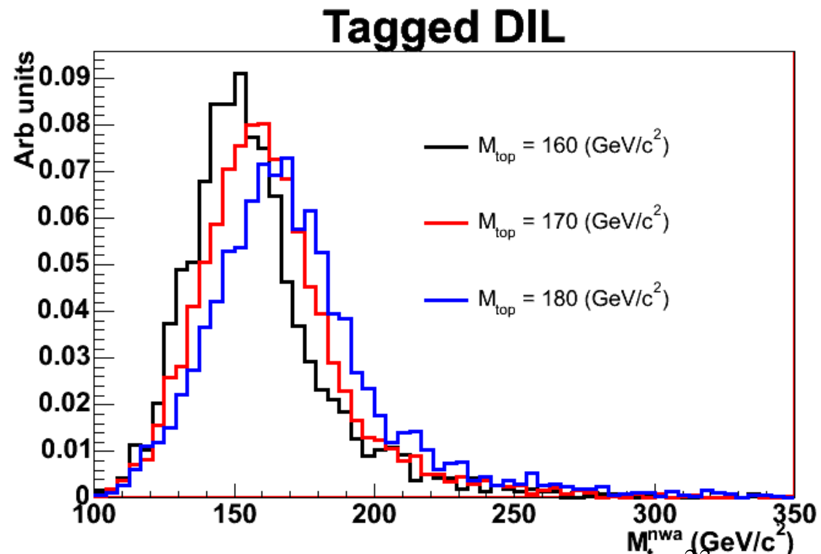
- Some over neutrino rapidities

$$W(m_t) = \int d\eta_1 \int d\eta_2 P(\eta_1) P(\eta_2) \sum_j \sum_i w(m_t)_{i,j}$$

- We have maximum weight m_t as reconstructed mass (m_t^{NWA})

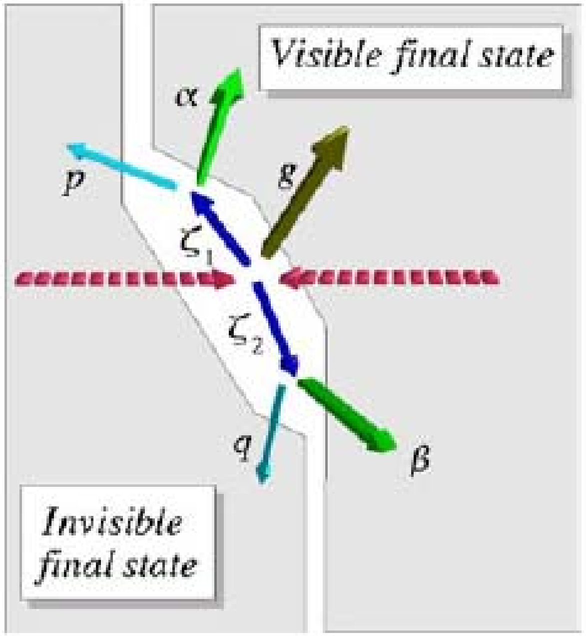
- We scan mt with 3GeV size and then decrease the step size upto 0.15GeV near the peak
- We have gaussian fit in the near of peak to get m_t continuously





m_{T2}

- Introduced to measure the mass of new physics particle)
  - Most of new physics predict long-live stable particle – dark matter candidate
  - We expect missing particle at the final state
  - If we consider pair production of new physics particle, it will have two missing particle



- Top dilepton channel have exactly same final state

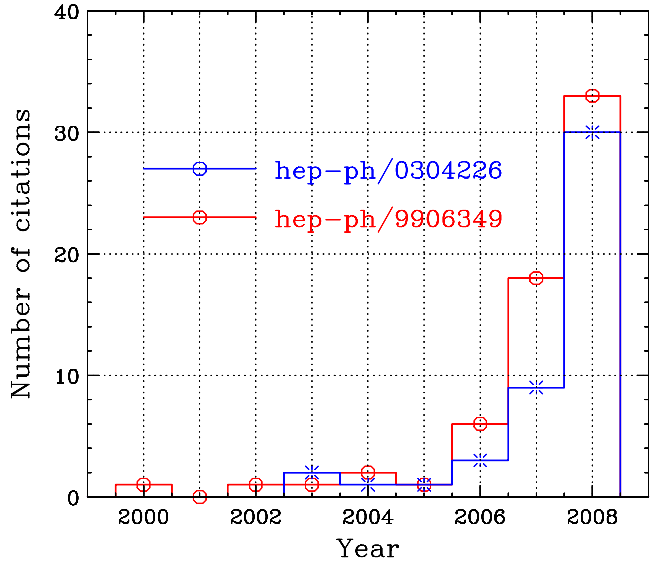
$$m_{T2} = \min[\max(m_{T(1)}, m_{T(2)})]$$

$$q_T + p_T = \text{missing } p_T$$

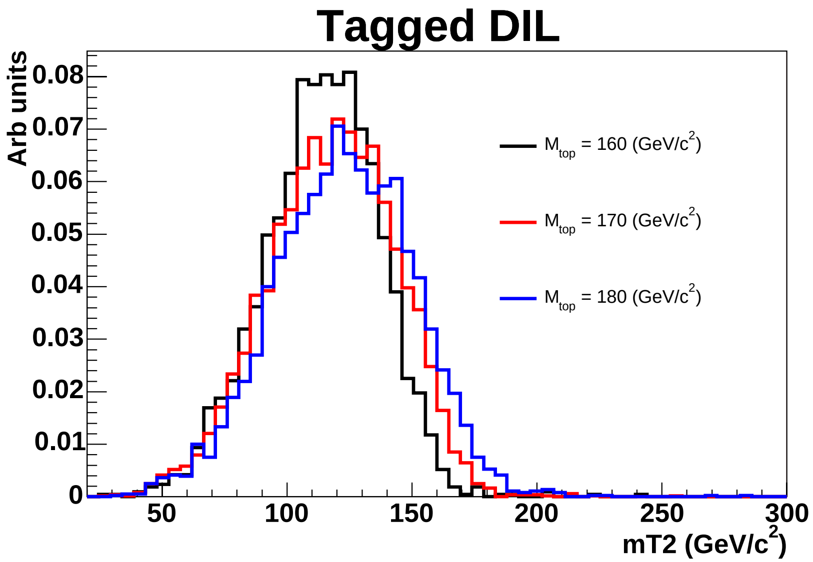
Alan Barr, Christopher Lester and Phil Stephens
J. Phys. G: Nucl. Part. Phys. **29** (2003) 2343–2363

m_{T2}

- Transverse mass of two missing particle system
  - Similar with m_T for W mass
- Can be useful to determine the mass of new physics particle
  - One of the most stringent variable



- Top dilepton channel is good example of m_{T2} variable (standard candle)
- We can use real data
  - First application in the real data



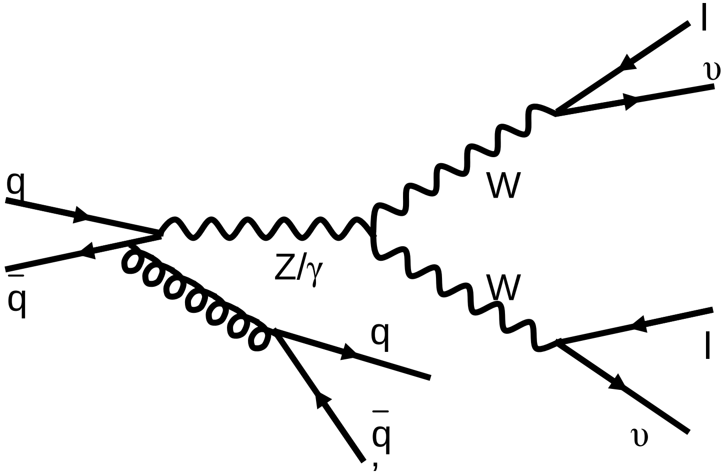
Estimated uncertainty in the dilepton channel

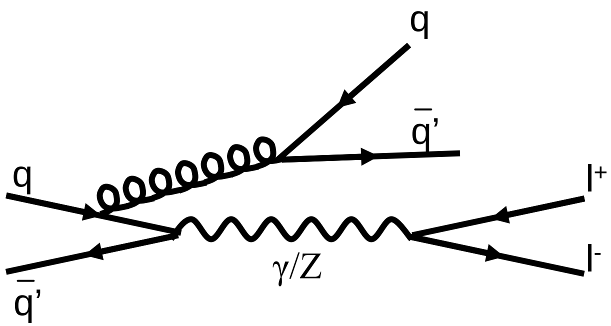
175 GeV/c^2 top mass assumed

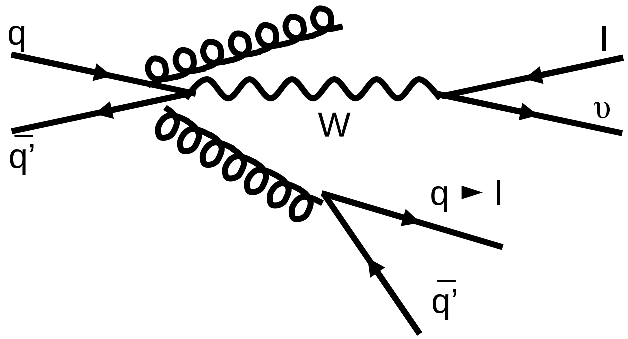
Unit (GeV/c^2)		m_{T2}	m_t^{NWA}	H_T	$m_t^{NWA}+m_{T2}$	$m_t^{NWA}+H_T$
Statistical		4.0	3.4	5.4	2.9	3.2
Systematic	Jet Energy Scale	2.6	3.5	3.7	3.0	3.4
	Generator	0.3	1.0	2.6	0.5	1.3
	Parton distribution functions	0.5	0.6	1.8	0.5	0.8
	b-JES	0.2	0.3	0.2	0.2	0.3
	Background shape	0.4	0.3	0.7	0.1	0.3
	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} give the best performance
between single observables

Dilepton background

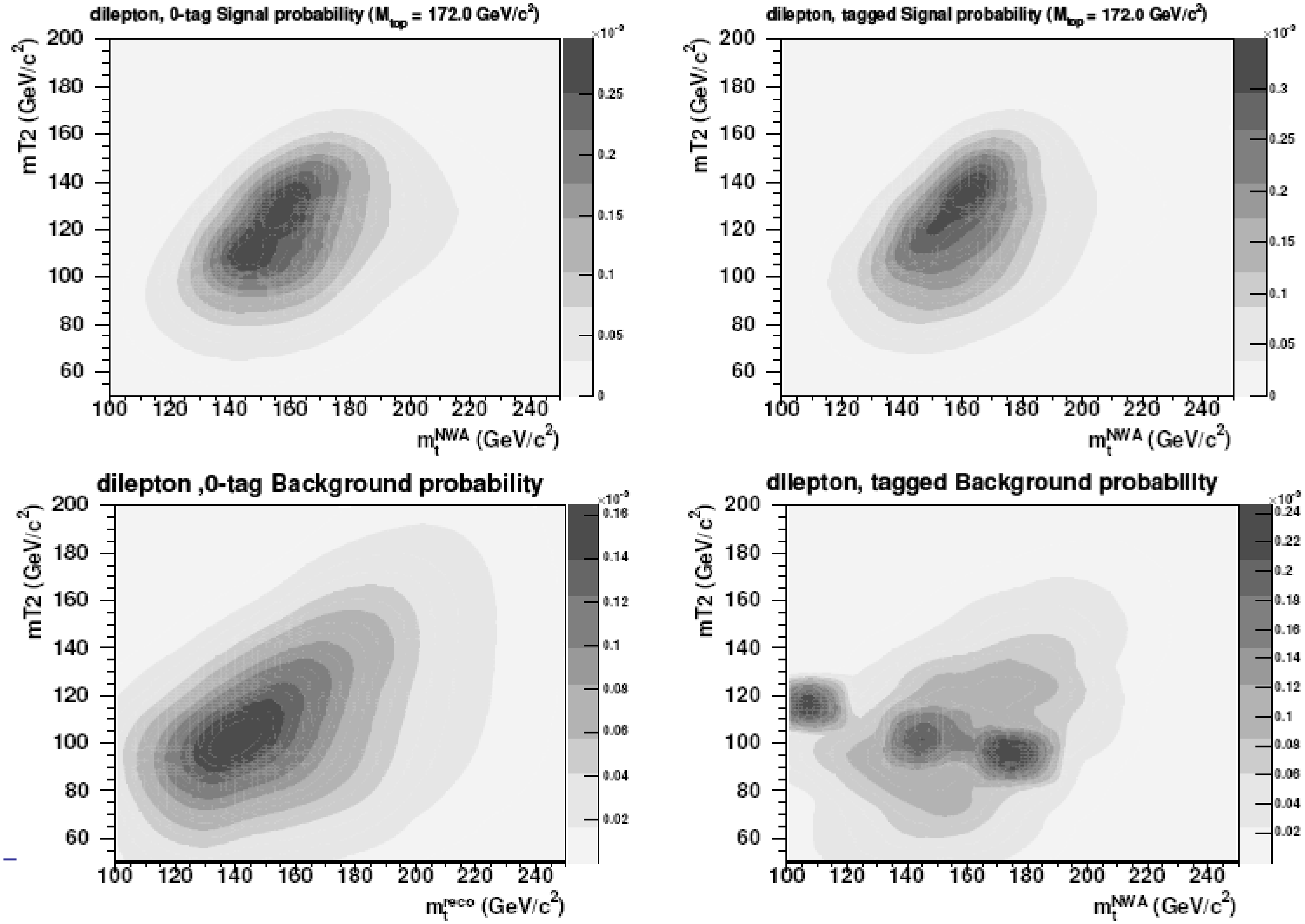






	non-tagged	tagged
Diboson	15.2 ± 2.3	0.6 ± 0.1
Drell-Yan	31.1 ± 3.5	1.7 ± 0.2
Fake	31.2 ± 8.7	4.5 ± 1.3
Total Background	71.3 ± 10.5	6.8 ± 1.3
$t\bar{t}$ (6.7 pb)	68.7 ± 6.8	88.4 ± 8.2
Observed (3.4 fb^{-1})	149	87

Signal and background probabilty



Combining lepton+jet and dilepton channel

$$\mathcal{L} = e^{\frac{\Delta_{\mathrm{JES}}{}^2}{2}} \times \mathcal{L}_{\text{lepton+jets,1-tag}} \times \mathcal{L}_{\text{lepton+jets,2-tag}} \times \mathcal{L}_{\text{dilepton,non-tagged}} \times \mathcal{L}_{\text{dilepton,tagged}}$$

$$\mathcal{L}_{\mathrm{bg}} = e^{\frac{(n_{b0}-n_b)^2}{2\sigma^2_{n_{b0}}}}$$

a-priori background constraint

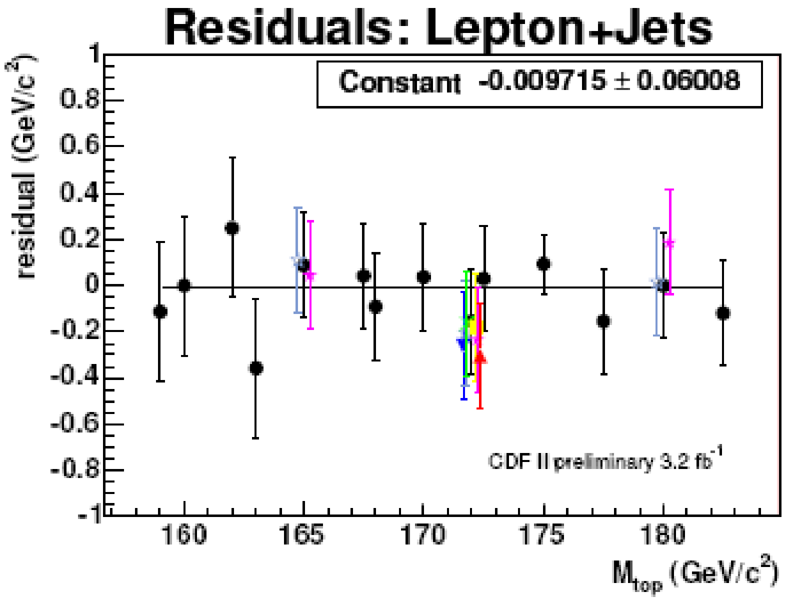
$$\mathcal{L}_{\mathrm{shape}} = \frac{\exp(-(n_s+n_b))(n_s+n_b)^N}{N!} \prod_{i=1}^{N} \frac{n_s P_s(m^{\mathrm{reco}}_{\mathrm{t},i}, v_i; \mathrm{M}_{\mathrm{top}}, \Delta_{\mathrm{JES}}) + n_b P_b(m^{\mathrm{reco}}_{\mathrm{t},i}, v_i; \Delta_{\mathrm{JES}})}{n_s + n_b}$$

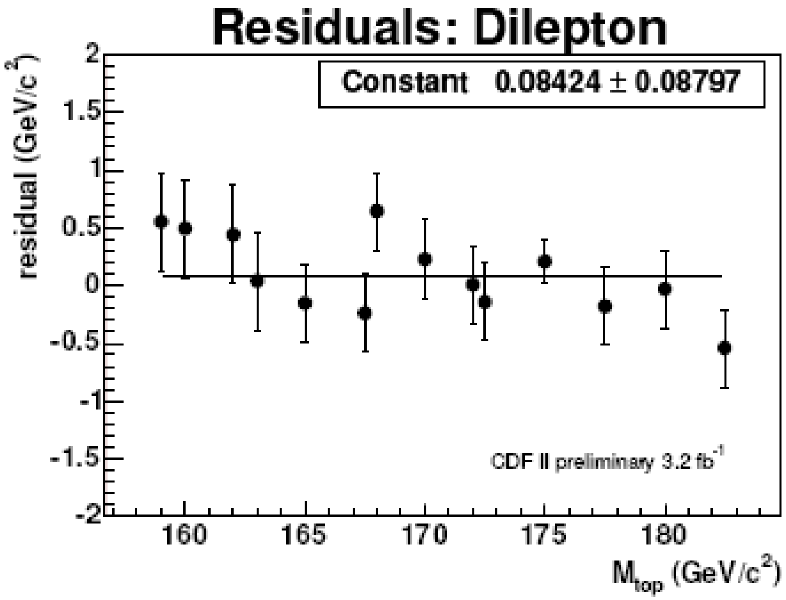
Extended likelihood:
fit for signal and
background expectation

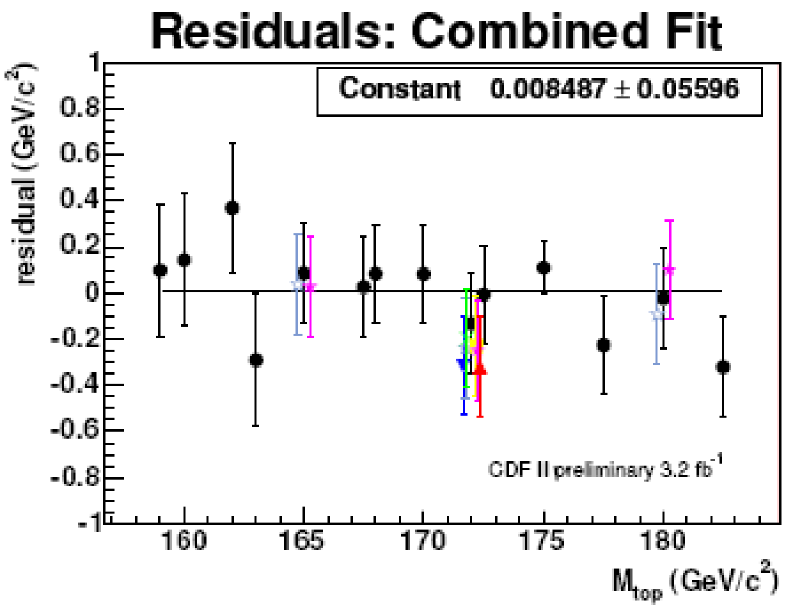
Why we measure simultaneously multiple channels

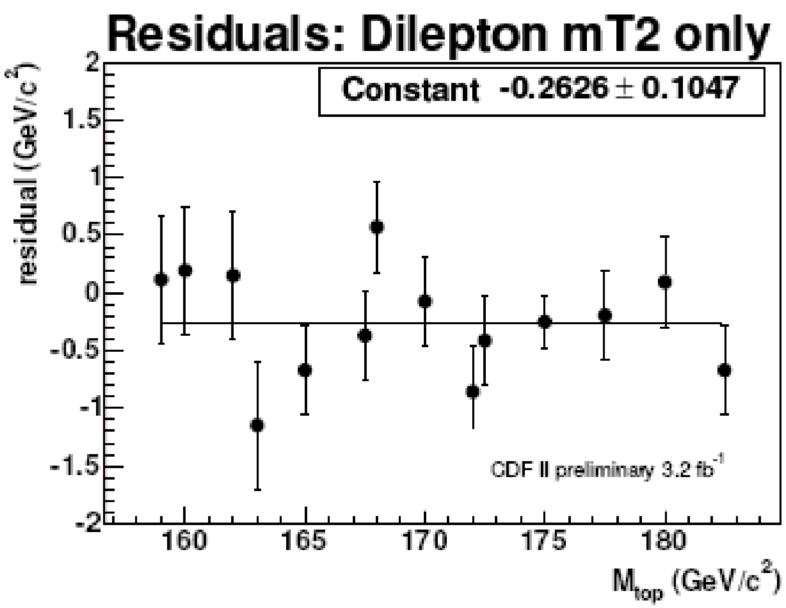
- Robust combination
- No assumptions about correlations for systematics
- Have cross check crossing different channel with one machinery
- Dileptons make use of the Lepton+Jets *in situ* JES calibration

Mass residual from MC pseudoexperiments

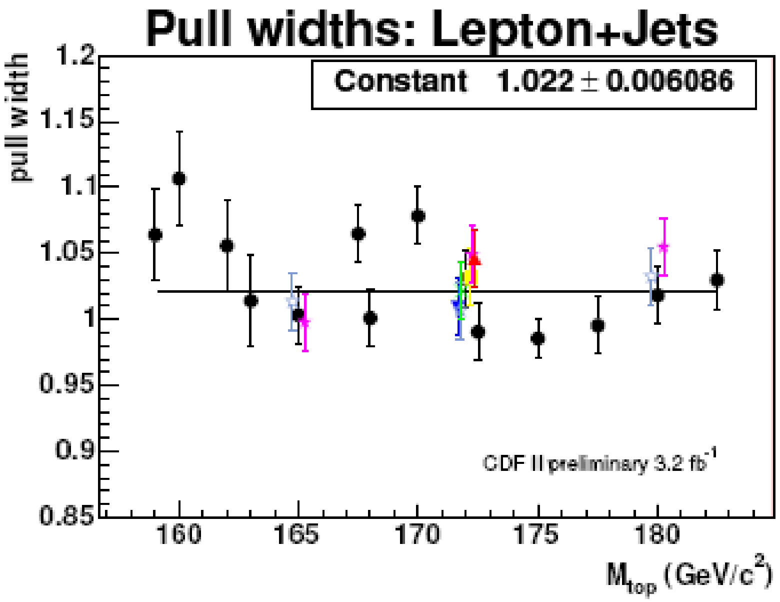


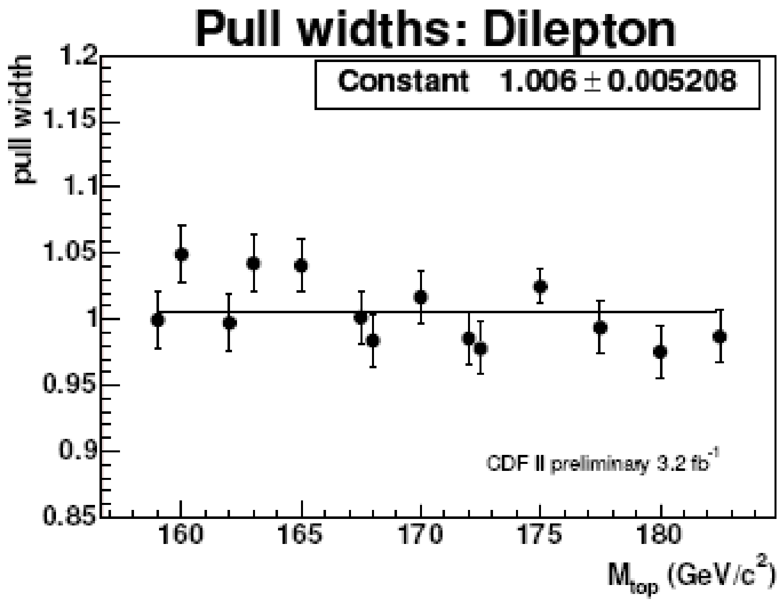


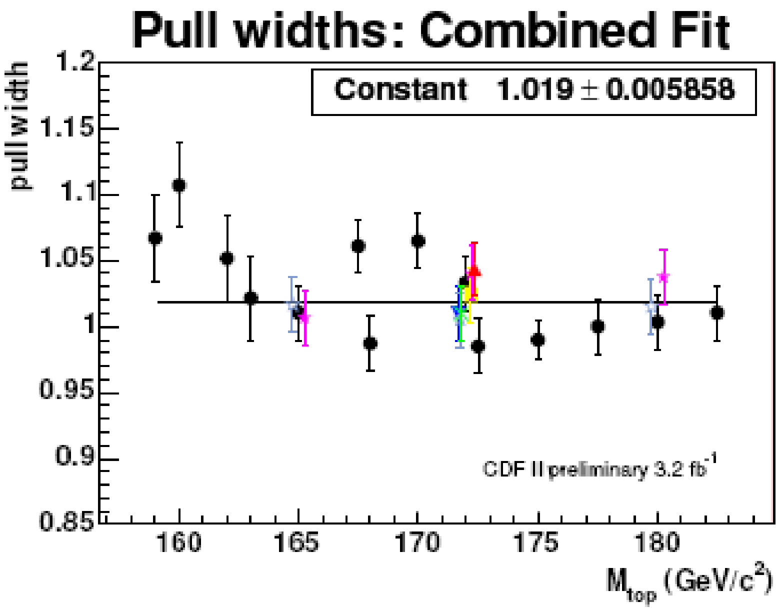


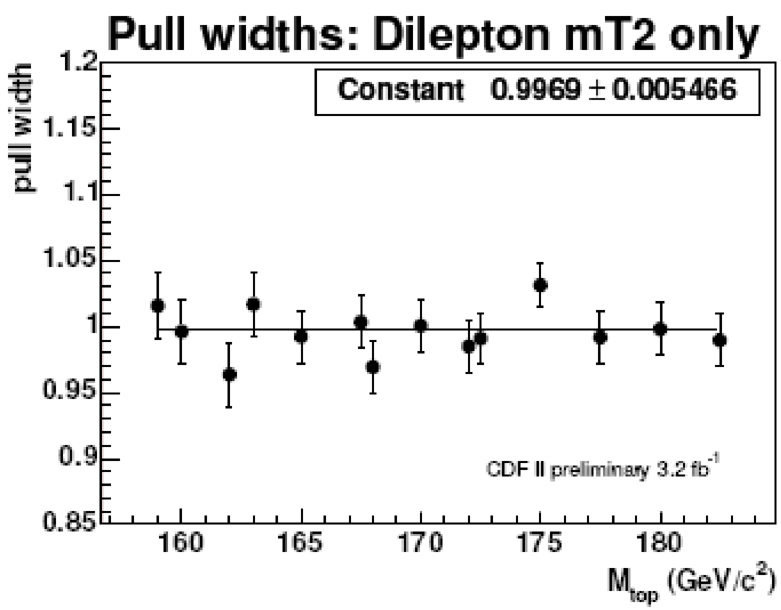


Pull width


Pull widths: Lepton+Jets
Constant 1.022 ± 0.006086
pull width
M_top (GeV/c^2)
CDF II preliminary 3.2 fb^-1


Pull widths: Dilepton
Constant 1.006 ± 0.005208
pull width
M_top (GeV/c^2)
CDF II preliminary 3.2 fb^-1

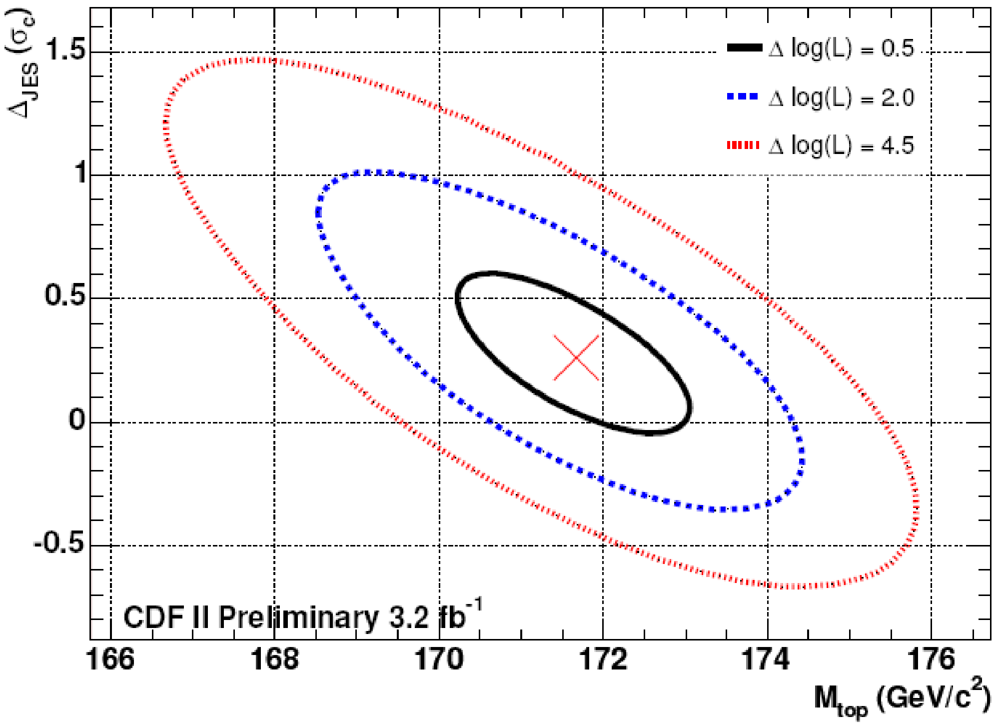

Pull widths: Combined Fit
Constant 1.019 ± 0.005858
pull width
M_top (GeV/c^2)
CDF II preliminary 3.2 fb^-1

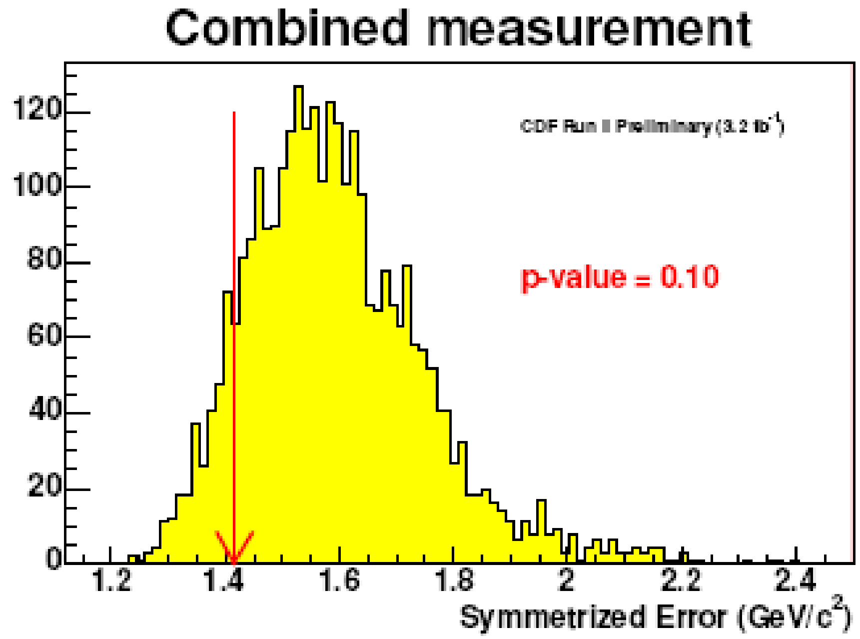

Pull widths: Dilepton mT2 only
Constant 0.9969 ± 0.005466
pull width
M_top (GeV/c^2)
CDF II preliminary 3.2 fb^-1

Systematics

Systematic (GeV/c^2)	Combination	LJ	DIL	DIL-m_{T2} only
Residual JES	0.68	0.66	3.04	2.58
Generator:	0.74	0.72	0.46	0.31
PDFs	0.19	0.17	0.48	0.47
b jet energy	0.17	0.18	0.21	0.21
Background shape	0.21	0.25	0.12	0.36
gg fraction	0.04	0.00	0.01	0.32
Radiation	0.13	0.14	0.34	0.57
MC statistics	0.10	0.09	0.34	0.34
Lepton energy	0.06	0.03	0.28	0.56
Multiple Hadron Interactions	0.19	0.24	0.34	0.18
Color reconnection	0.34	0.38	0.55	0.68
Total systematic	1.14	1.14	3.24	2.92

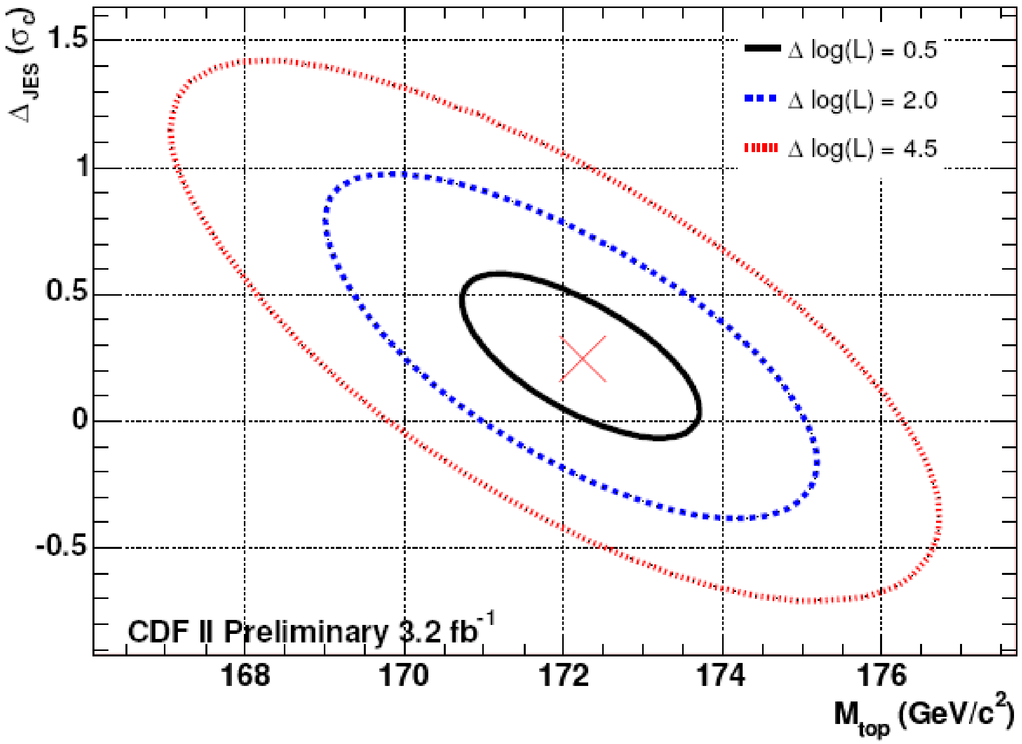
Result (combined LJ+DIL)

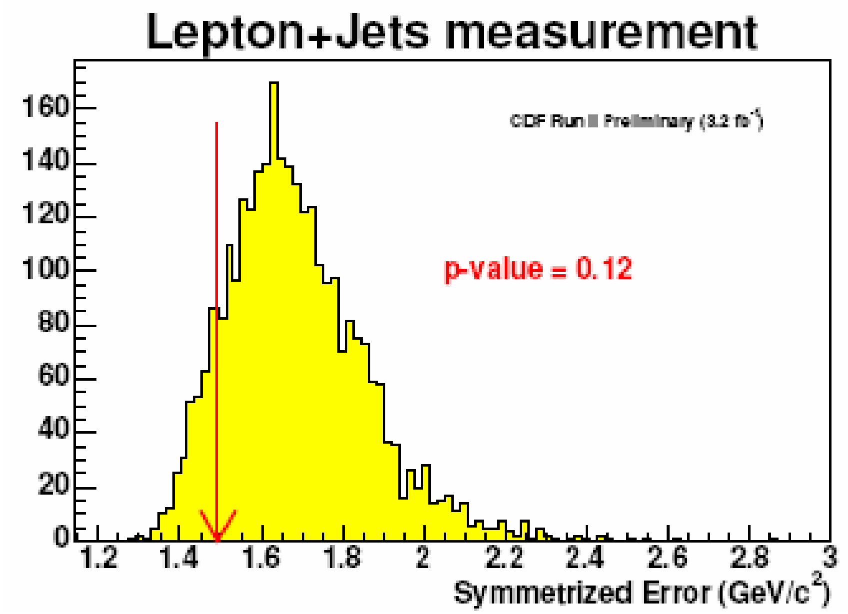




$$171.7\,^{+1.4}_{-1.5}\,(\text{stat.}) \pm 1.1\ \text{GeV/c}^2(\text{syst})$$
$$= 171.7\,^{+1.8}_{-1.9}\,\text{GeV/c}^2$$

Result (Lepton Jet)

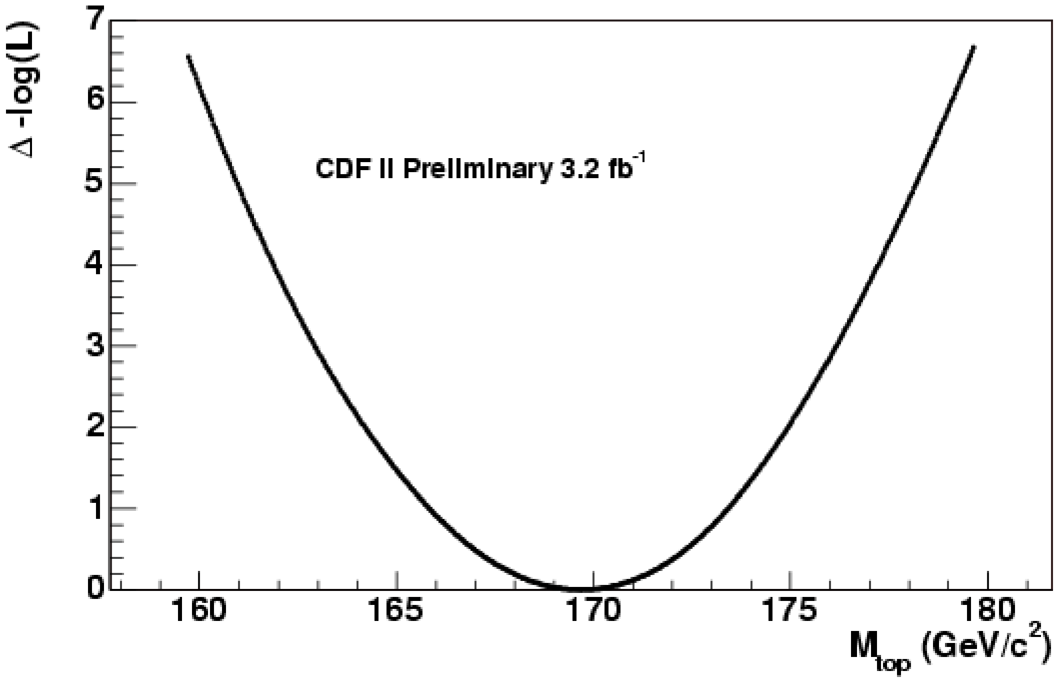


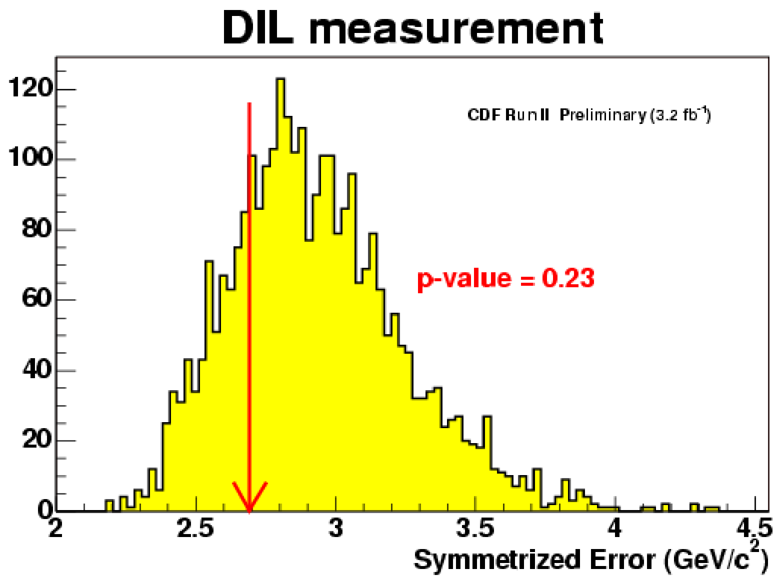


$$172.2 ^{+1.5}_{-1.6} \text{(stat.)} \pm 1.1 \text{ GeV/c}^2 \text{(syst)}$$

$$= 172.2 ^{+1.9}_{-1.9} \text{GeV/c}^2$$

Result : Dilepton (two observables)

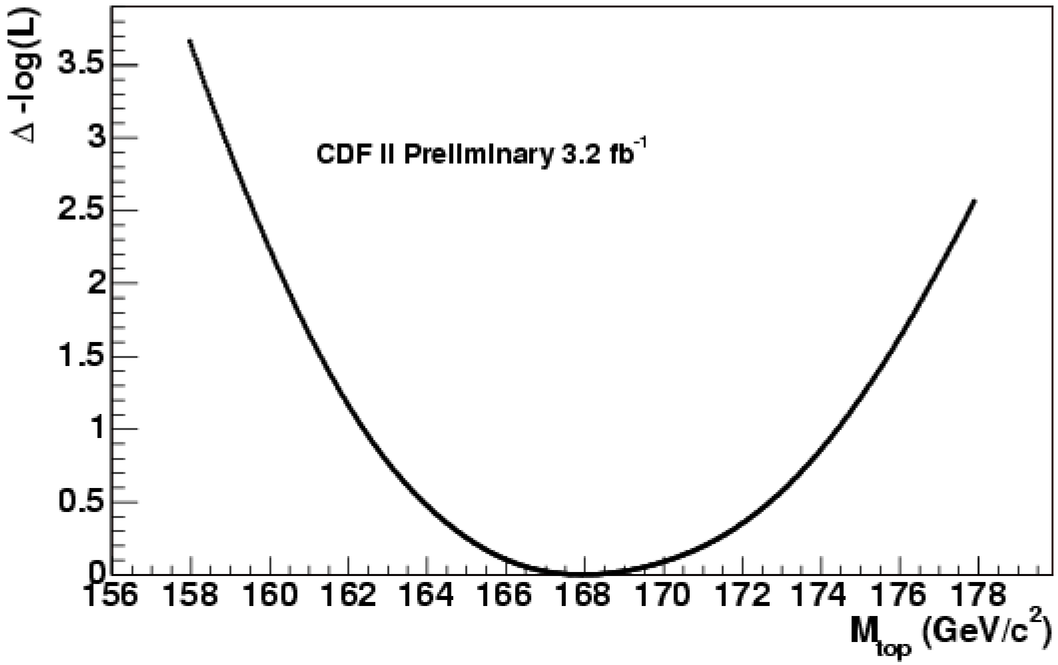


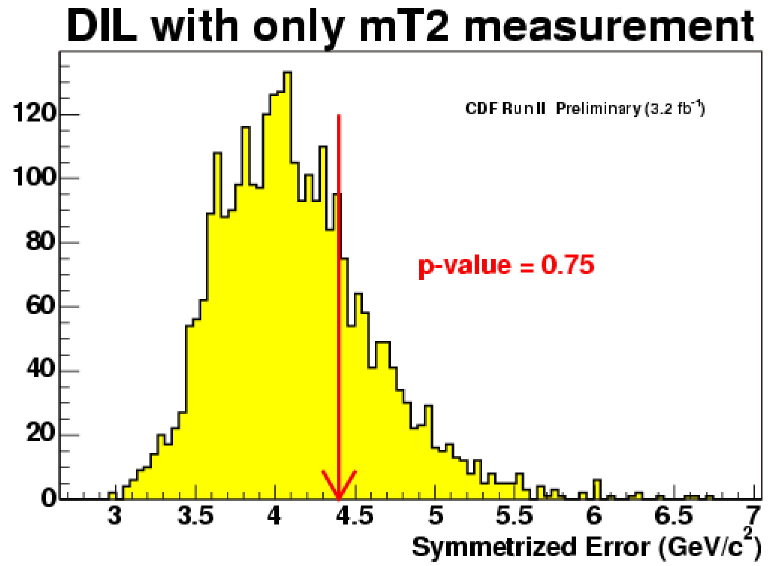


169.3 ± 2.7 (stat.) $\pm$ 3.2 GeV/c^2(syst.)

$= 169.3 \pm 4.2$ GeV/c^2

Result : Dilepton (m_{T2} only)



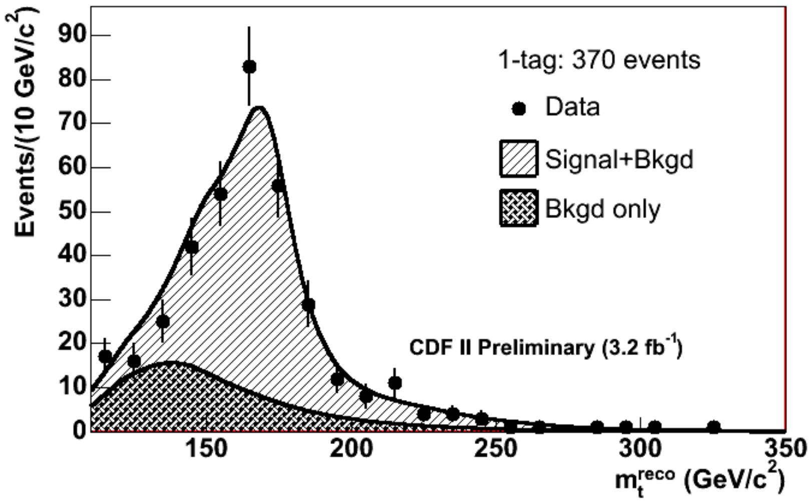


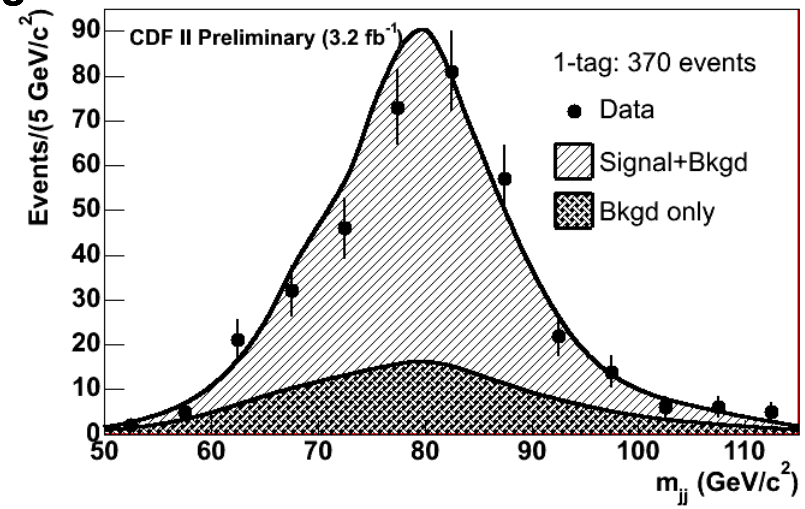
$$168.0^{+4.8}_{-4.0}\text{(stat.)} \pm 2.9\ \text{GeV}/c^2\text{(syst)}$$

$$= 168.0^{+5.6}_{-5.0}\ \text{GeV}/c^2$$

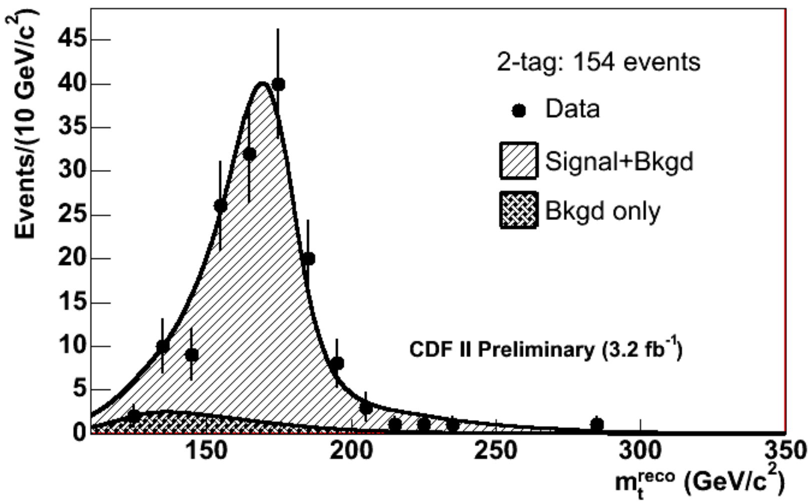
Data distribution (lepton jet channel)

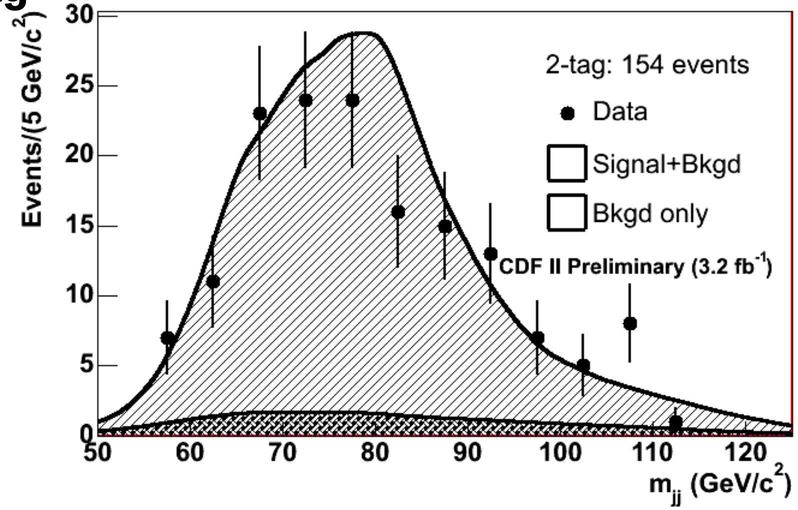
1 tag





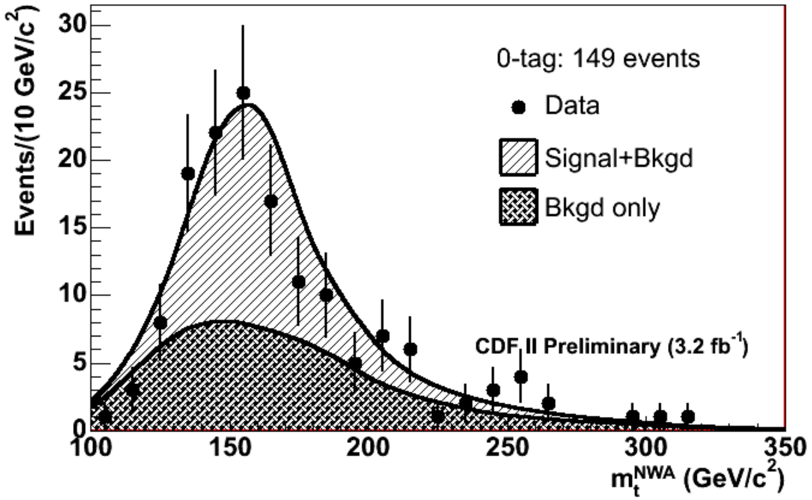
2tag

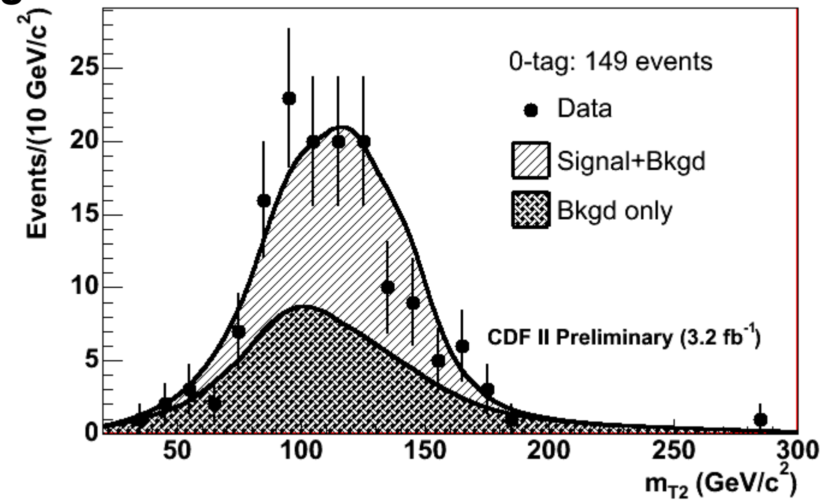




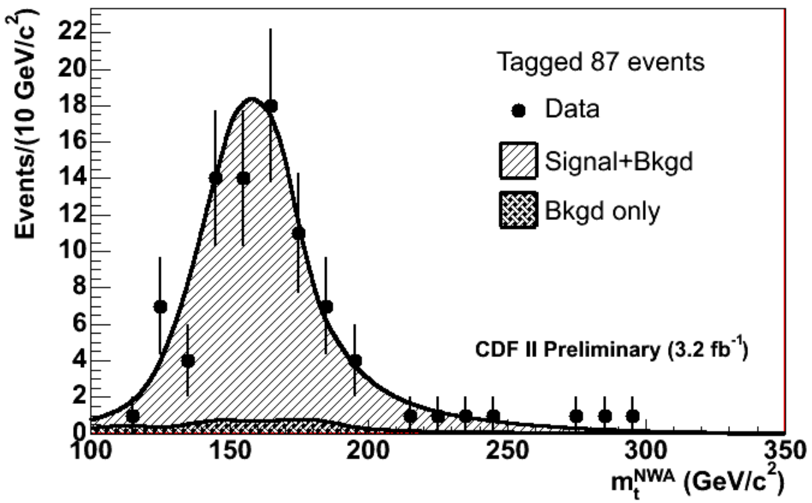
Data distribution (dilepton channel)

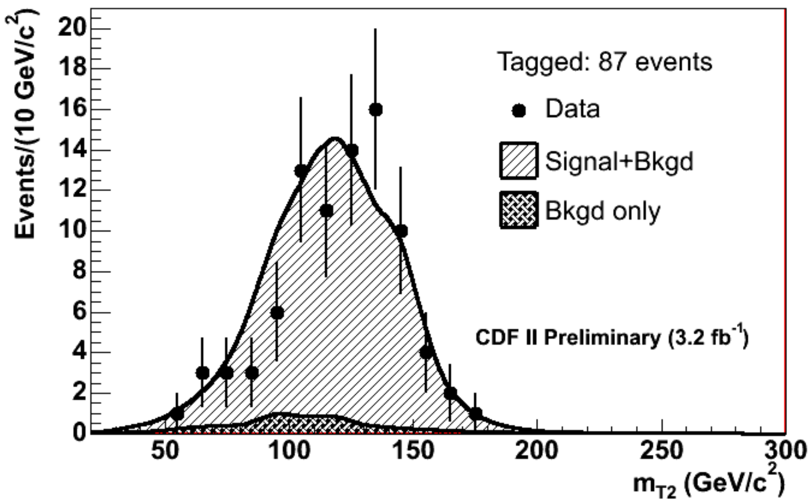
0 tag





Tagged





Conclusion

- We measured top mass from dilepton channel using m_{T2} observable
  - Measurement in LJ+Dilepton

$$171.8^{+1.8}_{-1.9}\ \text{GeV}/c^2$$

  - Measurement in Lepton Jet

$$172.2^{+1.9}_{-1.9}\ \text{GeV}/c^2$$

  - Measurement in Dilepton Channels

$$169.3 \pm 4.2\ \text{GeV}/c^2$$

  - Measurement in Dilepton Channel with m_{T2} only

$$168.0^{+5.6}_{-5.0}\ \text{GeV}/c^2$$

Backup

Future direction

- We are now trying to expand the dimension of KDE

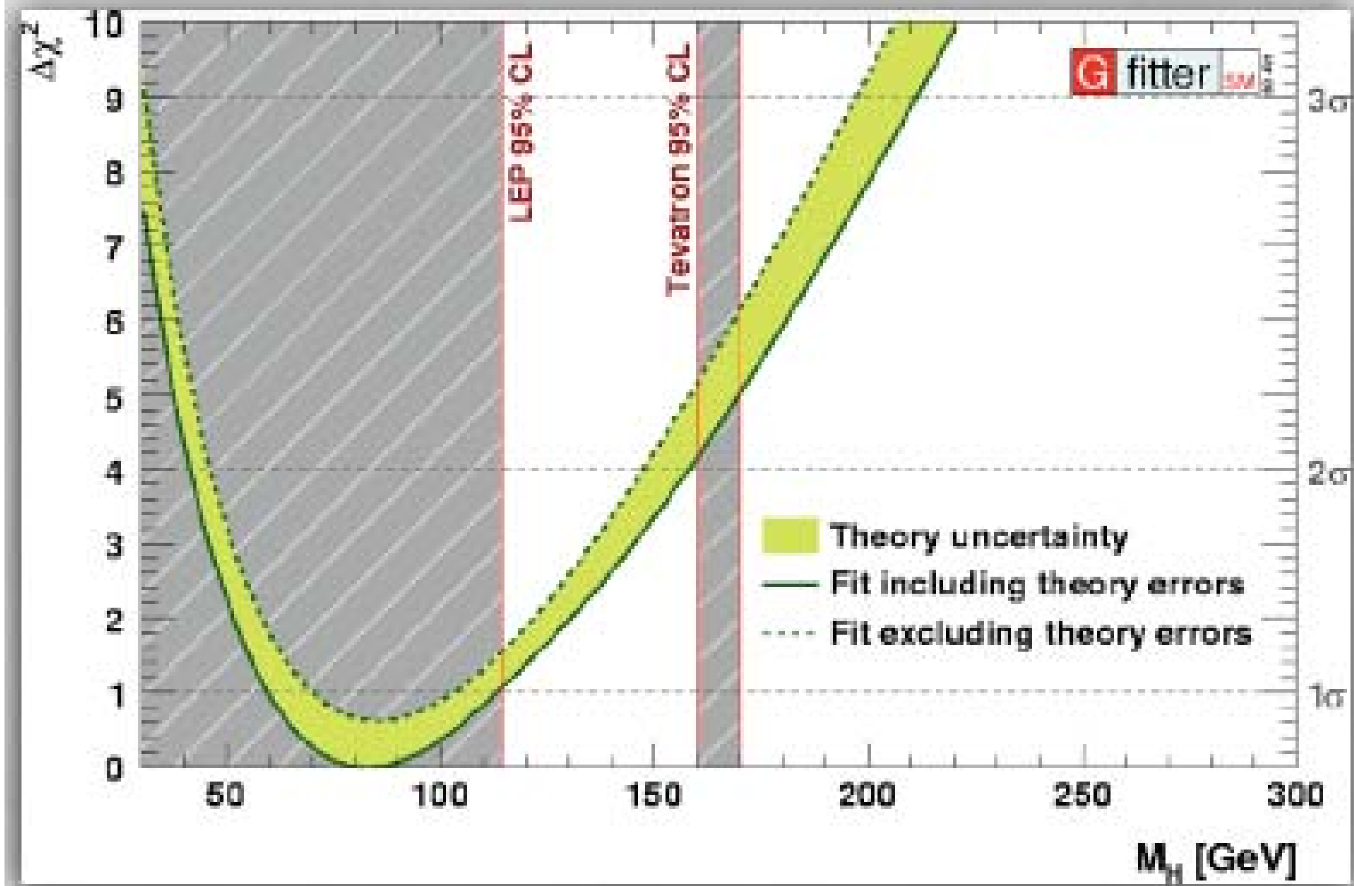
$$P(m_t^{NWA}, m_{T2}; M_{top}) \rightarrow P(m_t^{NWA}, m_{T2}, H_t; M_{top}) \rightarrow P(\vec{x}; M_{top})$$

  - Our resolution can be improved

- Possible application
  - M. Burns, K. Kong, K. T. Matchev, M. Park, Using Subsystem MT2 for complete mass determination in Decay chains with missing energy at Hadron collider - arXiv:0810.5576
  - They introduce the way to determine the mass of whole particle in the decay
  - We can expand our dimension for the likelihood fit then determine all particle mass simultaneously $P(\vec{x}; \vec{M})$

- It is generic method. Many other application

Higgs indirect search



- Central value ±1σ: $M_H = 82.8^{+30}_{-23}$ GeV
- 2σ and 3σ interval: [41, 158] and [28, 211] GeV