Top Physics at D0

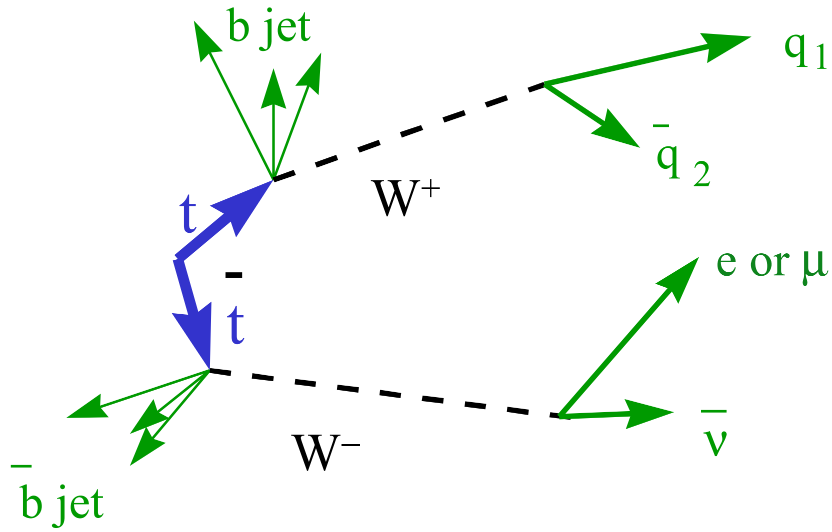
M. Strovink
University of California, Berkeley / LBNL
for the D0 Collaboration

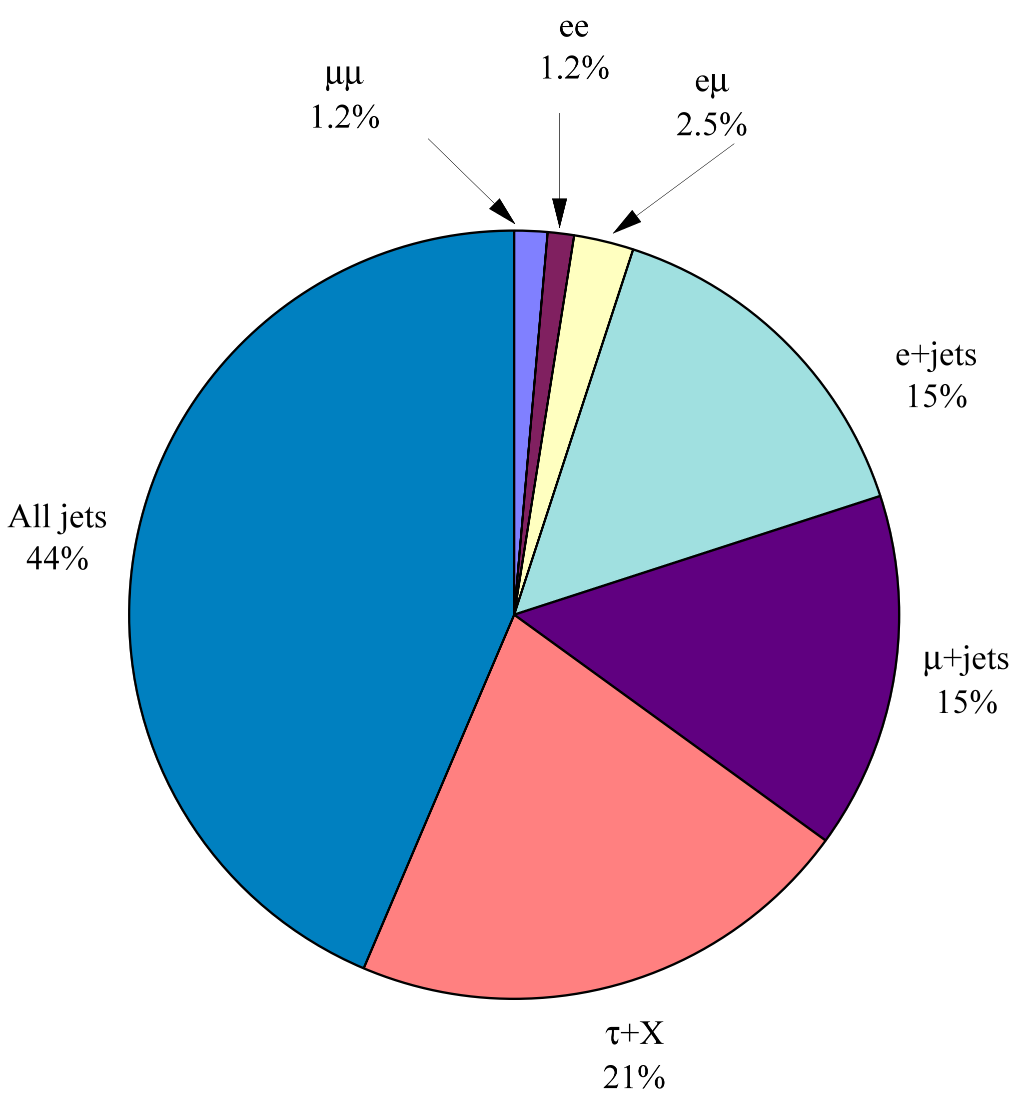
August 1997 at HEP'97 (Jerusalem)

Decay channels, event selection, and backgrounds
Cross section for top pair production
Top mass in the lepton+jets channel
Top mass in the dilepton channel
Kinematic properties of top pairs
Top in the all jets channel
Top disappearance search via $t \to bH^+$, $H^+ \to \tau^+\nu$, $\overline{s}c$.
Summary

URL for this talk: `http://www-d0.fnal.gov/~strovink`

Top pair decay channels are set by *W* branching ratios





Top pair event selection and backgrounds

Top events are characterized by

- # of jets (from $t \to Wb$ and $W \to \bar{q}q$)
- # of isolated leptons (from $W \to l\nu$)
- # of (non-isolated) tag muons (from $b \to \mu$ and $b \to c \to \mu$)
- missing E_T (from $W \to l\nu$, also from b and τ decay)

The "$e\nu$" channel is sensitive to ee and $e\mu$ events with 1 lepton undetected, to e+jets events with up to 2 jets undetected, and to $e\tau$ events with $\tau \to$ hadrons.

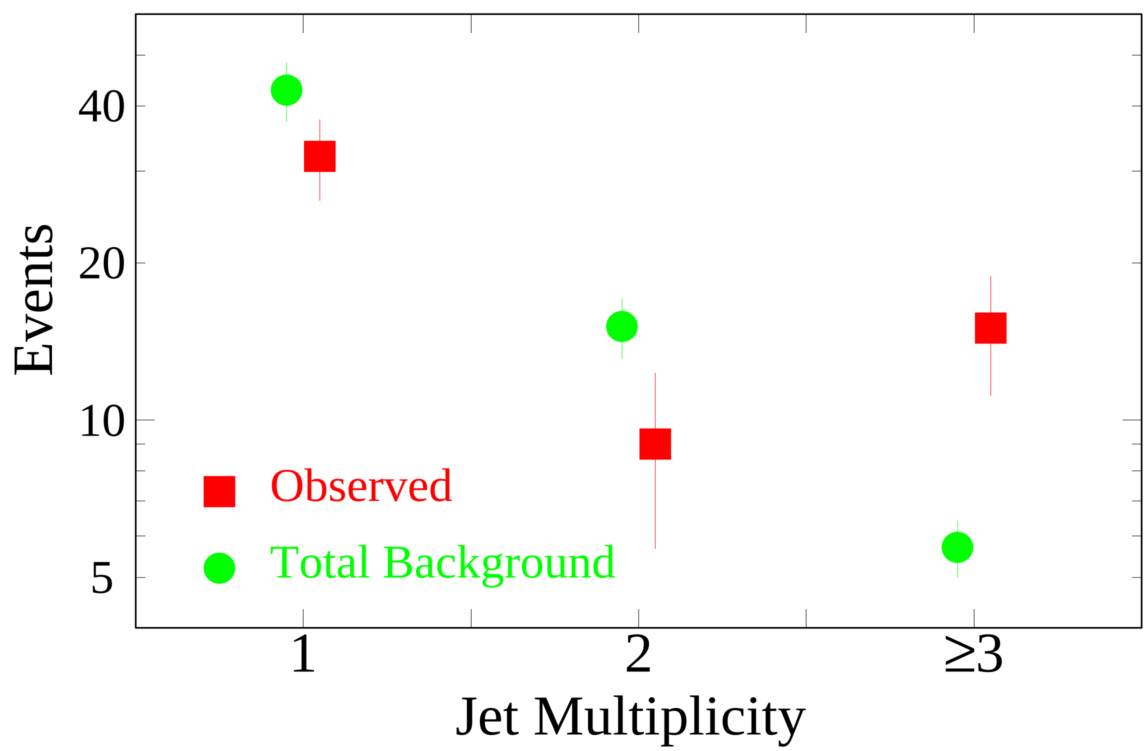
After the standard selection is applied, the S/N in the l+jets channel (~1/5) and particularly in the all jets channel (~1/2500) is small. Special analysis is needed.

General top pair analysis:	l+jets	l+jets/μ	dilepton	$e\nu$	all jets
# of jets	≥ 4	≥ 3	≥ 2	≥ 2	≥ 6
jet E_T (GeV)	≥ 15	≥ 20	≥ 20	≥ 30	≥ 10
H_T=sum of jet $\lvert E_T \rvert$ (GeV)	—	—	$\geq$ ~100	—	—
# of isolated leptons (l)	1	1	2	1 (e)	0
# of tag muons (in jet cone)	0	≥ 1	—	—	≥ 1
lepton E_T (GeV)	≥ 20	≥ 20	≥15 or 20	≥ 20	—
missing E_T (GeV)	≥20 or 25	≥ 20	≥20 or 25	≥ 50	—
$e\nu$ M_T (GeV/c^2)	—	—	—	≥ 115	—
$l+\nu$ E_T (GeV)	≥ 60	—	—	—	—
$\lvert\eta\rvert$ of W	< 2	—	—	—	—
major background	W+jets	W+jets	Z decay	W+jets	QCD
Mass analysis: # of jets	—	≥ 4	—		
Cross section analysis:					
H_T (GeV)	≥ 180	≥ 110	$\geq$ ~100	—	*NN*
A = aplanarity (jets+W)	≥ 0.065	≥ 0.04	—	—	*NN*

Jet multiplicity for μ tagged events

The background for tag muons is particularly low in D0, with its short flight path and thick muon filter. Generic QCD jets have a low μ tag rate (typically 0.005 per jet), which is parametrized as a function of jet E_T and η.

The *l*+jets/μ data show an excess over background in the signal region (≥3 jets), while following expectation for lower jet multiplicities.

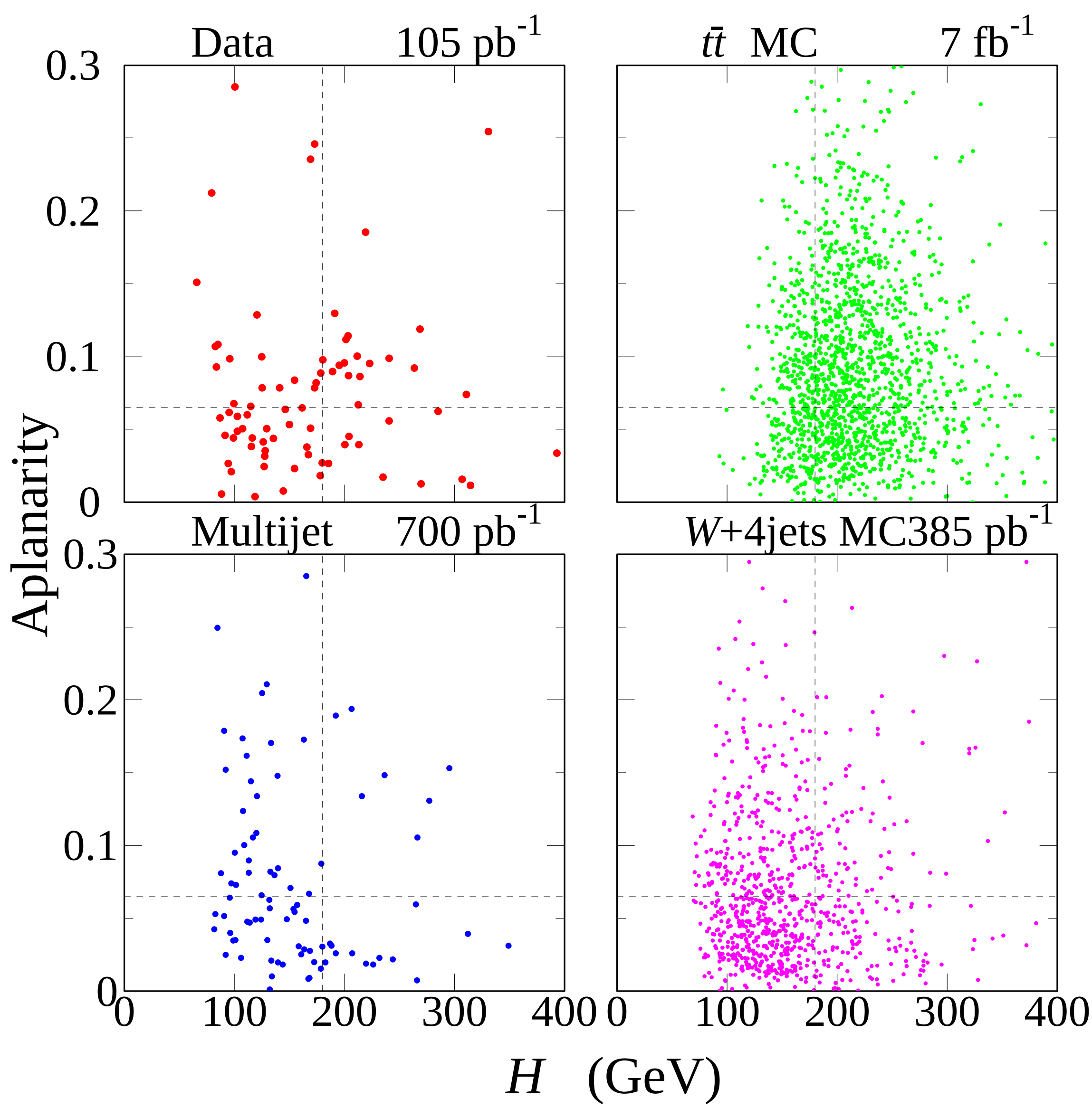


The background depends logarithmically on the number of jets. In the *l*+jets channel, where the lack of a μ tag allows larger backgrounds, this dependence is used to estimate the background level before stringent kinematic cuts are applied. For *W*+jets events, the effect of these cuts is modeled by VECBOS. For the minor background, QCD multijets with fake leptons, it is fixed by data.

Aplanarity vs. summed jet E_T for untagged events

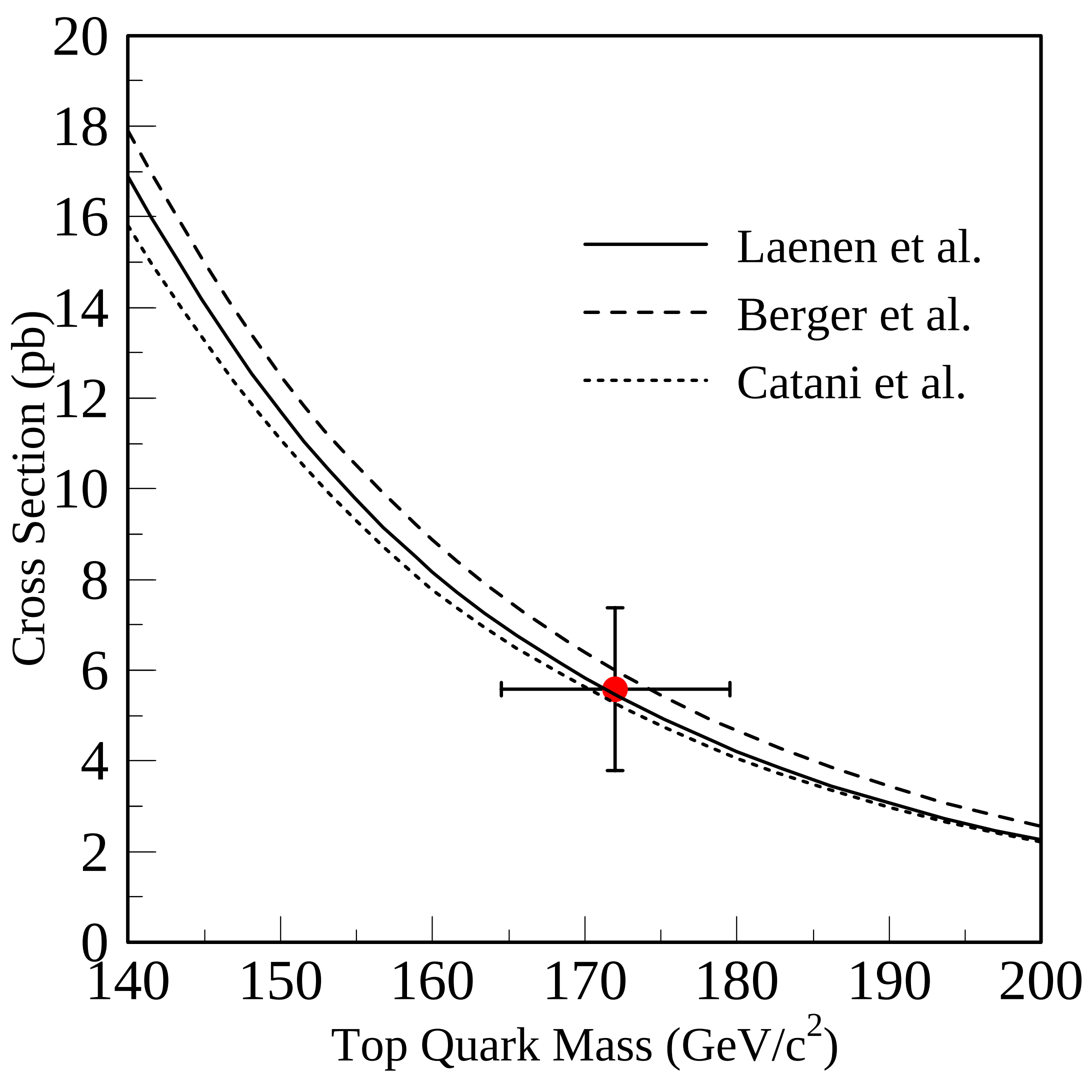
For the *l*+jets channel, without a μ tag, a stringent final cut is made on aplanarity ($A > 0.065$) and on summed jet E_T ($H_T > 180$ GeV). (A is 3/2 the smallest eigenvalue of the normalized laboratory momentum tensor, including the jets and the *W*).

Shown is the distribution in A vs. H_T for data, top, *W*+jets background, and QCD multijet background. In each panel, only the events in the top right sector pass the cut.



Top cross section results [PRL **79**, 1203 (1997)]

Channel	--- Expected (Laenen *et al.*) ---			Observed	σ(pb)
	background	top 170	sum		(m_t=173.3)
Dilepton	1.4±0.4*	4.1±0.7	5.5	5 (3 $e\mu$, 1 ee, 1 $\mu\mu$)	
$e\nu$	1.2±0.4	1.7±0.5	2.9	4	
Sum (dilepton + $e\nu$)					6.3±3.3
l+jets	8.7±1.7	14.1±3.1	22.8	19	4.1±2.0
l+jets/μ	2.4±0.5	5.8±1.0	8.2	11	8.2±3.5
Total	**13.7±2.2**	**25.7±4.6**	**39.4**	**39**	**5.5±1.8**†
	*0.2±0.2 ($e\mu$ only)		5.5±1.4(stat)±0.9(syst)±0.6(gen)†		



Top mass analysis in the l+jets channel

With 1 final state variable not measured [$p_z(\nu)$], and with 3 equations enforced (both top masses must be the same; each of the W decay pairs must have the W mass), a 2C kinematic fit is made to events in this channel.

If more than 4 jets are present (due to gluon radiation), we fit the 4 highest E_T jets within $|\eta| < 2$. For ~50% of the events, these jets do correspond to the 4 quarks to be fit.

There are 6 (12) ways to assign the jets to the quarks, when there are 1 (0) μ tags; and there are 2 solutions for $p_z(\nu)$. We choose the permutation with lowest fit χ^2 (events with minimum $\chi^2 > 10$ are rejected). When the 4 jets do correspond to the 4 quarks to be fit, this choice has ~40% probability to be correct.

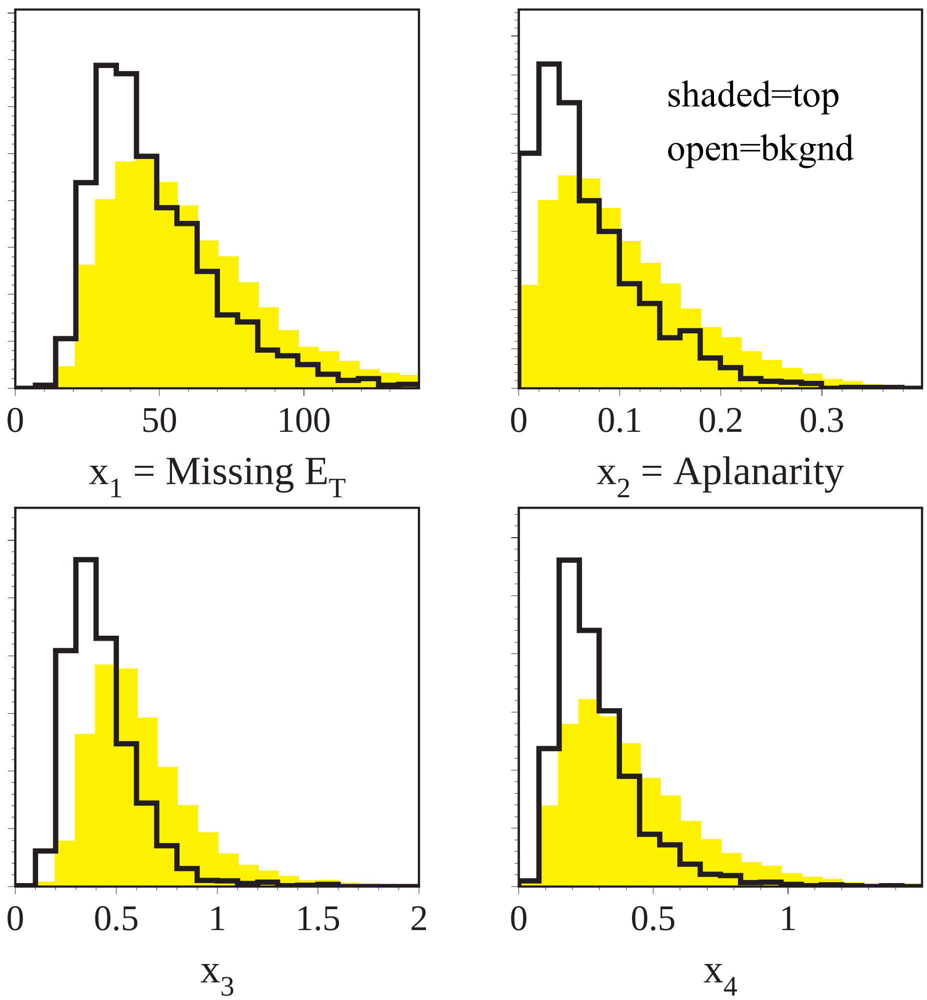
These ambiguities cause the 2C fit mass m_{fit} to be different from (but highly correlated with) the true top mass $m(t)$. $m(t)$ is extracted by comparing distributions in m_{fit} of data to those expected from a mixture of top and background.

After basic cuts, the sample consists of 77 events, of which ~51 are background. To improve the sensitivity to $m(t)$, data are binned in m_{fit} vs. a discriminant $0 < D < 1$, where D is larger for events that are more likely to be top.

We form D using multivariate techniques. The inputs are kinematic variables x_1- x_4 , where x_1 is missing E_T and x_2 is aplanarity.

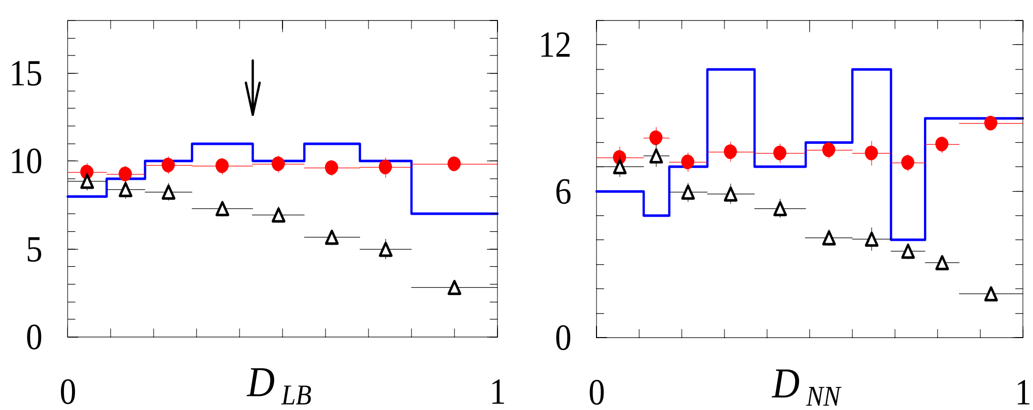
For best efficiency in estimating $m(t)$ and to avoid systematic bias, it is important that D and m_{fit} not be correlated.

Discriminant variables for top l+jets mass analysis



x_1- x_4 are weakly correlated with top mass. x_3 is sensitive to event centrality, x_4 to ΔR of jet pairs. D_{LB} and D_{NN} are likelihood and neural net discriminants based on x_1- x_4 .

(histogram=data, points=background+top, triangles=background)

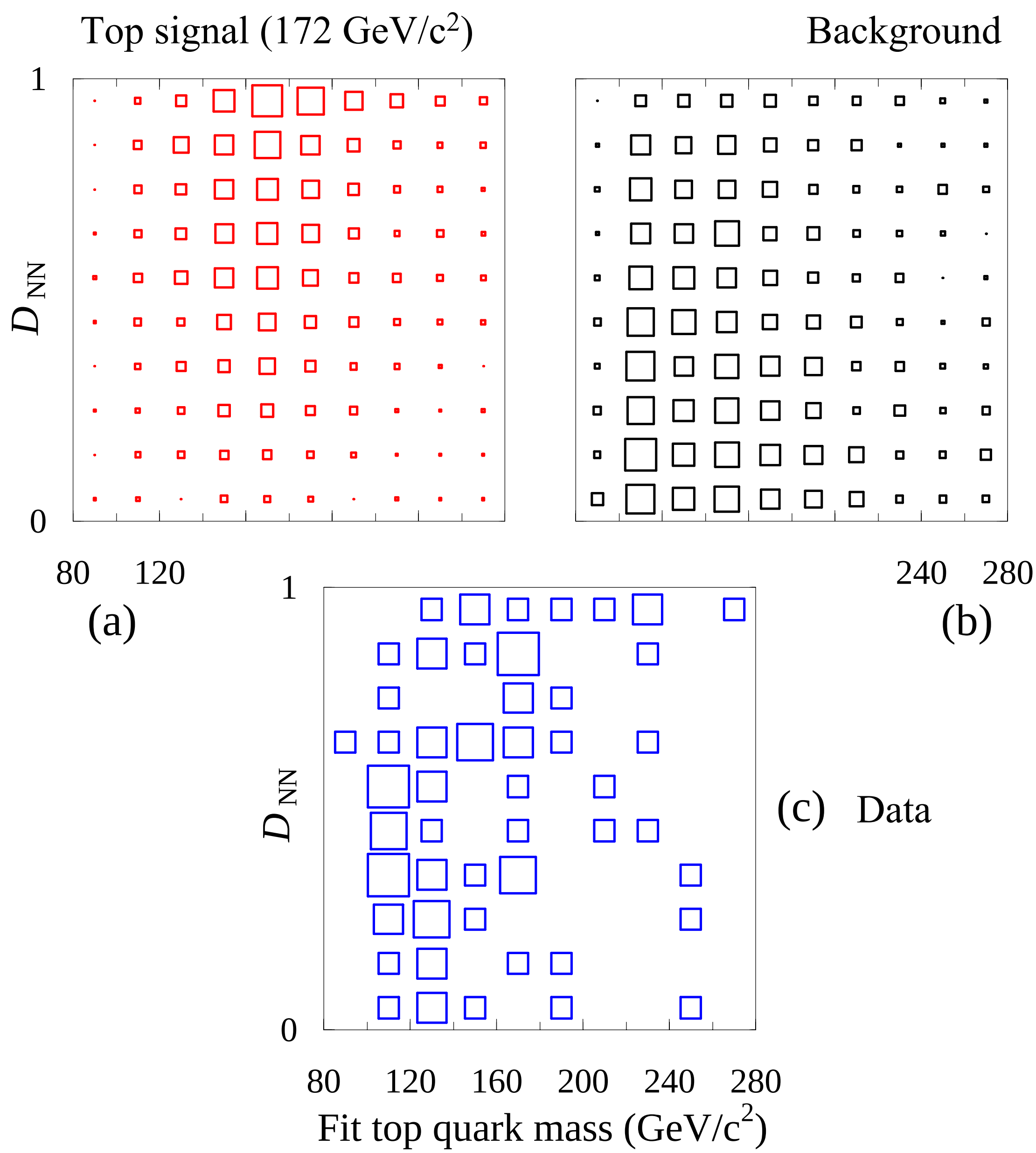


Neural network discriminant vs. fit top mass

Shown is the number of events per bin (∝ areas of boxes) vs. neural net discriminant D_{NN} (ordinate) and fit mass m_{fit} (abscissa) for (a) top signal, (b) background, and (c) data.

Little correlation between D_{NN} and m_{fit} is observed, verifying the freedom of x_1 - x_4 from significant top mass bias.

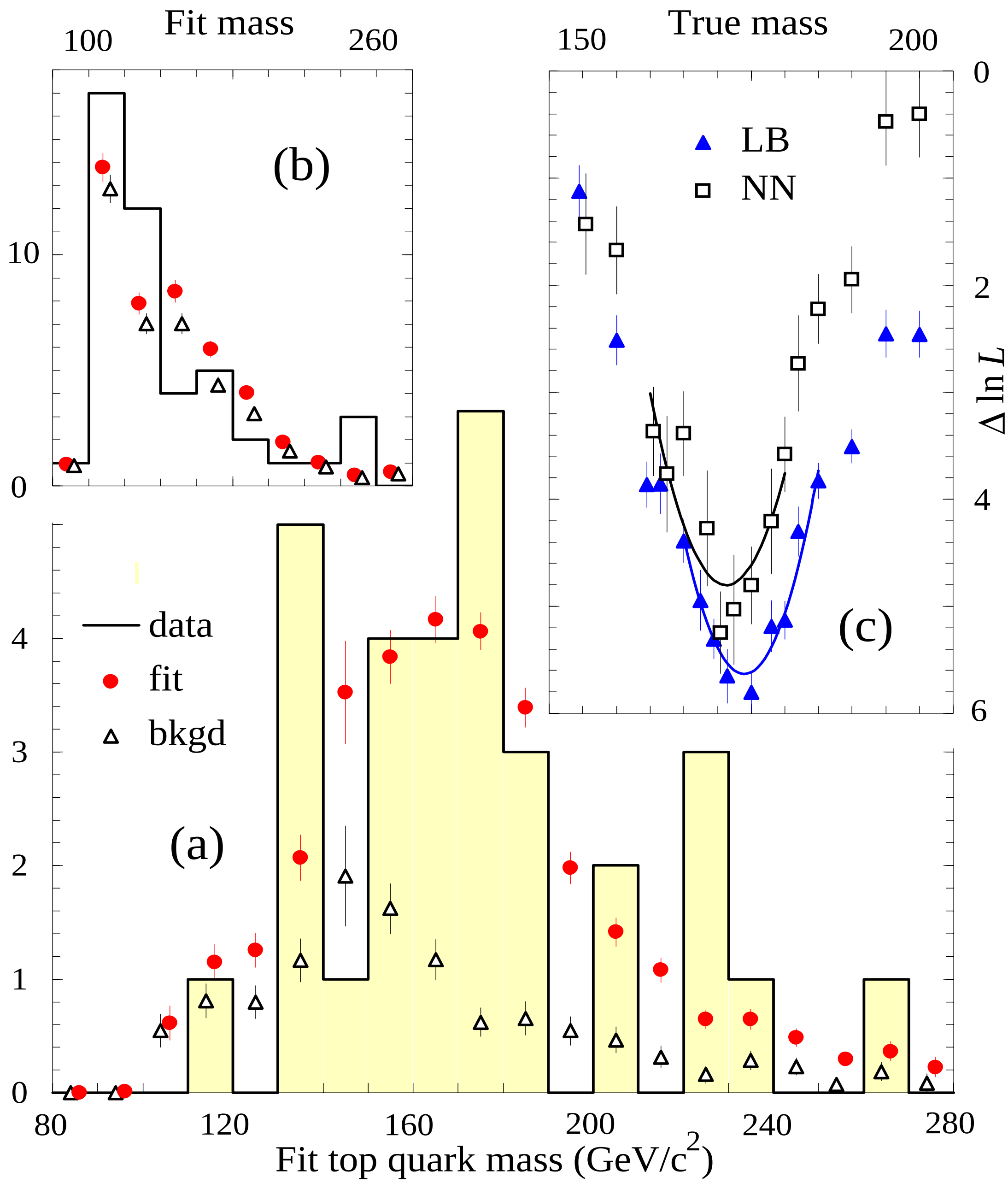
The data are seen to require contributions from both top signal and background sources.



Fit top mass distribution and true mass likelihood

Data are plotted in (a) top-rich or (b) background-rich regions. Events with μ tags are assigned to (a); otherwise events are put in (b) if $D_{LB} < 0.43$ or $H_T < E_T$(jet 1) + 90 GeV.

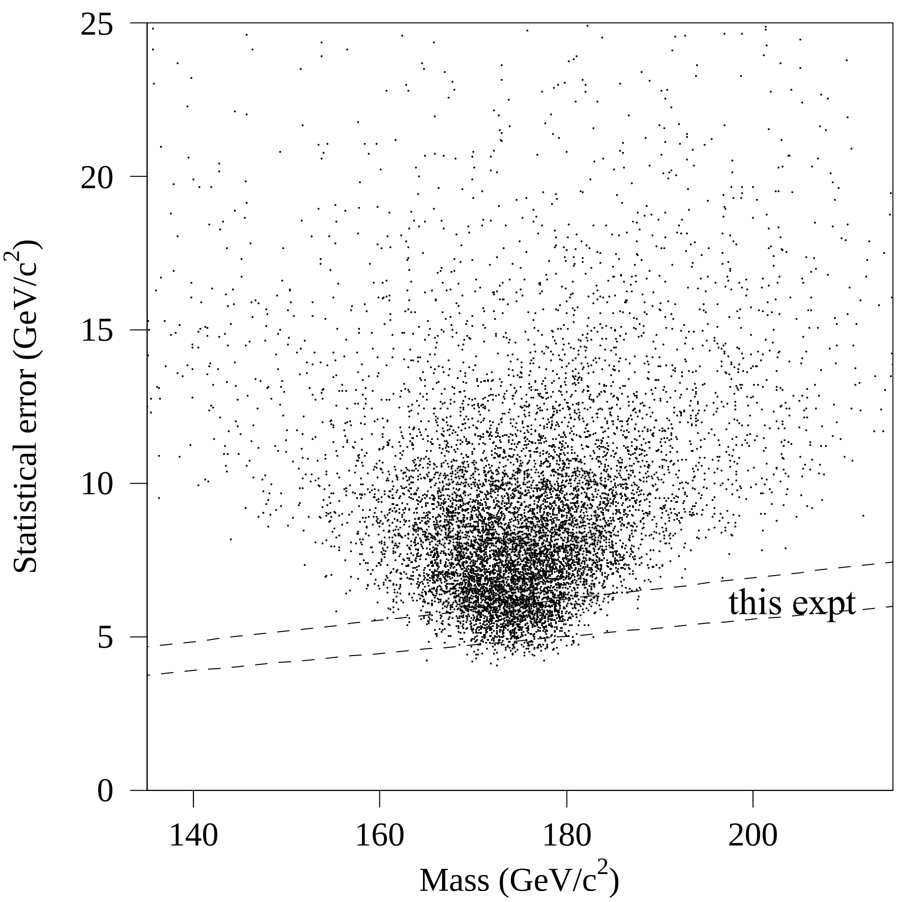
For each true top mass plotted in (c), a likelihood fit to data is made for a free mixture of top signal and background, binned in top richness vs. m_{fit}. A parabolic fit to the likelihood curve yields the true top mass and its error.



Top l+jets mass results [PRL **79**, 1197 (1997)]

Fits to data	---LB fit---		---NN fit---	
Quantity fit	value	σ(stat)	value	σ(stat)
m_t (GeV/c^2)	174.0	± 5.6	171.3	± 6.0
n_s	23.8	$^{+8.3}_{-7.8}$	28.8	$^{+8.4}_{-9.1}$
n_b	53.2	$^{+10.7}_{-9.3}$	48.2	$^{+11.4}_{-8.7}$
m_t correlation of LB and NN fits	(88 ± 2)%			
Systematic error on m_t	energy scale	± 4.0		
	generator	± 4.1		
	other	± 2.2		
Resulting m_t (GeV/c^2)	**173.3 ± 5.6 (stat) ± 6.2 (syst)**			

Smaller fit statistical error ⇔ smaller MC mass spread:



Alternate technique using 3C fits to a set of top masses for each *l*+jets event

As a cross-check, we have determined the top mass from l+jets events using a substantially different technique:

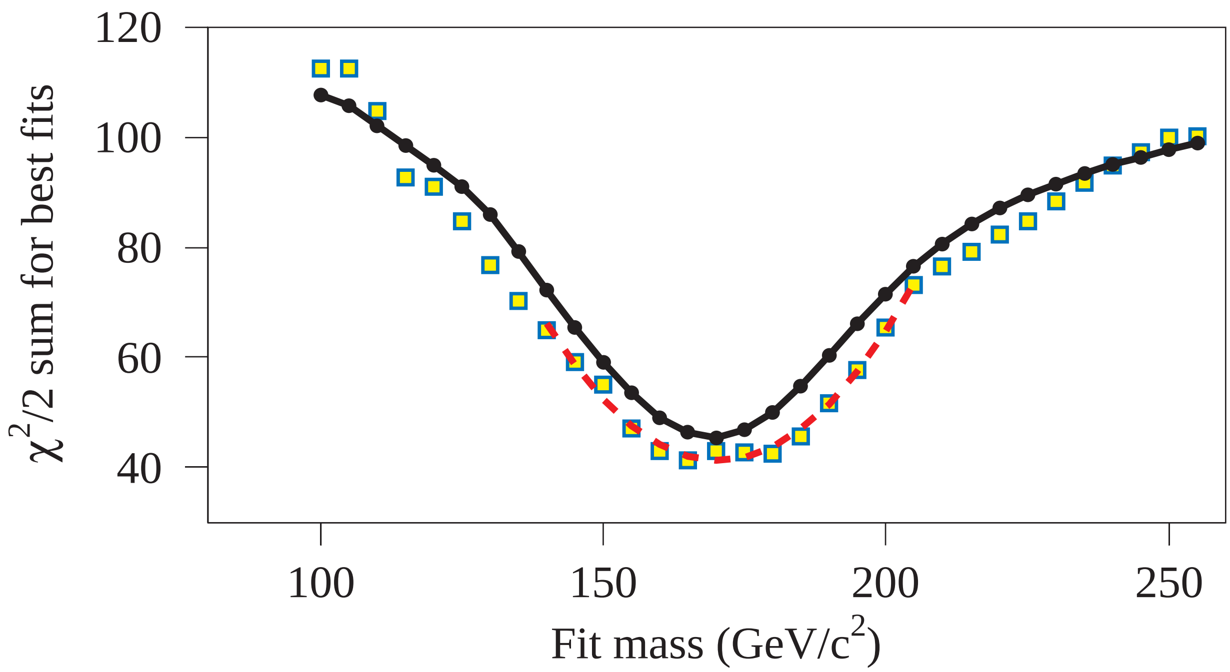
- Make 3C fits at a fixed top mass set $\{m_{fit}\}$ to each event.
- At each m_{fit} choose the jet permutation that minimizes $\chi^2/2$.
- Sum the set of $\{\chi^2/2\}$ over all events and plot vs. m_{fit} .
- Subtract the same plot for (32%) calculated background.
- Fit a parabola to obtain raw $m(t)$ and its error $\sigma(m(t))$.
- Use MC experiments to map raw to true $m(t)$ and $\sigma(m(t))$.

This method takes account of possible multiple local minima in fit top mass due to permutations in jet assignment.

In agreement with the main analysis, it yields

$$m(t) = 176.0 \pm 7.9 \text{ (stat) GeV/c}^2.$$

Shown is the $\chi^2/2$ sum vs. m_{fit} for the expected top signal (connected points) and for background subtracted data (squares). The dashed line is the parabolic fit.



Top mass analysis in the dilepton channel

In the dilepton channel, with 3 additional final state variables (the 2nd ν momentum) not measured, the system has -1 rather than +2 kinematic constraints. If a top mass is assumed, the system can be reconstructed via a quartic equation with 0, 2, or 4 real solutions.

Usually solutions exist for a wide range of $m(t)$. More discrimination can be gained by asking "if $m(t)$ had a certain value, how likely is it that the top decay products would appear in the detector as they did?"

The factors* in this likelihood $L(m(t))$ are:

A. $(1/\sigma)$ $(d\sigma$ / d LIPS) for $t\bar{t}$ production.

B. Probability density for energy of l in t rest frame.

C. Jacobian $|\partial$ LIPS / $\partial\{o\}|$ [$\{o\}$ = observed variables].

D0 makes two independent approximations to $L(m(t))$:

- Matrix element weight (*M*WT)

  Ignores **C**, includes **B**, approximates **A** using product of proton pdf's with empirical $m(t)$ dependent factor.

  Extension of Kondo; Dalitz & Goldstein ideas.

- Neutrino phase space weight (νWT)

  Ignores **A** and **B** and approximates **C.** Predicts missing E_T after fixing both ν rapidities to many different values. Compares to measured missing E_T and increments a likelihood sum.

To obtain their final weight, both methods sum over

Quartic solutions

Jet assignments (including isr and fsr)

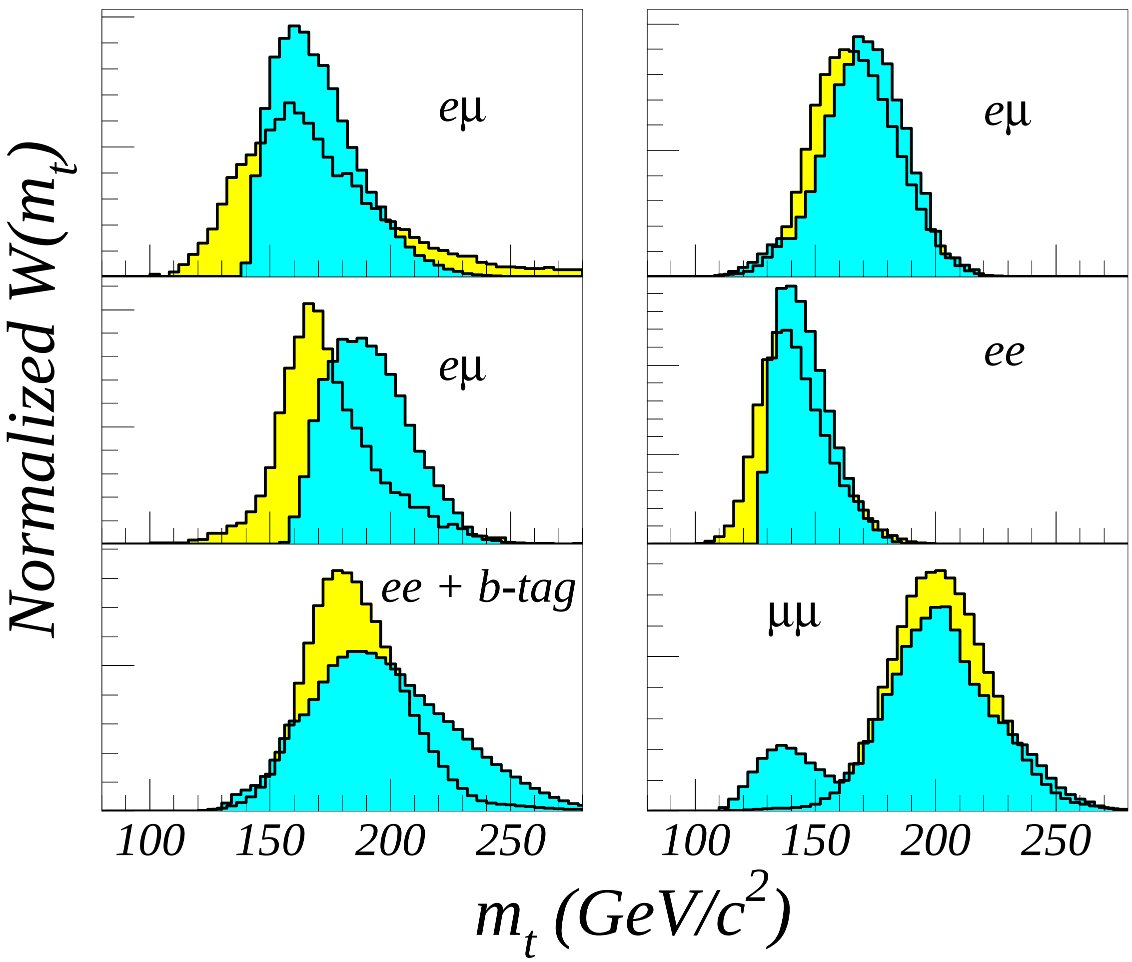
Many resolution-smeared versions of the same event

* details in D0's top quark contribution to EPS/HEP93 proceedings

Top mass weight distributions for dilepton candidates

Shown is the weight vs. top mass for 6 dilepton events (dark = matrix element method; light = ν phase space method).

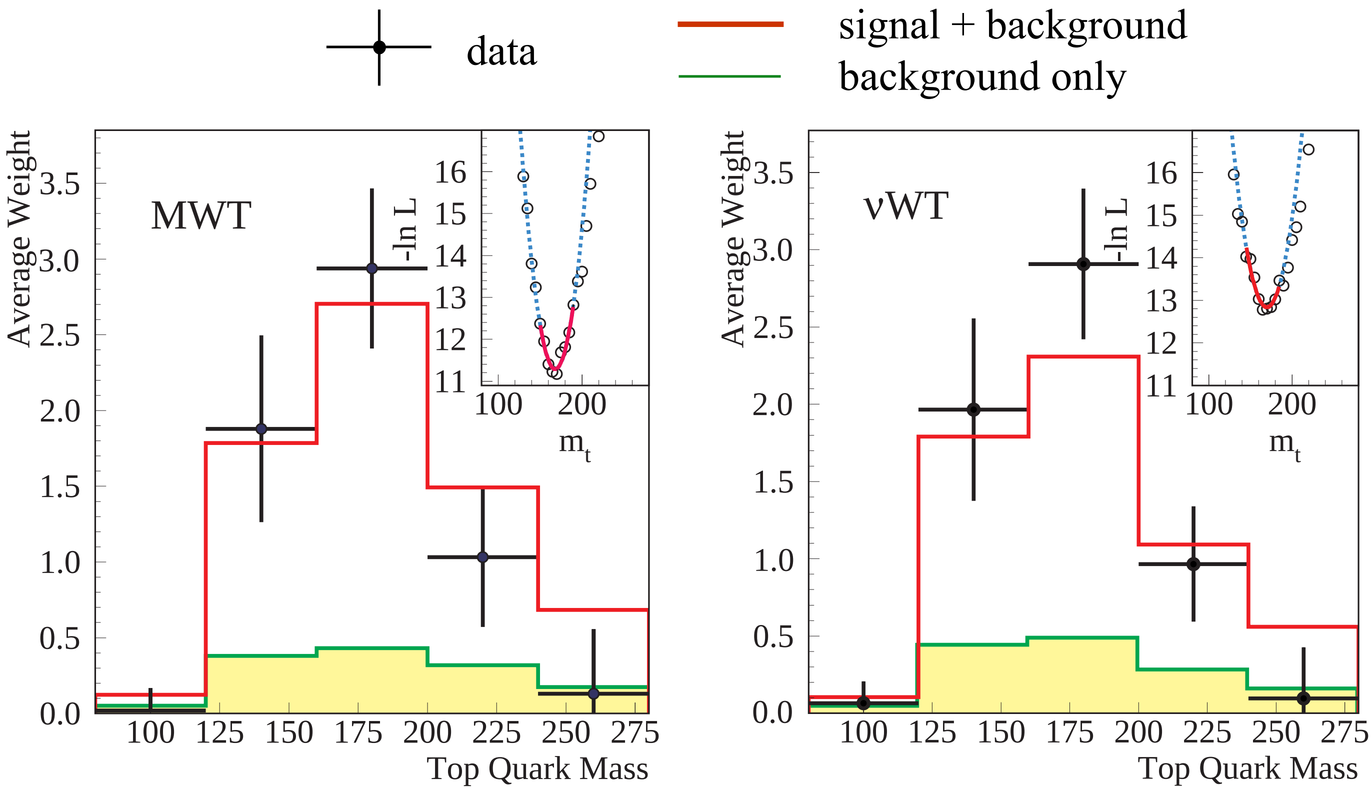
These distributions are not top mass probability densities. To extract the top mass, we compare them to distributions from a mixture of expected signal and background for many MC top masses, using a likelihood fit.



To represent the weight distribution for each event, we store a vector whose components are the fraction of the weight found in each of (5) 40 GeV/c^2 bins from 80 to 280 GeV/c^2.

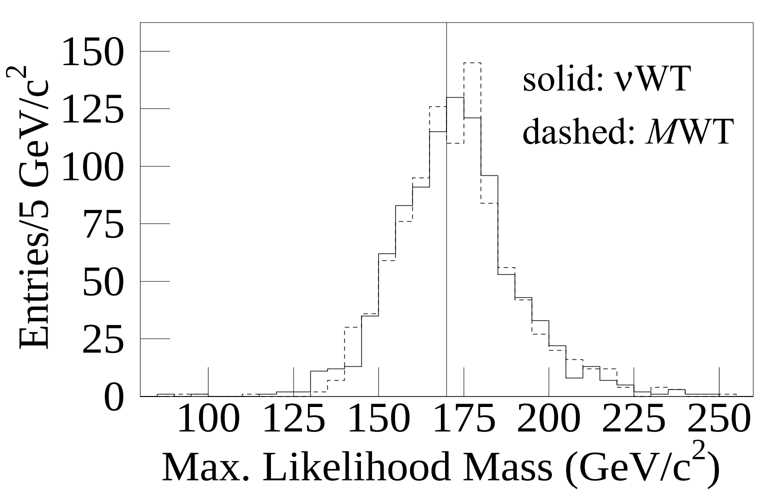
To estimate the probability densities for signal and background in this vector space, we accumulate a Gaussian kernel for each event in the modeled sample.

Average weight vector for dilepton sample, and results of simulated experiments



The results of the likelihood fits for (a) matrix element and (b) ν phase space weighting methods are shown in the insets. Plotted for each of 5 regions is the average weight for data, best fit mixture, and background.

Results (below) of MC experiments support the statistical errors that are assigned.



Top dilepton mass results (hep-ex/970614, submitted to PRL)

Analysis method	Result of fit to m_t (GeV/c^2)
Matrix element weighting	168.1 ± 12.4(stat)
Neutrino weighting	169.9 ± 14.8(stat)
Combined (77% correlated)	168.4 ± 12.3(stat)
Source of systematic error	$\sigma(m_t)$ (syst) (GeV/c^2)
jet energy scale	2.4
signal model	1.8
multiple interactions	1.3
background model	1.2
likelihood fit	1.3
Total systematic error	3.7
m_t (GeV/c^2) from dileptons	**168.4 ± 12.3(stat) ± 3.7(syst)**

Combined top mass results

We combine the results of the l+jets and dilepton top mass analyses by propagating the systematic uncertainties in each channel with correlation coefficients of either 0 (for MC statistics and background model) or 1 (for jet energy scale, multiple interactions, and $\bar{t}t$ production models).

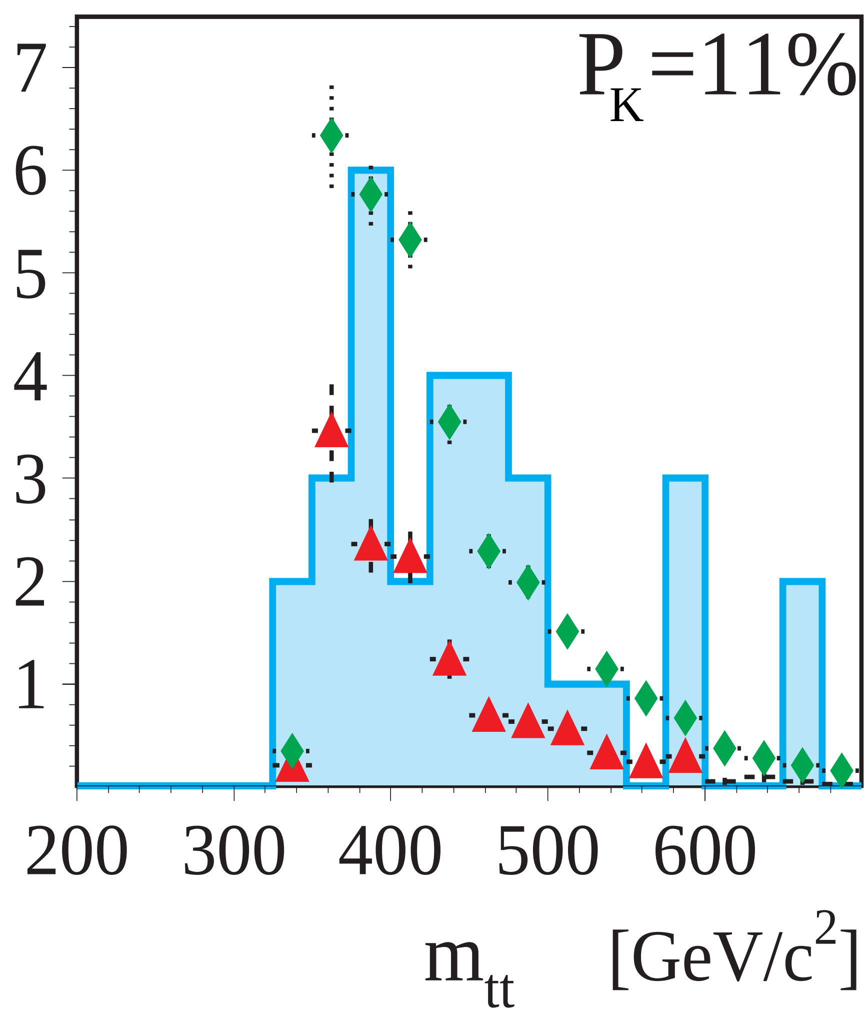
The combined result is

$$\boldsymbol{m(t) = 172.0 \pm 5.1(\text{stat}) \pm 5.7(\text{syst})\ \text{GeV}/c^2}\,.$$

Kinematic properties of l+jet top pairs

Kolmogorov-Smirnov comparisons between data and top + background model were made for variables including # of jets, $M_T(l\nu)$, mass and p_T of top pair, top p_T, top η, and $\Delta\eta$ and $\Delta\phi$ between the top pair. Results from both 2C and 3C fits (fixing $m(t)$=173.3 GeV) were studied. The mean K-S probability (P_K) is 53% and the lowest is 9%.

Shown as an example of these comparisons is the top pair invariant mass from the 3C fit. The expected background (triangles) is 12.5 events; the expected total is 31 events (diamonds). The data (histogram) include 7 (5) events above 500 (550) GeV/c^2, with 4.0 (2.9) events expected.



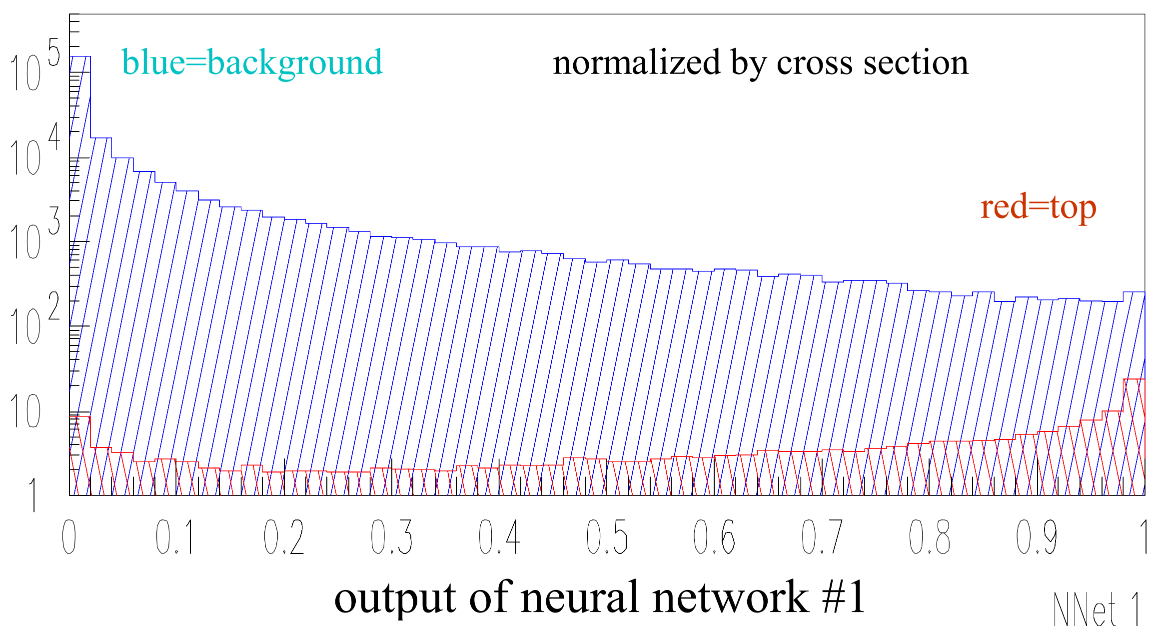
Kinematic variables for all jets neural network #1

Compared to QCD multijets, top events are:

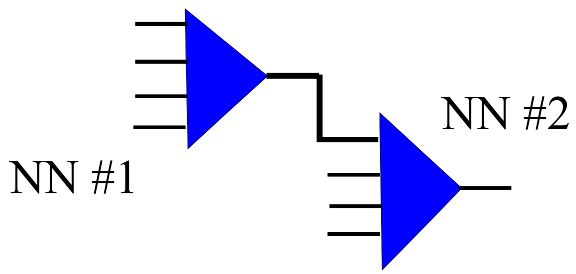
- harder
- more spherical (less planar)
- more central
- with stiffer non-leading jets {i.e., jets 3,4,5…}

Use ~two kinematic variables for each property:

H_T	Total scalar E_T
$\sqrt{s}$	Total invariant mass
$E_T(1)/H_T$	E_T fraction carried by leading jet
A	Aplanarity
S	Sphericity
C	Centrality = $H_T / \Sigma E$
η_{RMS}	Weighted RMS in η
$\lvert\eta_5\,\eta_6\rvert$	$\langle\eta^2\rangle$ of 5th & 6th jets
$H_T\,(3j)$	Total scalar E_T of non-leading jets
$\langle N_j\rangle$	Threshold weighted number of jets
$\sqrt{[E_T(5)\,E_T(6)]}$	$\langle E_T\rangle$ of 5th & 6th jets

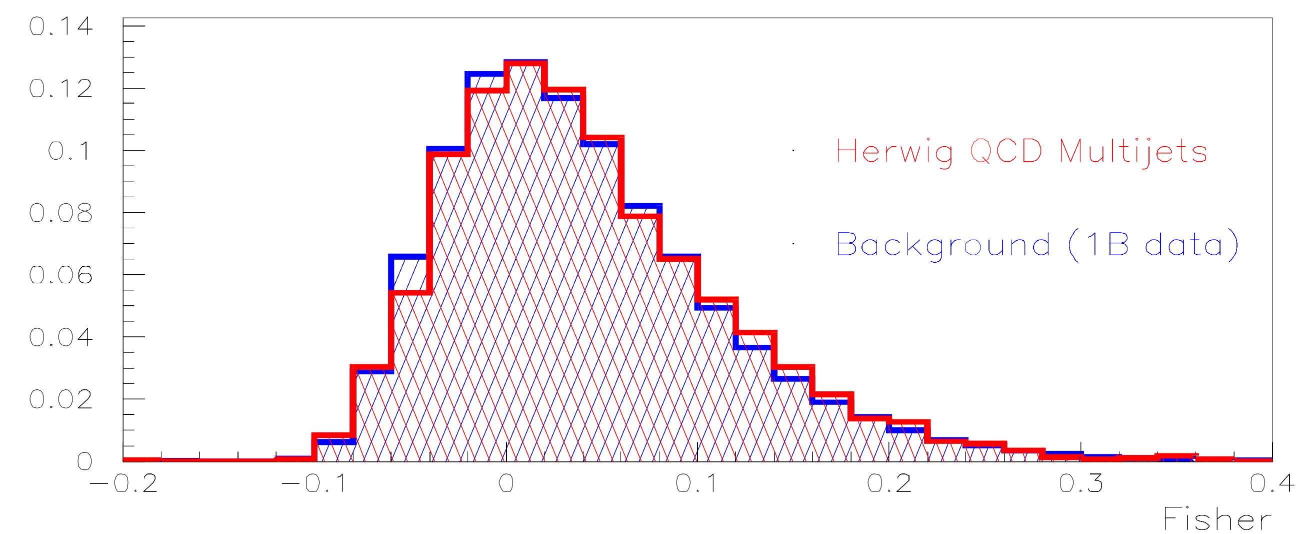


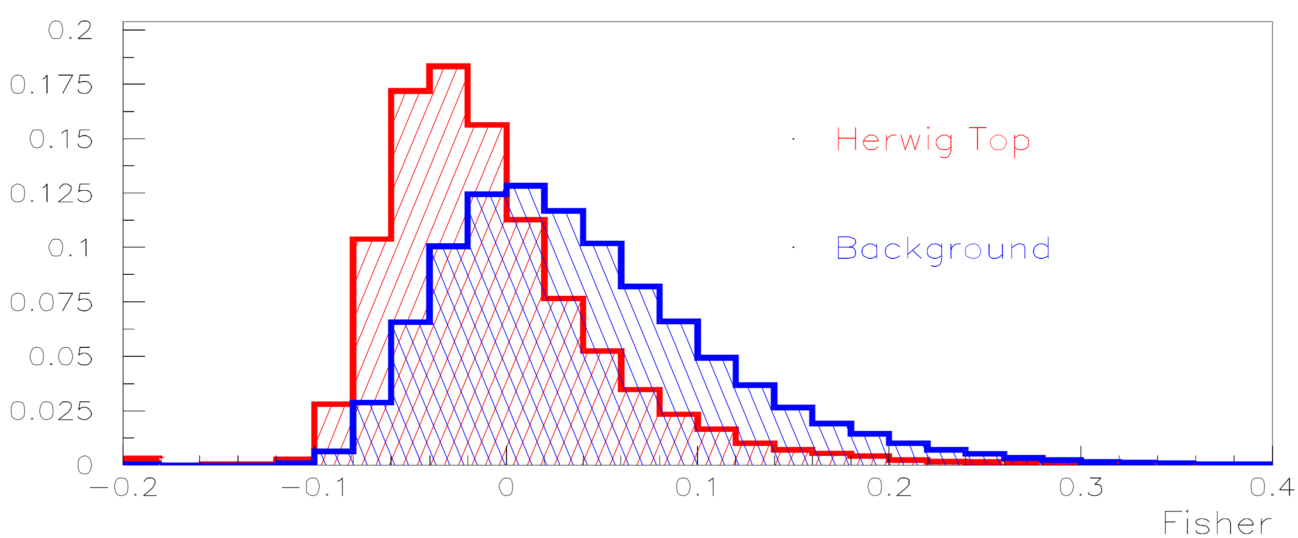
Variables for all jets neural network #2



Inputs to neural network #2:

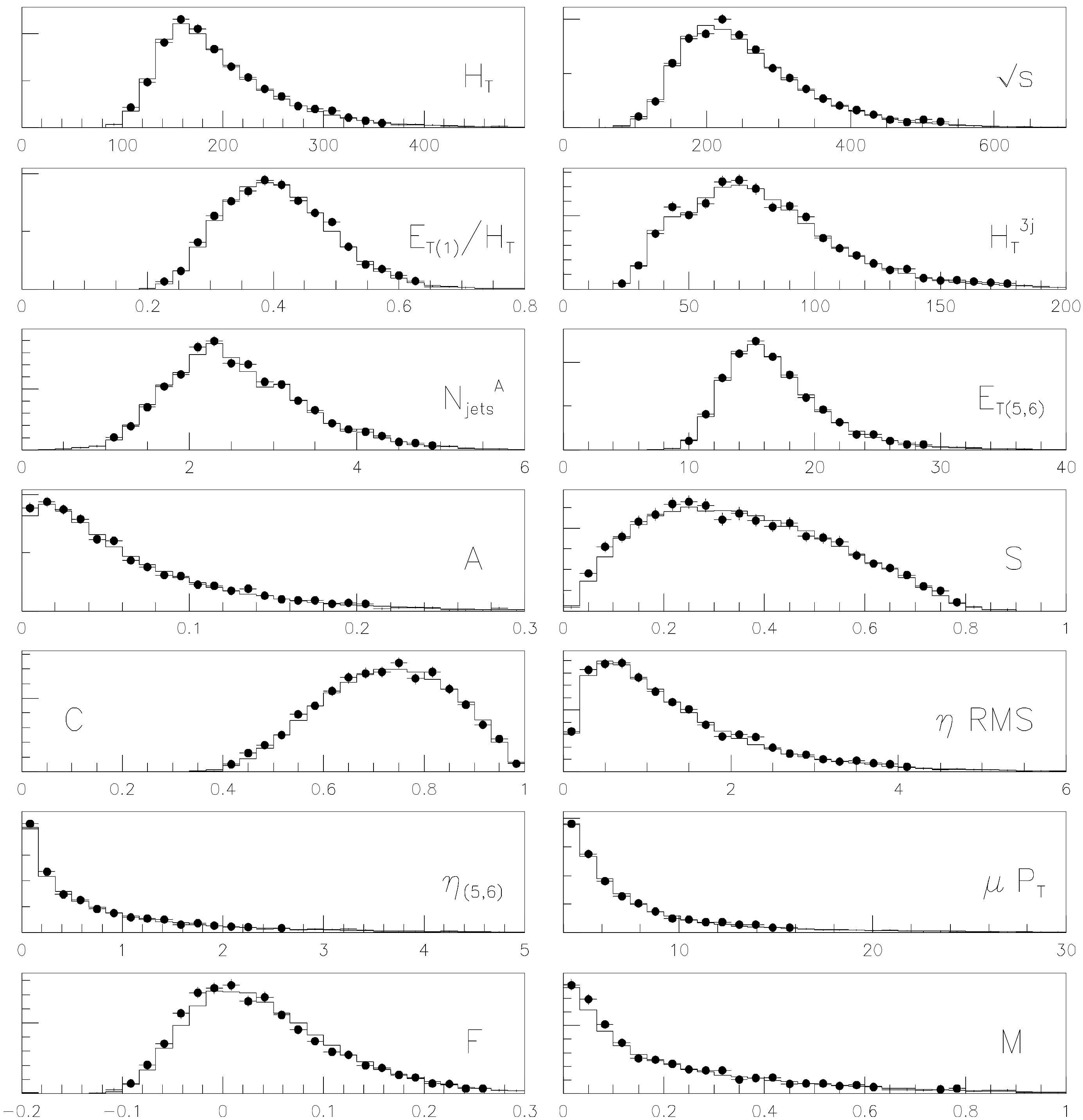
- Output of neural network #1
- p_T of muon
- Variable sensitive to quality of constrained fit to any top mass
- Fisher discriminant sensitive to jet width (signal=q, bkgnd=g):





Consistency of distributions in observed and modeled all jets neural net parameters

In all 14 neural net parameters, observed (•) [μ tagged data] distributions agree with those of model (*histogram*) [untagged data × (μ tag rate = $f(\mu\ p_T$, jet E_T , detector η))].



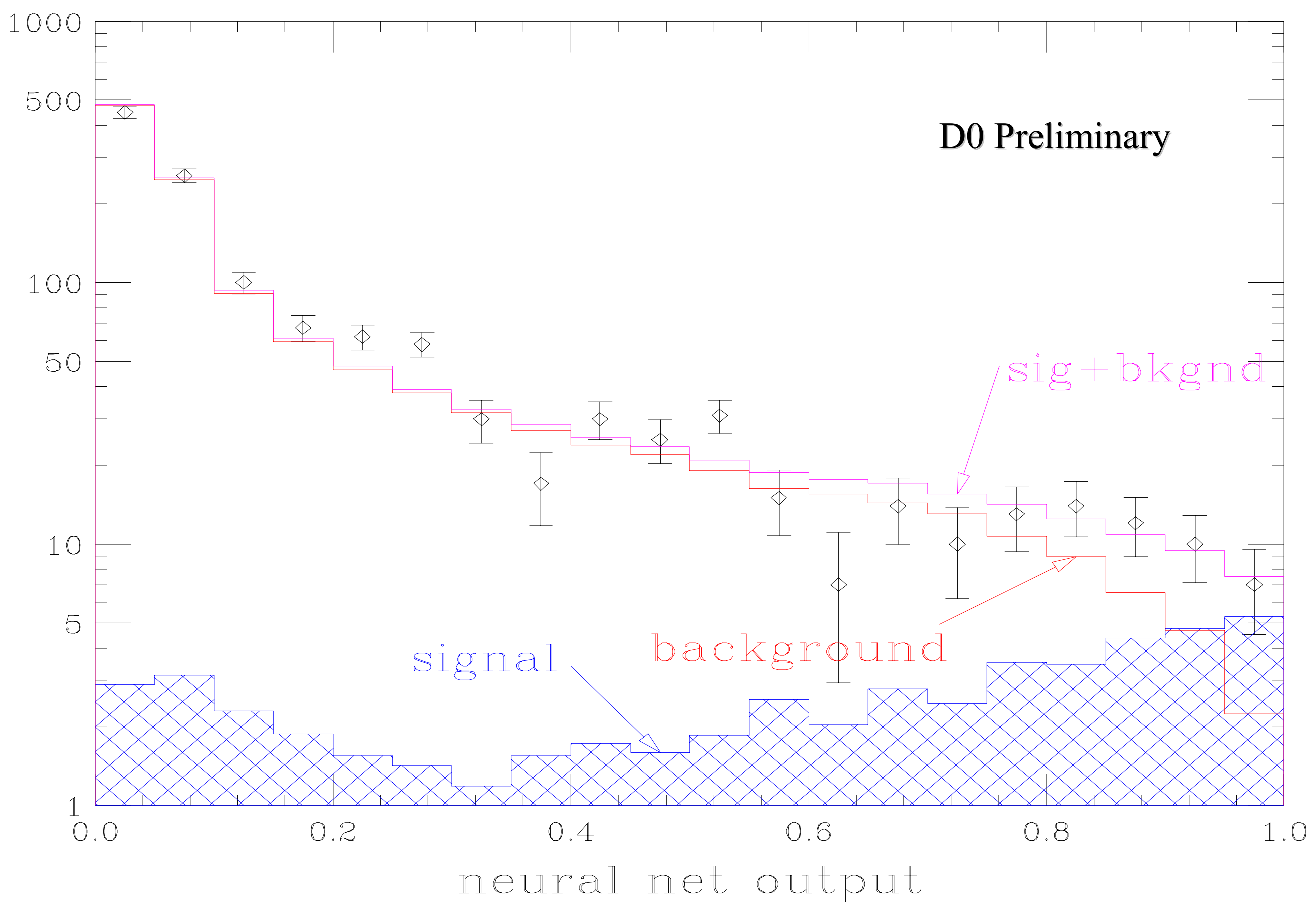
Combined fit to all jets data -- cross section and significance

Fit the NN #2 output for tagged data (points) to a sum of signal (μ tagged Herwig top) and background (untagged data × μ tag rate), with background normalization and top cross section as free parameters, obtaining

$\sigma_{t\bar{t}}$ = 7.9 ± 3.1 (stat) ± 1.7 (syst) pb (preliminary)

at a top mass of 172 GeV/c^2. The largest systematic uncertainties are in the background model (11%), μ p_T spectrum (7%), μ efficiency (7%), and μ tag parametrization (7%), with 9 smaller sources.

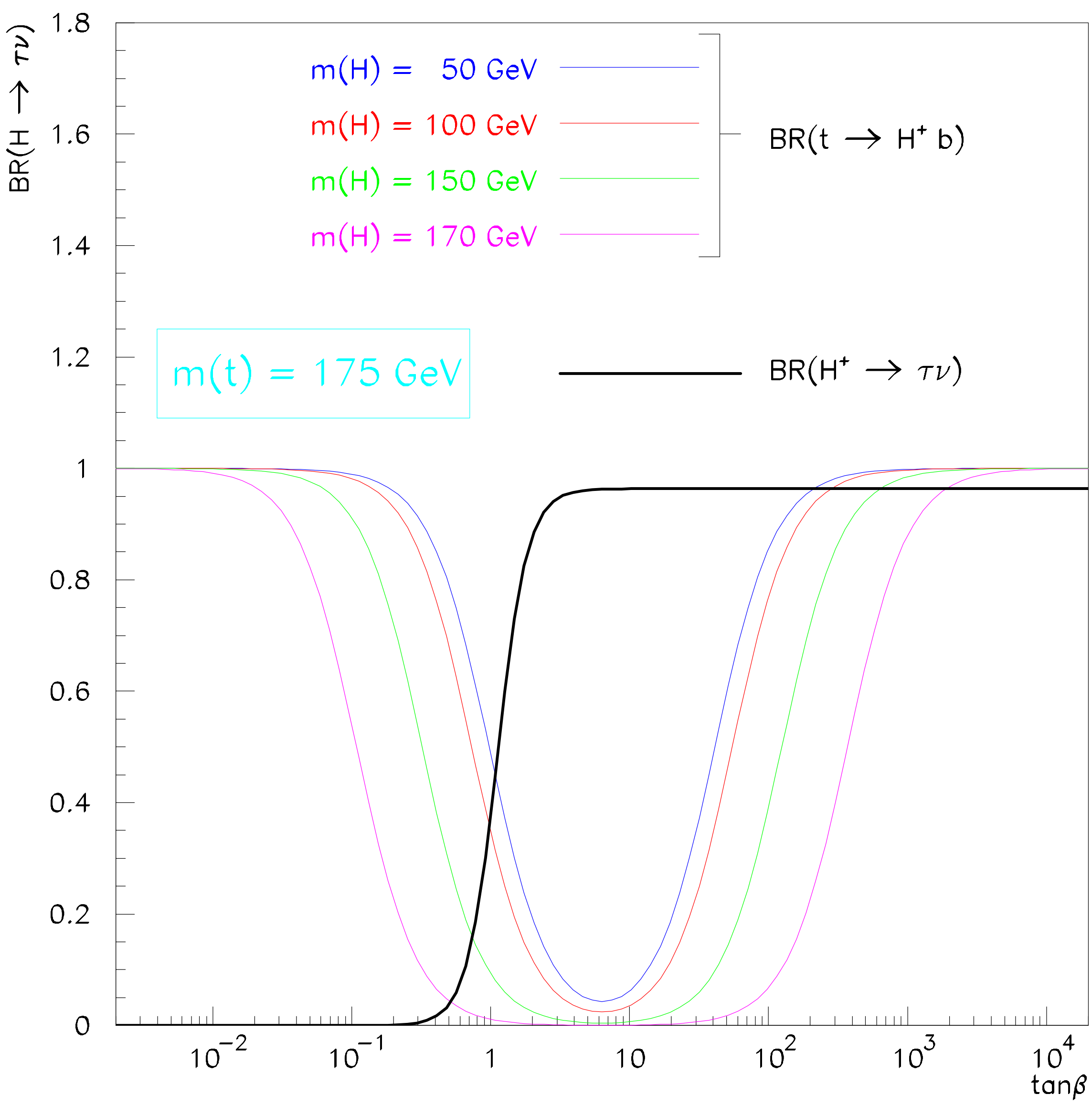
Requiring NN #2 output > 0.78, obtain 44 events with an expected background of 25.3±7.3 and an expected top signal of 11.6±4.5. This excess corresponds to a Gaussian equivalent fluctuation of ~3 sigma.



Top disappearance search via $t \to bH^+$, $H^+ \to \tau^+\nu$, $\bar{s}c$.

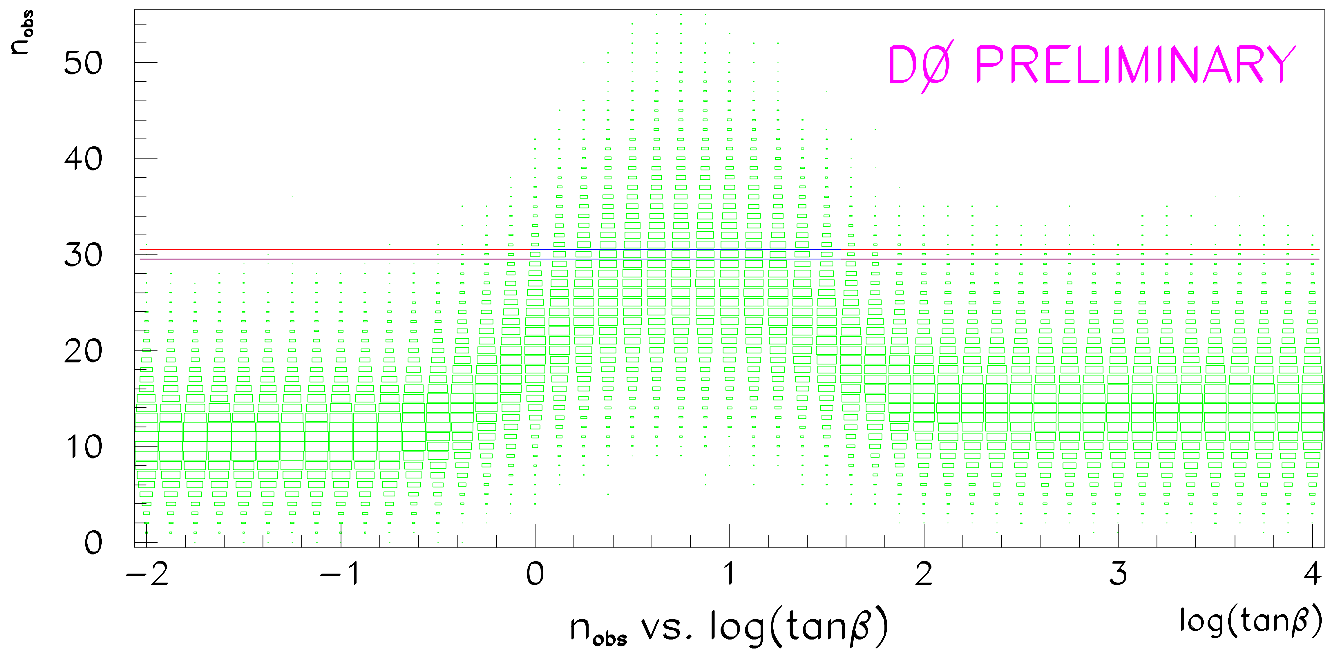
If one or both of the top quarks were to decay to H^+b rather than W^+b, the same l+jets analysis used for D0's top cross section measurement would be less efficient, causing a shortfall in the measured cross section relative to the SM.

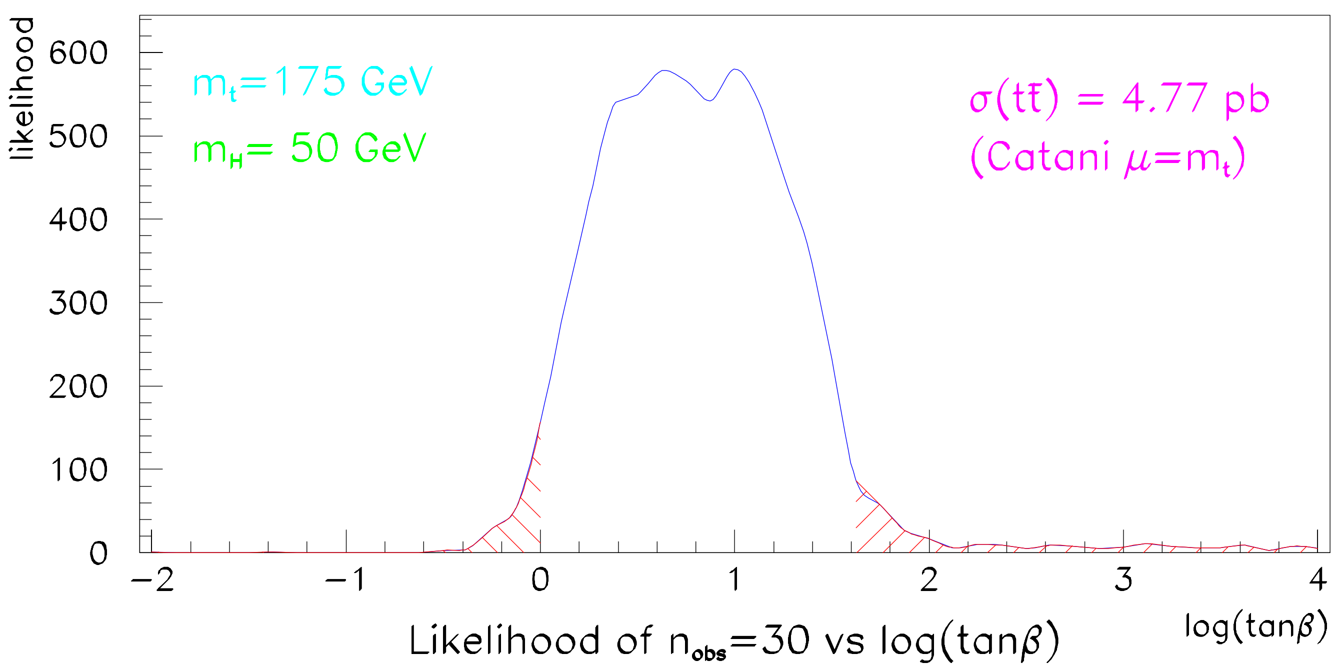
Within the MSSM, $t \to H^+b$ primarily at low and high tan β. The shortfall occurs both at low tan β, where $H^+ \to \bar{s}c$, and at high tan β, where $H^+ \to \tau^+\nu$, due mainly to a lack of energetic isolated leptons. It leads to exclusion regions at the tan β extremities of the $M(H^+)$ vs. log tan β plane.



Example: using MC experiments to set exclusion region

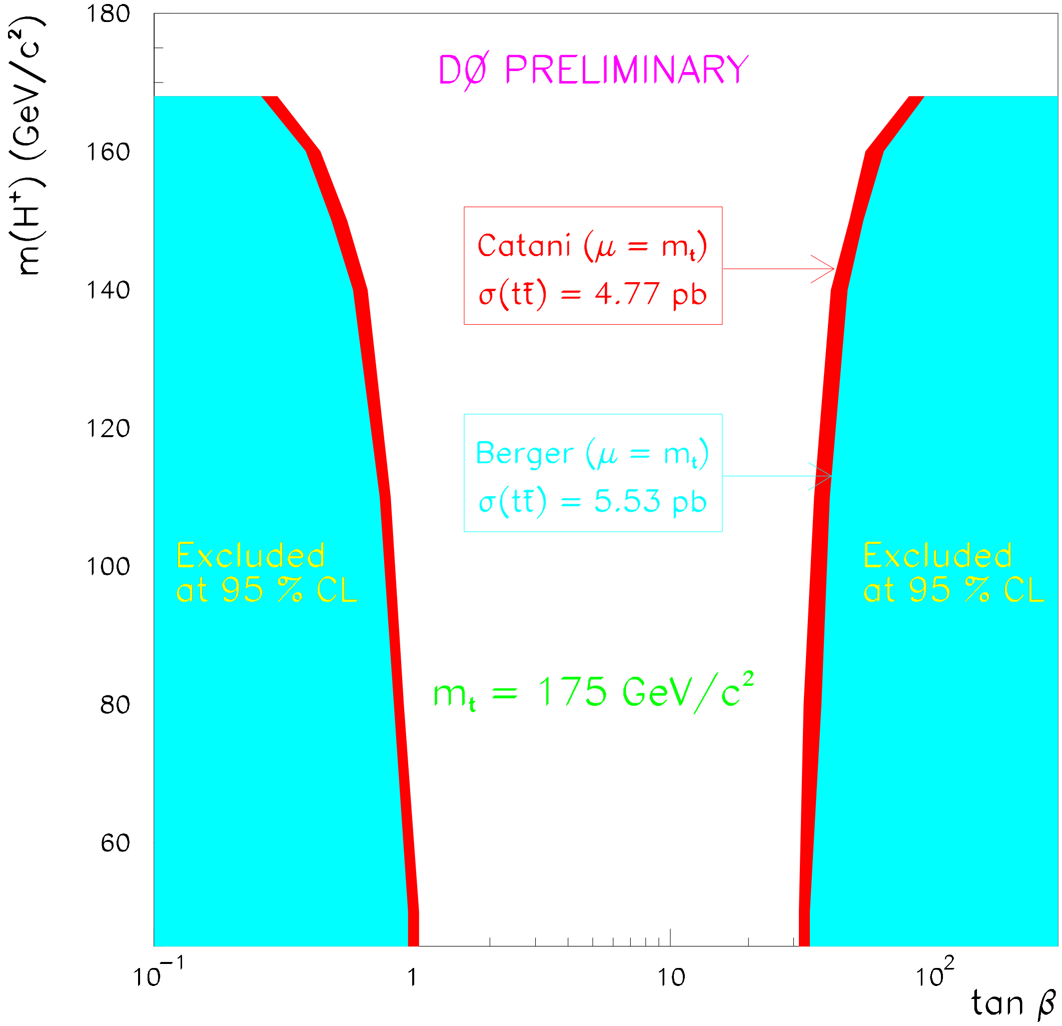
(a) No. of events obtained in MC experiments vs. log tan β, for $M(H^+)$=140 GeV/c^2, $m(t)$=175 GeV/c^2, $\sigma_{t\bar{t}}$ = 6.48 pb.

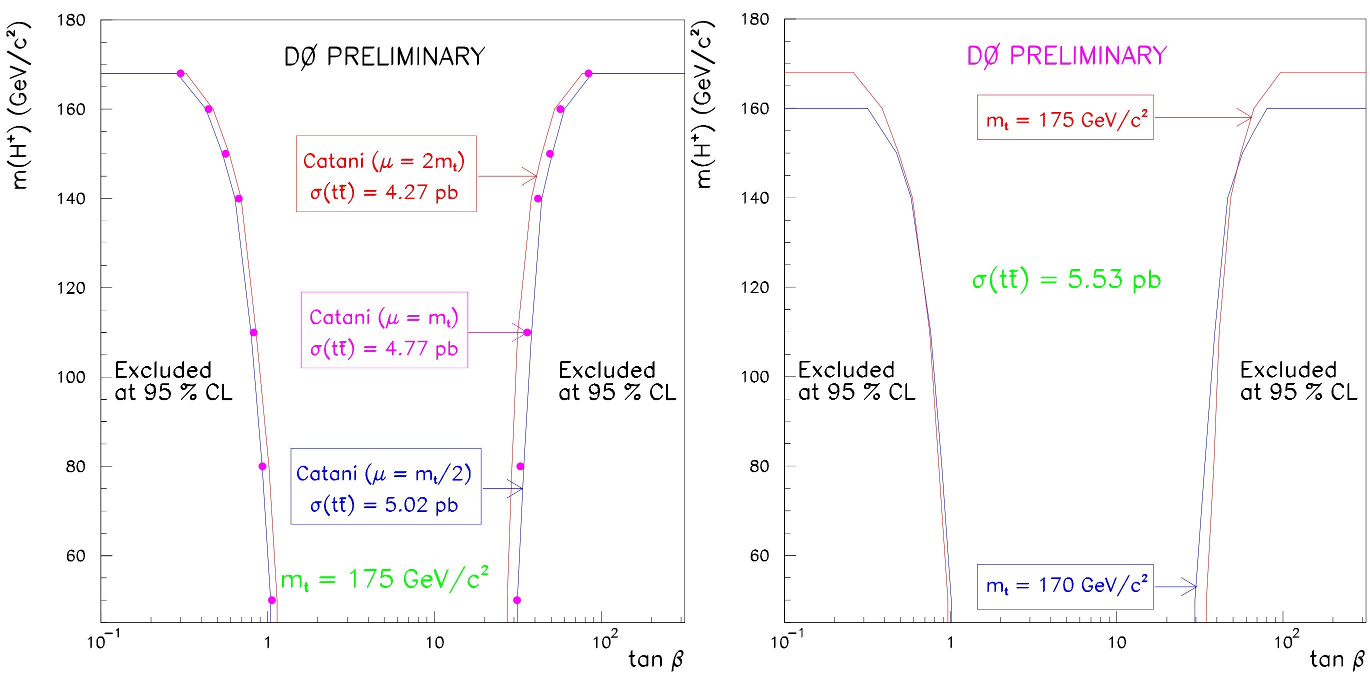




(b) Relative likelihood vs. log tan β for observing 30 events, with the same parameters. The hatched regions are approximately those excluded at 95% CL.

Exclusion region in the $M(H^+)$ vs. tan β plane, varying calculated $\sigma(t\,t)$, μ, and top mass





Summary

- Within the MSSM, taking $m(t) = 175$ GeV/c^2 and $\sigma_{t\bar{t}} =$ 5.53 pb, the l+jets disappearance experiment requires

$$0.96 < \tan\beta \text{ for } M(H^+) = 50 \text{ GeV/c}^2$$
$$0.26 < \tan\beta \text{ for } M(H^+) = 168 \text{ GeV/c}^2 \quad \text{(95\% CL)}$$

-- or --

$$\tan\beta < 35 \quad \text{for } M(H^+) = 50 \text{ GeV/c}^2$$
$$\tan\beta < 96 \quad \text{for } M(H^+) = 168 \text{ GeV/c}^2 \quad \text{(95\% CL)}$$

(preliminary). The dependence on $m(t)$, $\sigma_{t\bar{t}}$, and μ is modest over most of the $M(H^+)$ vs. $\tan\beta$ plane.

- From a neural net analysis of top pairs decaying to all jets, the preliminary top pair production cross section is

$$\sigma_{t\bar{t}} = 7.9 \pm 3.1 \text{ (stat)} \pm 1.7 \text{ (syst) pb.}$$

- The top quark mass is measured to be

$$m(t) = 173.3 \pm 5.6 \text{ (stat)} \pm 6.2 \text{ (syst) GeV/c}^2 \; (l\text{+jets})$$
$$= 168.4 \pm 12.3 \text{ (stat)} \pm 3.7 \text{ (syst)} \; \text{(dilepton)}$$
$$m(t) = 172.0 \pm 5.1 \text{ (stat)} \pm 5.7 \text{ (syst) GeV/c}^2 \; \text{(combined)}$$

- The top pair production cross section is

$$\sigma_{t\bar{t}} = 5.5 \pm 1.8 \text{ pb.}$$

- These results, and many kinematic distributions of top quark pairs, are in embarrassing agreement with the Standard Model.

Summary of D0 Top Physics

- Within the MSSM, taking $m(t) = 175$ GeV/c^2 and $\sigma_{t\bar{t}} = 5.53$ pb, the l+jets disappearance experiment requires

  $0.96 < \tan\beta$ for $M(H^+) = 50$ GeV/c^2
  $0.26 < \tan\beta$ for $M(H^+) = 168$ GeV/c^2 (95% CL)

  -- or --

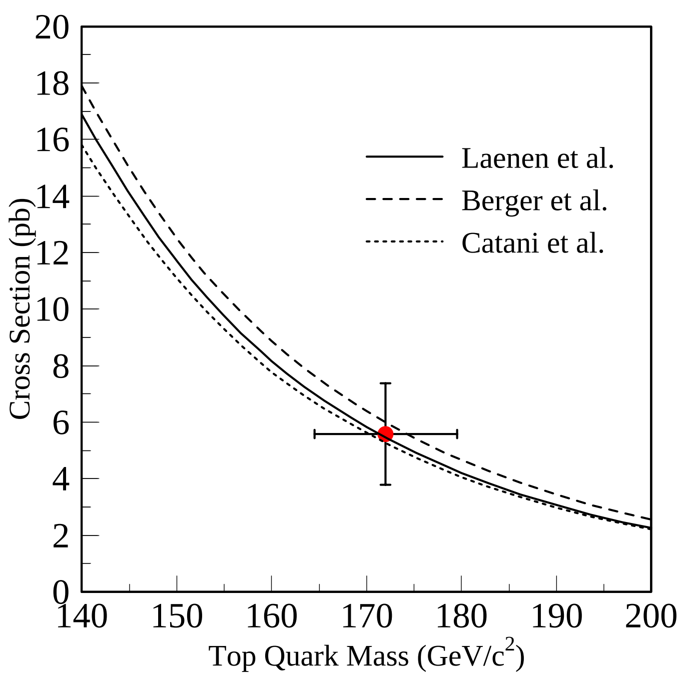
  $\tan\beta < 35$ for $M(H^+) = 50$ GeV/c^2
  $\tan\beta < 96$ for $M(H^+) = 168$ GeV/c^2 (95% CL)

  (preliminary). The dependence on $m(t)$, $\sigma_{t\bar{t}}$, and μ is modest over most of the $M(H^+)$ vs. $\tan\beta$ plane.
- From a neural net analysis of top pairs decaying to all jets, the preliminary top pair production cross section is

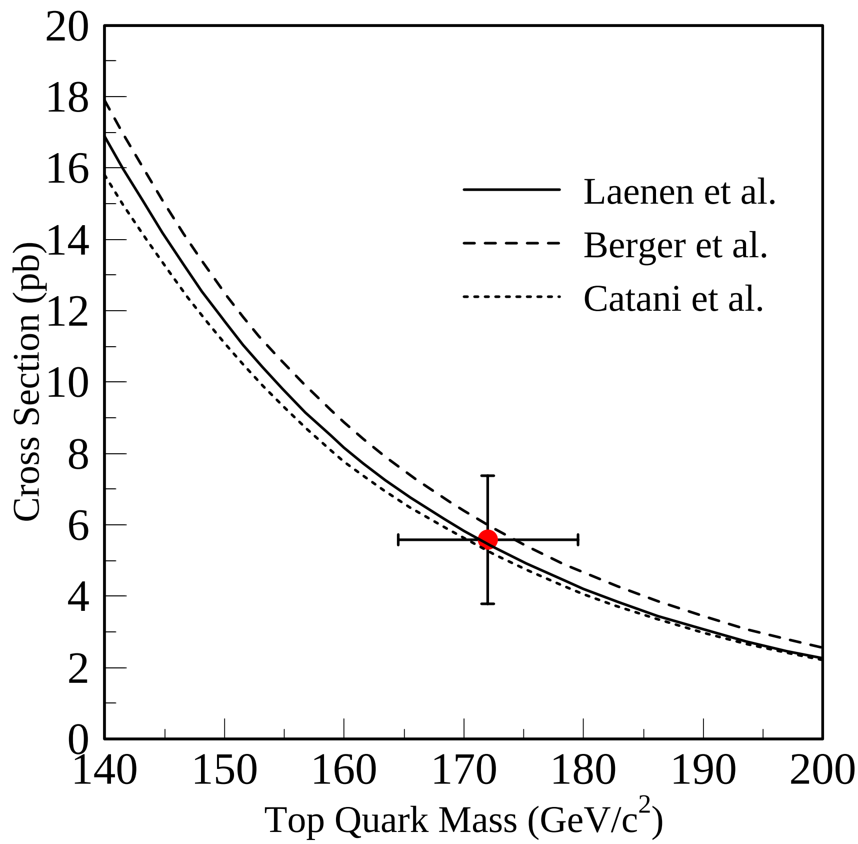
  $\sigma_{t\bar{t}} = 7.9 \pm 3.1$ (stat) ± 1.7 (syst) pb.
- The top quark mass is measured to be

  $m(t) = 173.3 \pm 5.6$ (stat) ± 6.2 (syst) GeV/c^2 (l+jets)
  $= 168.4 \pm 12.3$ (stat) ± 3.7 (syst) (dilepton)
  $m(t) = 172.0 \pm 5.1$ (stat) ± 5.7 (syst) GeV/c^2 (combined)
- The top pair production cross section is

  $\sigma_{t\bar{t}} = 5.5 \pm 1.8$ pb.
- These results, and many kinematic distributions of top quark pairs, are in agreement with the Standard Model.



Summary of D0 Top Physics



- $m(t)$ = 173.3 ± 5.6(stat) ± 6.2(syst) GeV/c^2 (l+jets)
  = 168.4 ± 12.3(stat) ± 3.7(syst) (dilepton)
  $m(t)$ = 172.0 ± 5.1(stat) ± 5.7(syst) GeV/c^2 (combined)
- $\sigma_{t\bar{t}}$ = 5.5 ± 1.8 pb (modes with ≥1 isolated lepton)
  = 7.9 ± 3.1(stat) ± 1.7(syst) pb (all jets, preliminary)
- For $m(t)$ = 175 and $\sigma_{t\bar{t}}$ = 5.53 pb in MSSM (preliminary):
  0.96 < tan β for $M(H^+)$ = 50 GeV/c^2
  0.26 < tan β for $M(H^+)$ = 168 GeV/c^2 (95% CL)
  -- or --
  tan β < 35 for $M(H^+)$ = 50 GeV/c^2
  tan β < 96 for $M(H^+)$ = 168 GeV/c^2 (95% CL)
- These results, and many kinematic distributions of top quark pairs, are in agreement with the Standard Model.