Measurement of the Top Quark Mass at DØ

Serban Protopopescu
Brookhaven National Laboratory
For the DØ Collaboration

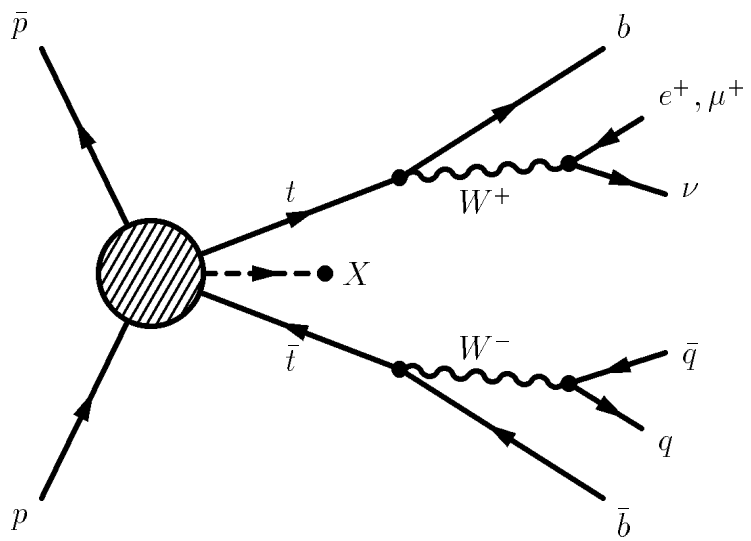
ICHEP Warsaw
July, 1996

DØ TOP MASS ANALYSIS

- Introduction
- Event Selection
- Jet corrections
- Mass calculation
- Top Likelihood Discriminant
  - Likelihood variables
  - Construction of Likelihood Discriminant
- Fit for top mass
- Systematic errors
- Neural Network fits
- Conclusion

Lepton + Jets Mass Analysis

- Use lepton + jets events: $t\bar{t} \rightarrow (bl\nu) + (bjj)$.



- Apply jet corrections.
- Calculate a top mass for each event using constrained kinematic fit.
- Fit top mass distributions to Monte Carlo expectations to extract m_t.
- Signal and background models:
  - Signal: Herwig (v5.7).
  - W+jets background: Vecbos (v3.0).
  - non-W background: data.

Event Selection

- An isolated e or μ, $E_T^l > 20$ GeV and $|\eta_e| < 2.0$ or $|\eta_\mu| < 1.7$
- At least 4 jets found (cone $R = 0.5$), $E_T > 15$ GeV, and $|\eta(jets)| < 2.0$
- $\not{E}_T^{cal} > 25$ GeV for e+jets, $\not{E}_T^{cal} > 20$ for μ+jets
- $\not{E}_T > 20$ GeV

μ-tag events require (8 events):

- non-isolated μ, $p_T > 4$ GeV and $|\eta| < 1.7$

Non μ-tag events require (85 events):

- $E_T^W (\equiv \not{E}_T + E_T^l) > 60$ GeV
- $|\eta_W| < 2.0$

Integrated luminosity ≈ 115 pb^{-1}

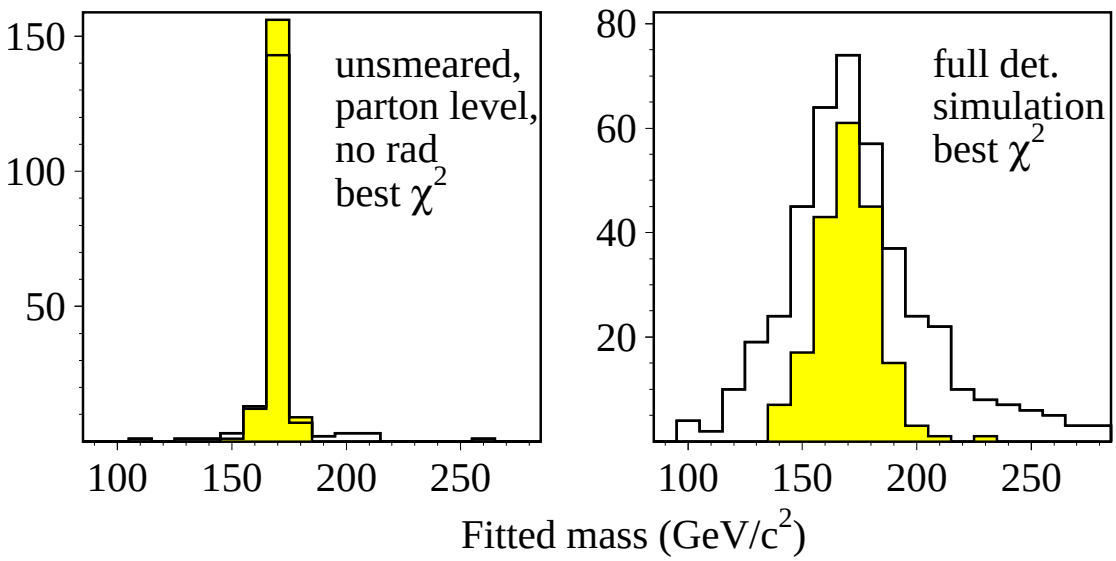
Mass Calculation

- Measured variables (17):
  - Lepton (E, ϕ, η) 3
  - 4 jets (E, ϕ, η) $4 \times 3 = 12$
  - $\vec{\not{E}}_T$ 2
- Unmeasured variables (1):
  - $p_z(\nu)$.
- Constraints 3:
  - $m(l\nu) = M_W$
  - $m(jj) = M_W$
  - $m(l\nu j) = m(jjj)$
- Three constraints and one unmeasured variable gives a 2C fit.

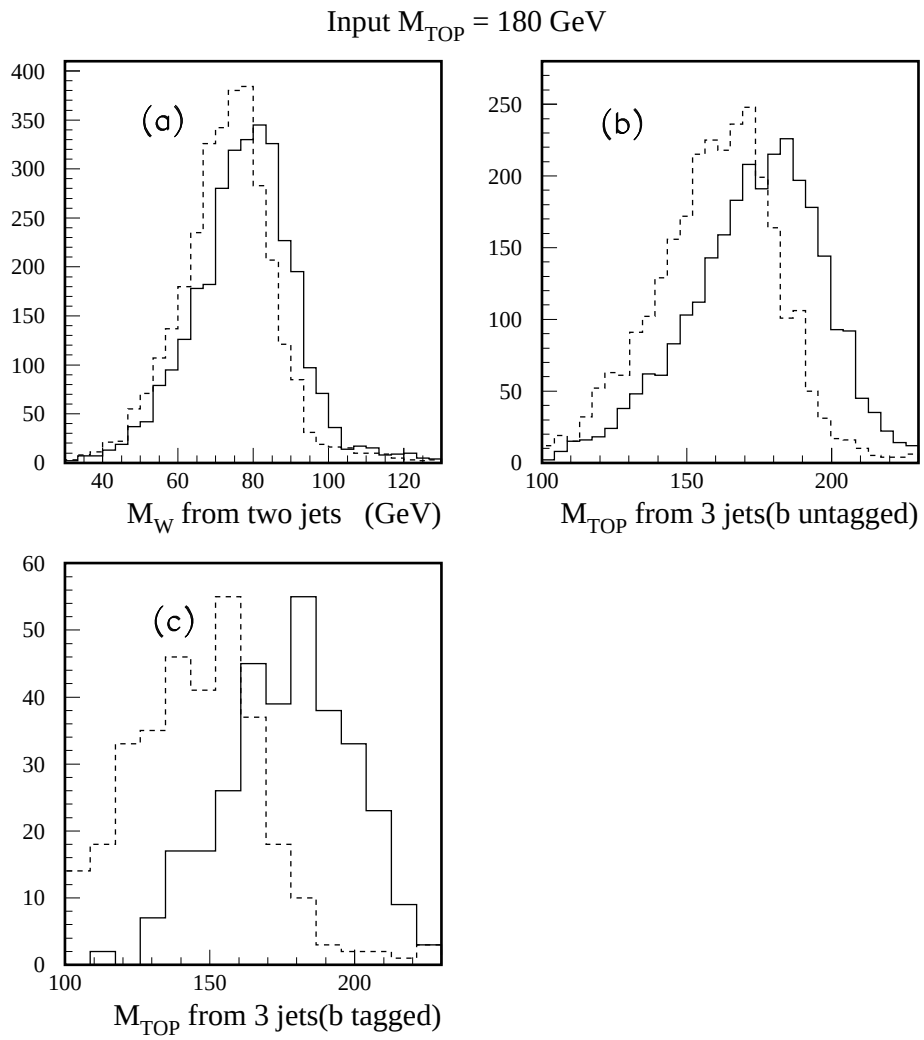
Effects of Radiation

- Width of the fitted mass distribution is due primarily to effects of jet combinatorics and QCD radiation.

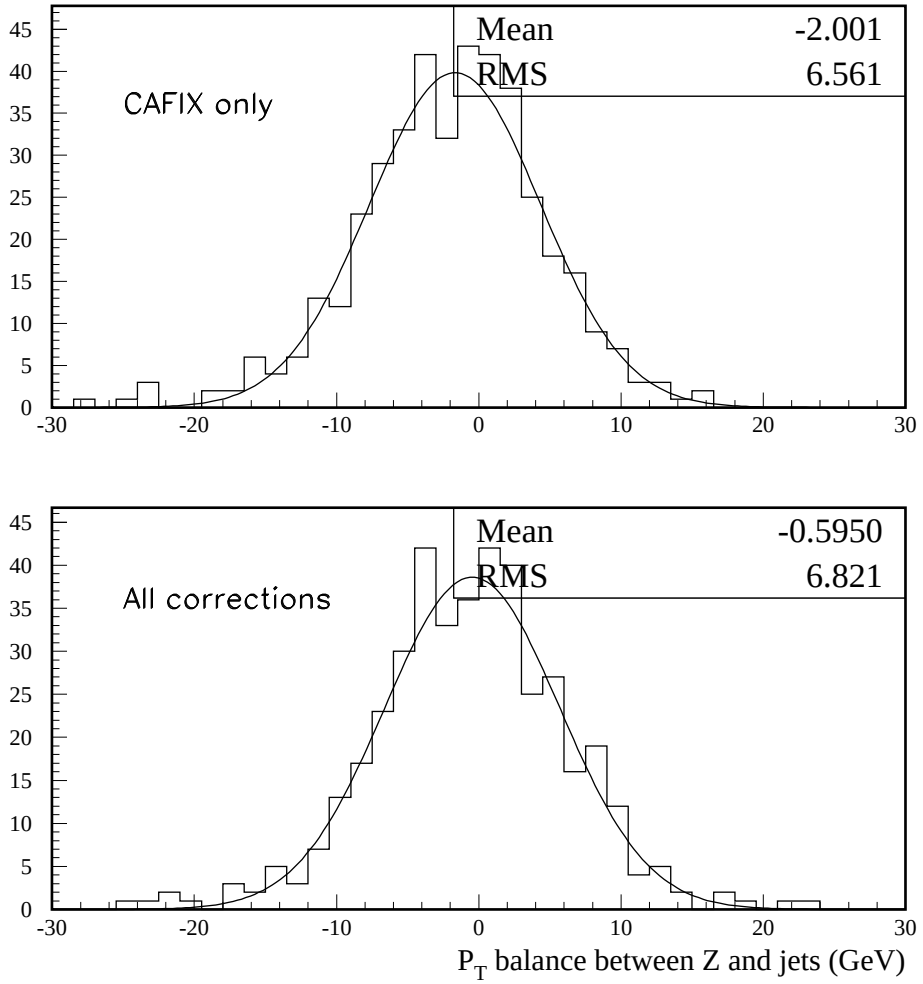
HERWIG $t\bar{t}$ MC with $m_t = 170$ GeV/c^2
(Shaded plots show correct jet permutations)



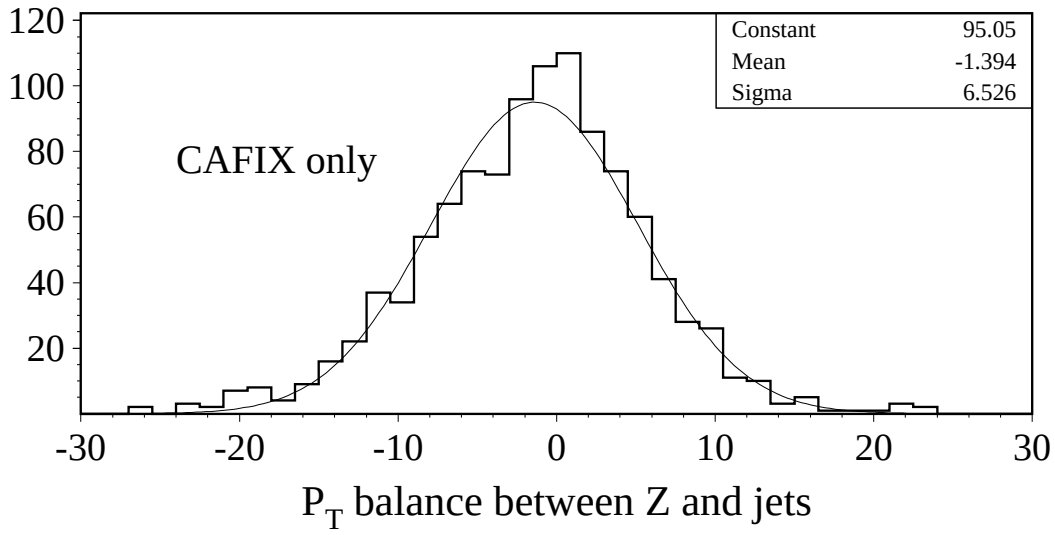
Masses before and after out-of-cone corrections (OOC).
OOC are functions of E_T and η.

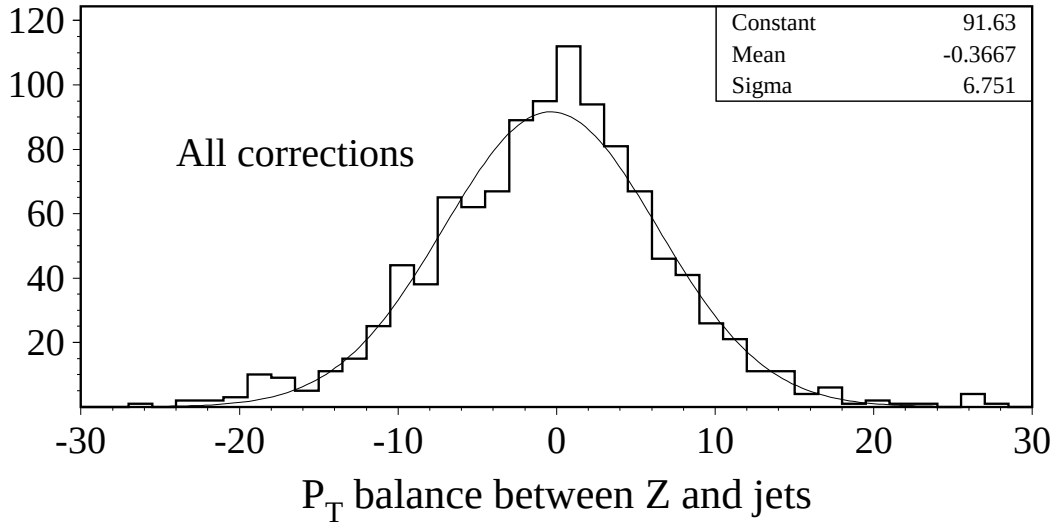


Z+jets balance (DATA)



Z+jets balance (Monte Carlo)





Likelihood variables

Four variables are defined using jets and W:

$v_1 \equiv \not{E}_T$

$v_2 \equiv \mathcal{A} \equiv \frac{3}{2} \times$ least eigenvalue of $\mathcal{P}$

$$v_3 \equiv \frac{H_{T2} \equiv H_T - E_T^{j1}}{H_{||}}$$

$$v_4 \equiv \frac{K_{T\min} \equiv (\text{min of 6 } \Delta\mathcal{R}_{jj}) \cdot E_T^{\text{lesser } j}}{E_T^W} .$$

Where

$\mathcal{A} \equiv$ aplanarity

$\mathcal{P} \equiv$ normalized lab. mom. tensor $H_T \equiv$ scalar sum of jets E_T

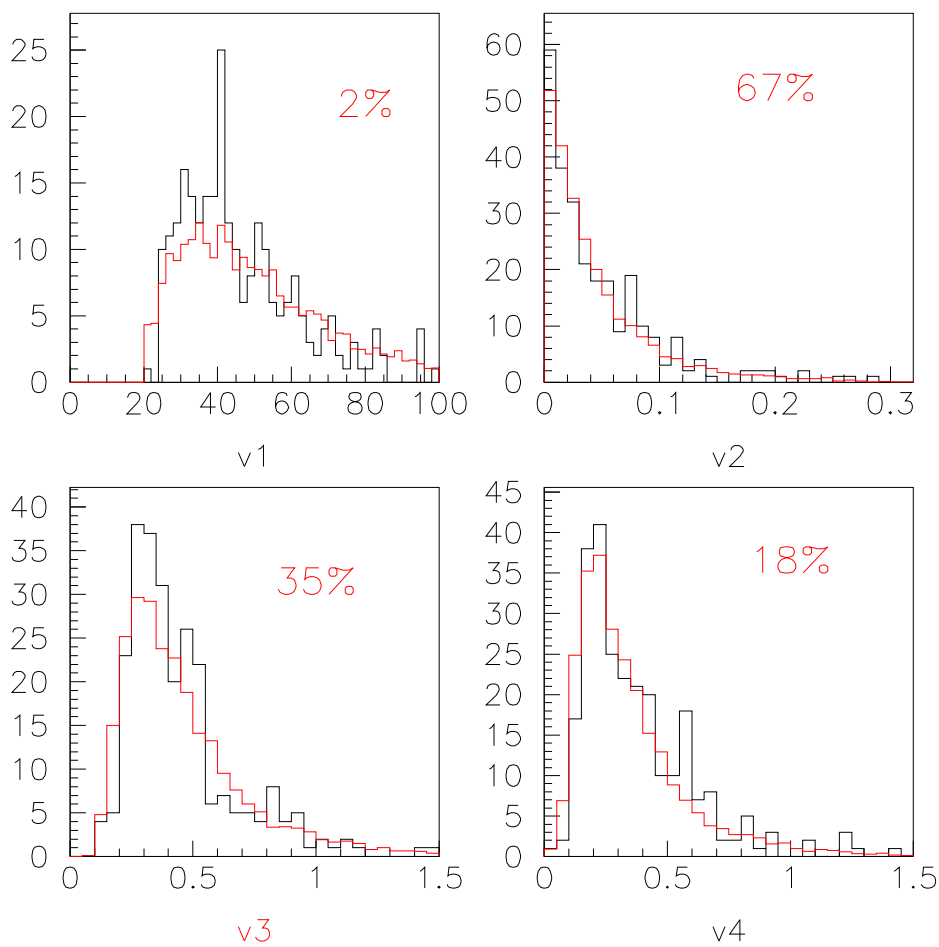
$H_{||} \equiv$ scalar sum of jets & leptons p_z

$E_T^W \equiv$ scalar sum of leptons E_T

Events with 3 Jets

Distributions of $v_1 - v_4$ for data (solid) and prediction (dashed)

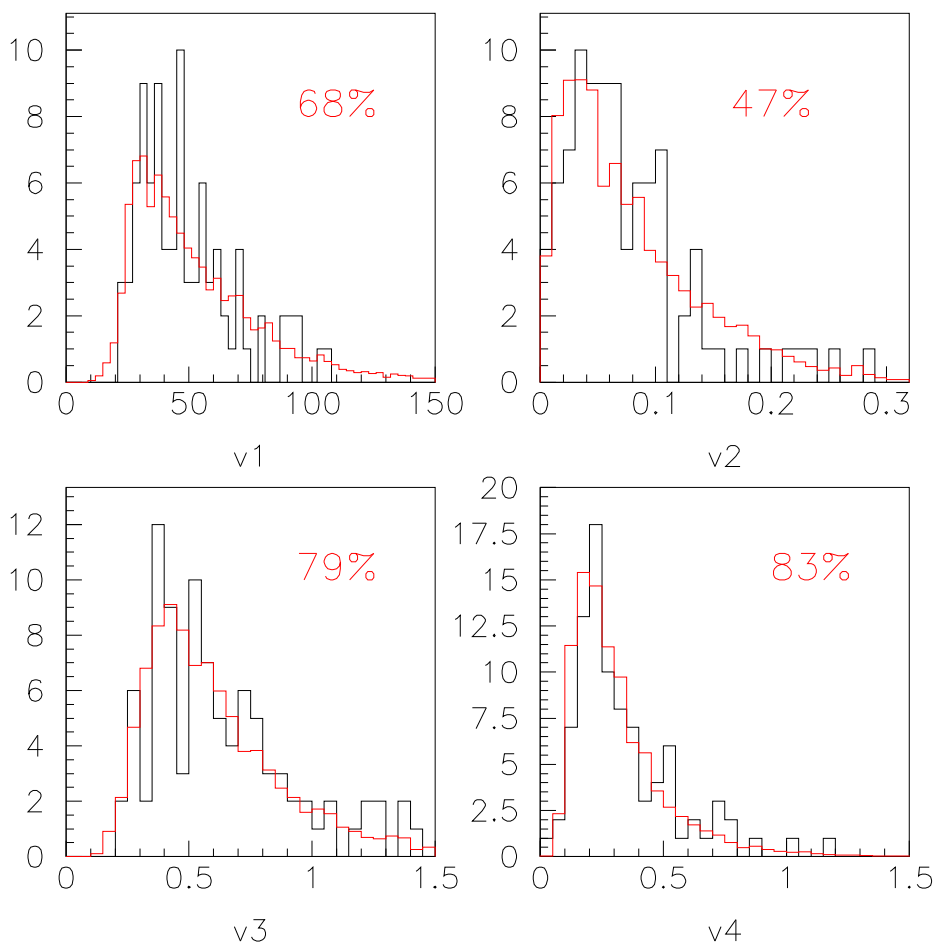
% is Kolmogorov-Smirnov probability



Events with 4 or more Jets

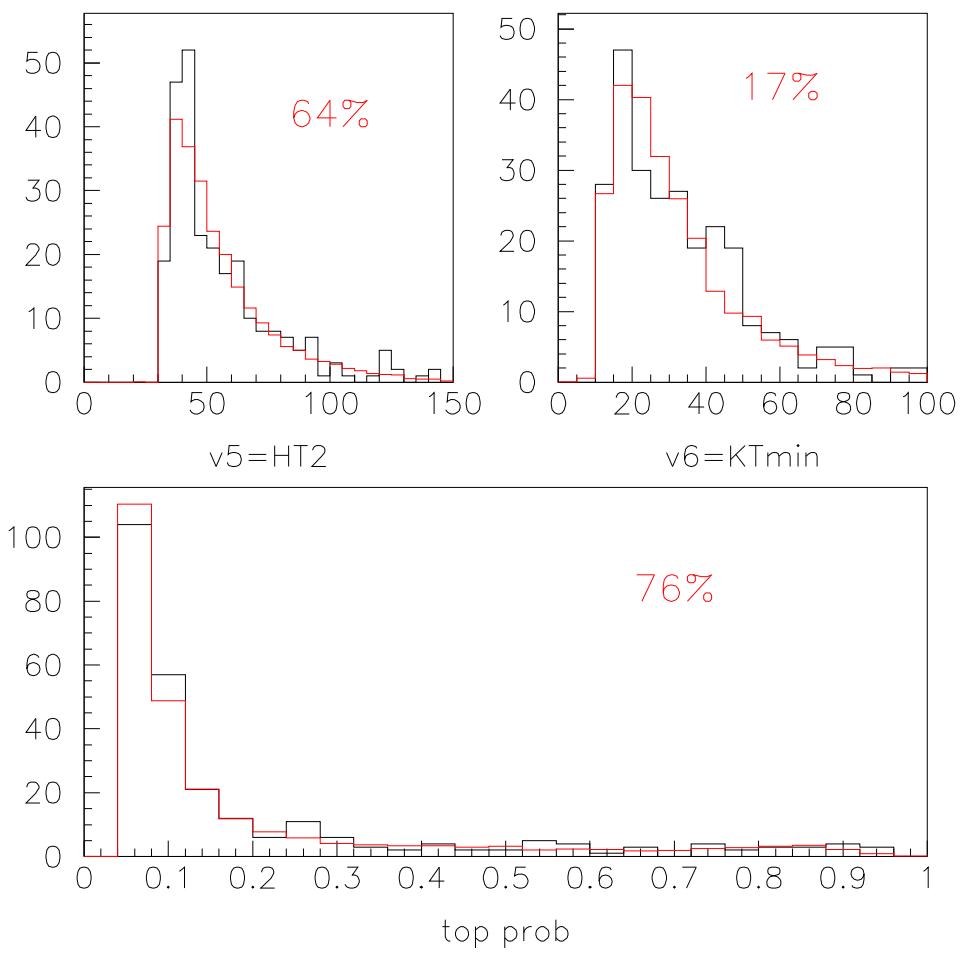
Distributions of $v_1 - v_4$ for data (solid) and prediction (dashed)

% is Kolmogorov-Smirnov probability



Events with 3 Jets

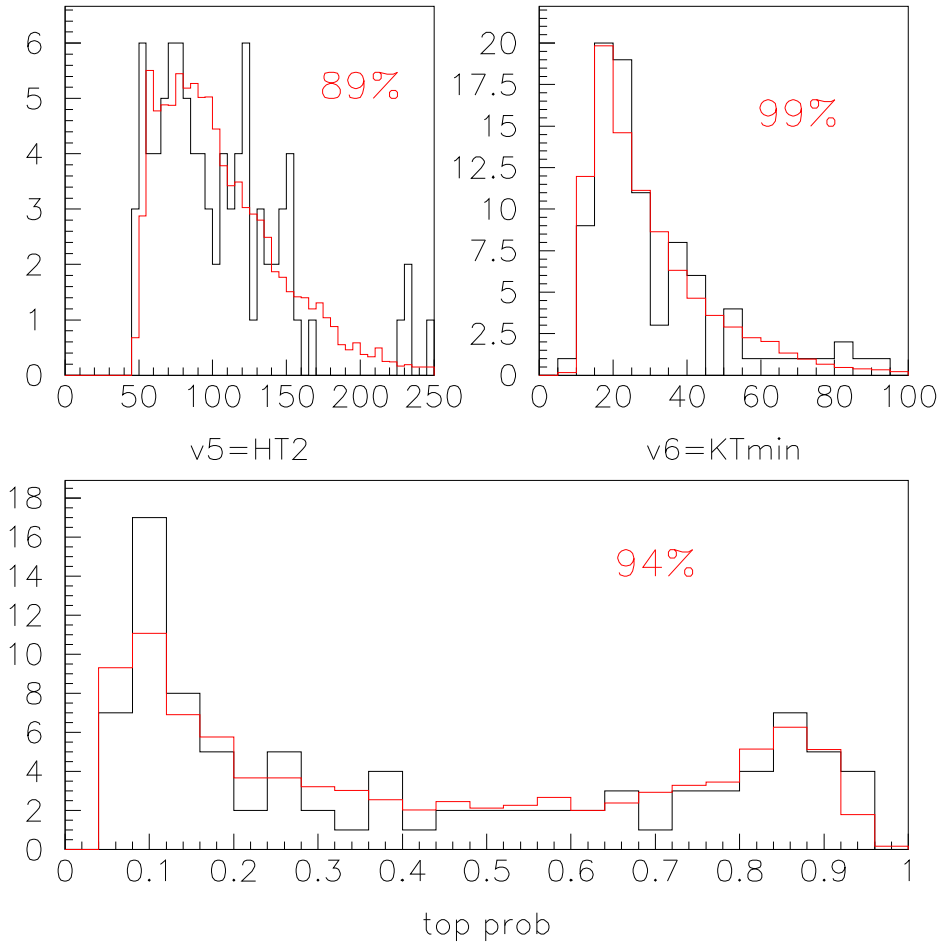
Distributions of v_5, v_6, top_{prob} for data (solid) and prediction (dashed)
% is Kolmogorov-Smirnov probability



Events with 4 or more Jets

Distributions of v_5, v_6, top_{prob} for data (solid) and prediction (dashed)

% is Kolmogorov-Smirnov probability



Top Likelihood Discriminant

For the four variables v_i parametrize with simple polynomials the normalized ratio of distributions:

$$\ln \text{top/bkgnd} \equiv \ln \mathcal{L}_i(v_i)$$

For each events form:

$$\ln \mathcal{L} \equiv \sum_i \omega_i \ln \mathcal{L}_i$$

Normally $\omega_i = 1$, adjusted to null the correlation between $\mathcal{L}$ and top mass.
v_i were chosen so that $\omega_i \approx 1$

We then form a top likelihood discriminant ($\mathcal{D}$):

$$\mathcal{D} \equiv \frac{\mathcal{L}}{1 + \mathcal{L}} \, .$$

Mass Fitting Algorithm

- Use four highest E_T jets for the fit.
- 12 possible ways of assigning jets as $b\bar{b}jj$ times 2 solutions for $p_z(\nu)$
- Enforce b-tag in tagged events
- For each permutation apply OOC jet corrections.
- For each permutation, form a χ^2 using the measured variables $\mathbf{x}^m$ and their inverse error matrix $\mathbf{G}$:

  $$\chi^2 = (\mathbf{x} - \mathbf{x}^m)^T \mathbf{G} (\mathbf{x} - \mathbf{x}^m).$$

  Find $\mathbf{x}$ to minimize this quantity subject to the kinematic constraints.
- Pick solution with smallest χ^2 and require $\chi^2 < 7$.

Top Mass Likelihood

Use binned Poisson-statistics for discrete top masses (m_t).
M number of bins,
N_j events in bin j,
A_j^s (A_j^b) number of Monte Carlo signal (backg.) events,
a_j^s (a_j^b) 'true' distribution for signal (backg.).

$L_j(D|A, a, p, m_t) =$
$q(N_j, p_s a_j^s + p_b a_j^b) q(A_j^s, a_j^s) q(A_j^b, a_j^b)$

$$L = \prod_{j=1}^{M} L_j$$

where $q(N, a)$ is the Poisson distribution $q(N, a) = e^{-a} a^N / N!$

Top Mass Likelihood (cont'd)

p_s and p_b are signal and background strengths.

$$p_s = n_s / (\sum_j A_j^s)$$

where n_s is number of signal events. To constrain n_s (or n_b) to some value μ with width σ multiply L by

$$e^{-(n_s-\mu)^2/2\sigma^2}$$

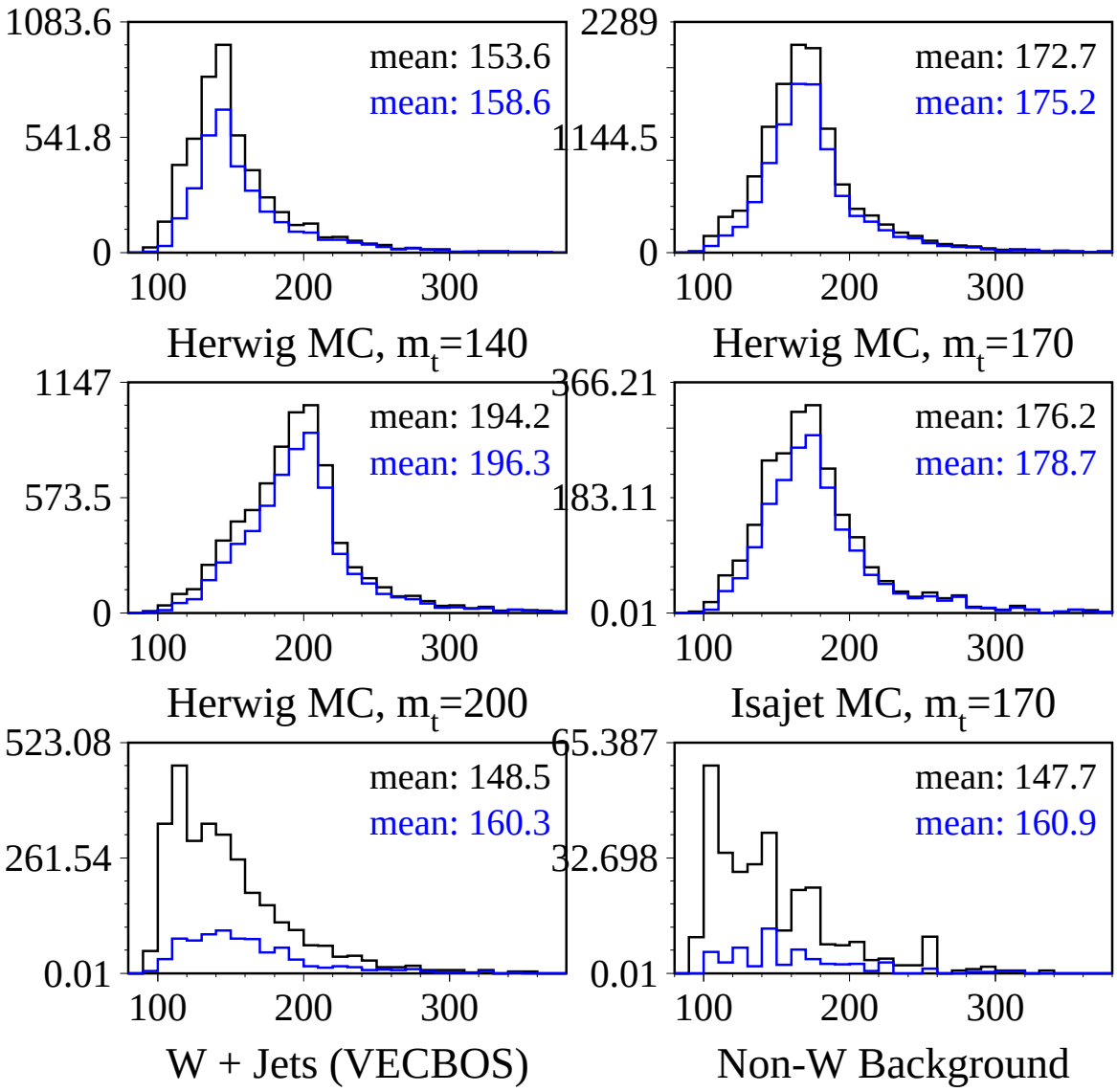
Fitting Procedure:

- find n_s and n_b that maximizes lnL for a given m_t.
- fit lnL to find best m_t and error on m_t.

MC Top Mass Distributions

Best χ^2 solution plotted

- PR - LB



LB cuts produce no significant mass bias

Likelihood Binning

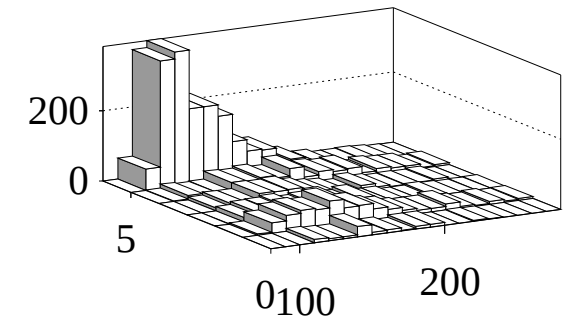
For 2D divide each mass bin into 6 bins:

Bin 1: Tagged events.
Bin 2: $H_{T2} \geq 90\,\mathrm{GeV}$, $0.59 \leq \mathcal{D}$
Bin 3: $H_{T2} \geq 90\,\mathrm{GeV}$, $0.43 \leq \mathcal{D} < 0.59$

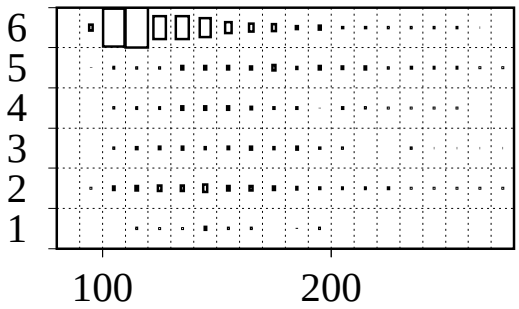
Bin 4: $H_{T2} \geq 90\,\mathrm{GeV}$, $0.27 \leq \mathcal{D} < 0.43$
Bin 5: $H_{T2} \geq 90\,\mathrm{GeV}$, $\mathcal{D} < 0.27$
Bin 6: $H_{T2} < 90\,\mathrm{GeV}$.

Bins 2-6 contain only untagged events
Bins 1-6 are pre-cut (PR) sample
Bins 1-3 are low bias (LB) sample

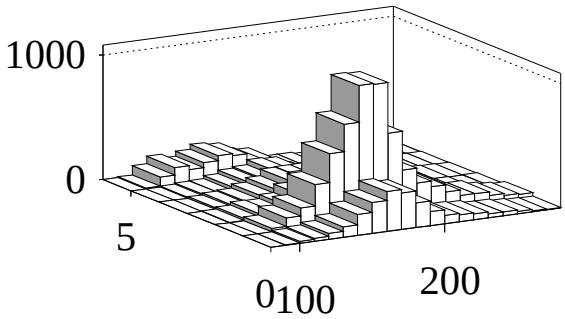
Distributions of top mass vs Likl. bins



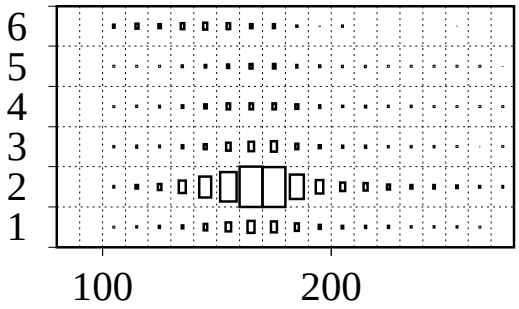
Background



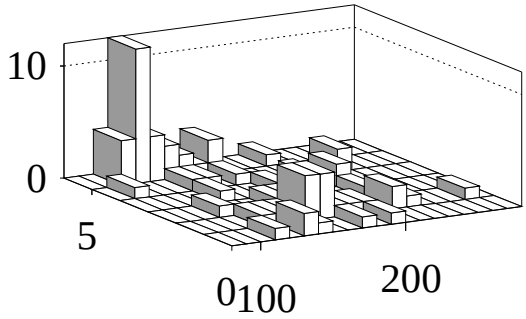
Background



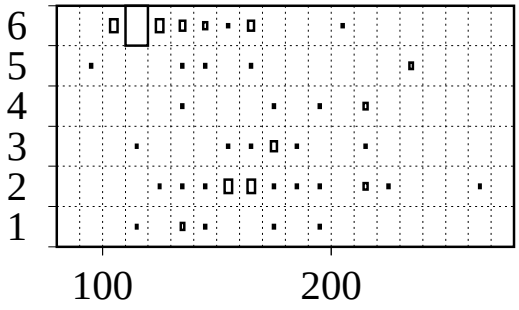
Herwig MC, m_t=170



Herwig MC, m_t=170



Data



Data

Monte Carlo samples

Generated high statistics sample in a fine grid of top masses and a large background sample:

- HERWIG v5.7
  - 50K events every 5 GeV between 130 and 220 GeV
  - 100K events at 165, 170 and 175 GeV
  - 50K events at 110, 120, 230 and 240 GeV
- ISAJET v7.15
  - 30K events every 10 GeV between 130 and 220 GeV
- VECBOS v3.0 (+ISAJET for fragm.)
  - 70K events at 2 different scales

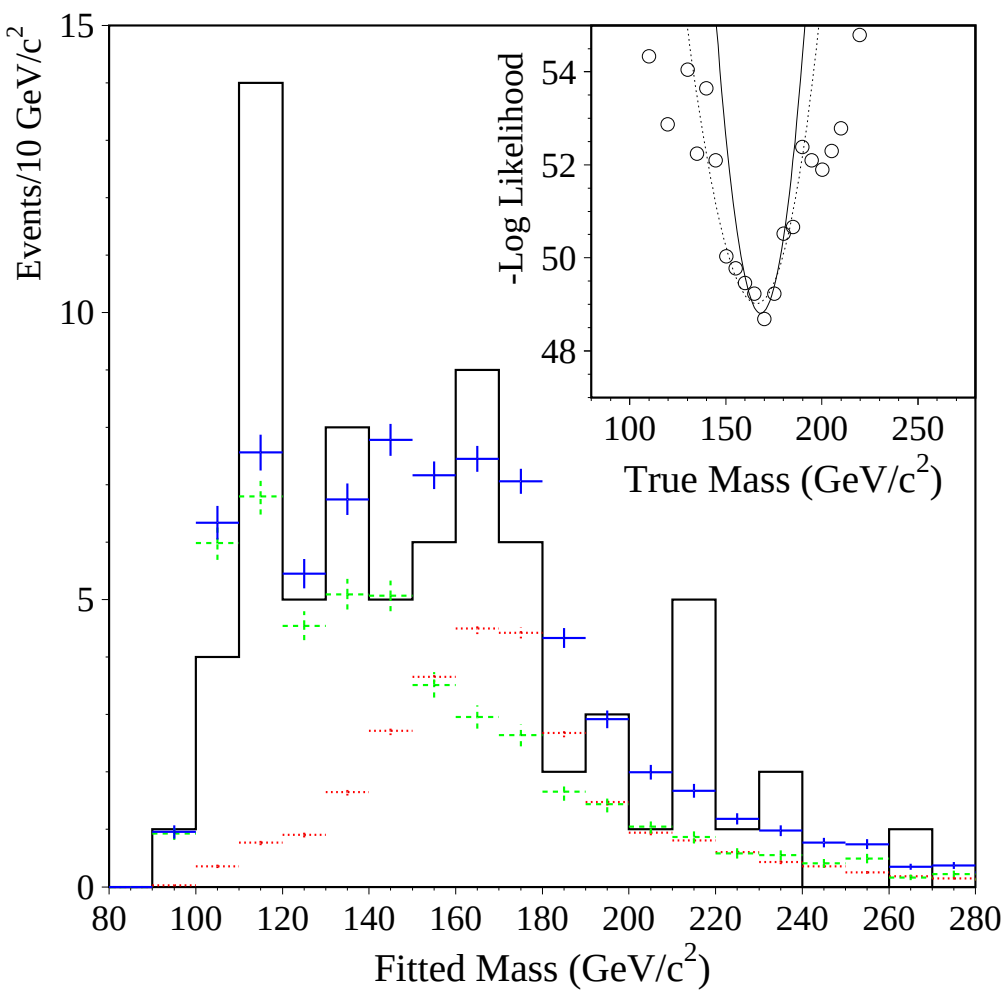
DØ Unconstrained Fit

2D fit to 74 events (PR sample)

$n_s = 27 \pm 7,\ n_b = 44 \pm 9,\ \chi^2/ndf = 49/39$

$m_t = 168 \pm 10(stat.)$

expected n_s from $\sigma_{t\bar{t}}$: 33 ± 9.



+ background, + signal, + sum

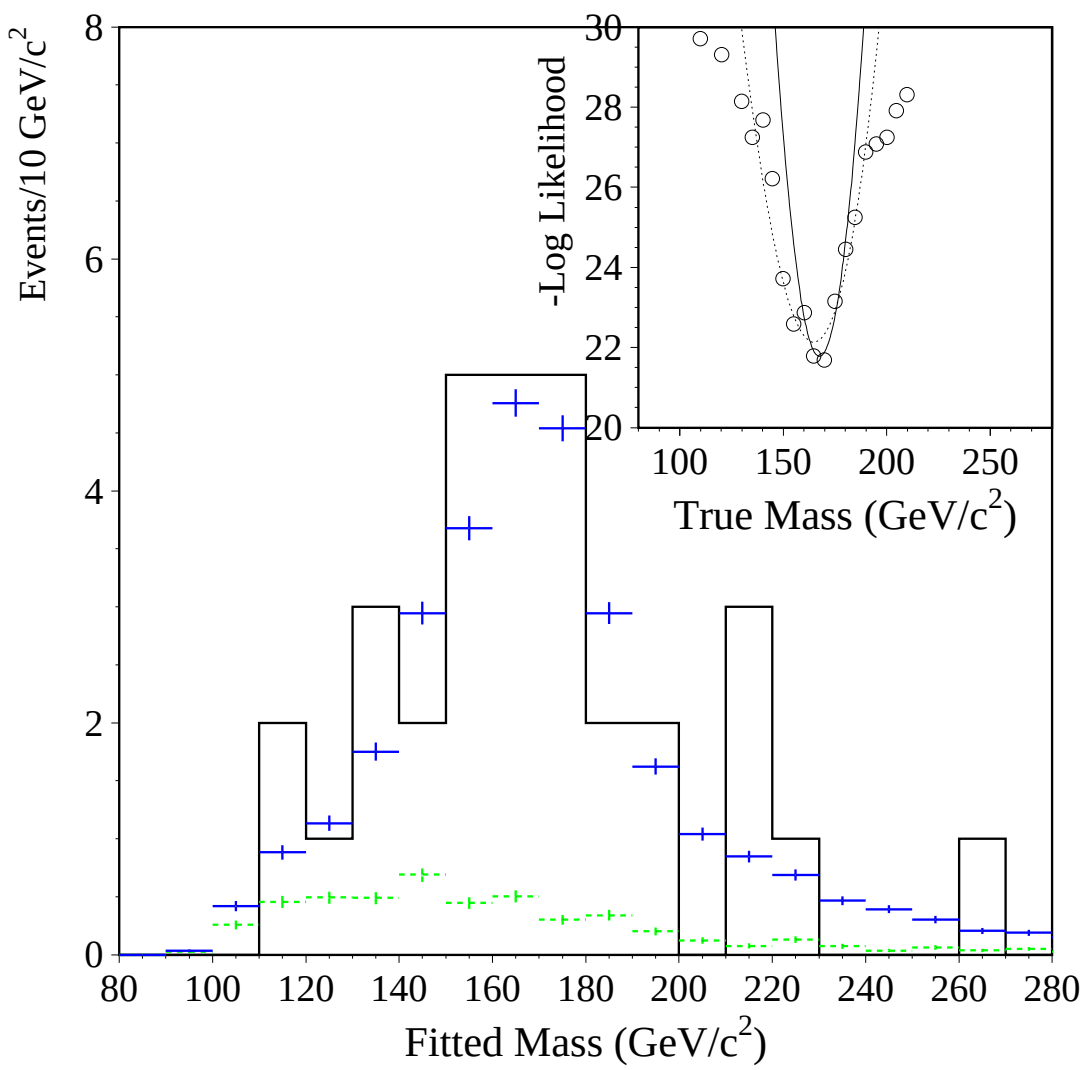
DØ Mass Fit

preliminary

2D fit to 32 events (LB sample)

constrain signal by PR fit

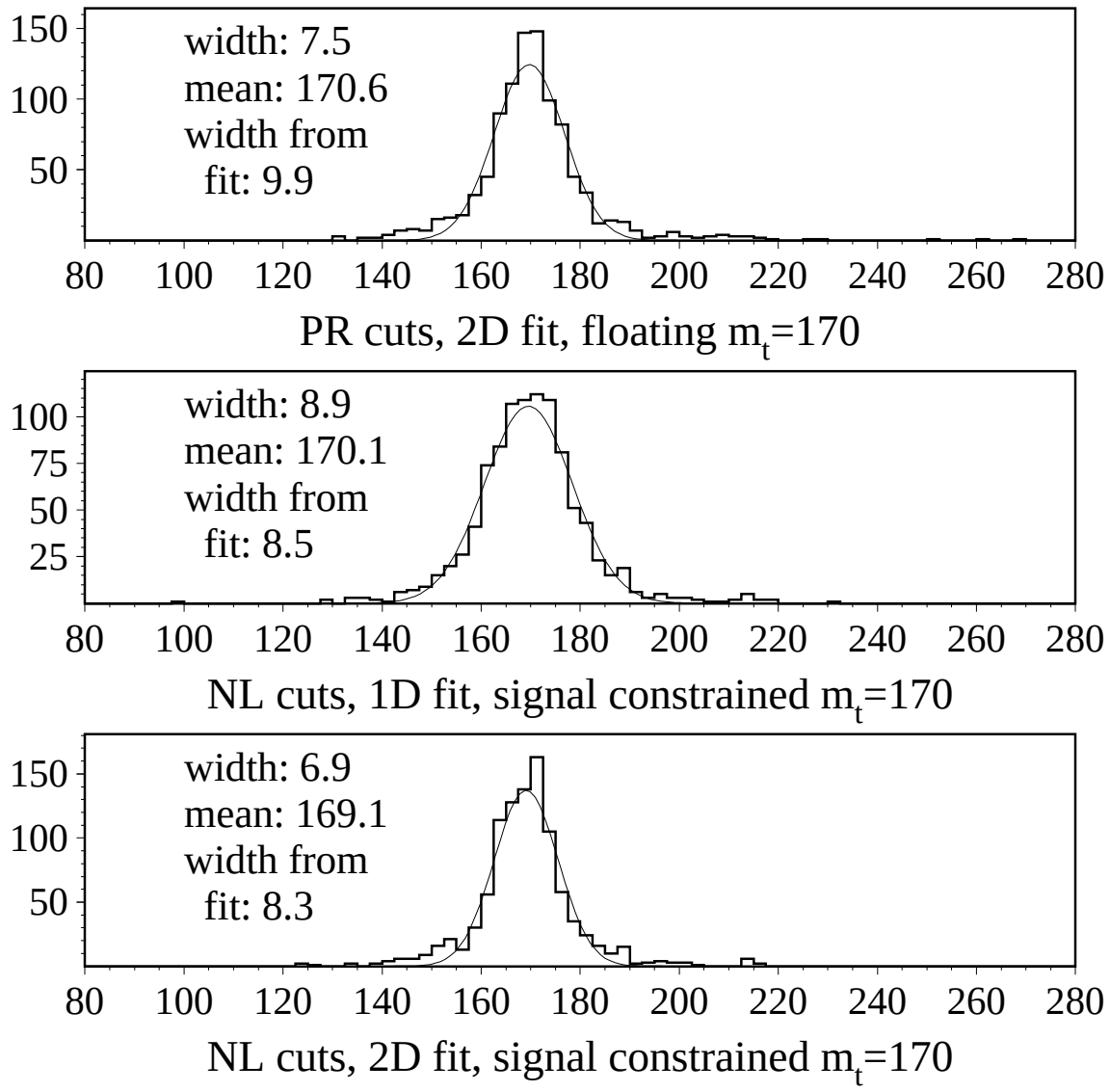
$m_{top} = 168 \pm 8(stat.),\ \chi^2 = 21.7/22$



+ background, + signal+background

DØ Mass Fit

Ensembles Fit Distributions



Systematic Errors

The dominant systematic errors come from uncertainties in jet energy corrections and Monte Carlo model.

- jet energy corrections ± 7.3 GeV/c^2
- Monte Carlo model ± 3.3 GeV/c^2
  - HERWIG- ISAJET ± 3.3 GeV/c^2
  - VECBOS scale ± 0.5 GeV/c^2
- **Fitting method** ± 2.0 **GeV/c^2**

Total systematic error ± 8 GeV/c^2

Add 1 GeV to m_t for difference between ensemble average and input m_t.

$$\boxed{m_t = 169 \pm 8(\text{stat}) \pm 8(\text{syst})\ \text{GeV}/c^2}$$

Neural Network Analysis

Use two neural networks to construct a top probability (top_{prob}).

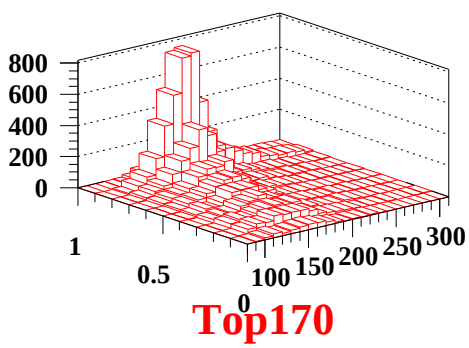
- NN1 uses $v_1 - v_4$, output $n1$
- NN2 uses $v_5 - v_6$, output $n2$
  - $v_5 \equiv H_{T2}$
  - $v_6 \equiv K_{Tmin}$

  define $x = n1 \times n2$

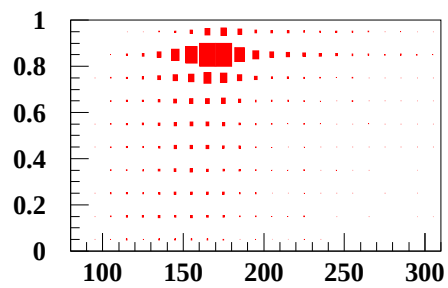
$$top_{prob} \equiv \frac{s(x)}{s(x) + b(x)}$$

where $s(x)$ and $b(x)$ are signal and background distributions of x.
Use 10 bins in top_{prob} for every mass bin.

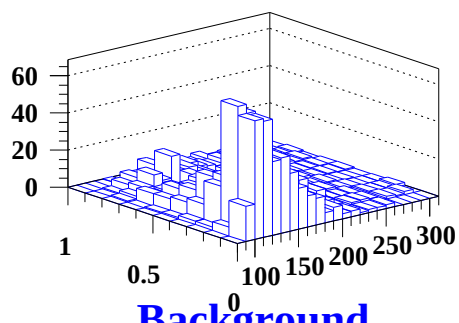
Top masses vs top_{prob}



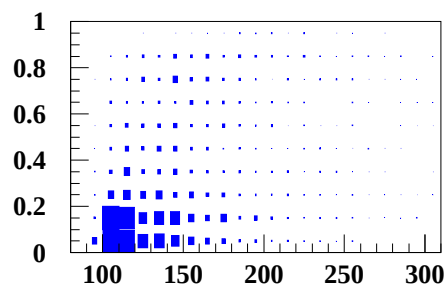
Top170



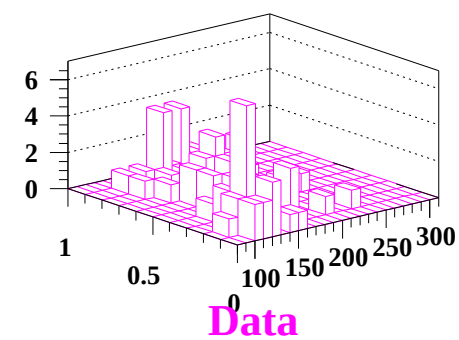
Top170



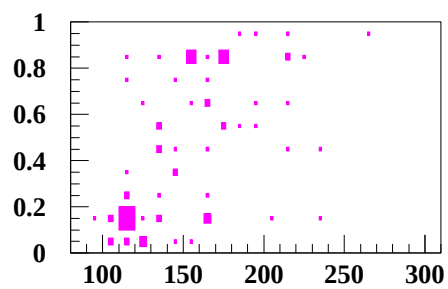
Background



Background

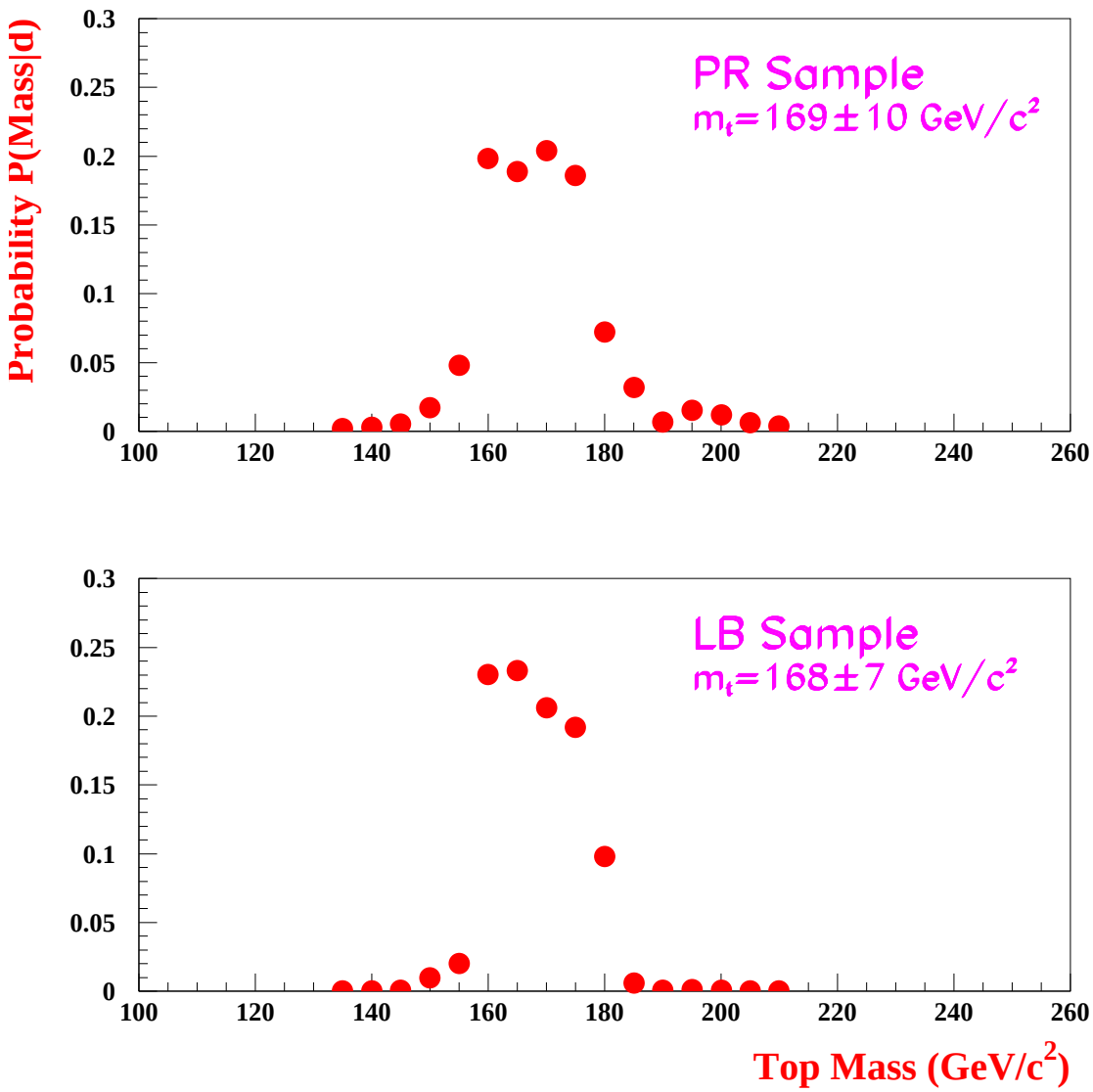


Data



Data

Top mass probability as function of true top mass



CONCLUSION

We used multivariate techniques to construct discriminants between signal and background that are weakly correlated with top mass. This allowed us to:

- Fit for background/signal ratios
- Select an event sample with small background and small mass bias
- Significantly reduce the statistical error on the mass determination.

Estimated top mass

$$\boxed{m_t = 169 \pm 8(\text{stat}) \pm 8(\text{syst})\,\text{GeV}/c^2}$$

(preliminary)