Measurements of the $t\bar{t}$ Cross Section and Top Mass at CDF

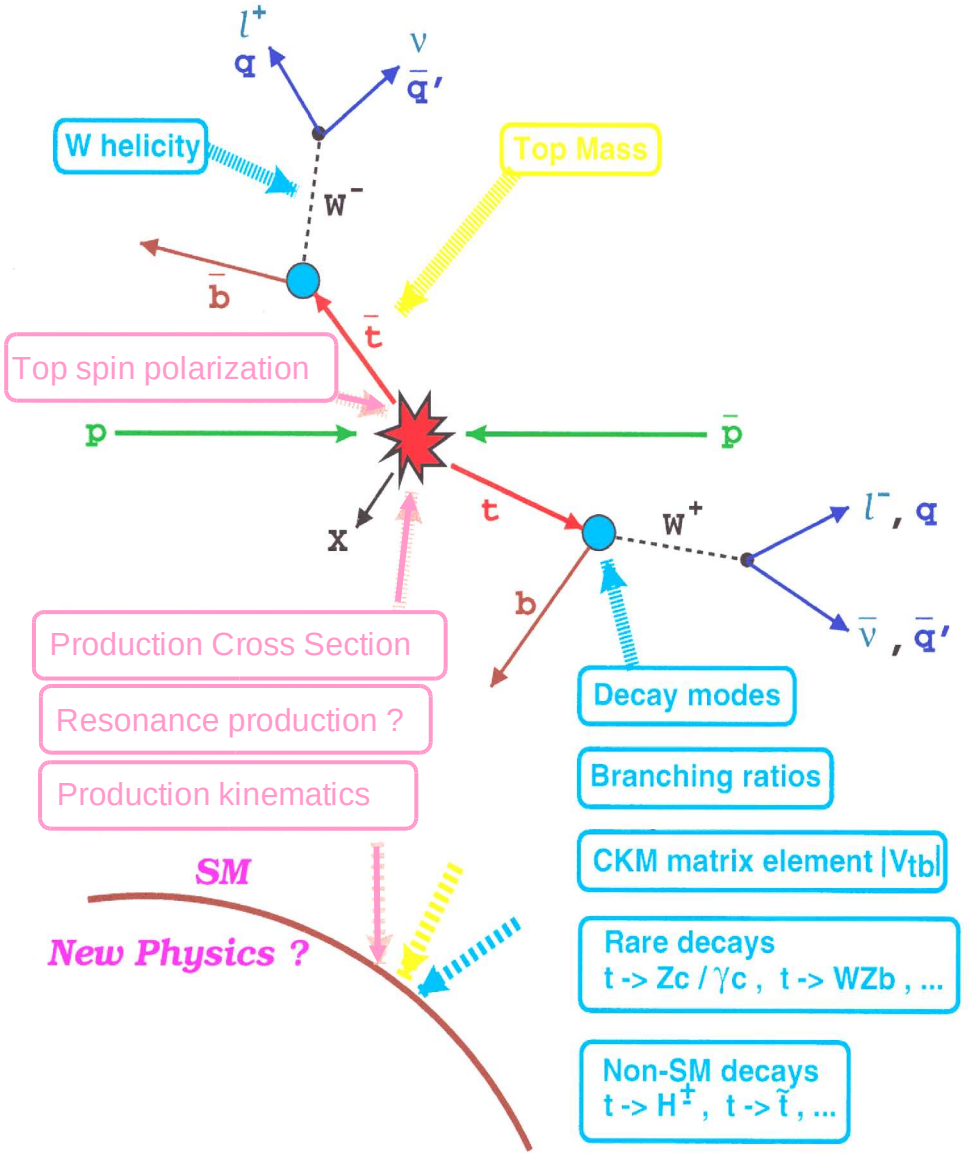
Igor Volobouev

University of Chicago / LBNL

For the CDF Collaboration

Studying the $t\bar{t}$ Cross Section

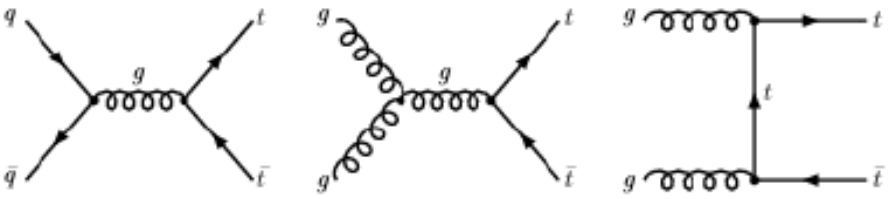
- Main question: is top quark adequately described by the Standard Model?
- Resonant production, non-SM decays?
- Background for other processes (single top, exotics, SUSY, ..)

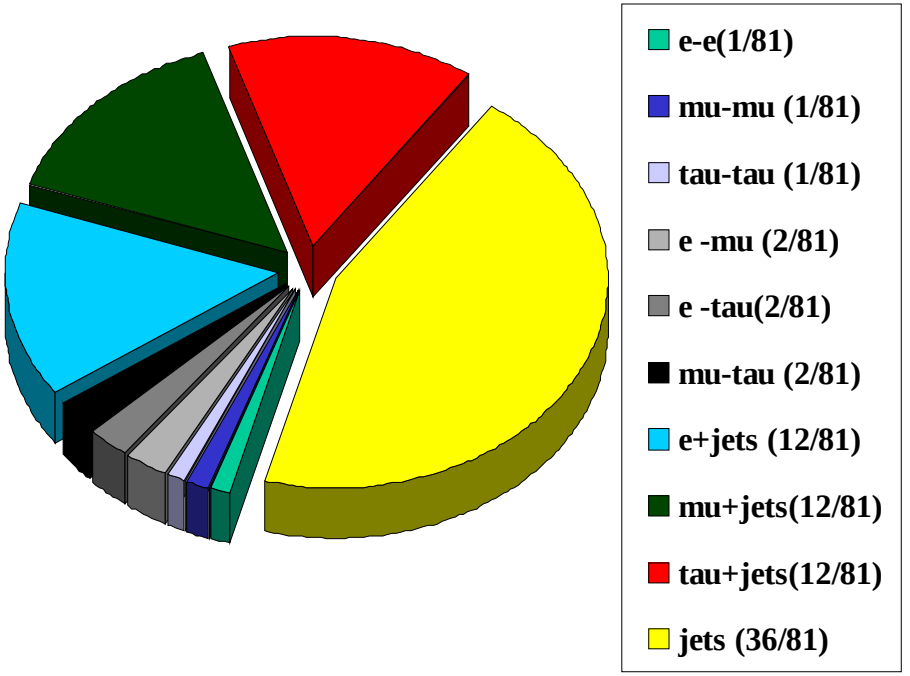


Igor Volobouev

Top Production and Decay Basics

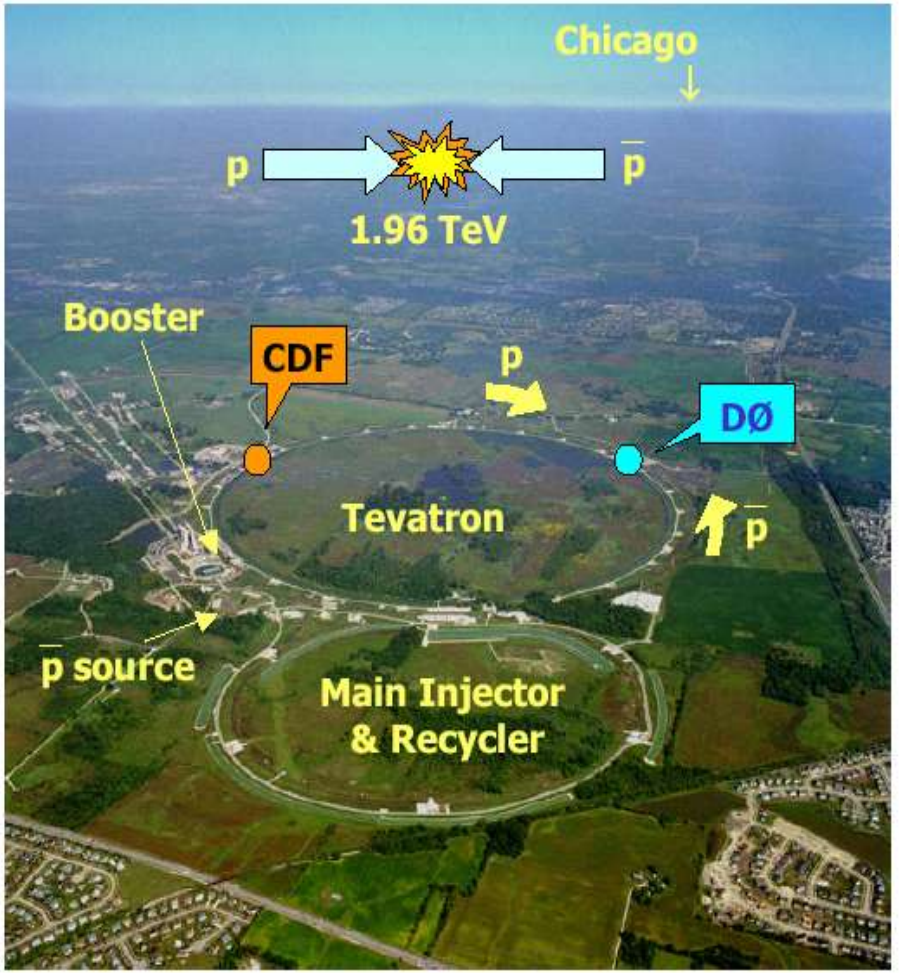
- At Tevatron, top quarks are produced predominantly in pairs (90% $q\bar{q}$ annihilation, 10% gluon fusion at 1.8 TeV)
- $\sigma_{t\bar{t}}$(1.8 TeV) ≈ 5 pb (theory), 6.2 ± 1.2 pb (experiment)
- Single top production cross section is about 40% of $\sigma_{t\bar{t}}$. Single top has not been observed yet.
- Top quark decays into Wb in ~99.9% of the cases (SM). Observed $t\bar{t}$ final states are classified according to subsequent decays of the W.





Tevatron Run 2 Upgrade

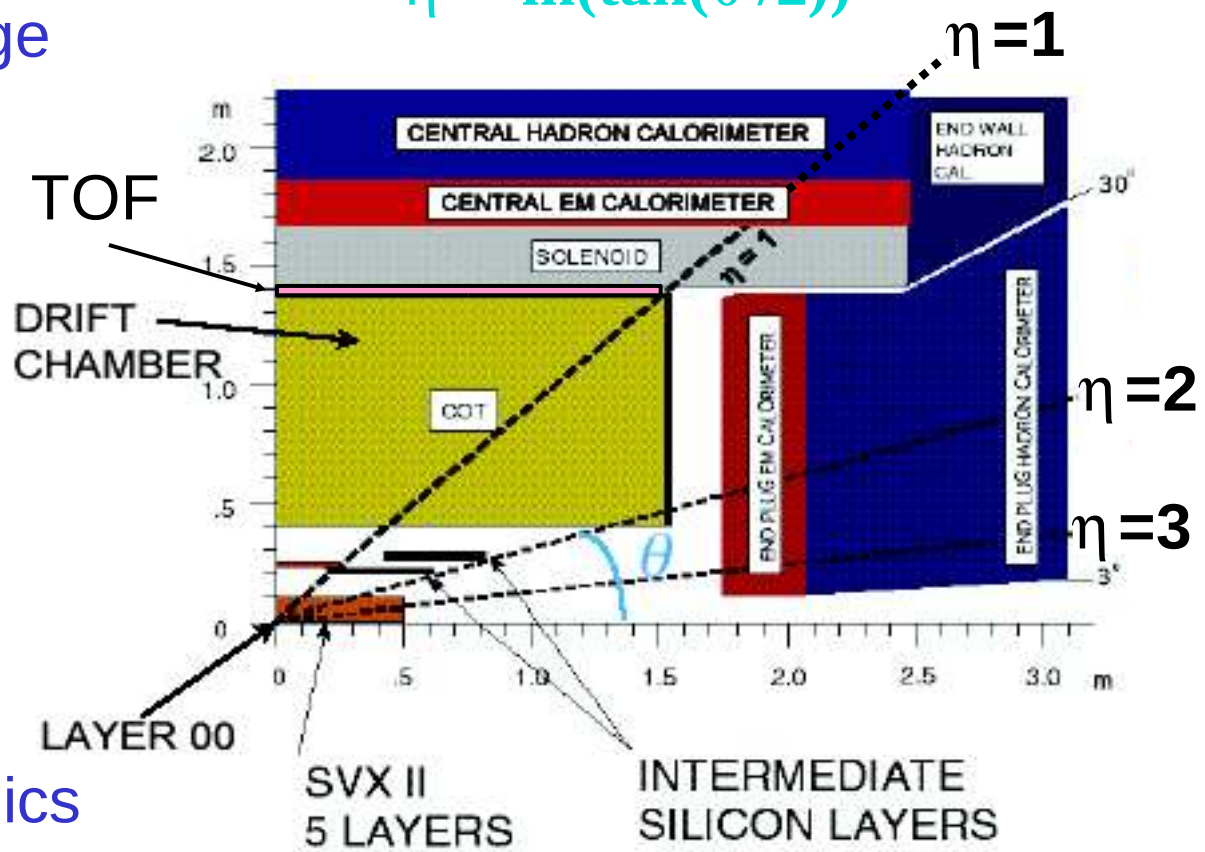
- New Main Injector & Recycler
- Improved antiproton source
- CM energy increased from 1.8 TeV to 1.96 TeV ($t\bar{t}$ cross section up by ~35%)
- 36x36 bunches, 396 ns between bunch crossing (was 6x6 with 3.5 μ s in Run 1)
- Increased luminosity. Goals by the end of FY09:
  - 4.4 fb^{-1} "base"
  - 8.5 fb^{-1} "design"



CDF Upgrade

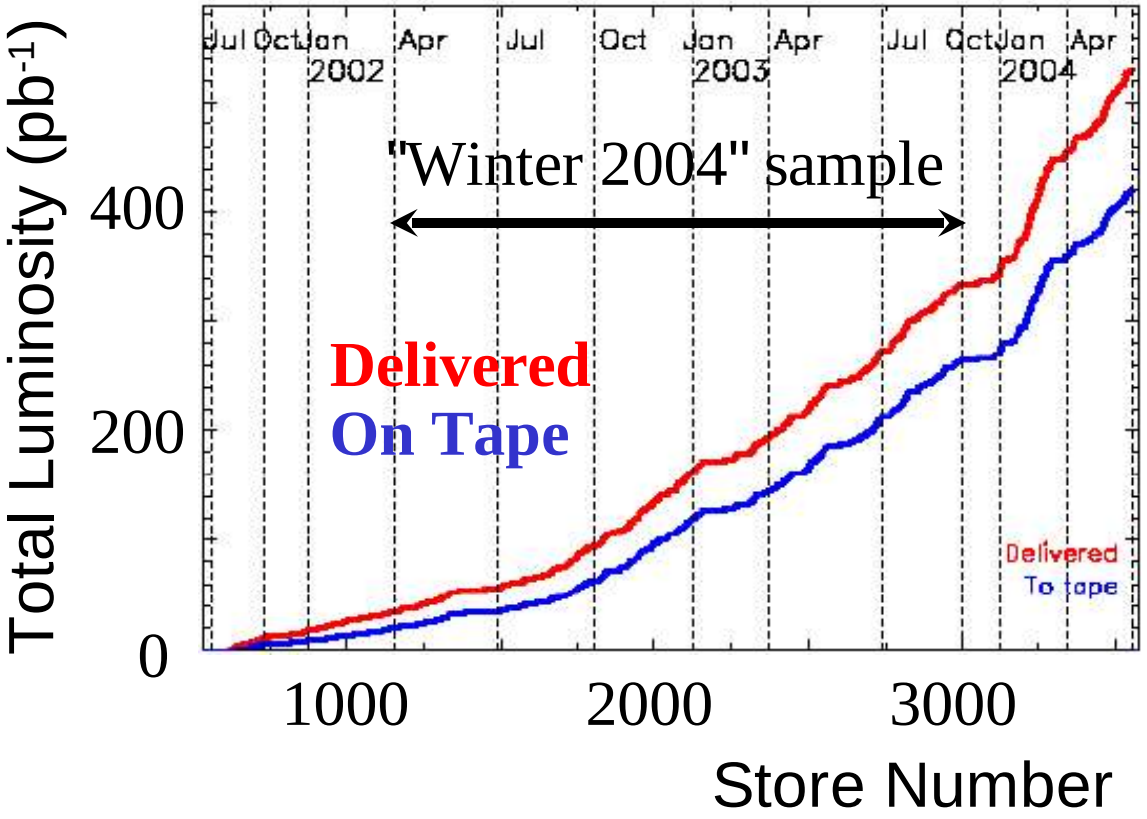
- Improved Si coverage
  - $|\eta| < 2$
  - up to 8 layers
- New central tracker
  - 96 layers
- Time of Flight
- Expanded muon system
- Forward calorimeter
- Trigger and electronics

$\eta = -\ln(\tan(\theta/2))$



Run 2 Data Sample

- Total current sample on tape: ~420 pb^{-1}
- "Winter 2004" analysis sample: 160-200 pb^{-1}
- 8-12 pb^{-1}/week ~85% efficiency

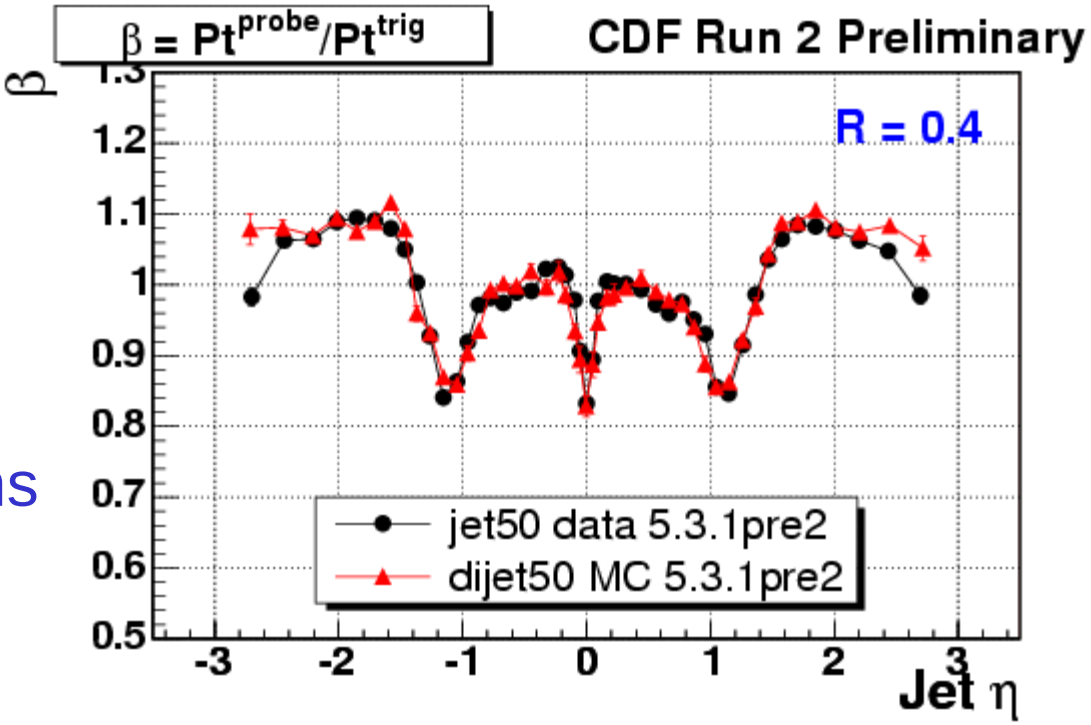


Top Reconstruction

- $t\bar{t}$ events have been successfully reconstructed in all channels (dilepton, lepton+jets, all hadronic)
- Main signatures
  - High p_T leptons and/or jets
  - Missing energy due to escaping neutrinos
  - Two b jets in the final state
  - Production near threshold implies spherical topology
- Lepton+jets channel is the best for initial top mass and cross section measurements
  - Lepton in the final state reduces the QCD background
  - Manageable jet combinatorics, especially with one or two b tags
  - 5 kinematic constraints (momentum conservation in the transverse plane, two W masses, $M_t = M_{\bar{t}}$), 3 unknowns (neutrino momentum)
  - Although exceptionally clean, the dilepton channel has smaller branching fraction than l+jets by factor of 6. There are 6 unknowns, so full event reconstruction is impossible.

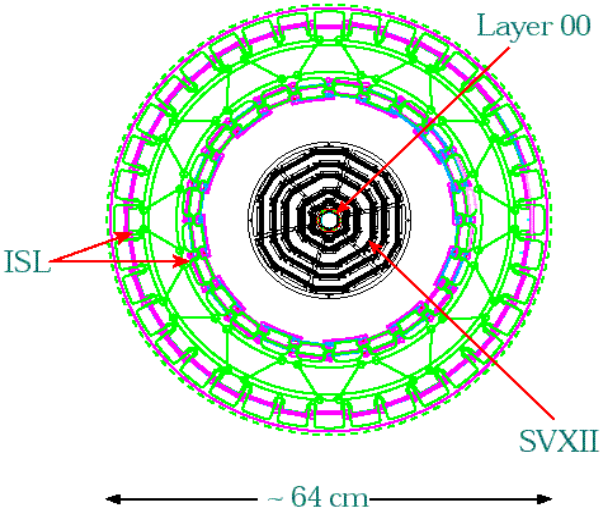
Jet Energy Calibration

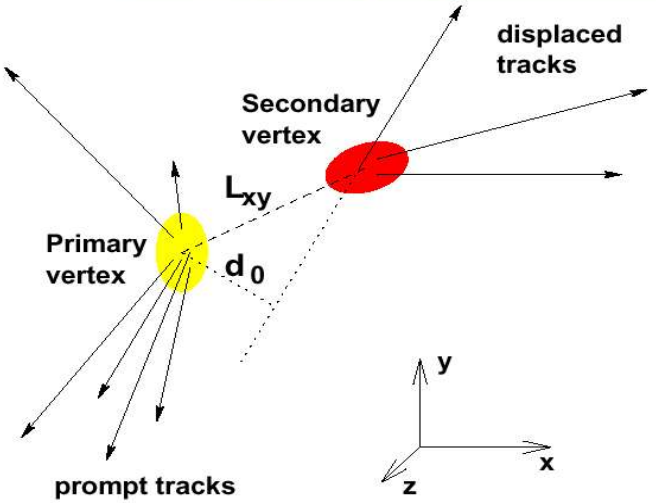
- Extremely important for the top mass measurement
- Electromagnetic calorimeter is calibrated using $Z \rightarrow e^+e^-$
- Hadronic calorimeter is calibrated by monitoring MIP response from muons and referencing to test beam data
- Jet response is studied using photon-jet and dijet balance



B Tagging with Silicon

- At least two well-reconstructed tracks with ≥ 3 silicon hits
- Secondary vertex L_{XY} significance at least $+3\sigma$ (typical $\sigma \sim 150\mu$ m)
- Efficiency to tag a $t\bar{t}$ event is 55%
- $t\bar{t}$ tag fake rate: $\sim 0.5\%$

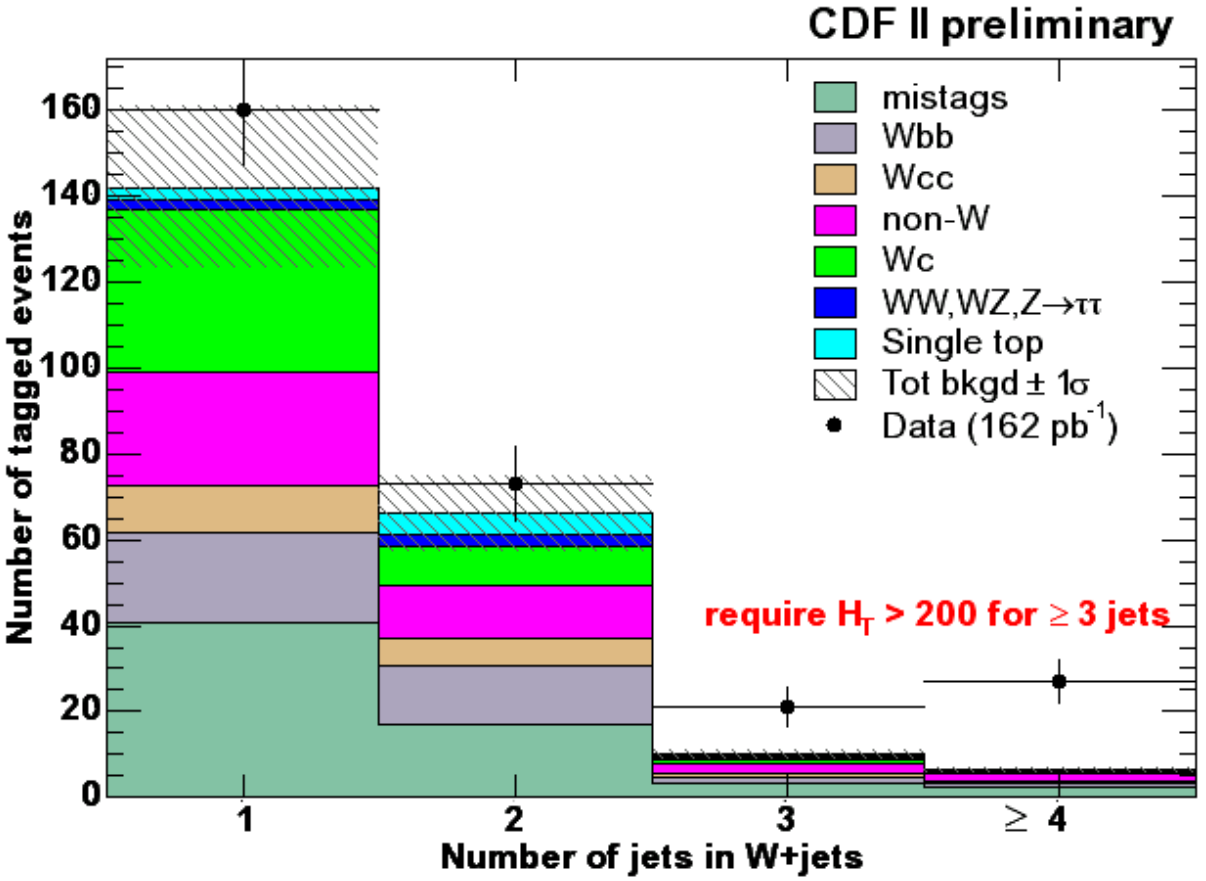




L + Jets $t\bar{t}$ Cross Section Using B Tag

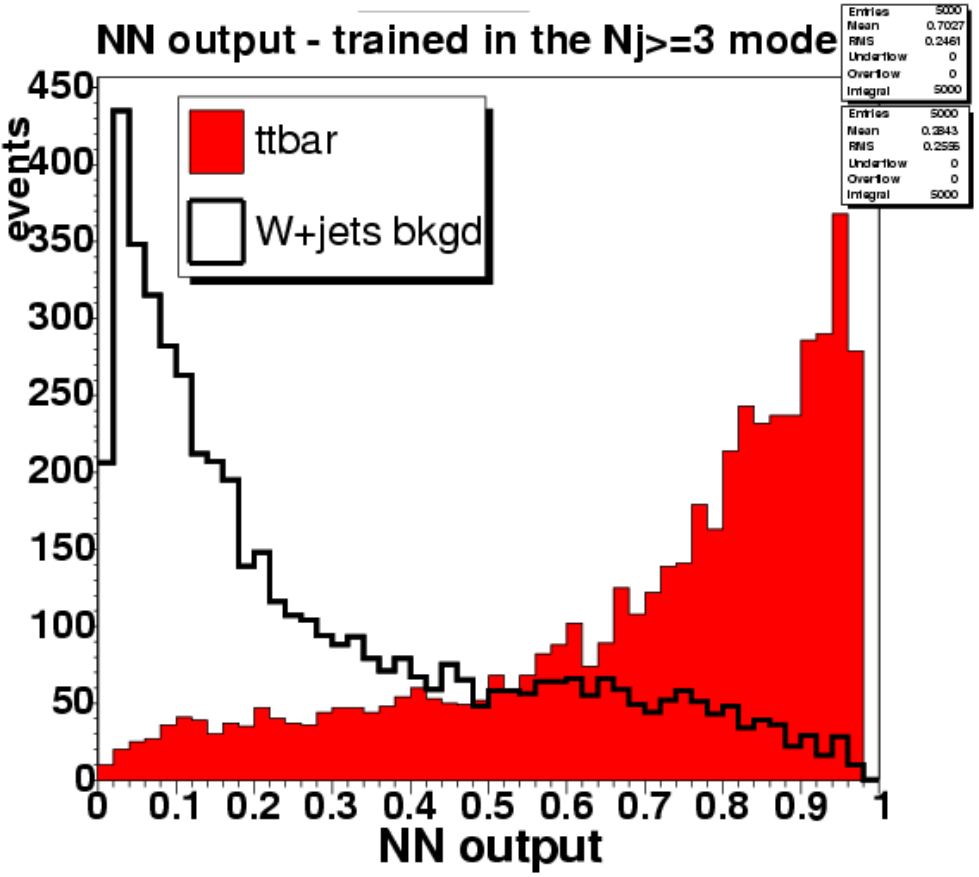
- Counting experiment: $$\sigma_{t\bar{t}} = \frac{N_{obs} - N_{bgd}}{\epsilon_{t\bar{t}} * \int L dt}$$
- Requires careful evaluation of efficiency and expected backgrounds
- Various efficiencies are normally determined from MC and scaled to data. The overall acceptance is about 2.4% for electron + jets and 1.8% for muon + jets.
- Backgrounds:
  - W + HF – from ALPGEN+HERWIG MC
  - W + light jets – from "mistag matrix" based on the negative tag rates in jet data
  - "QCD" – from data ("Isolation vs MET" method)

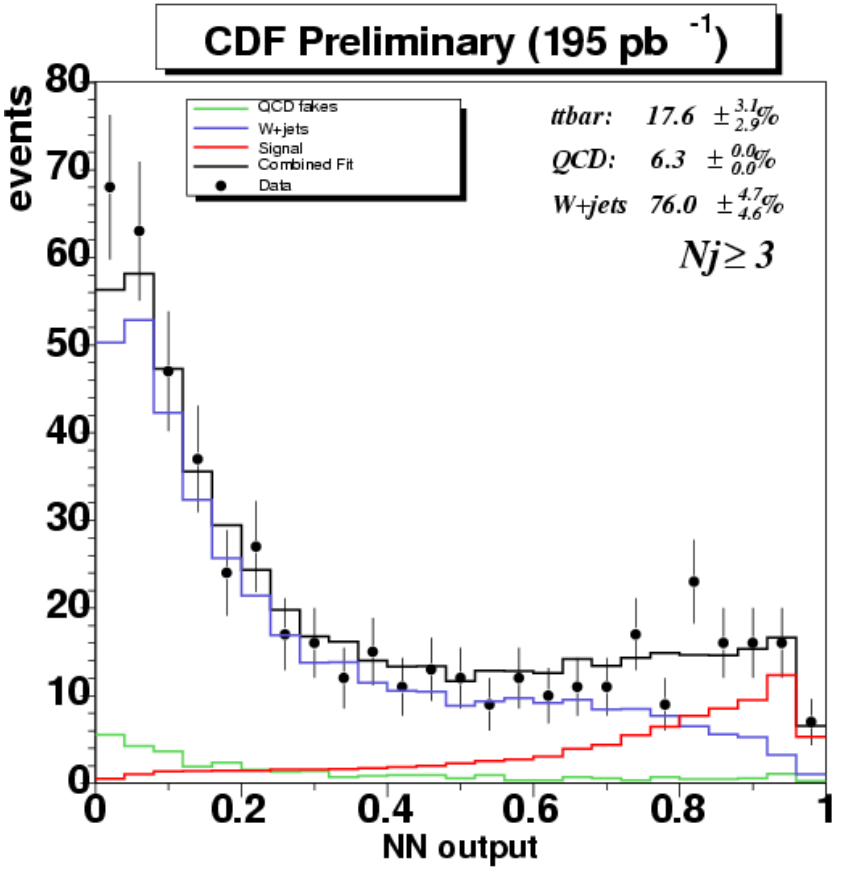
L + Jets B Tag Cross Section Result



- $\sigma_{t\bar{t}} = 5.6^{+1.2+1.0}_{-1.1\ -0.7}$ pb
- On track for publication

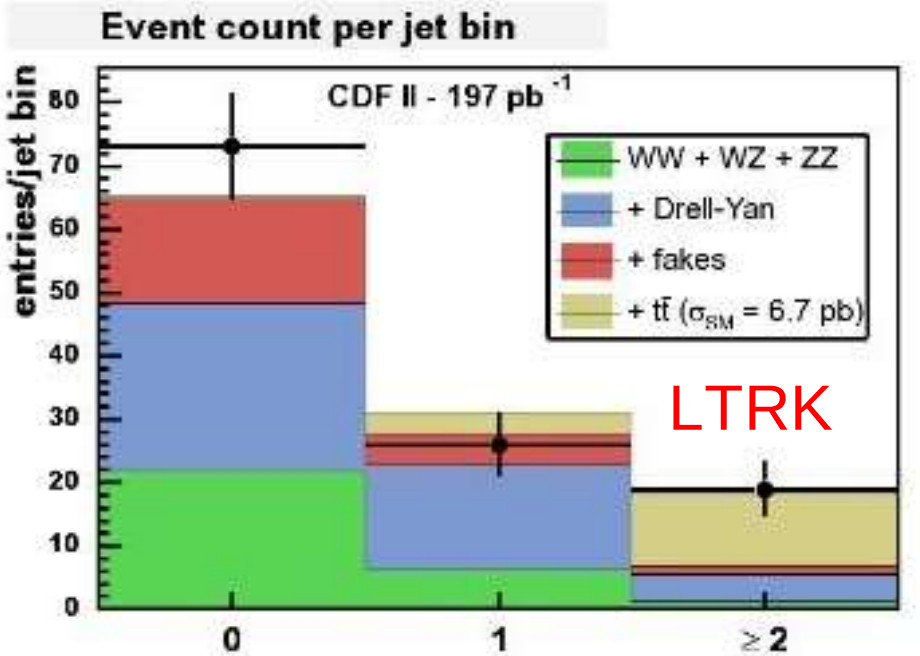
L + Jets $t\bar{t}$ Cross Section Using a NN

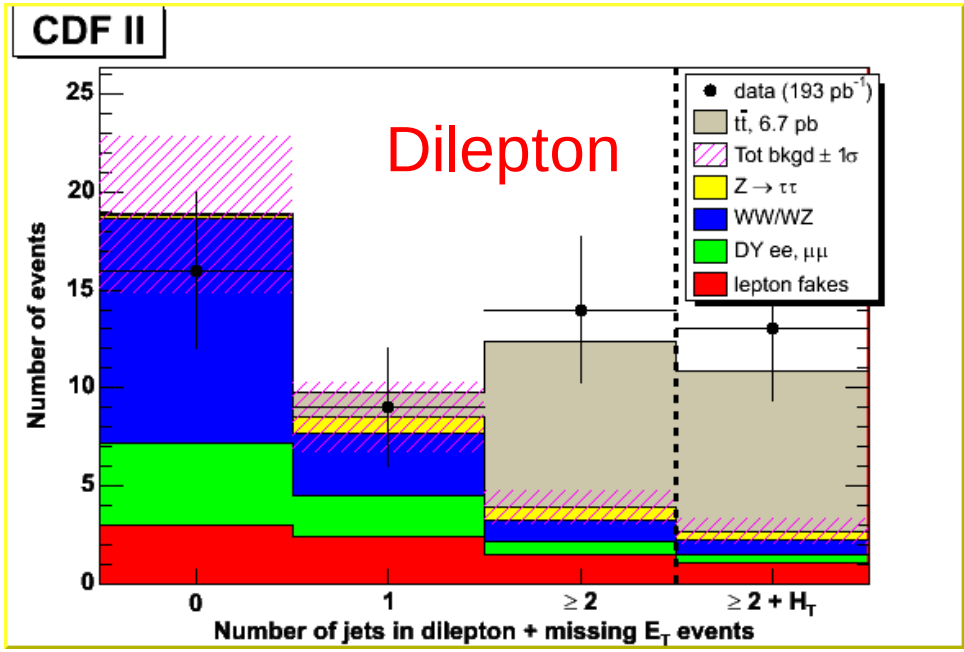




- Kinematic fit uses a neural network with 7 input variables
- No b tag, so the data sample is larger
- $\sigma_{t\bar{t}} = 6.7 \pm 1.1 \pm 1.6$ pb

Dilepton Channel Cross Section

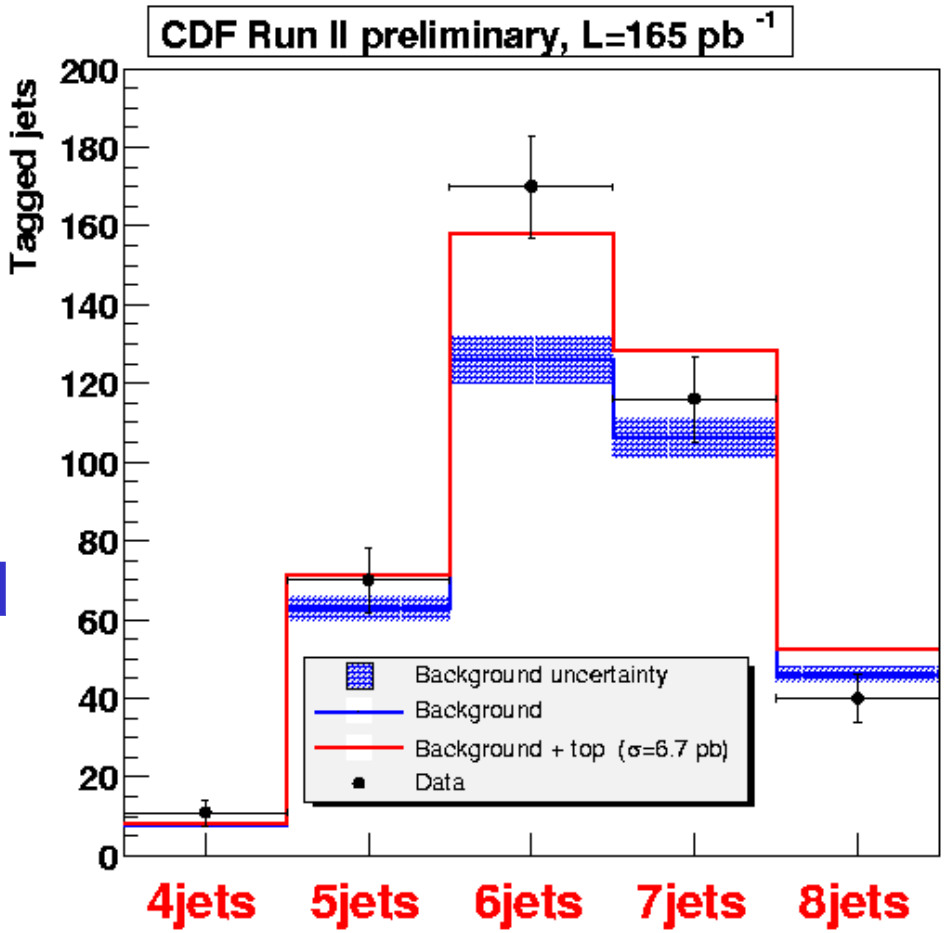




- Two complementary analyses: "dilepton" and "lepton + track"
  - Dilepton: tight second lepton, high purity
  - Lepton + track: second "lepton" is just an isolated track in the drift chamber. Can be, *e.g.*, from 1-prong hadronic τ decays. Acceptance is 0.88% which is about 40% larger than for dilepton.
- Combined result is submitted for publication in PRL (also hep-ex/0404036) – the first Run 2 top paper!

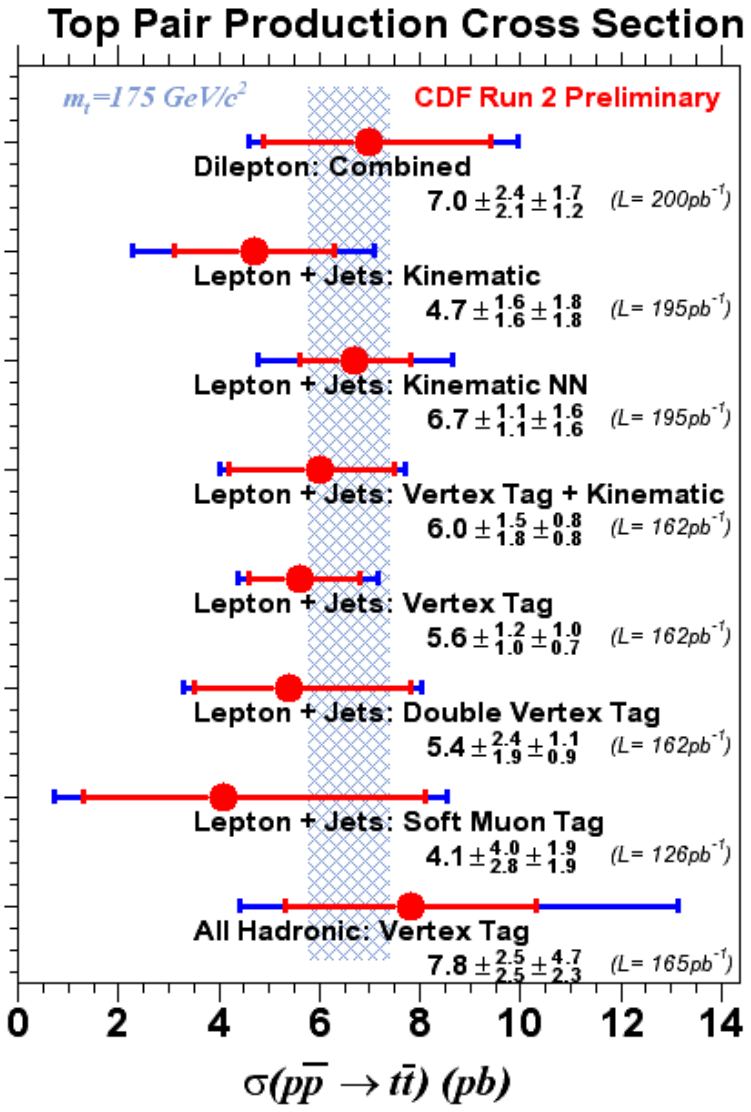
All Hadronic Cross Section

- Multijet trigger
- S/B ratio is improved by cuts on ΣE_T, aplanarity, centrality
- Count b tagged jets. Excess over expected background is attributed to $t\bar{t}$.
- $\sigma_{t\bar{t}} = 7.8 \pm 2.5^{+4.7}_{-2.3}$ pb



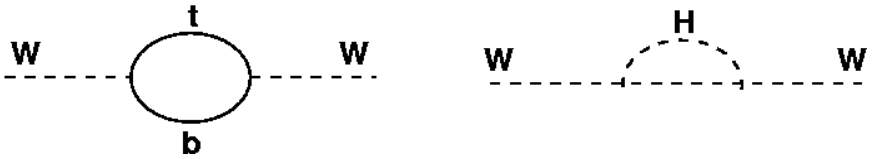
Summary of $t\bar{t}$ Cross Section Results

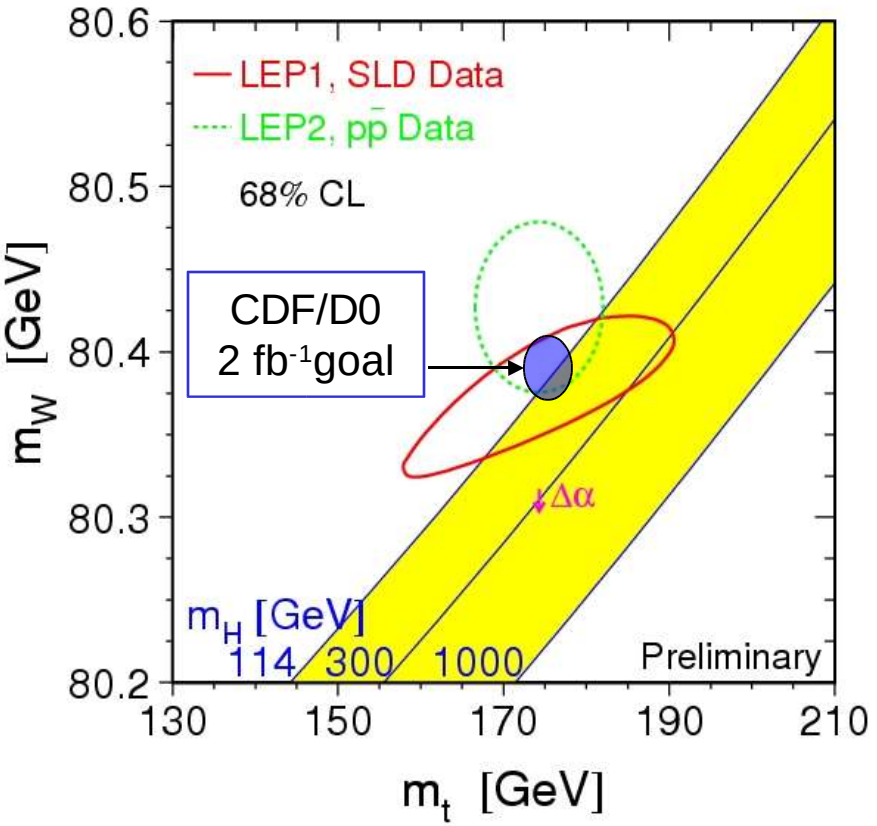
- Many different approaches
- Tested a variety of top production aspects
- All compatible with the SM prediction (mild eμ anomaly seen in Run 1)
- Will combine all channels for better precision



Top Mass in the Standard Model

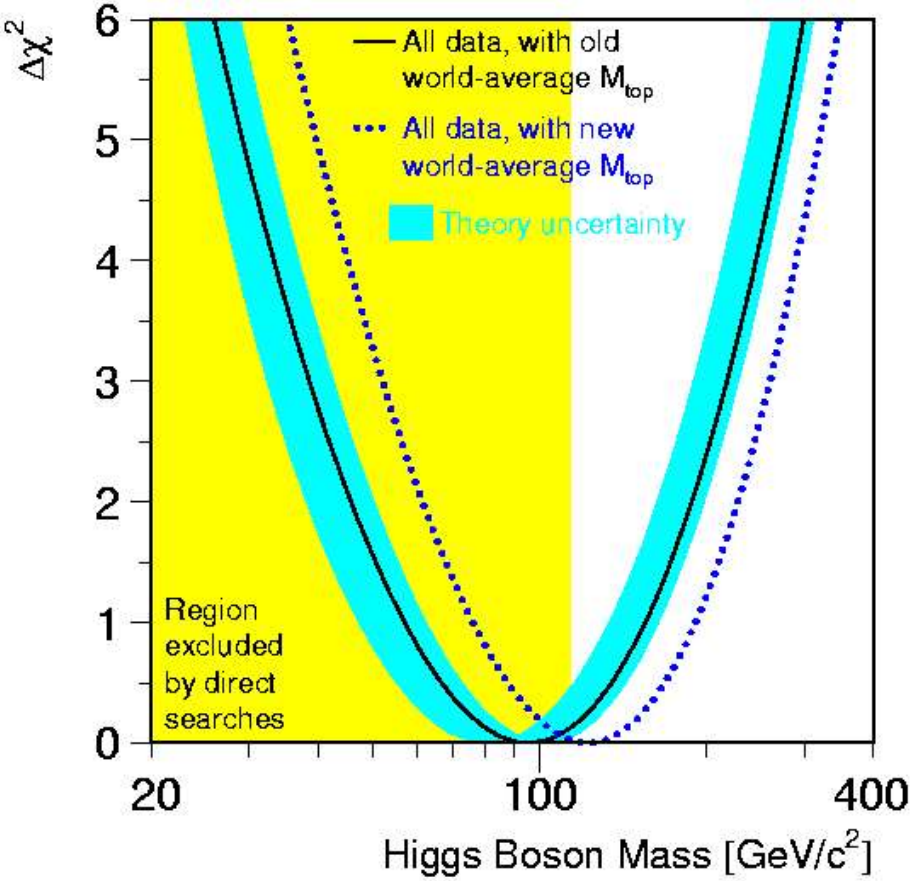
- Fundamental parameter
- Enters into a variety of electroweak calculations at one loop level
- Example: W mass receives quantum corrections proportional to M_t^2 and $\log(M_H)$
- Highly correlated with M_H in the current precision SM fit





Top Mass and Higgs Constraints

- Old Standard Model fit: $M_H = 96^{+60}_{-38}$ GeV/c^2 used $M_t = 174.3 \pm 5.1$ GeV/c^2
- New M_t world average: 178 ± 4.3 GeV/c^2 (hep-ex 0404010). Corresponding $M_H = 113^{+62}_{-42}$ GeV/c^2
- 95% CL upper bound on M_H is now at 237 GeV/c^2



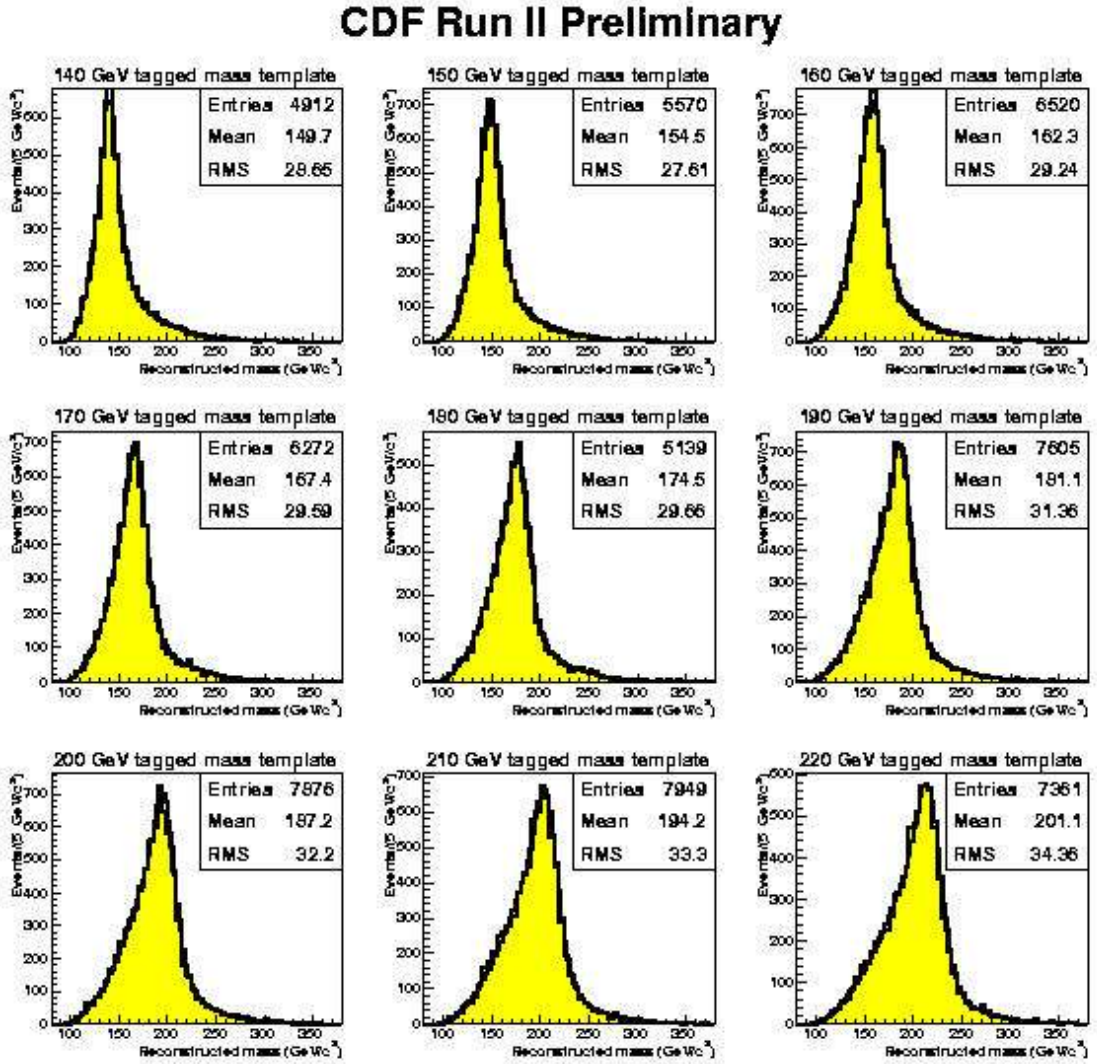
Mass Reconstruction by Run 1 Method

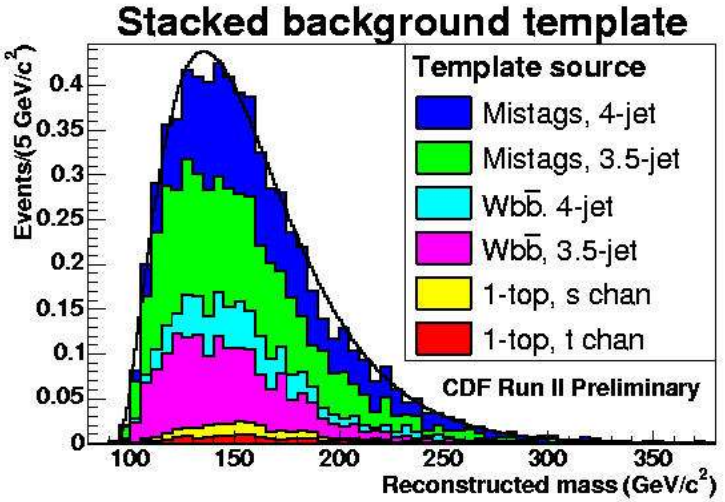
- Simplified χ^2 expression is constructed using transverse momenta of the jets and $t\bar{t}$ recoil, as well as kinematic constraints:

$$\chi^2 = \sum_{\ell,jets} \frac{\left(\hat{P}_T - P_T\right)^2}{\sigma^2_{P_T}} + \sum_{i=x,y} \frac{\left(\hat{U}'_i - U'_i\right)^2}{\sigma^2_{U'_i}} + \frac{(M_{\ell\nu} - M_W)^2}{\sigma^2_{M_W}} + \frac{(M_{jj} - M_W)^2}{\sigma^2_{M_W}} + \frac{(M_{\ell\nu j} - M_t)^2}{\sigma^2_{M_t}} + \frac{(M_{jjj} - M_t)^2}{\sigma^2_{M_t}}.$$

- Solution with the best χ^2 value is found (up to 24 solutions possible due to jet/neutrino combinatorics, less if there is one or more b tags). This solution is used as the reconstructed top mass in the event.
- MC samples generated with different M_t are used to populate mass templates. Background templates are added later.
- Templates are continuously parameterized as a function of M_t.
- Value of M_t is found for which likelihood of the data sample is maximized using parameterized templates as probability density.

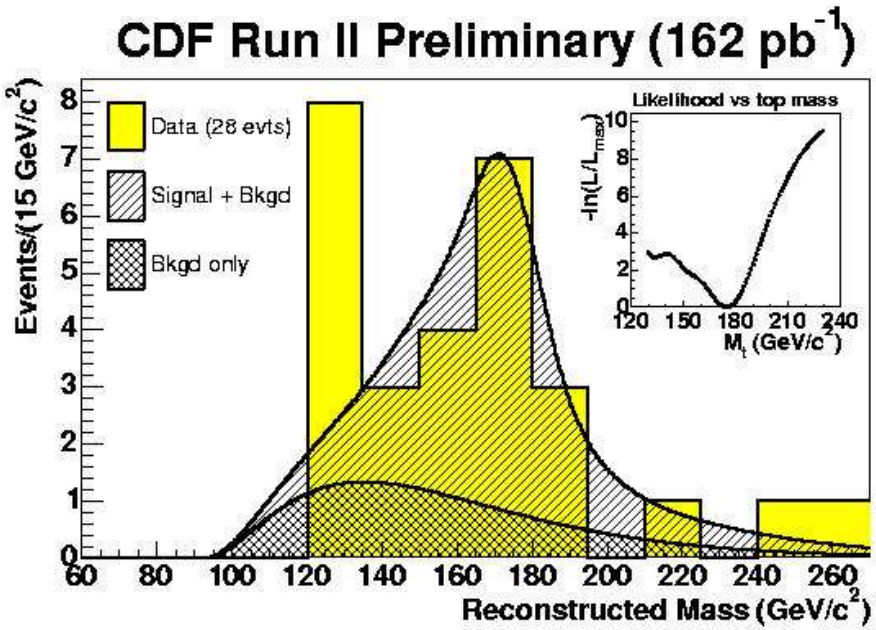
Mass Templates





- Top mass templates are obtained from MC and parameterized by continuous functions

Run 1 Method Result



Systematic Errors

Jet Energy Scale	6.3
Initial State Radiation	0.4
Final State Radiation	0.9
Parton Distribution Functions	0.2
Generators	0.4
Other MC Modeling	0.7
Background Shape	0.8
B-tagging	0.1
Total	6.5

- Background is constrained in the fit to its expected value using the cross section measurement
- From 28 events with $\chi^2 < 9$:

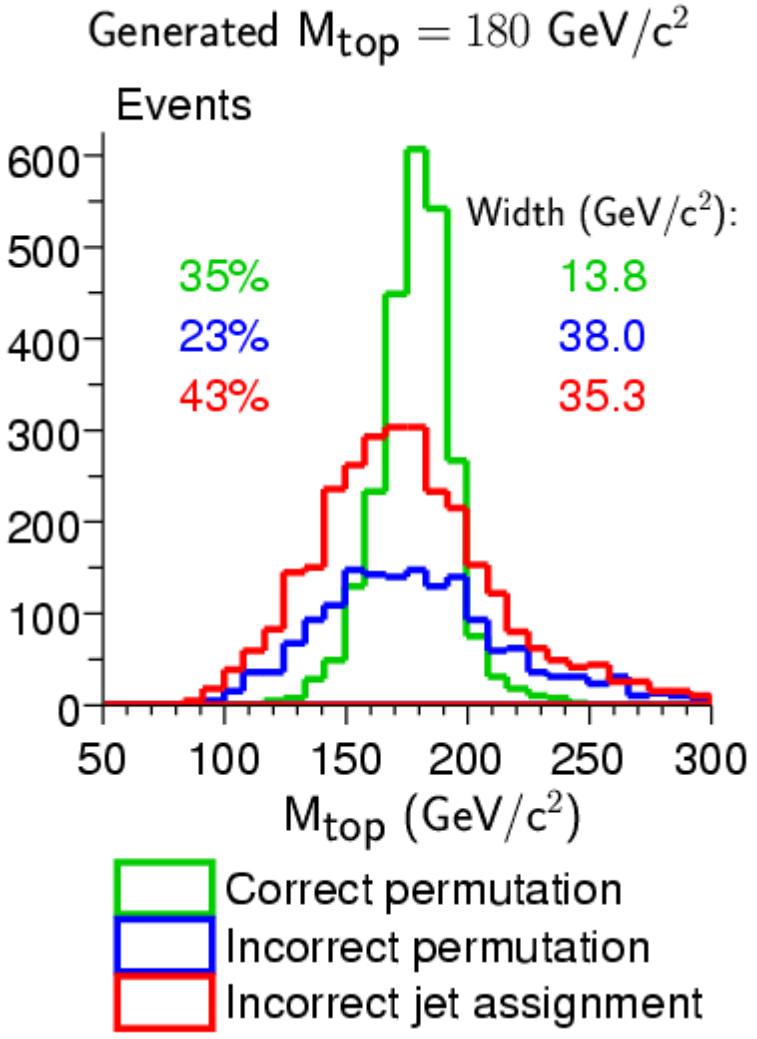
  $M_{top} = 174.9\ ^{+7.1}_{-7.7}$ (stat.) ± 6.5 (syst.) GeV/c^2

Mass Reconstruction – Run 2

- Two new methods have emerged in the lepton+jets channel:
  - Multivariate Template Method (MTM)
  - Dynamic Likelihood Method (DLM)
- D0 has reanalyzed Run 1 data using a matrix element approach. Promising for Run 2.

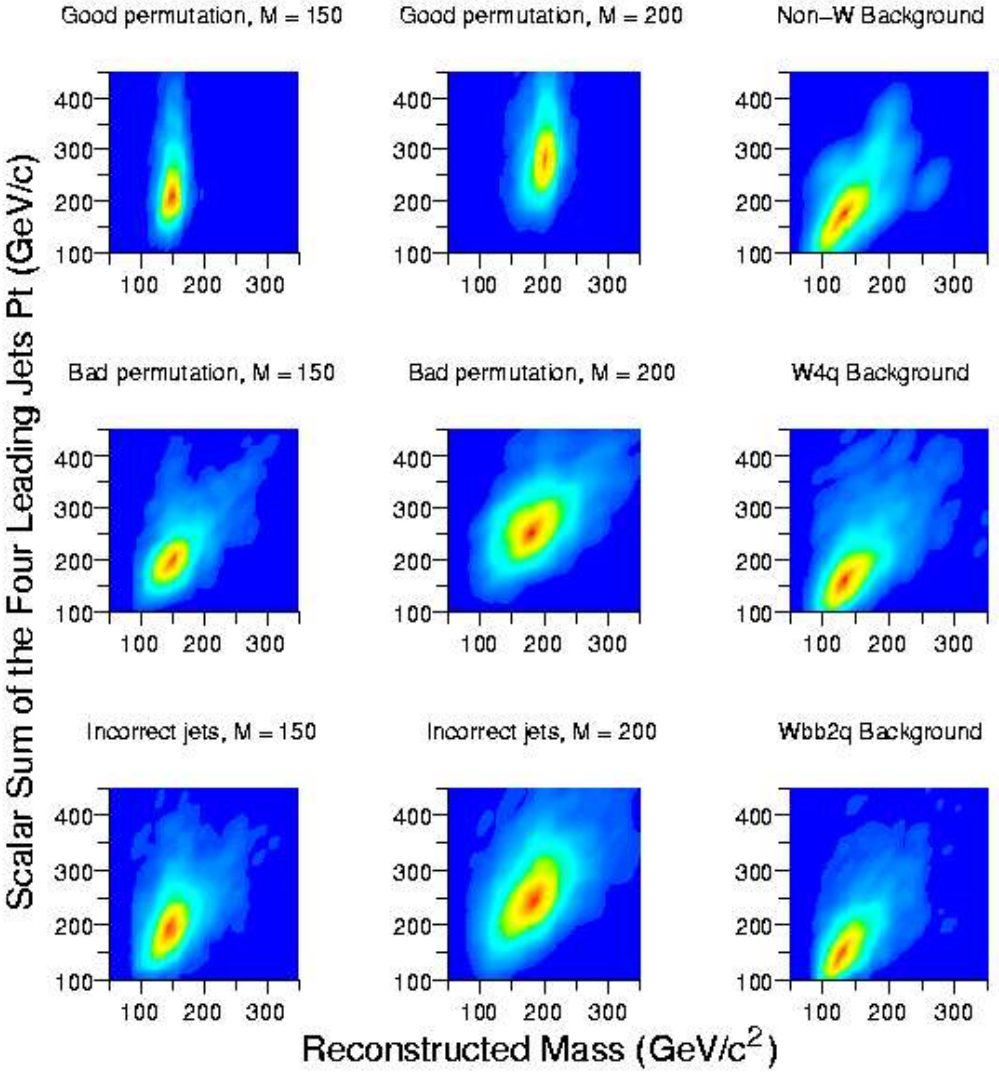
Closer Look at the Mass Templates

- Correct permutation template has much better resolution
- In case of negligible background exact knowledge of the signal subsample would improve the mass resolution by factor of ~1.7
- The probability to pick the correct jet permutation can be estimated using χ^2 differences between permutations. MTM weights templates accordingly.



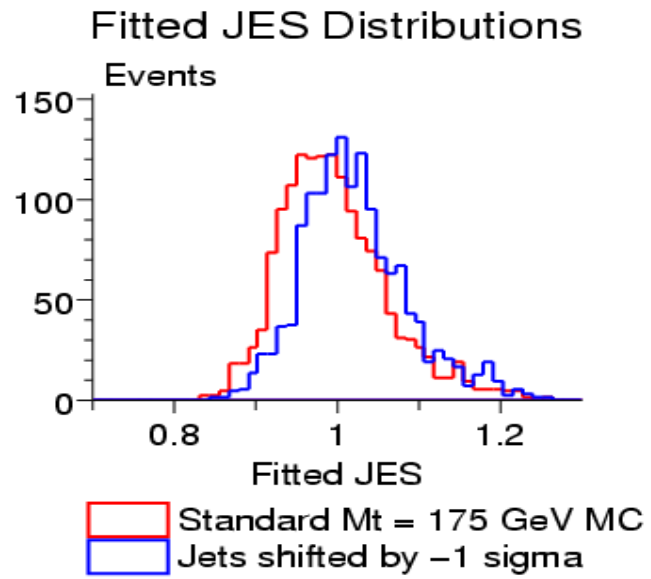
Multivariate Templates

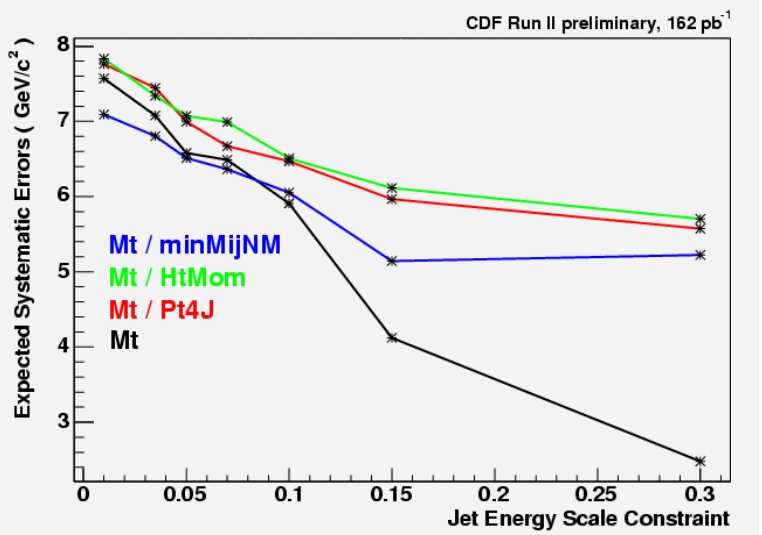
- Templates can use several variables
- Kernel density estimation method is used to create multivariate signal and background templates
- Inverse of a robust covariance matrix is used as a metric. Standard plug-in algorithm determines global bandwidth.



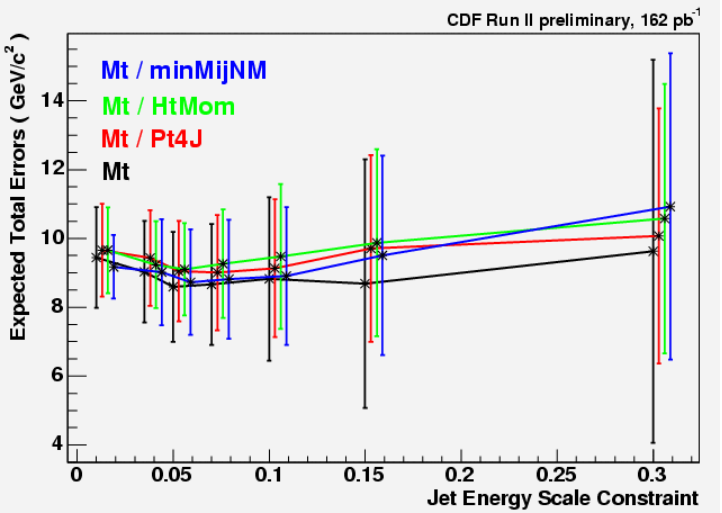
MTM Kinematic Fit

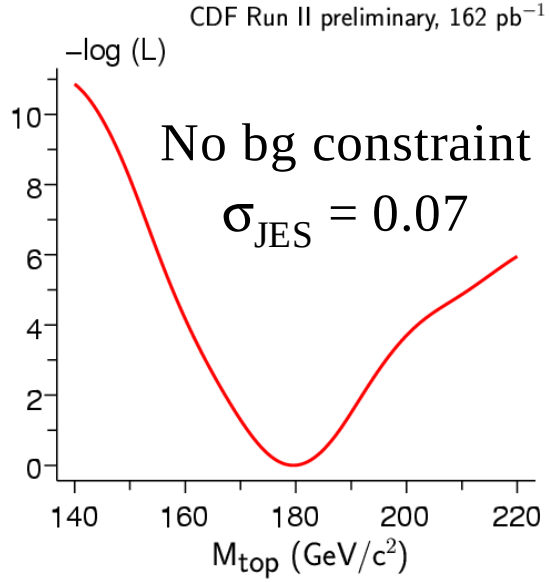
- Jet energy scale factor is included into the W mass kinematic fit with a Gaussian constraint. The constraint is a tunable parameter.
- All jets in the event are multiplied by the jet energy scale value obtained in the W mass fit.
- Fitted jet energy scale is different from one jet permutation to another. For the correct permutation, scale shifts due to the W mass constraint compensate on average systematic shifts.
- Global energy scale fit in the top events is possible but difficult due to background and combinatorics.

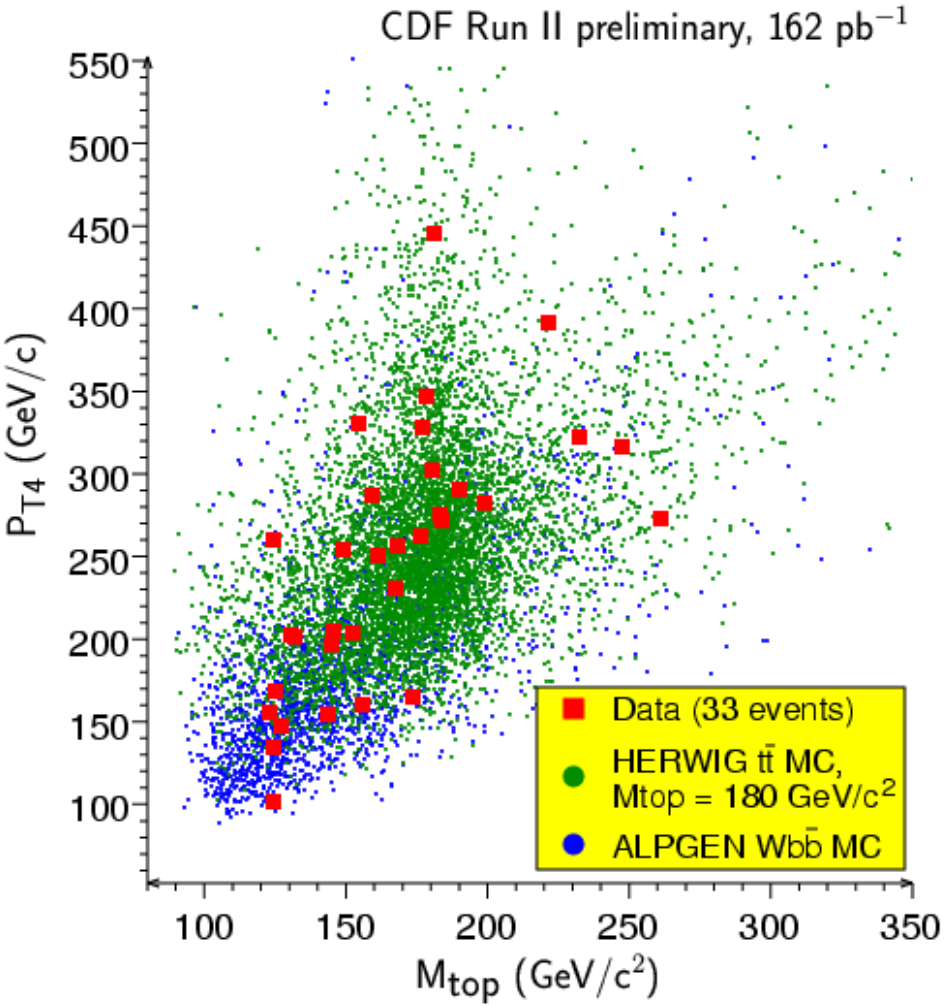




Applying MTM to the Data







$$M_{top} = 179.6^{+6.4}_{-6.3} \pm 6.8 \text{ GeV/c}^2$$

Dynamical Likelihood Method

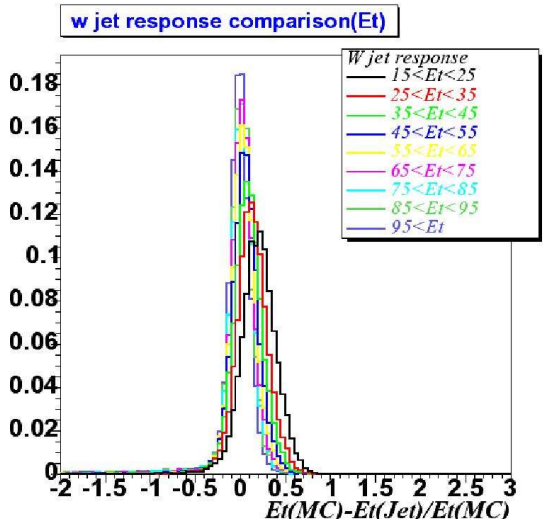
- For each event, likelihood of $\alpha = M_{top}$ is

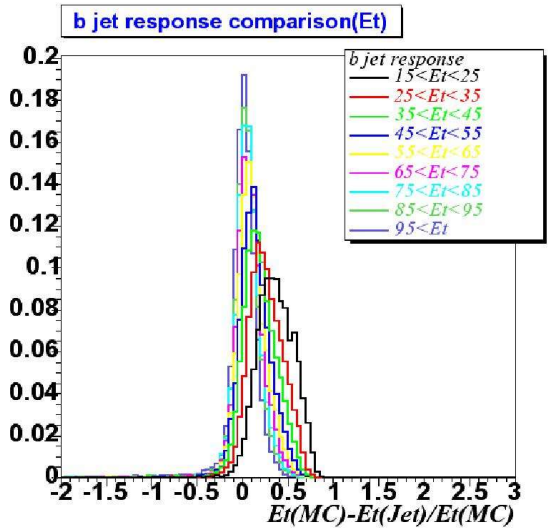
$$L^i(M_{top}) = \int \sum_{combnsol} \sum \frac{2\pi^4}{Flux} |M|^2 F(z_1, z_2) f(p_t) w(\mathbf{x}, \mathbf{y}_t; \alpha) d\mathbf{x}$$

- M – production and decay matrix element (function of $\mathbf{x}$ and α)
- F – parton distribution function
- f – probability density of the $t\bar{t}$ system p_T
- $\mathbf{x}$ – parton momenta (nuisance parameters)
- $\mathbf{y}$ – observed momenta
- w – probability that $\mathbf{x}$ was generated in a $t\bar{t}$ MC when $\mathbf{y}$ was reconstructed. w is a good prior for $\mathbf{x}$.

Calorimeter Transfer Functions

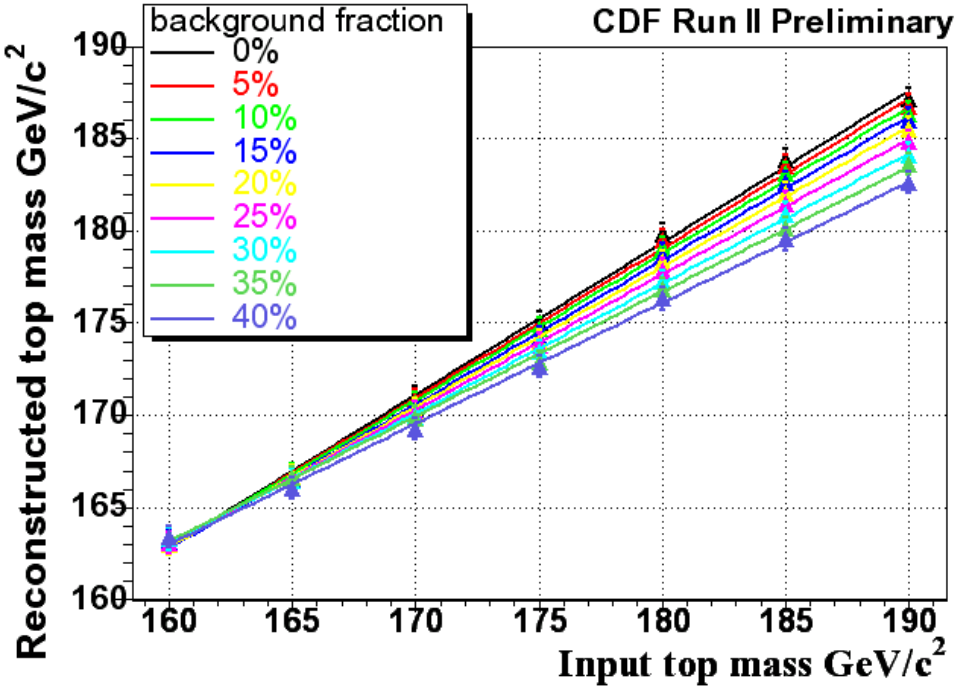
- Obtained from MC
- Expressed as functions of (E(parton) – E(jet))/E(parton)
- 9 bins in E_T, 3 in η
- Checked using different generators (HERWIG and PYTHIA) and by reconstructing the W mass



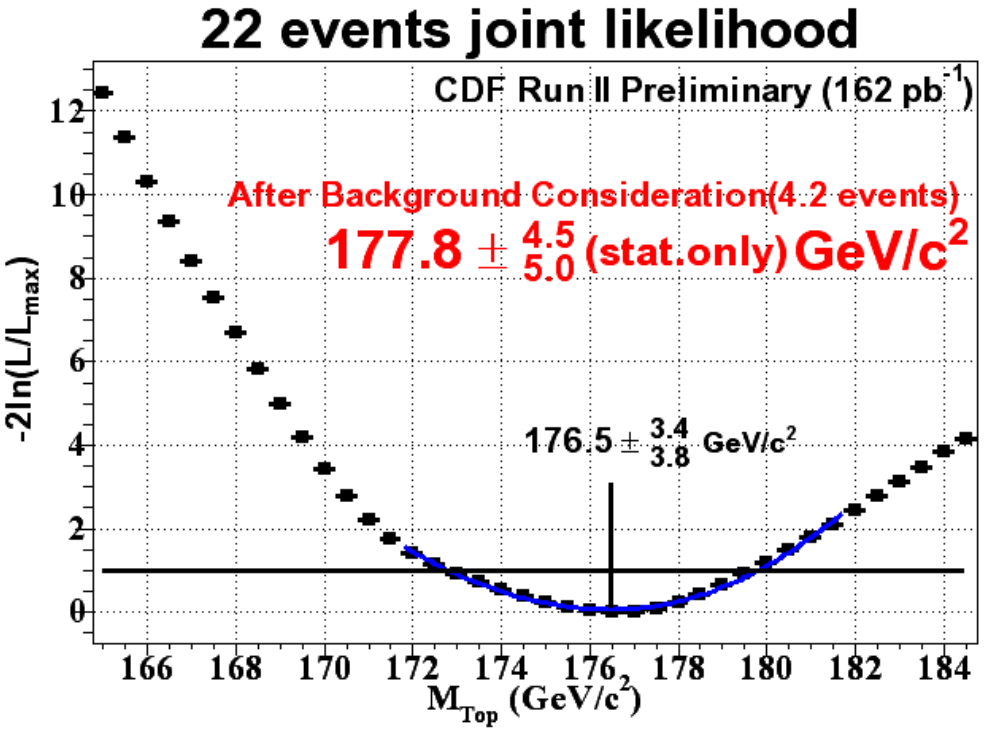


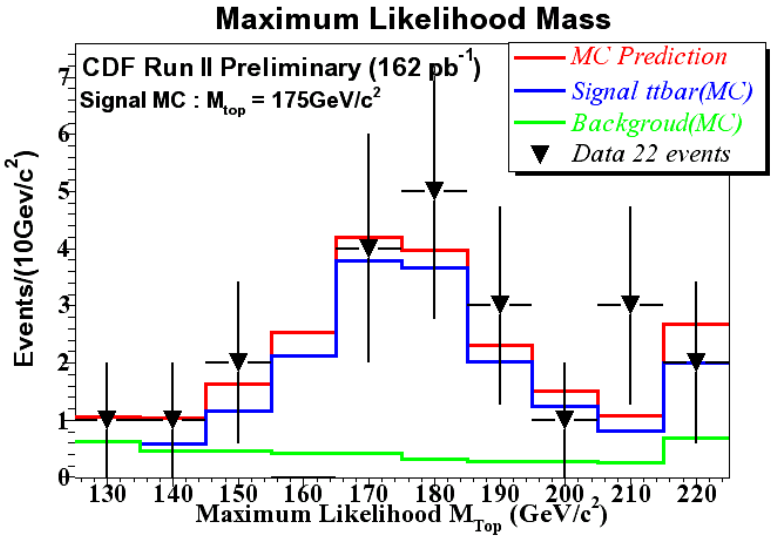
Background Treatment in DLM

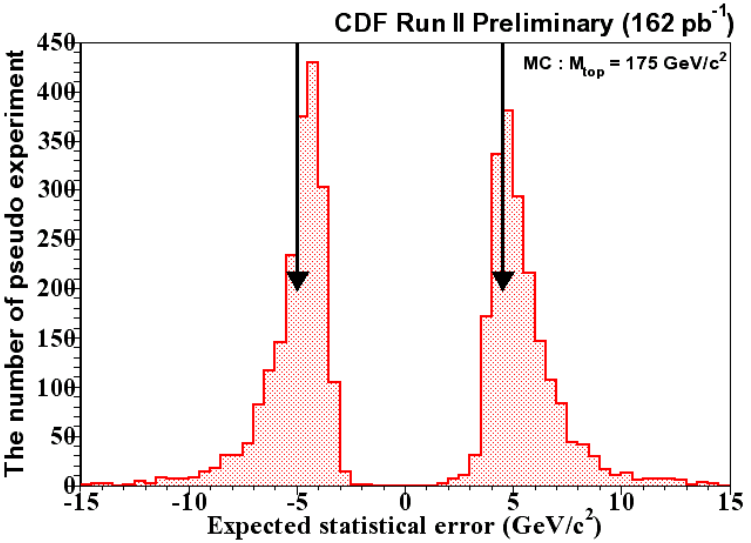
- Background fraction is minimized by choosing events with exactly 4 jets
- Maximum likelihood mass is remapped using expected background from the cross section measurement
- After mapping the estimator is unbiased and pull distributions are unit Gaussians



DLM Data Likelihood





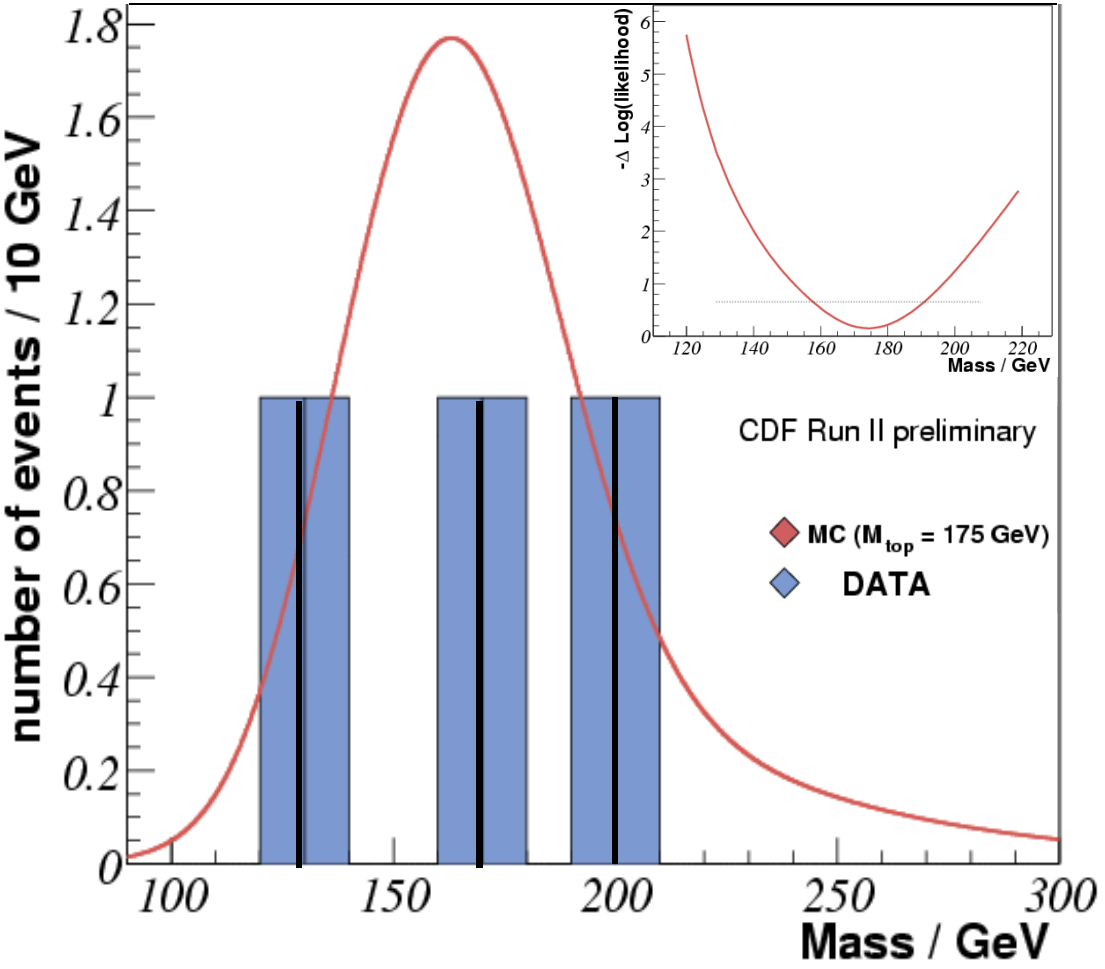


$$M_{top} = 177.8^{+4.5}_{-5.0} \pm 6.2 \text{ GeV/c}^2$$

Best CDF Run 2 top mass measurement

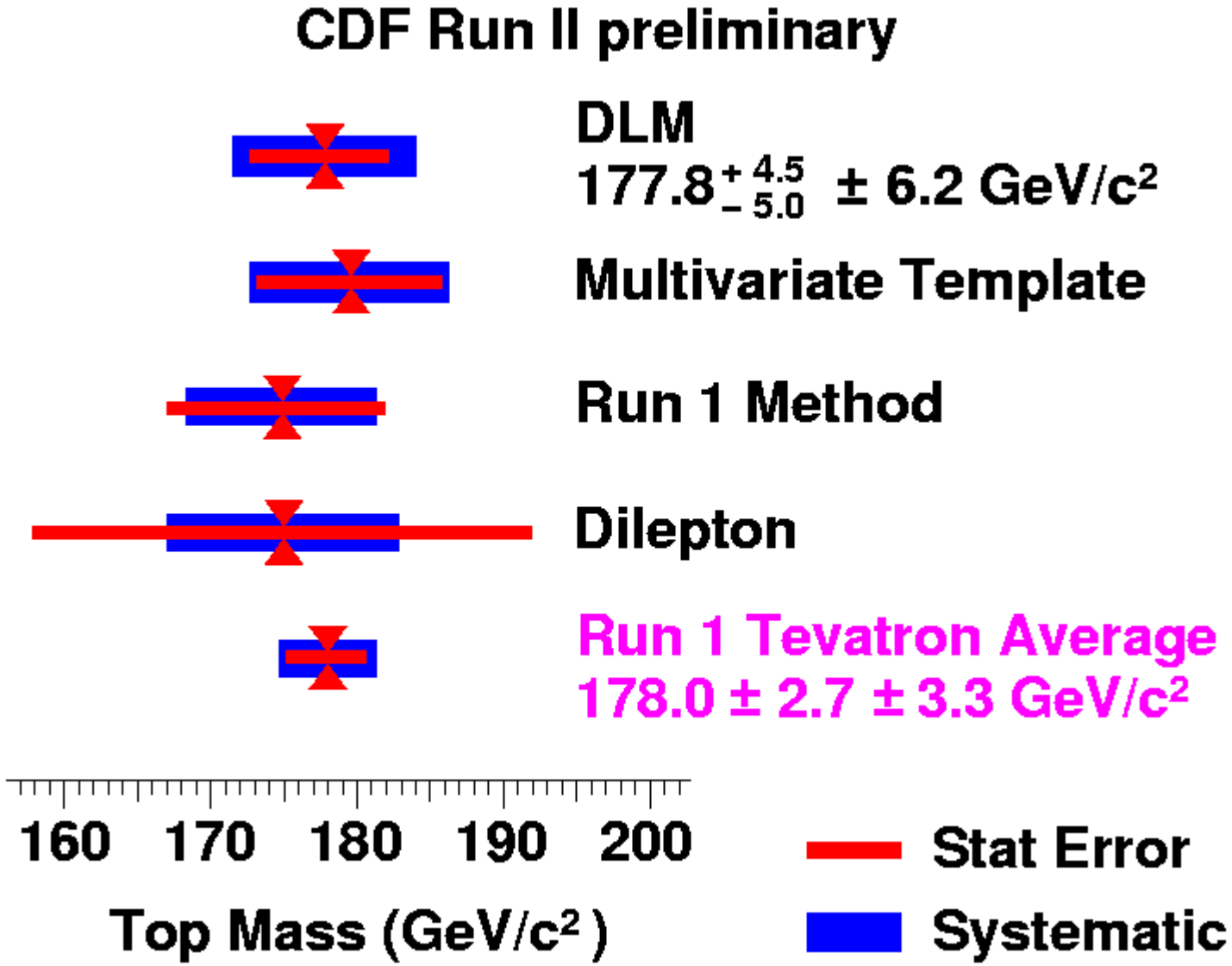
Top Mass in the Dilepton Channel

- Based on 126 pb^{-1}
- Templates are built by sampling the z momentum of the $t\bar{t}$ system to get the most probable mass for each event. Use permutation/neutrino solution with the smallest $t\bar{t}$ mass.
- Background is ~0.5 events.



$M_{top} = 175 \pm 17 \pm 8$ GeV/c^2

Summary of Top Mass Results



Future Directions for M_t at CDF

- Expect a significant improvement in the systematic error in the next pass of top mass measurements (aim for publications in Fall 2004)
- Balance statistical and systematic uncertainties. Optimize expected total error.
- Add soft lepton tagger
- Include l+jets events without b tags. Verify background model.
- Fully explore dilepton and all hadronic channels
- Improve jet energy resolution by taking into account jet fragmentation and differences between quark and gluon jets
- Switch to a better clustering algorithm

Conclusions

- CDF Run 2 top mass and cross section measurements have established the basis for a variety of top studies
- Up to now all measurements are consistent with the Standard Model top with $M_t \sim 180$ GeV/c^2.
- Precision Tevatron top physics is coming. New methods are being developed, and improvements in calibration and simulation are on the way.

Backup Slides Follow

High P_T Lepton Triggers

- Electron trigger
  - Requires central EM cluster with E_T > 18 GeV and E_{HAD}/E_{EM} < 0.125
  - A good quality track with P_T > 9 GeV/c must be matched to the cluster
  - About 96% efficient for " triggerable" electrons with E_T > 20 GeV in the W ? eν sample. Inefficiency is dominated by tracking.
- Muon trigger
  - Requires a match between a good quality track and a muon chamber stub
  - About 95% efficient for " triggerable" muons in the Z ?$\mu^+\mu^-$ sample

Electron Identification

- Good quality track with $p_T > 10$ GeV/c
- Track $|z0| < 60$ cm
- CEM transverse energy $E_T > 20$ GeV
- $E_T/p_T < 2.0$ when $p_T < 50$ GeV
- Cluster $E_{HAD}/E_{EM} < 0.055 + 0.00045 * E$
- Track-to-shower match ~ 3 cm
- Fractional calorimeter energy isolation < 0.1
- Shower profile consistent with electron
- Fiducial to CES
- Conversion veto

Muon Identification

- Good quality track with $p_T > 20$ GeV/c
- Track $|z_0| < 60$ cm
- Cosmic ray veto
- Track impact parameter < 0.02 cm with silicon hits, 0.2 cm without
- $E_{EM} < 2 + \max(0, 0.0115 * (p - 100))$ GeV
- $E_{HAD} < 6 + \max(0, 0.0280 * (p - 100))$ GeV
- Fractional calorimeter energy isolation < 0.1
- Track match to a muon chamber stub: 3, 5, and 6 cm for CMU, CMP, and CMX, respectively

JES Uncertainty Details

CDF Run II Preliminary

Source	Uncertainty (GeV/c^2)	
	$\geq$ 3.5 jets	$\geq$ 4 jets
Relative to Central	3.0	3.2
Central Calorimeter Response	4.6	4.7
Corrections to Hadrons (Absolute Scale)	2.2	2.3
Corrections to Partons (Out-of-Cone)	2.3	2.3
Total	6.3	6.6

MTM Likelihood

$$L(m_t) = \prod_{i=1}^{N} (f_b P_b(m_i, x_i) + (1 - f_b) P_s(m_i, x_i, m_t))$$

$$P_b(m, x) = \sum_{\text{bg types}} a_j B_j(m, x), \qquad \sum_{\text{bg types}} a_j = 1$$

$$P_s(m, x, m_t) = (p_{cp} S_{0,m_t}(m, x) + (1 - p_{cp}) S_{1,m_t}(m, x)) p_{cj} + (1 - p_{cj}) S_{2,m_t}(m, x)$$

N is the number of observed events

m_i is the top mass in the ith event. x_i symbolizes all other template variables.

P_s and P_b are the signal and background densities

f_b is the background fraction, treated as a nuisance parameter

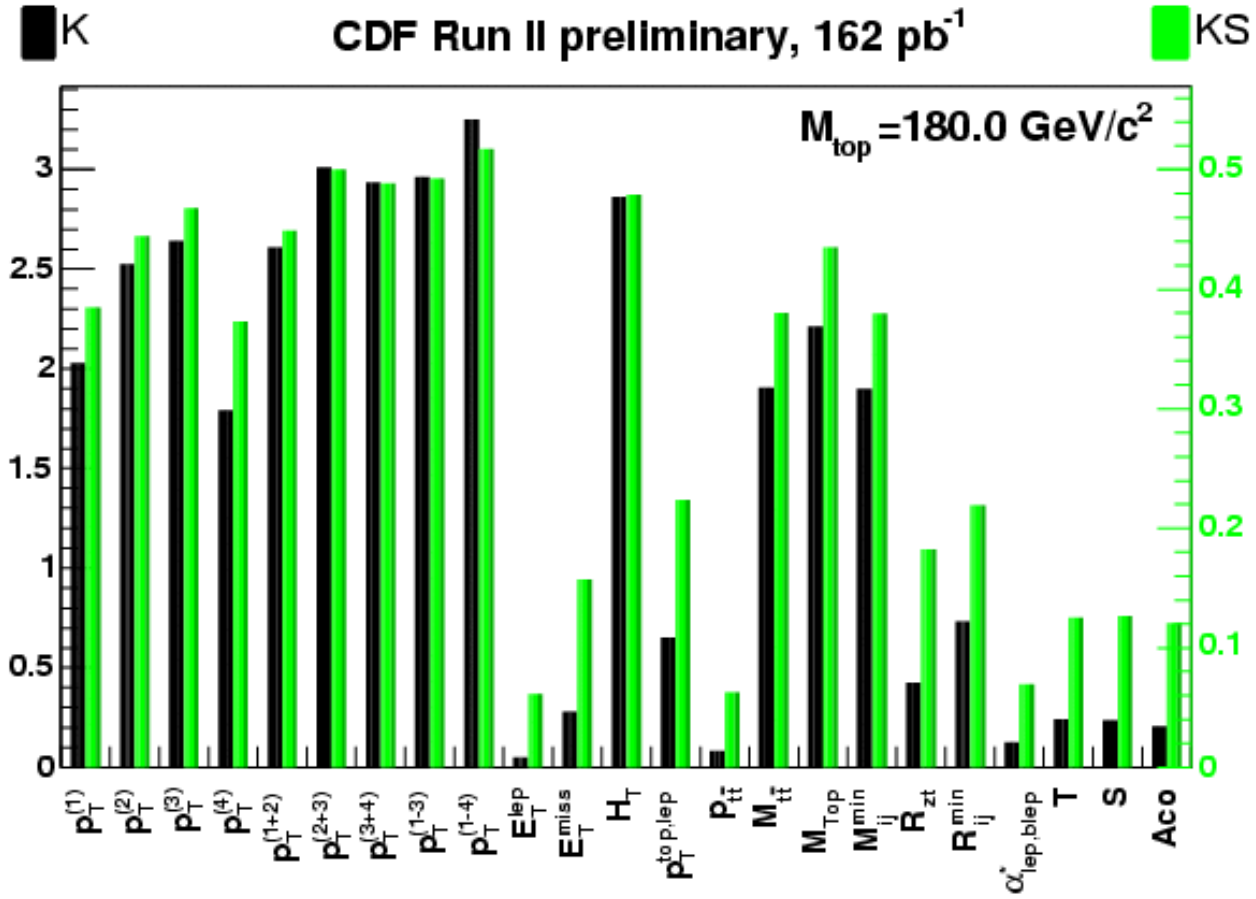
B_j are the templates for different background types. a_j are the background composition coefficients (assumed known).

S_{0,m_t}, S_{1,m_t}, and S_{2,m_t} are the three signal templates for the given generated m_t

p_{cj}, p_{cp} are the probabilities of correct jet and permutation assignments, respectively

MTM Signal / Background Separation

- Use statistical divergence measures to get an idea how useful a variable may be in separating signal from background



DLM Systematic Error Summary

Sources	δ M_{top}(GeV/c²)
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)	0.6
Transfer function	2.0
Total	6.2 GeV

Towards Ultimate M_t Measurement

- Tevatron/LHC: with current methods, the jet energy systematic error will eventually limit the M_t precision at 1-2 GeV
- A new method will be needed for hadron collider experiments to take advantage of very high luminosities
  - Measure M_t/M_W rather than M_t ?
  - Emphasize angular distributions over energies?
  - Be careful about potential non-SM contributions
- Threshold scan at a high energy e^+e^- linear collider can be used to measure M_t up to ~100 MeV