New preliminary measurement of the mass of the top quark at DØ using Run I data

Juan Estrada - Fermilab

for the DØ Collaboration

W&C Seminar

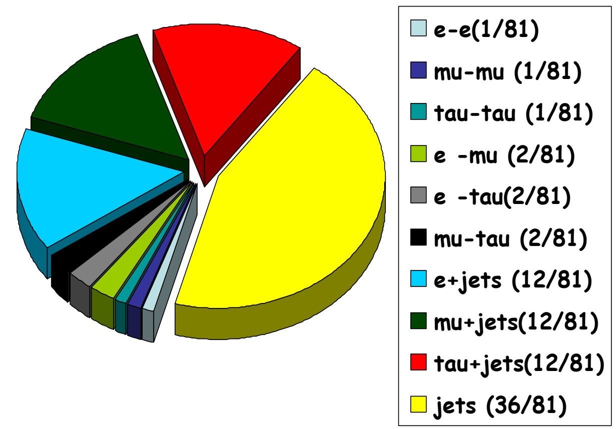
April 25, 2003

Overview

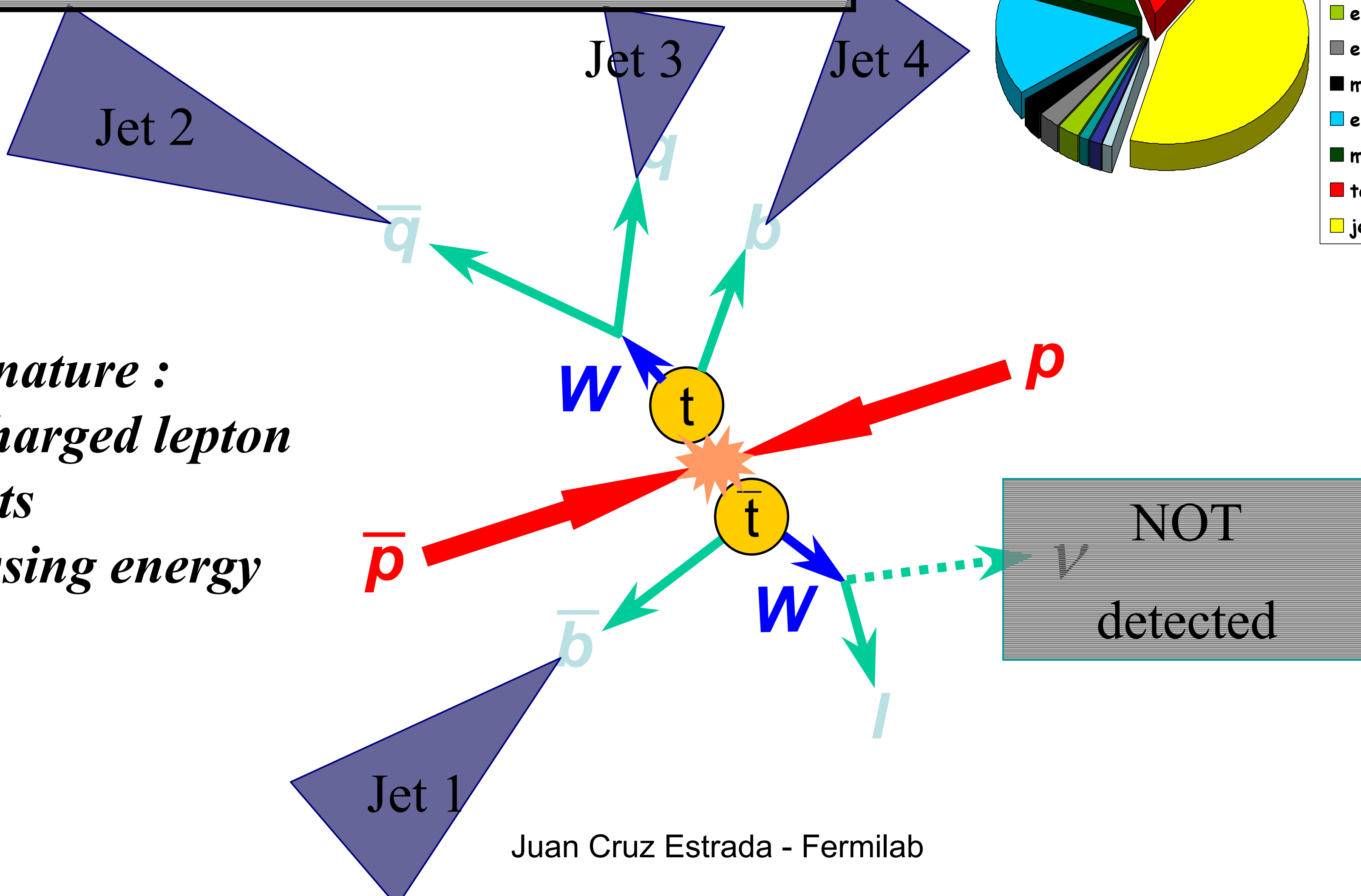
- The lepton+jets decays of the top quark
- Introduction to the measurement of M_t
- Method used for this re-analysis
- Testing the analysis in simulated samples
- New preliminary Run I M_t measurement
- The mass of the W boson in the same sample
- Systematic uncertainties (JES)
- Conclusion

Lepton + jets channel

30% of the total
significantly less background that the all jets channel.



Signature :
1 charged lepton
4 jets
Missing energy



Lepton+jets channel

DØ Statistics Run I : 125 pb^{-1}

Standard Selections :

- Lepton: E_t>20 GeV,$|\eta^e|<2$,$|\eta^\mu|<1.7$
- Jets: ≥4, E_T>15 GeV, $|\eta|<2$
- Missing E_T > 20 GeV
- "E_T^W"> 60 GeV ; $|\eta_W|<2$

→ **91 events**

Ref. PRD 58 (1998), 052001:

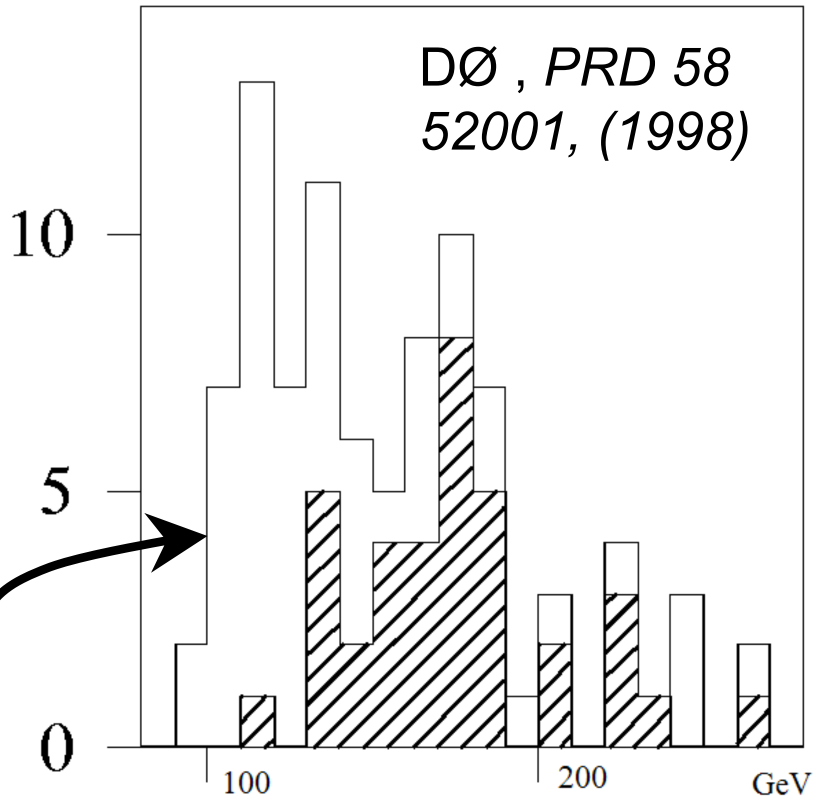
After χ^2(77 events): ~29 signal + ~48 backg.

(0.8 *W+jets* and 0.2 QCD)

Specific cuts for this analysis:

- **4 Jets only :** → **71 events**
- **Background Prob. :** → **22 events**

Top Mass



fitted mass

standard selection

For those of you who did not try to measure M_t…, this is how the mass distributions looks like.

It is a challenging problem and that is why we have been applying **sophisticated methods** making good use all the information that we have.

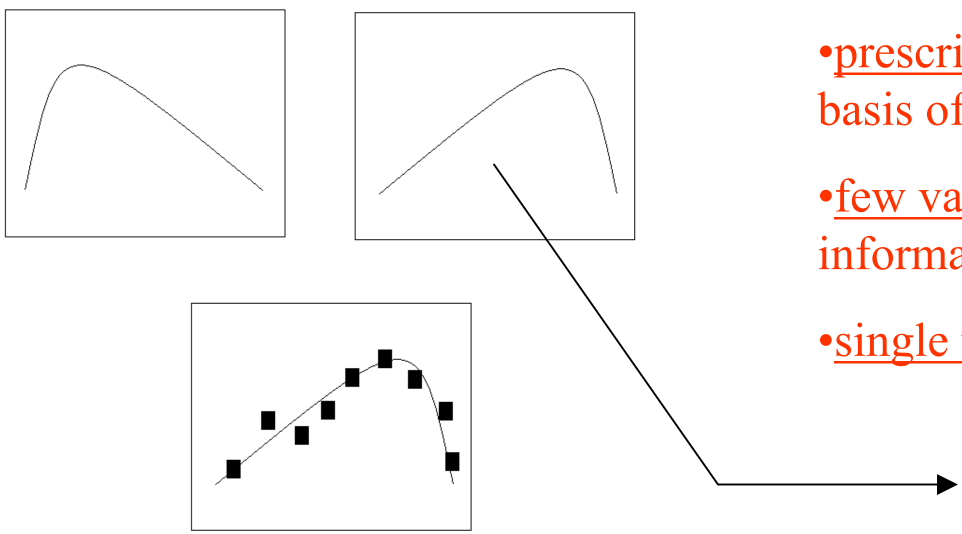
Template method

Previous DØ and CDF publications

<u>Reducing the dimensionality of the problem</u>

A multidimensional (x_i) template is obtained for each value of the input mass, and the data sample is then compared with those MC templates to find the most likely value for $\boldsymbol{M_t}$:

Template(x_i;$\boldsymbol{M_t}$=A) Template(x_i;$\boldsymbol{M_t}$=B)



Sample probabilities

Data => $\boldsymbol{M_t}$~B

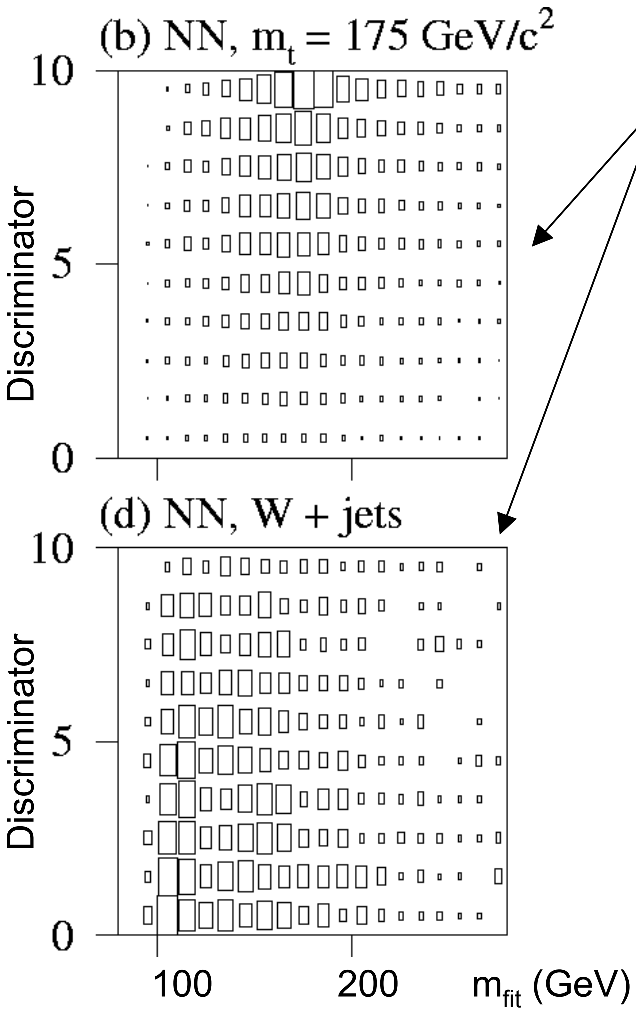
Some limitations:

•<u>prescribed permutation</u> is selected on basis of a kinematic fit.

•<u>few variables</u>, containing most of the information, are selected for the templates.

•<u>single template</u> fits the whole sample.

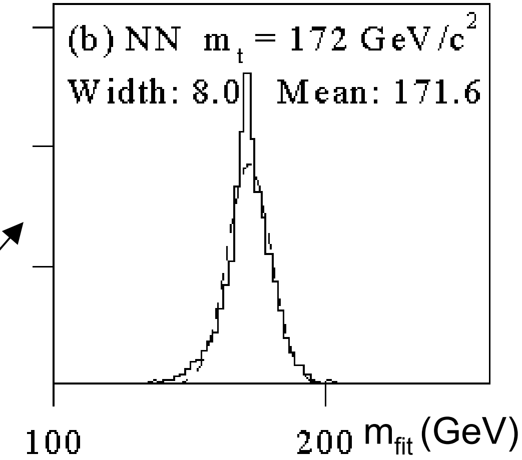
DØ measurement using templates
PRD 58 52001, (1998)

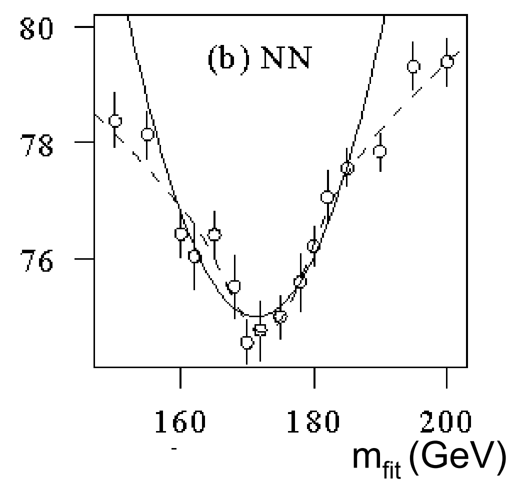


2D templates (NN, fitted mass)

MC studies give 8 GeV resolution

Fit to data sample





Measurement of M_t, two alternatives

Maximize:

$$L(M_t) = L(M_t; x_1 \ldots x_N)$$

using templates

$$L(M_t; x_1 \ldots x_N) = \prod_{i=1}^{N} T(M_t; m_i)$$

But, if possible, it will be better to calculate a probability for each individual event

$$L(M_t; x_1 \ldots x_N) = \prod_{i=1}^{N} P_i(M_t; x_i)$$

Simple case with several types of events, each type follows a Gaussian distribution with width σ_k and mean M_0.

$$\overline{M_t} = \sum W(m_i) \cdot m_i$$

event weight depends on m_i

$$\overline{M_t} = \frac{\sum (m_i / \sigma_i^2)}{\sum (1/\sigma_i^2)}$$

correct weighted average!!

Measurement of M_t, two alternatives

If you have to choose between templates and event probabilities, consider the following points:

- The **the greater the variation between the events the greater the difference in the two methods**

- For **templates you could actually introduce larger fluctuation if you add to your sample low quality events** (large σ_k), this will never happen for the correctly weighted average.

..note that there is no reason why analysis of same sample with two different methods should give you the same result, because methods are not totally correlated.

Measurement of M_t using event probability

…ok, the event probabilities are better.

How do we calculate the events probabilities?

(for templates, we just use MC simulation of the detector)

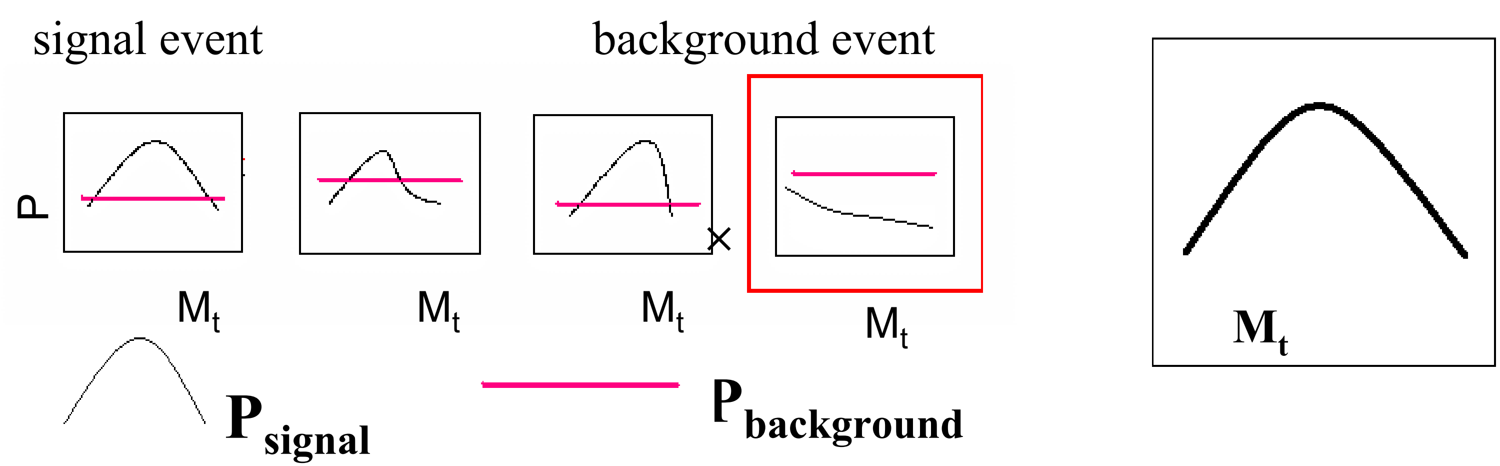
The standard model predicts these probabilities (differential cross section) in terms of the parton level quantities (four-vectors of all partons involved), *but we do not have access to parton level quantities for events in our data sample.*

We will use only LO calculations, and I will show how far we can get with that. I will also show some problems due to our LO approximation…

Measurement of M_t using event probability

(before we get into de details)

The probability for each event being signal is calculated as a function of the top mass. The probability for each event being background is also calculated. The results are combined in one likelihood for the sample. (Similar to the methods of Dalitz, Goldstein and Kondo, M_t measurement in the dilepton channel by DØ - PRD **60** 52001 (1999) and idea by Berends et al for W^+W^- production.)



Three differences between the two approaches

Template Method

1. All the events are presented to the **same template**. Average probability distribution.
2. The template corresponds to a probability distribution for the entire sample, using **selected variables** calculated from MC simulations.
3. The **features of individual events are averaged** over the variables not considered in the template.

This analysis

1. Each event has its **own probability distribution.**
2. The probability depends on **all measured quantities** (except for unclustered energy).
3. Each event contributes with its own specific features to the probability, which **depends how well is measured**.

Calculation of signal probability

If we could **access all parton level quantities** in the events (the four momentum for all final and initial state particles), then we would simply **evaluate the differential cross section as a function of the mass of the top quark for these partons**. This way we would be using our best knowledge of the physics involved.

Since we do not have the parton level information for data, we use the differential cross section and **integrate over everything we do not know**.

$$P_{t\bar{t}}(x) = \frac{1}{\sigma_{tot}} \int d\sigma(y) dq_1 dq_2 f(q_1) f(q_2) W(x,y)$$

Transfer function W(x,y)

W(x,y) probability of measuring **x** when **y** was produced (***x*** jet variables, ***y*** parton variables):

$$W(x,y) = \delta^3(p_e^y - p_e^x)\prod_{j=1}^{4} W_{jet}(E_j^y, E_j^x)\prod_{i=1}^{4}\delta^2(\Omega_i^y - \Omega_i^x)$$

where

E^y	energy of the produced quarks
E^x	measured and corrected jet energy
$p^y{}_e$	produced electron momenta
$p^x{}_e$	measured electron momenta
$\Omega^y{}_j$, $\Omega^x{}_j$	produced and measured jet angles

Energy of electrons is considered well measured, an extra integral is done for events with muons. Due to the excellent granularity of the D∅ calorimeter, angles are also considered as well measured. A sum of two Gaussians is used for the jet transfer function (W_{jet}), parameters extracted from MC simulation.

Probability for $t\bar{t}$ events ("dσ")

$$P_{t\bar{t}} = \frac{1}{\sigma_{tot}} \int d\rho_1 dm_1^2 dM_1^2 dm_2^2 dM_2^2 \sum_{comb,\nu} |M|^2 \frac{f(q_1)f(q_2)}{|q_1||q_2|} \phi_6 W_{jet}(x,y)$$

2(in) + 18(final) = 20 degrees of freedom

3(e) + 8(Ω1..Ω4) + 3(P_{in}=P_{final})+1(E_{in}=E_{final}) = 15 constraints

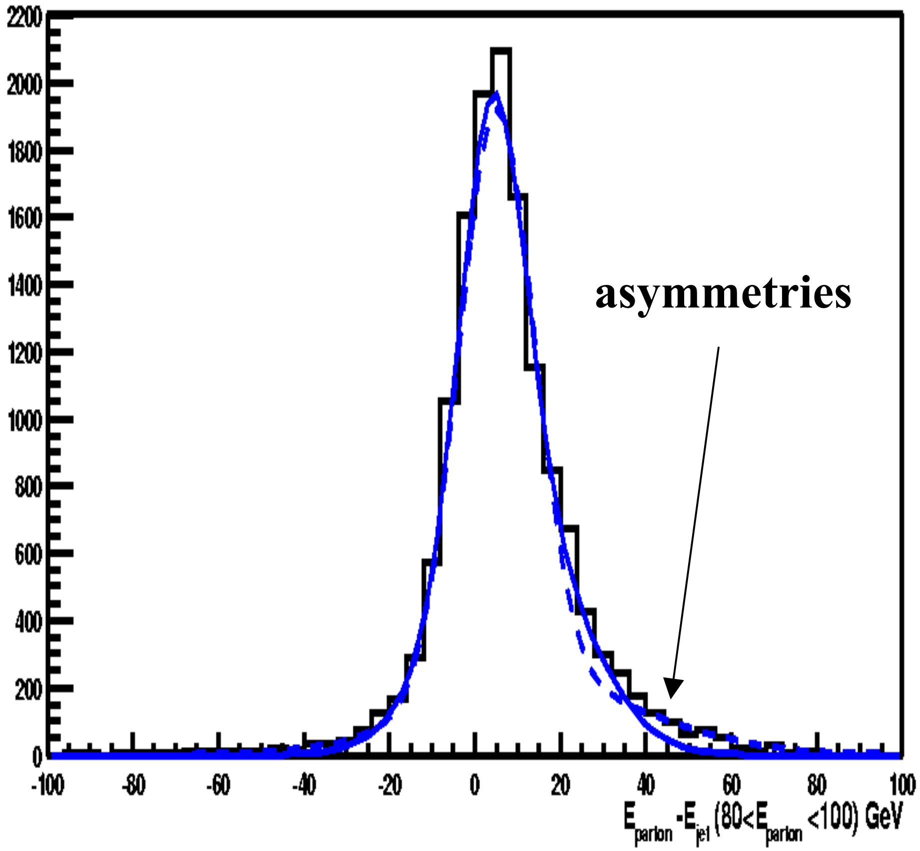
20 – 15 = 5 integrals

Sum over 24 combinations of jets, all values of the neutrino momentum are considered. Because it is L.O., we use only 4-jet events.

ρ_1	momentum of one of the jets	m_1,m_2	top mass in the event
M_1,M_2	W mass in the event	$f(q_1)$,$f(q_2)$	parton distribution functions (CTEQ4) for qq incident chann.
q_1,q_2	initial parton momenta	ϕ_6	six particle phase space
$W(x,y)$	probability of measuring x when y was produced in the collision		

We choose these variables of integration because $|M|^2$ is almost negligible, except near the four peaks of the Breit-Wigners within $|M|^2$.

Transfer function $W_{jet}(x,y)$



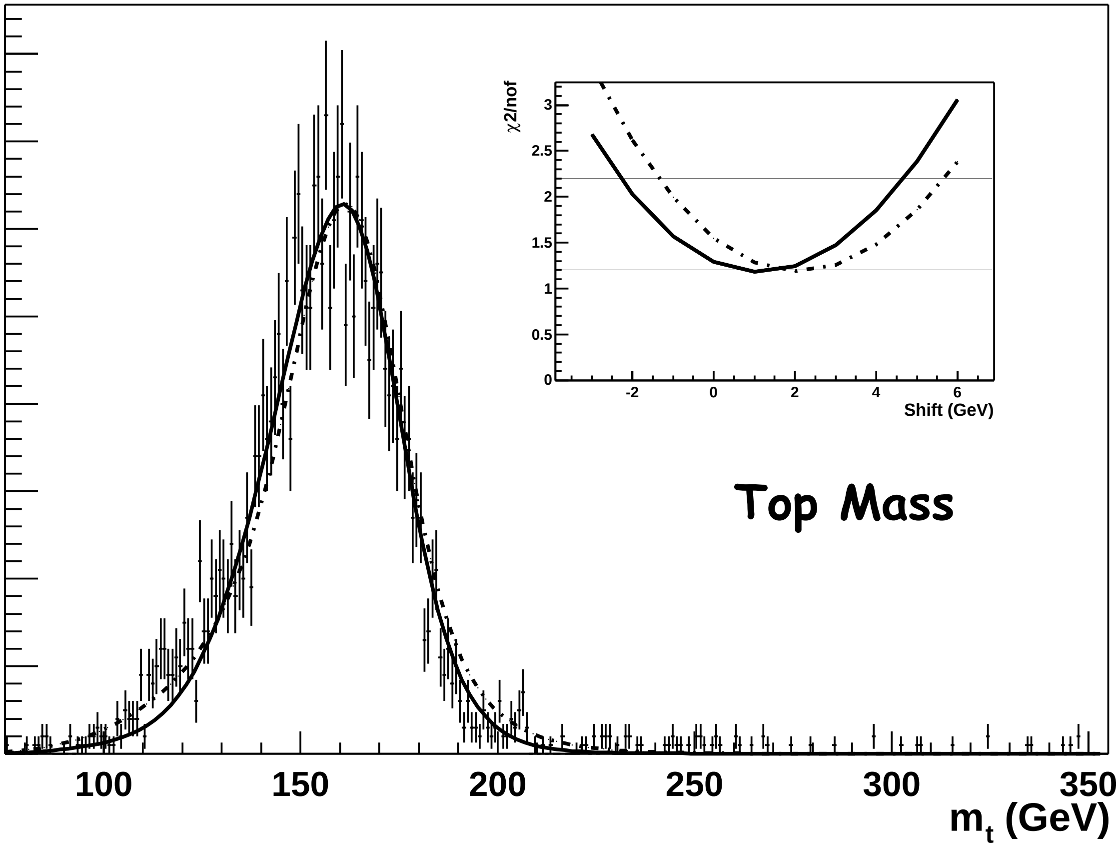
Models the smearing in jet energies from effects of **radiation, hadronization,** measurement **resolution** and jet **reconstruction algorithm**

Correcting on average, and considering these distributions to be single Gaussians, can underestimate parton energies

Use 2 Gaussians, one to account for the peak and the other to fit the asymmetric tails (light quarks and b quarks have separate parameters).

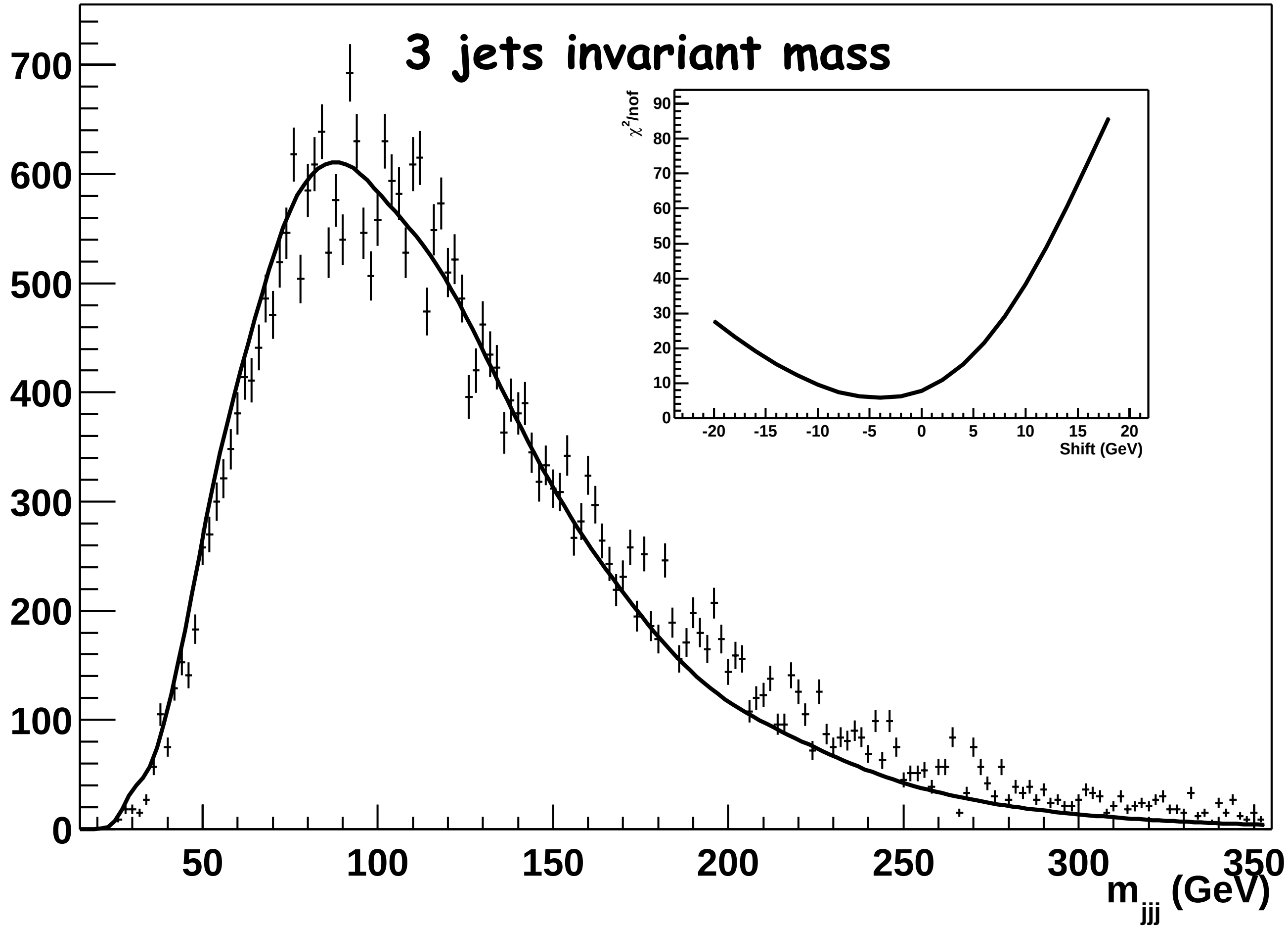
$$W_{jet}(x,y) = F(\delta_E) = \frac{1}{\sqrt{2\pi}(p_1 + p_2 p_5)}[\exp\frac{-(\delta_E - p_1)^2}{2p_2^2} + p_3 \exp\frac{-(\delta_E - p_4)^2}{2p_5^2}]$$

Transfer function in ttbar



Histogram: HERWIG events after **full DØ reconstruction**, using the standard criteria
Solid Line: Calculated by using the **transfer function on partons**
Dashed: Same as solid, but with a variant transfer function

Transfer Function in W+jets



Invariant mass of three jets using W+4-jets events from VECBOS + ISAJET (Smooth curve, same as before). The shape is reproduced…. any bias introduced by non perfect agreement is calibrated using MC simulations.

Acceptance Corrections

Likelihood

$$-\ln L(\alpha) = -\sum_{i=1}^{N} \ln P(x_i;\alpha) + N\int P(x;\alpha)dx$$

Detector Acceptance

$$P(x;\alpha) = Acc(x)P_0(x;\alpha)$$

Measured probability

Detector acceptance

Production probability

$$-\ln L(\alpha) = -\sum_{i=1}^{N} \ln P_0(x_i;\alpha) + N\int Acc(x)P_0(x;\alpha)dx$$

$$\int Acc(x)P_0(x;\alpha)dx = \frac{12V}{N_{gen}} \sum_{j=accep.}^{N} P_0(x_j;\alpha)$$

where $V = \int d^n\sigma_{MC}(y)dq_1dq_2 f_{MC}(q_1) f_{MC}(q_2)$, N_{gen}(N) is number of generated(observed) events

Signal and Background

$$-\ln L(\alpha) = N\int A(x)\left[c_1 P_{tt}(x;\alpha) + c_2 P_{bkg}(x)\right]dx$$

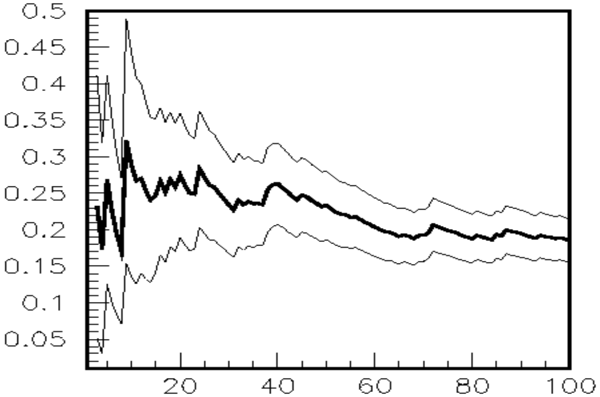
$$-\sum_{i=1}^{N}\left\{\ln\left[c_1 P_{tt}(x_i;\alpha) + c_2 P_{bkg}(x_i)\right]\right\}$$

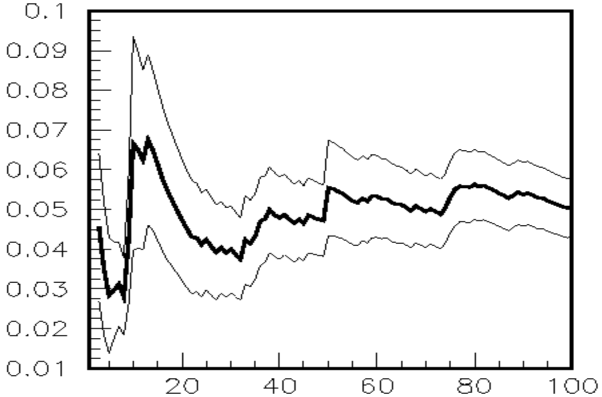
- **The background probability is defined only in terms of the main backgound** (W+jets, 80%) which proves to be also adequate for multijet background treatment in this analysis.
- The background probability for each event is calculated using **VECBOS subroutines for W+jets**.
- The values of **c_1 and c_2 are optimized**, and the likelihood is normalized automatically at each value of α.

Background Probability

We extracted from VECBOS events simulator the subroutines that calculate the matrix element for W+4 jets events. W. Giele helped showing us the way to use these subroutines. For this probability we use MC integration:

$$P_{bkg}(x) = \frac{1}{\sigma_{tot}^{bkg}} \left[\frac{1}{N} \sum_{i=1}^{N} P_{W+jet}(\Omega_1,\ldots,\Omega_4,\bar{p}_{lep},M_W,\underline{E_1^{part},\ldots,E_4^{part}}) \right]_{W(E_i^{jet},E_i^{part})}$$



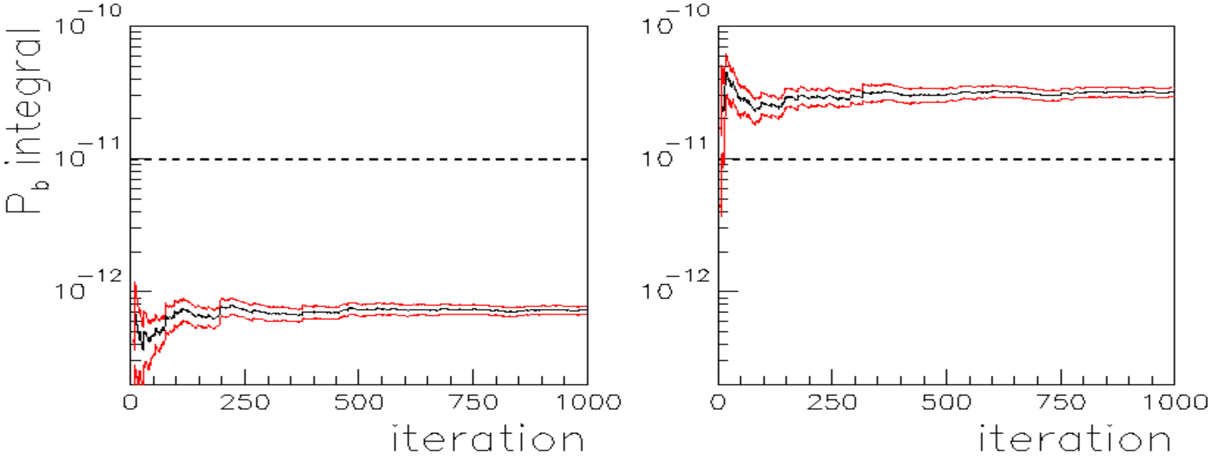


We integrate until we ensure the convergence.

Background Probability

We extracted from VECBOS events simulator the subroutines that calculate the matrix element for W+4 jets events. W. Giele helped showing us the way to use these subroutines. For this probability we use MC integration:

$$P_{bkg}(x) = \frac{1}{\sigma_{tot}^{bkg}} \left[\frac{1}{N} \sum_{i=1}^{N} P_{W+jet}(\Omega_1, \ldots, \Omega_4, \overline{p}_{lep}, M_W, \underline{E_1^{part}, \ldots, E_4^{part}}) \right]_{W(E_i^{jet}, E_i^{part})}$$

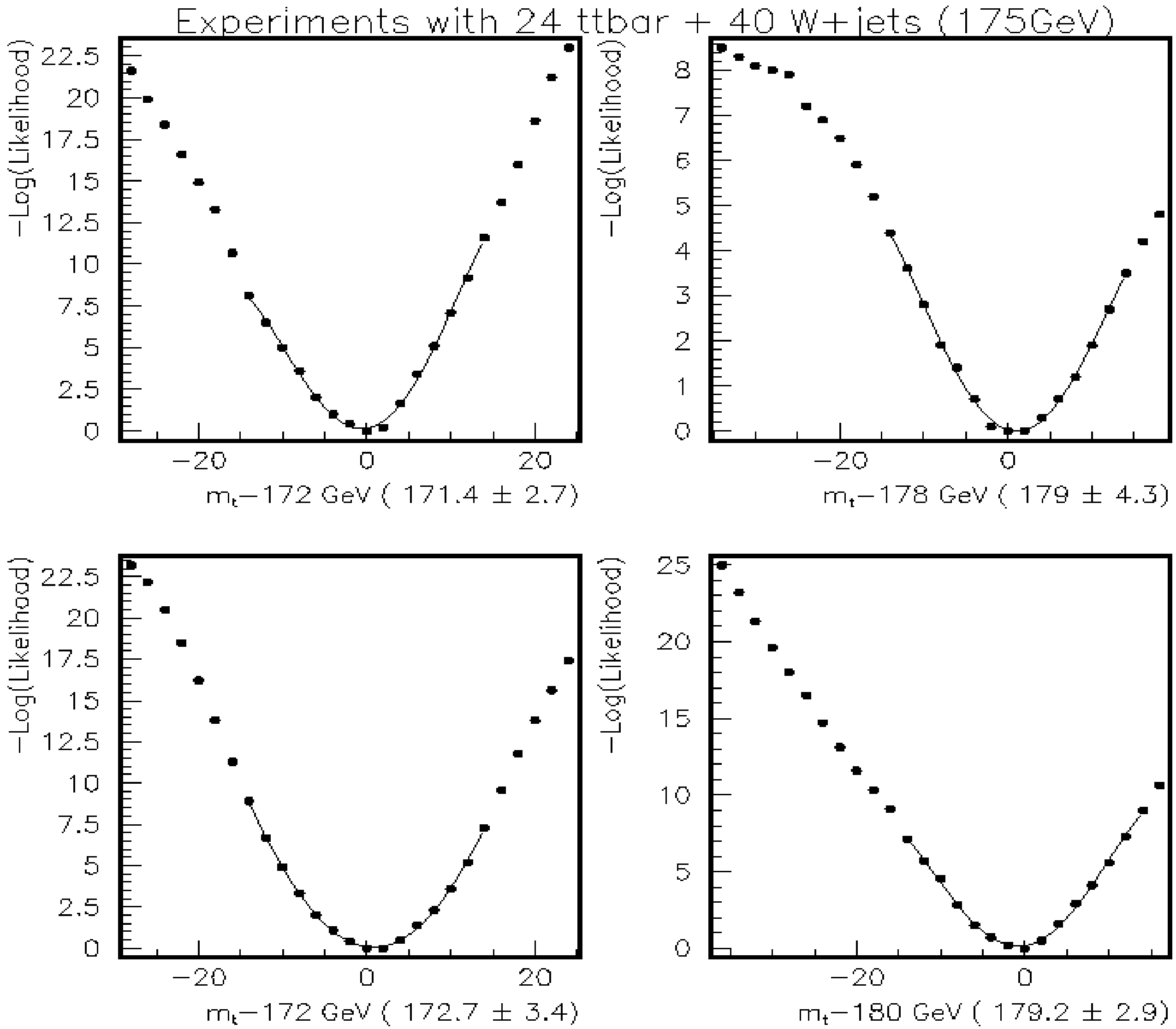


We integrate until we ensure the convergence.

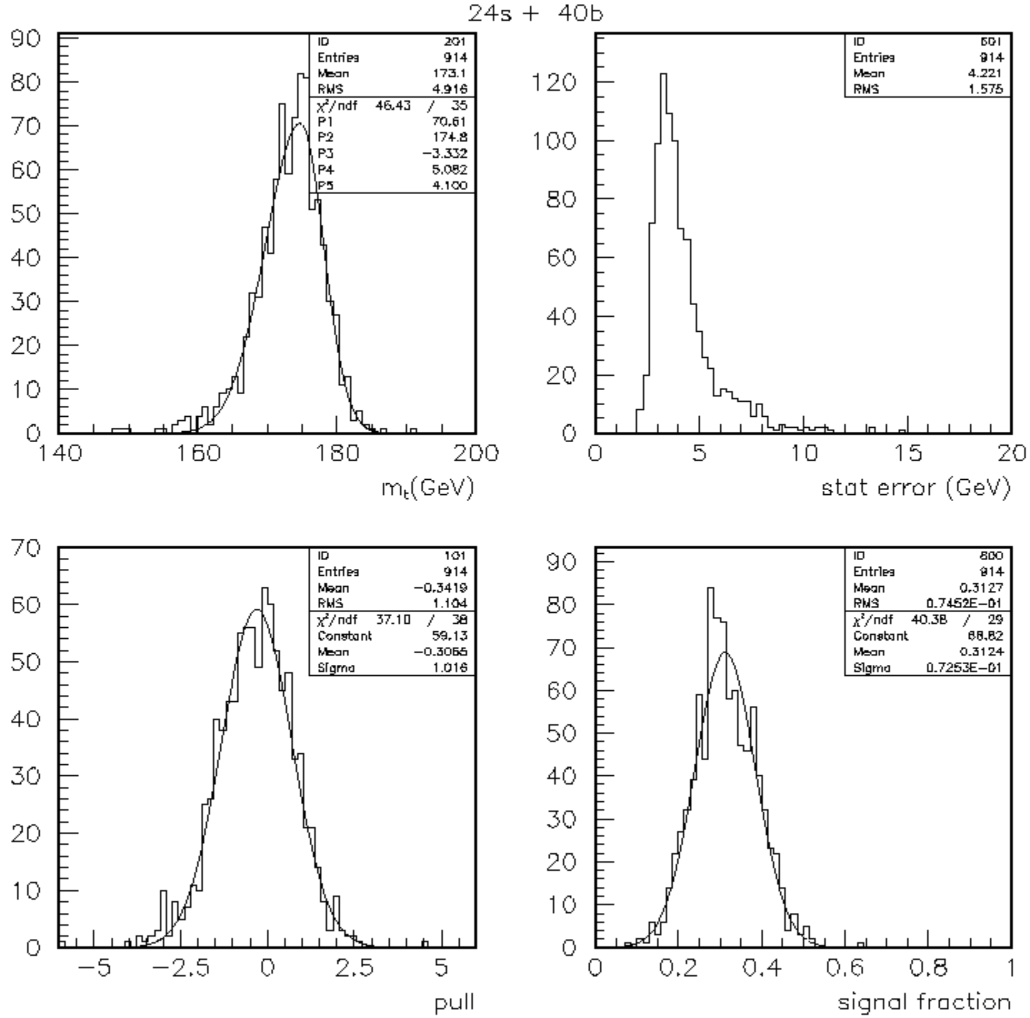
Testing this in Run I DØ using full MC

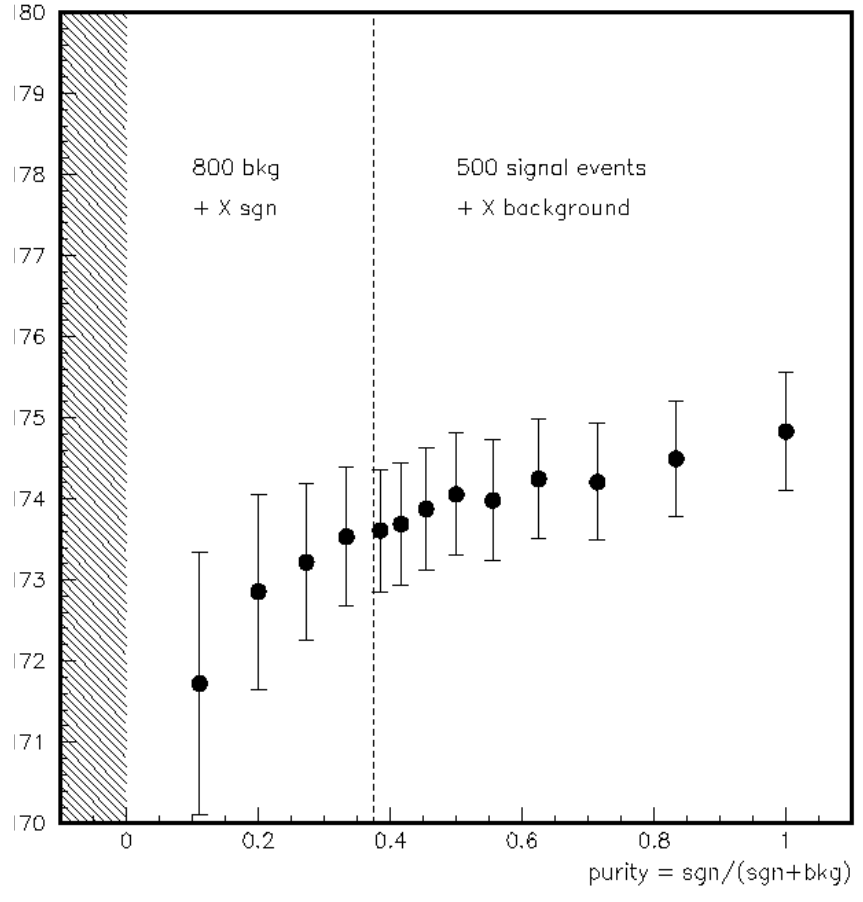
Examples of product likelihood functions. **Each example corresponds to one experiment with the statistics that DØ collected during Run I.**

The signal(HERWIG) and background (VECBOS) events were run through the full DØ Run I simulation.



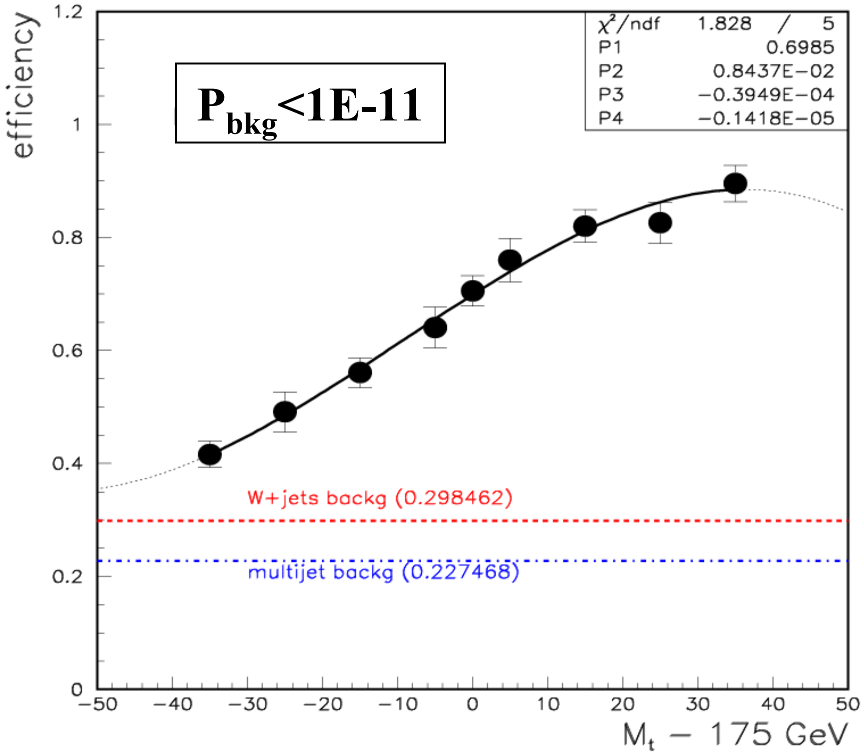
Blind Analysis

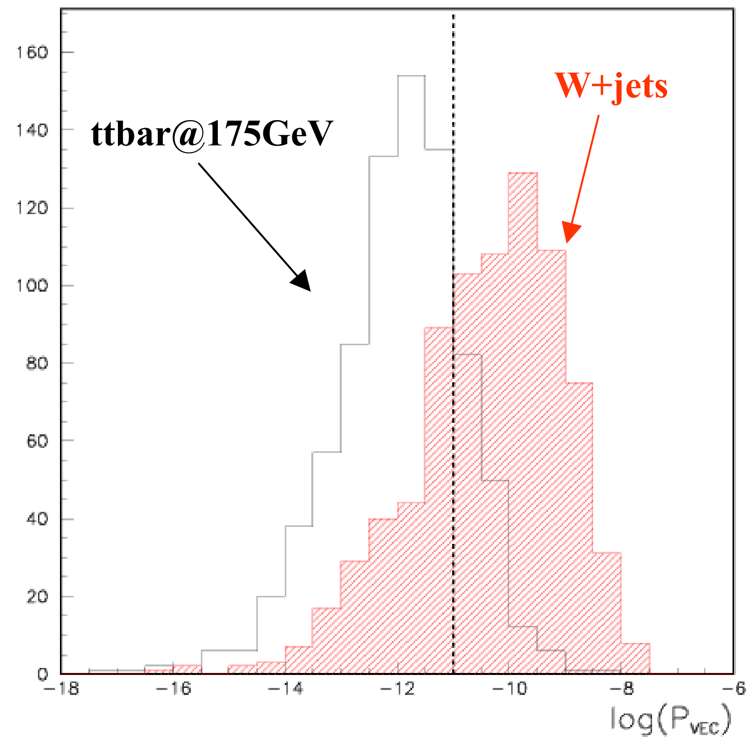




This analysis was defined by MC studies, without looking at the data sample. One of the checks indicated that there could be a shift introduced by background contamination.

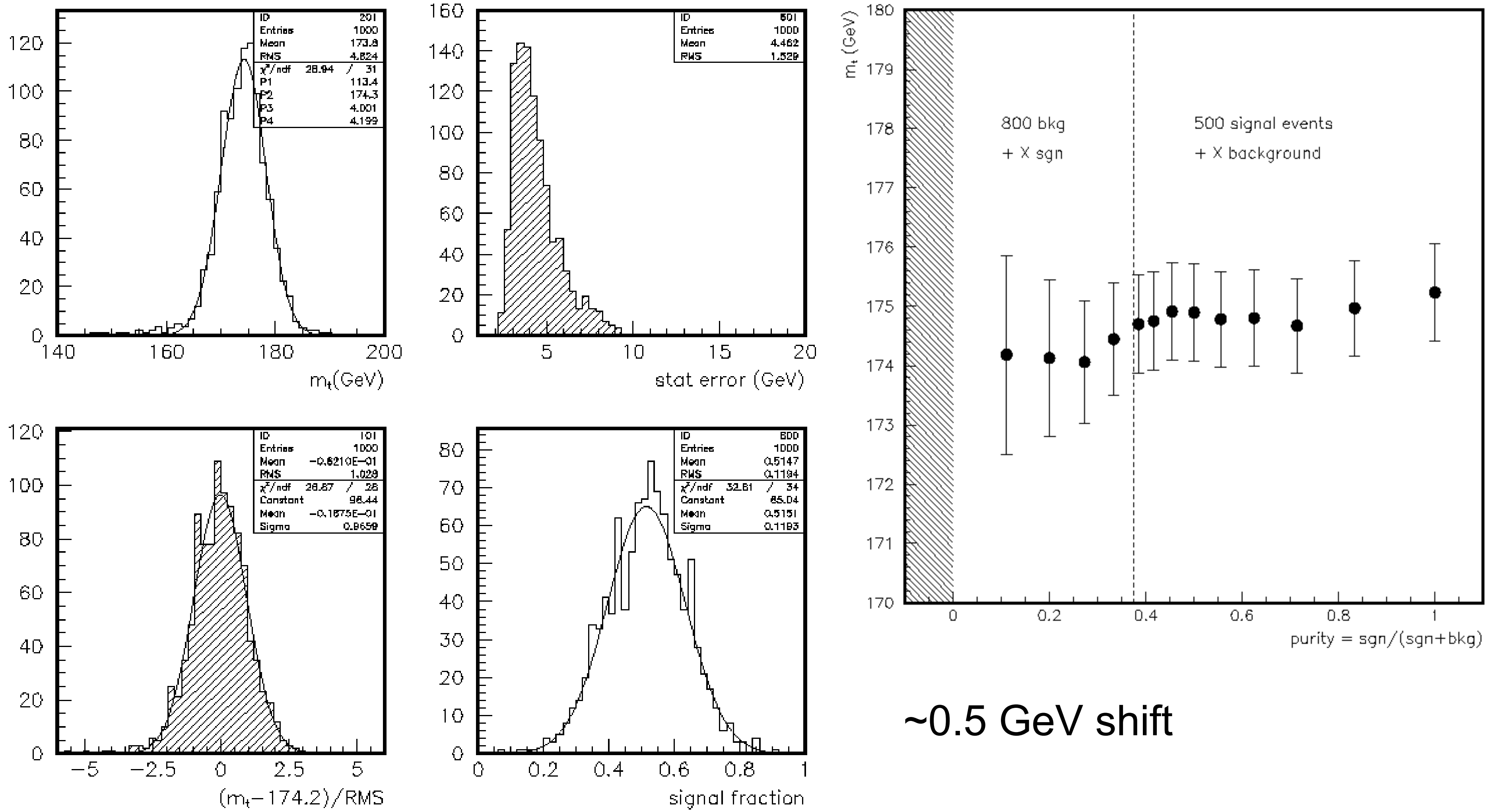
Extra selection in P_{bkg}





In order to increase the purity of signal, another selection is applied on P_{bkg}, with efficiencies: ε_{signal}=0.70, ε_{W+jets}=0.30, $\varepsilon_{multijets}$=0.23

Blind Analysis, purified sample

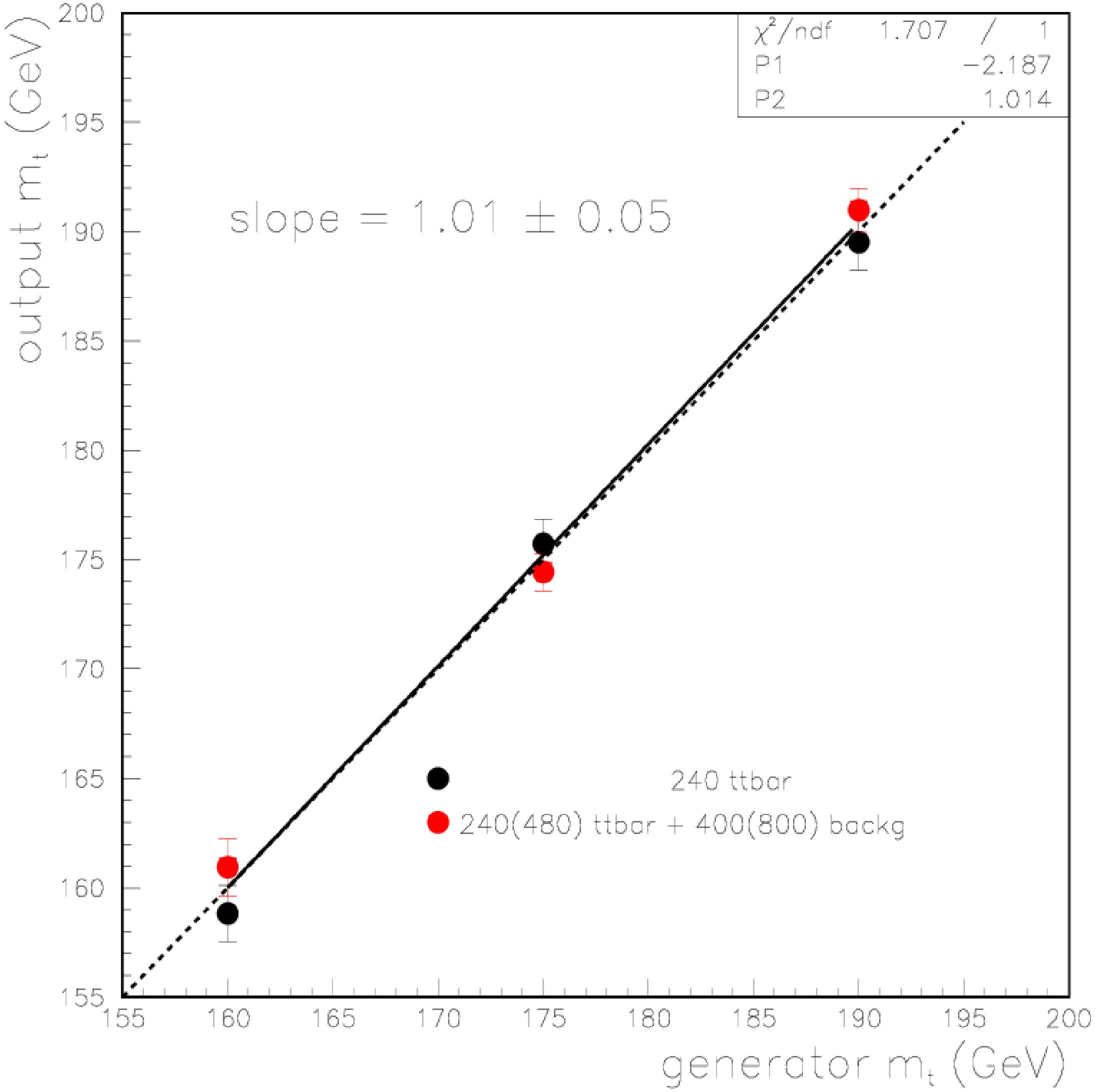


~0.5 GeV shift

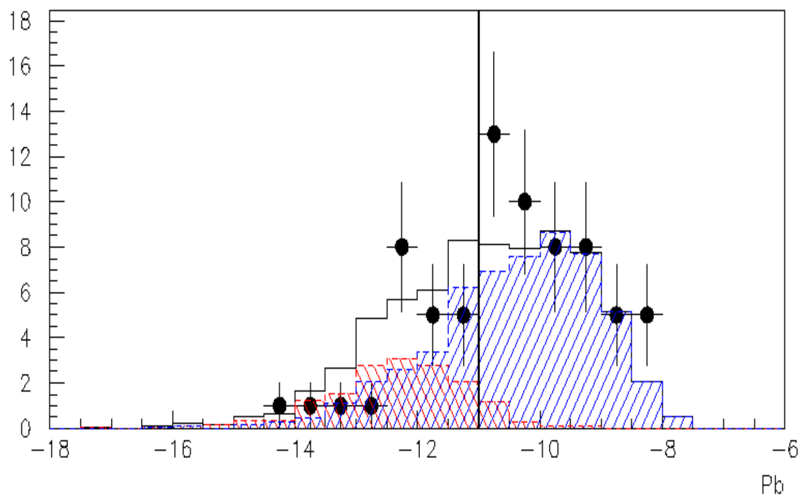
This analysis was defined by MC studies, without looking at the data sample. One of the checks indicated that there could be a shift introduced by background contamination.

Crosscheck of linearity of response

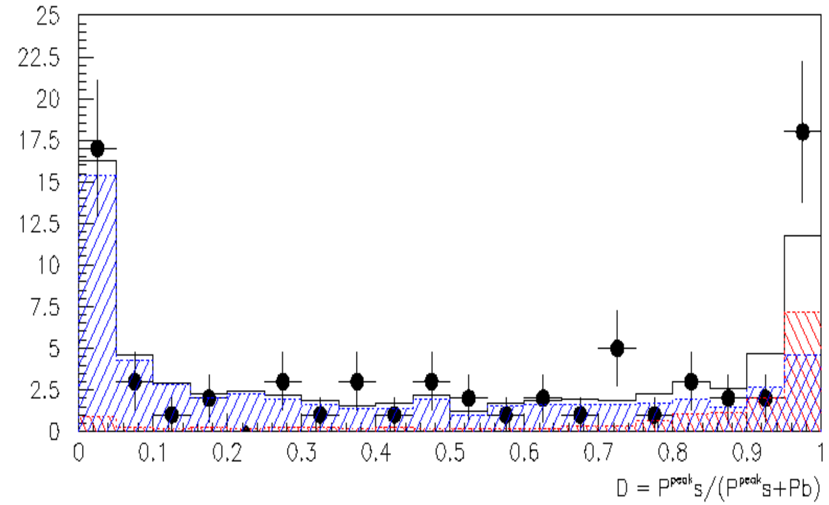
Test of linearity of response with MC samples containing large numbers of events.



Probabilities in Data



Background probability



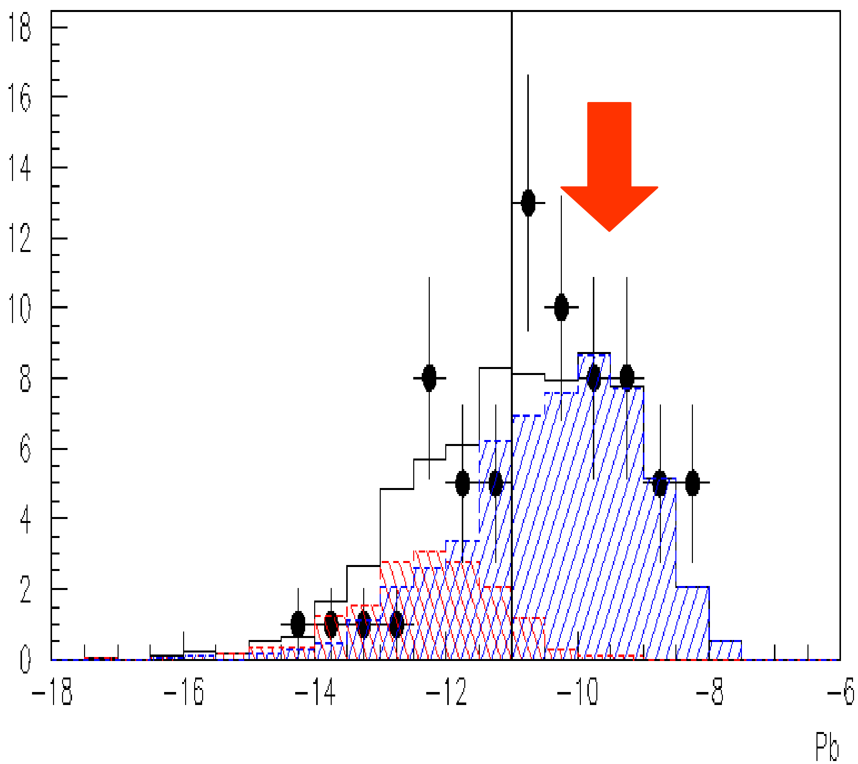
Discriminator

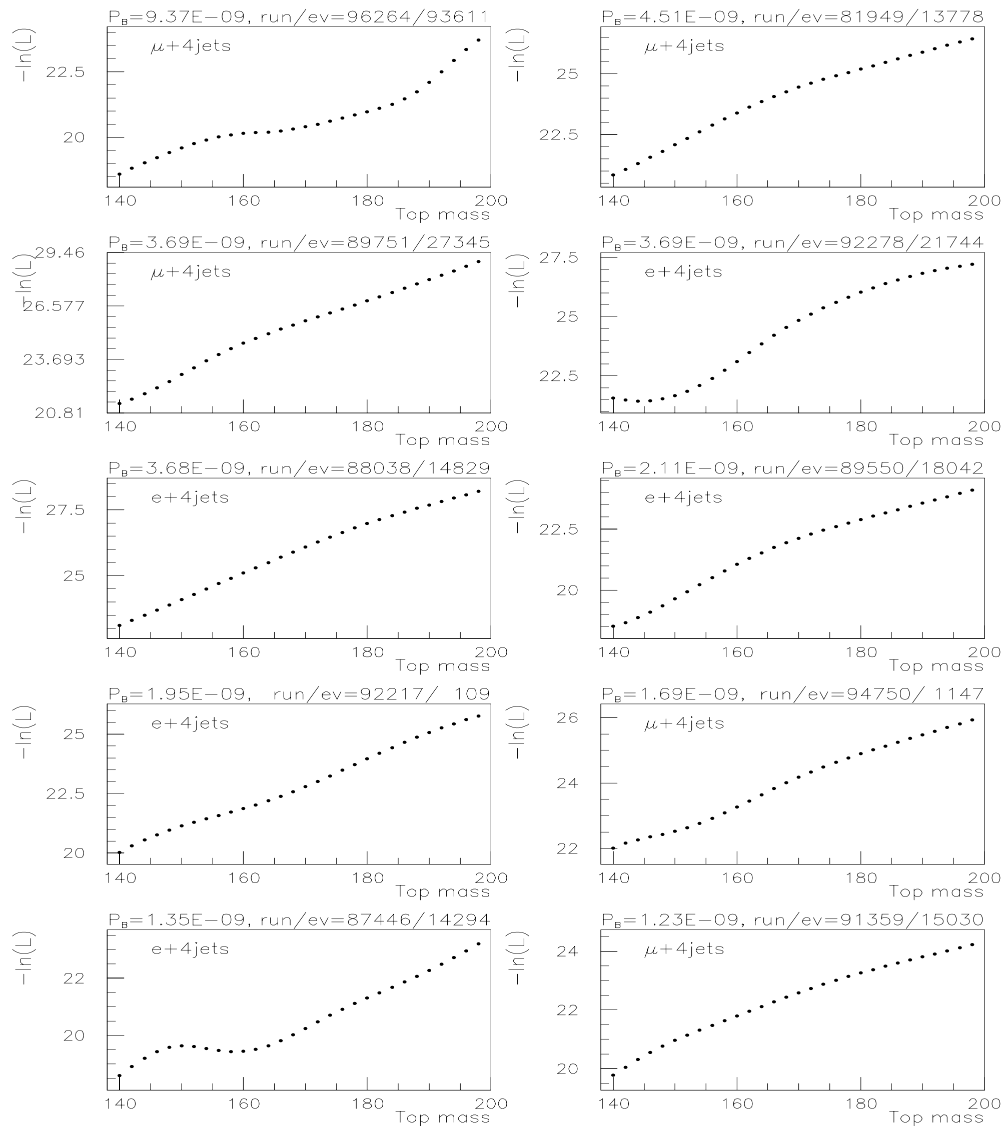
Comparison of (16 signal + 55 background) MC and data sample before the background probability selection.

Top probability in the background region

$-\ln(P_{tt})$ as a function of M_t

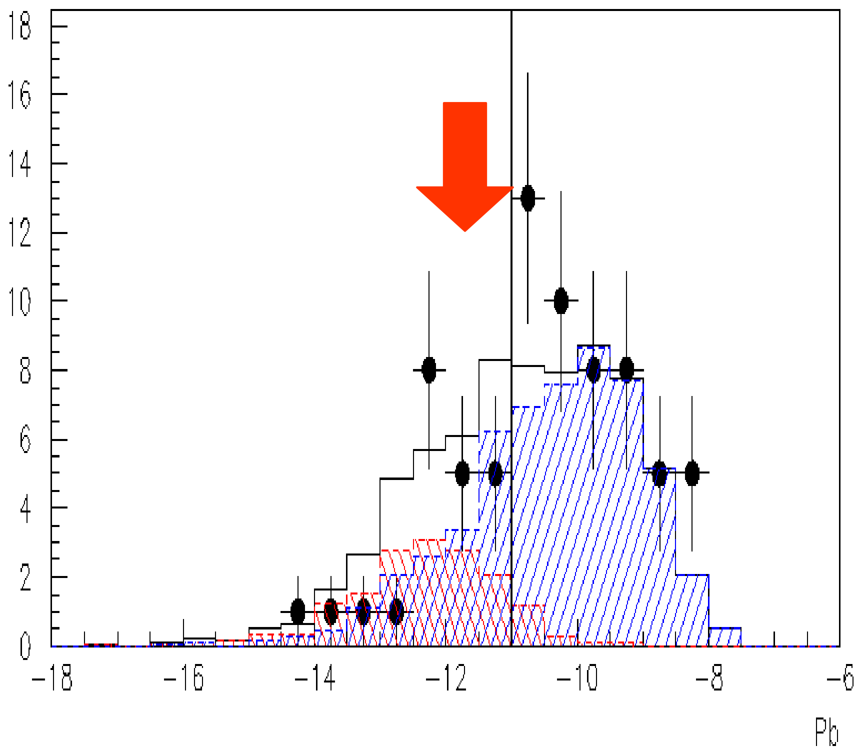
for $10^{-9} < P_{bkg} < 10^{-8}$

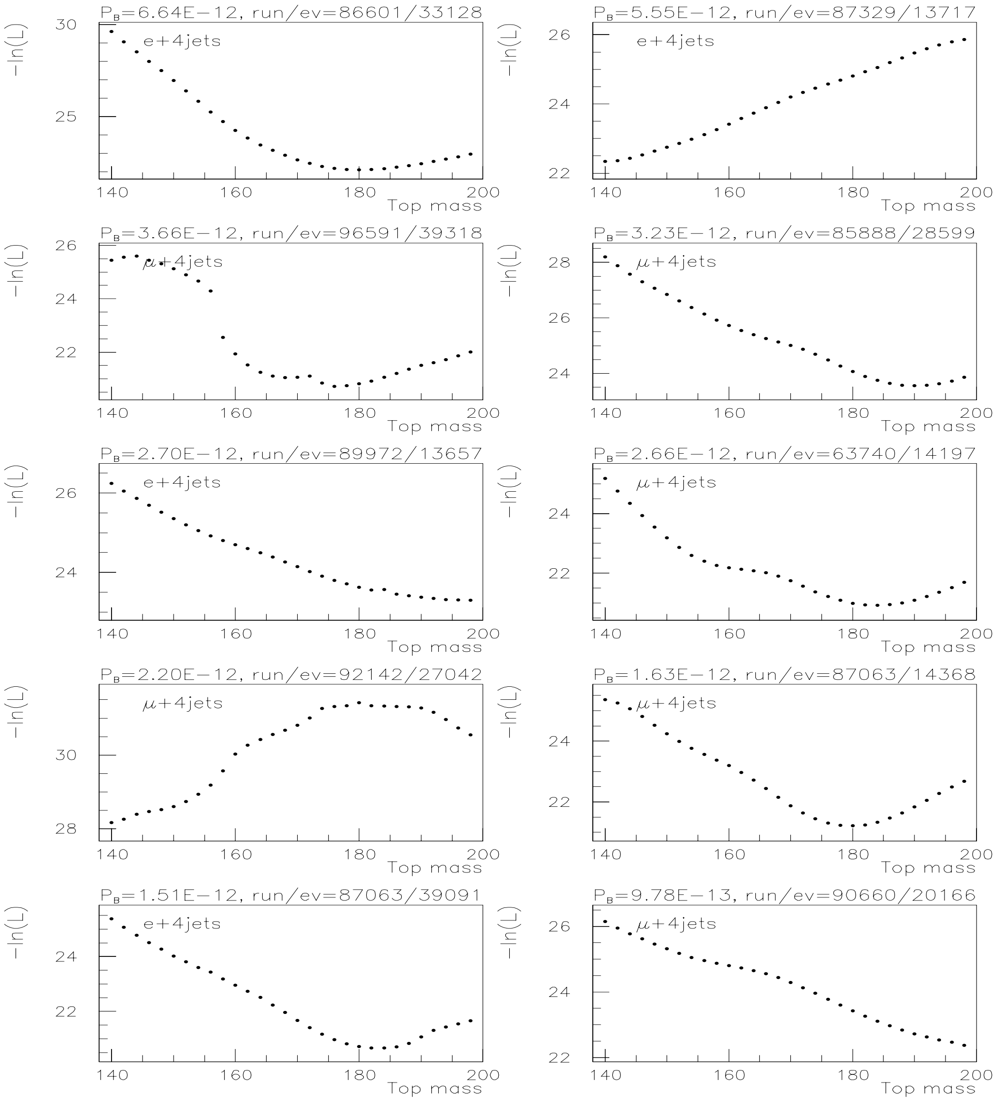




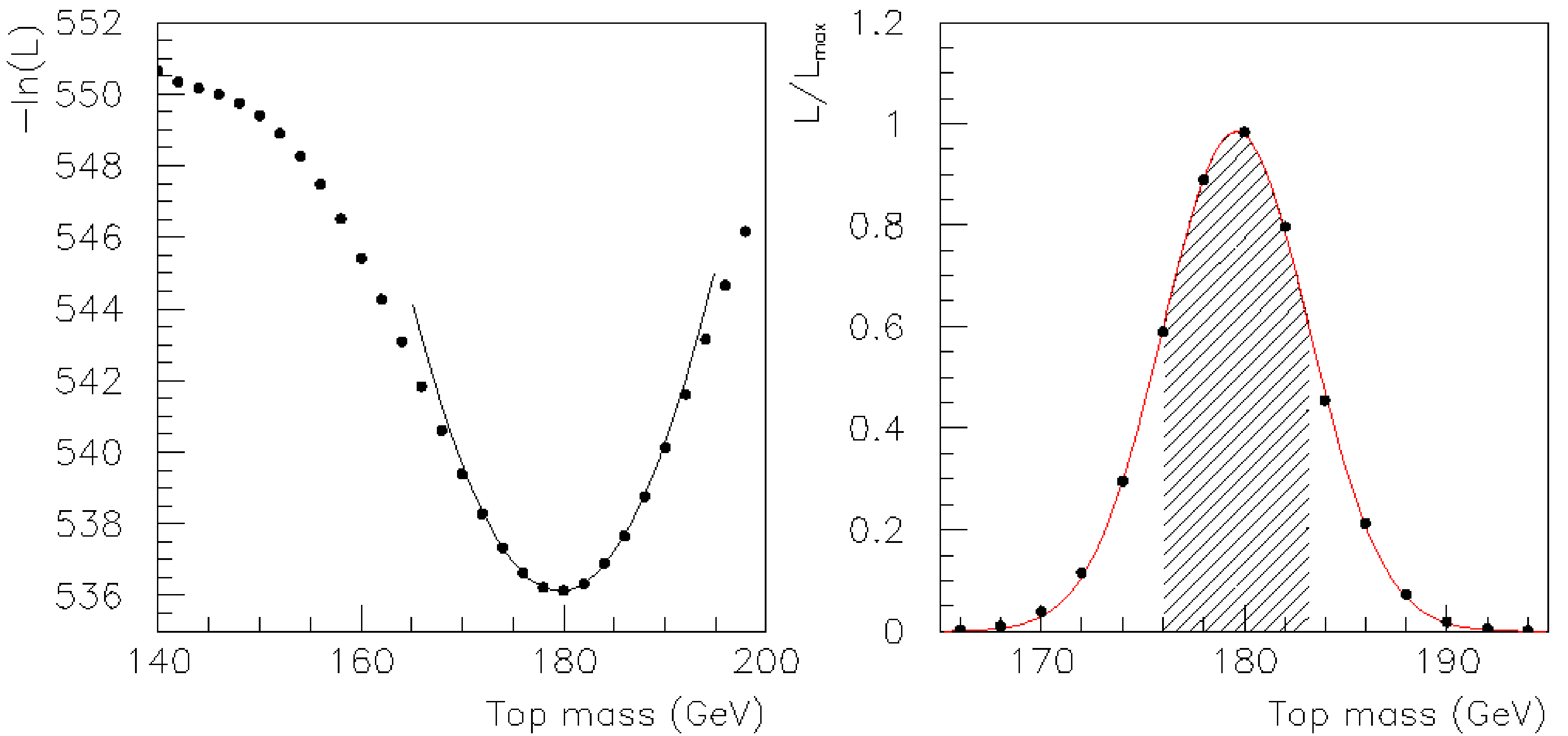
Top probability for signal region

-ln(P_{tt}) as a function of M_t for $10^{-12}<P_{bkg}<10^{-11}$





New Preliminary Result



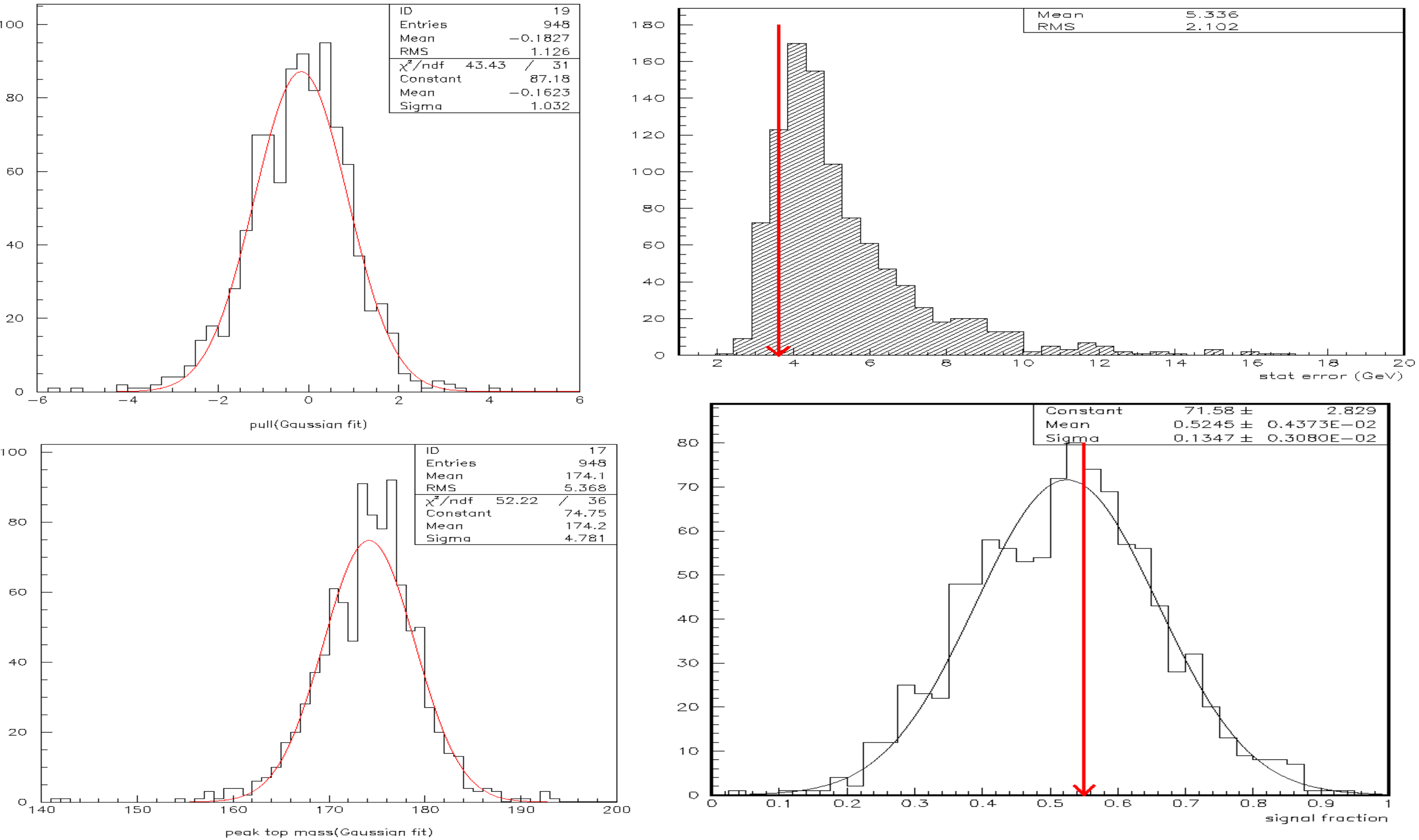
M_t= 180.1 ± 3.6 GeV ± SYST - preliminary

This new technique improves the statistical error on M_t from 5.6 GeV [PRD 58 52001, (1998)] to 3.6 GeV. <u>This is equivalent to a factor of 2.4 in the number of events.</u> 22 events pass our cuts, from fit: (12 s + 10 b)

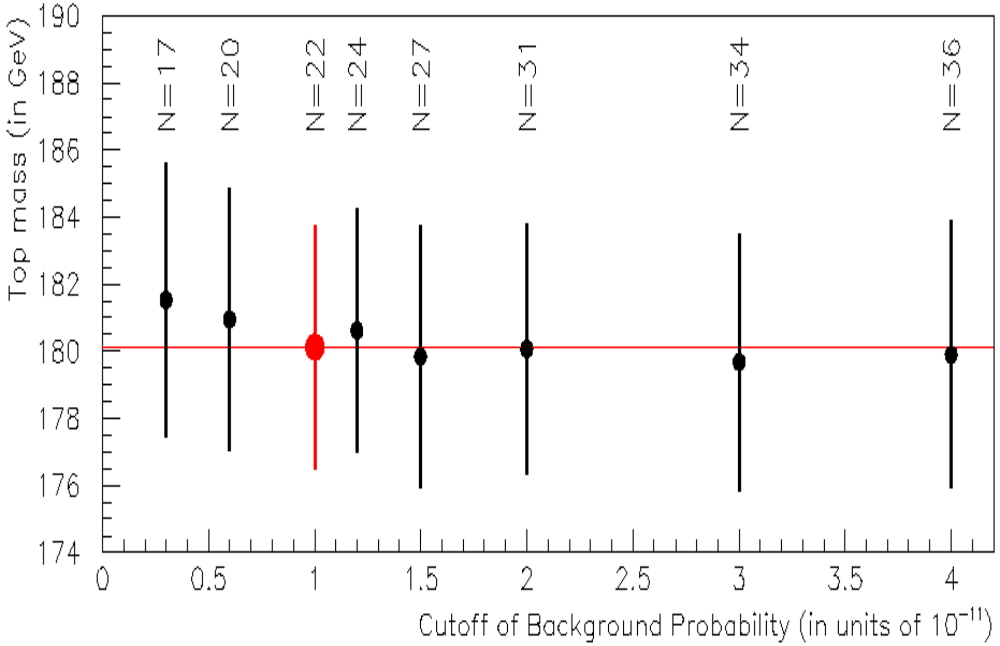
(0.5 GeV shift has been applied, from MC studies)

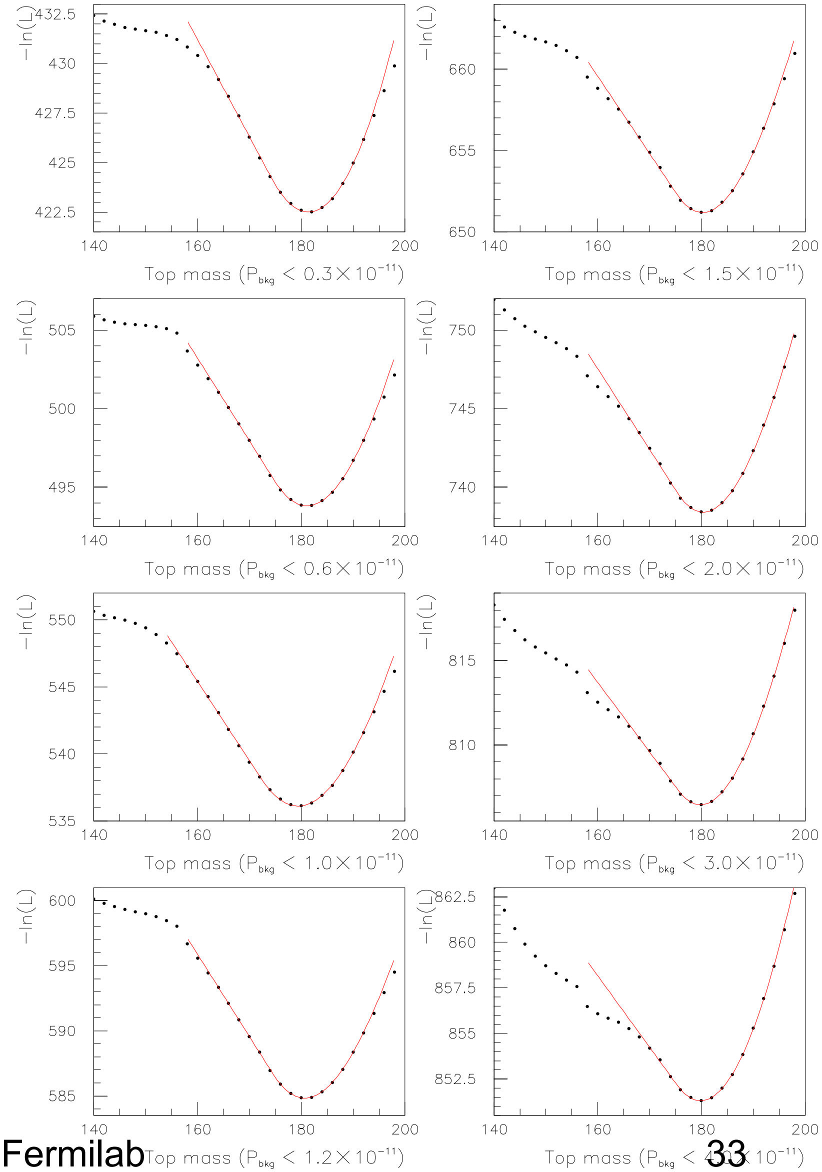
MC studies with 12s+10b

The MC simulations show that the results obtained are consistent with expectations.



Varying the background probability selection





The result is very stable with respect to the selection on background probability.

Comparing results in data and MC

	mass	stat.err. data	stat.err. expected		uncertainty from JES [GeV]
PRD **58 52001, (1998)**	173.3	5.6	**(NN)** 8.0	**(LB)** 8.7	4.0
(new)	**180.1**	**3.6**	**(12s+10b)** 5.4		3.3

Number of top decays

PRD **58 52001, (1998) : 71 events**

NN = 24 ± 8

LB = 29 ± 8

New : **22 events**

n_S = 12 ± 4 (measured sample)

corrections:0.70(4 jets)x0.71 (Pb)

N_S = 25 ± 7 (in the 1998 sample)

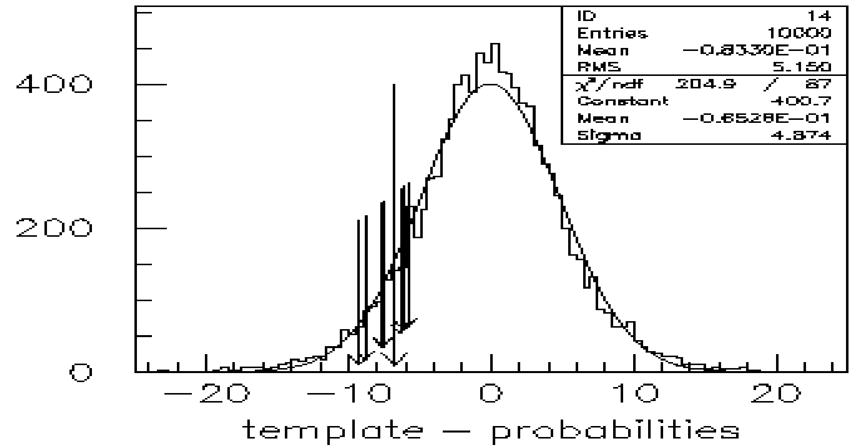
Compatibility between 1998 and 2003

Assume there is **no systematic difference between the two methods** and that both methods give an unbiased measurement of M_t .

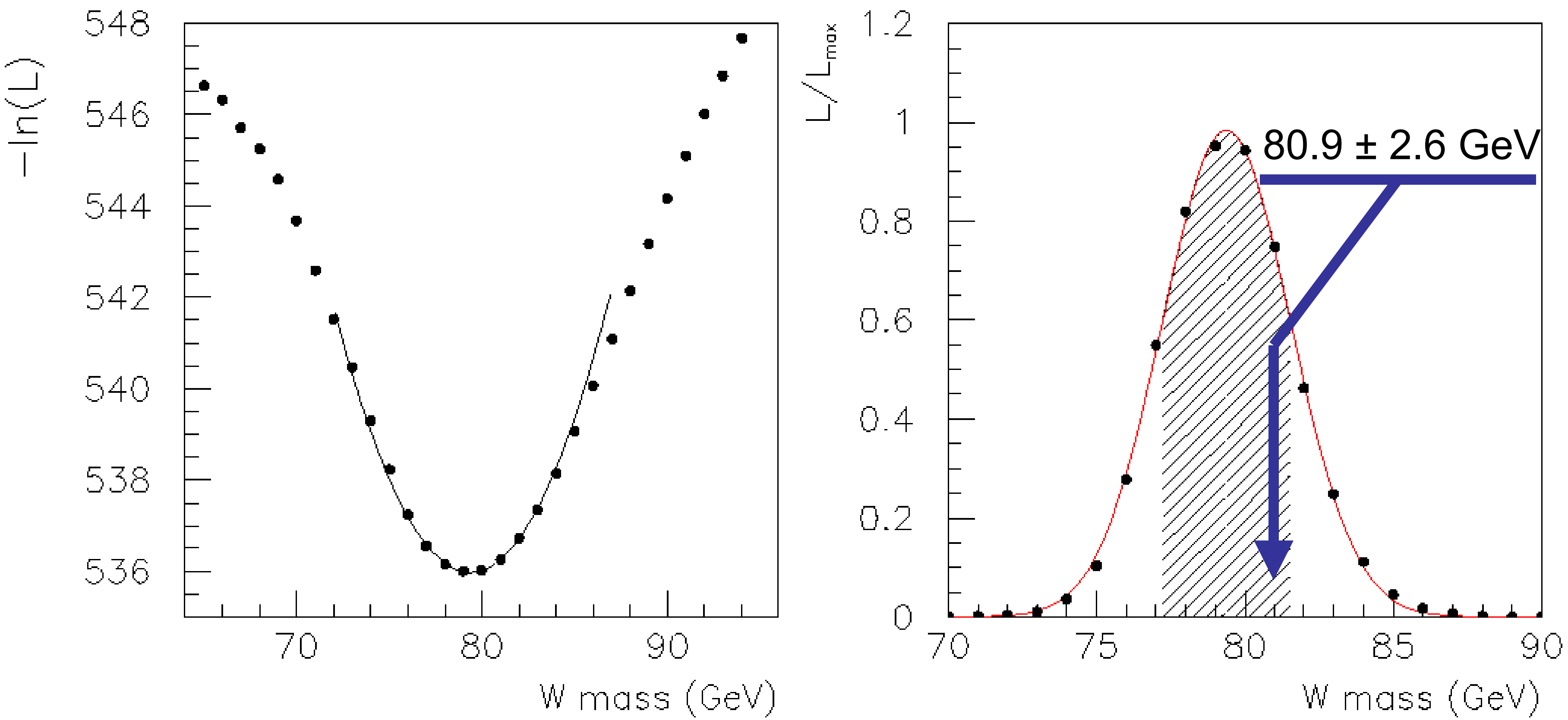
1998 (24 events)	$\rightarrow$(173.3 ± 5.6) GeV
2003 (12 ev. subsample)	$\rightarrow$(180.1 ± 3.6) GeV

1. If the methods were the same and **all the events had the same resolution**: σ_{24}=5.6 $\rightarrow$ σ_{12}=7.9, the fluctuations is 24 events when you vary only 12 is $\sigma_{12/24}$=7.9/2=4.0 $\rightarrow$ the difference is therefore 1.7 σ

2. Both methods are not the same (see introduction), and all the events have their own width, no reason to expect the same result… even with the same sample. Simple case, $\sigma_{12/24}$=5.0, $\rightarrow$ the difference is 1.4 σ

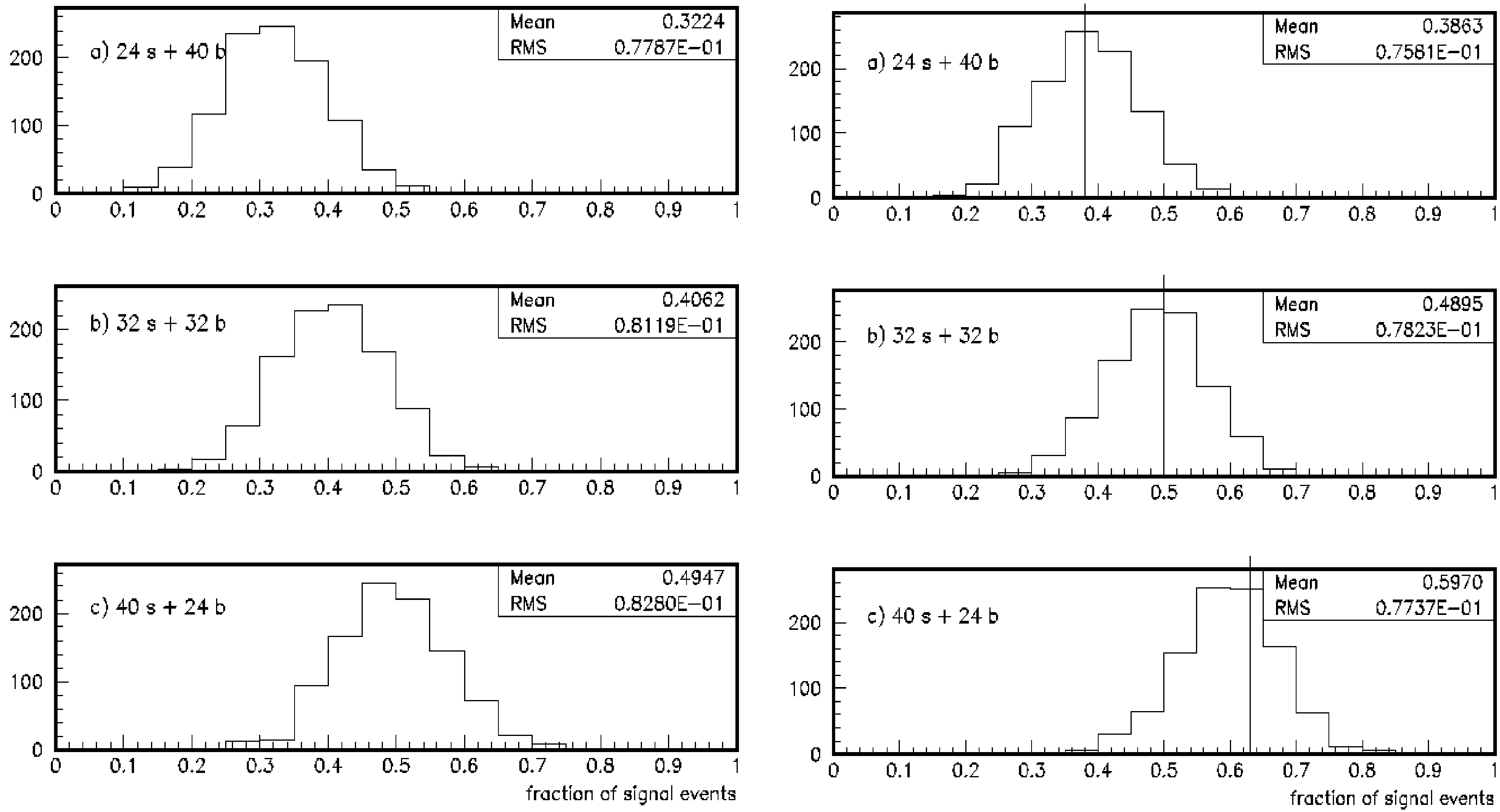


Check of M_W with DØ Run I Data



Can <u>help reduce the uncertainty in the jet energy scale (JES)</u>
see:<u>http://dpf2002.velopers.net/talks_pdf/120talk.pdf</u> (DPF2002 proceedings)
1.5 GeV shift is applied and 20% increase in the error, from MC studies. We associate this shift to effects from our L.O. approximation.

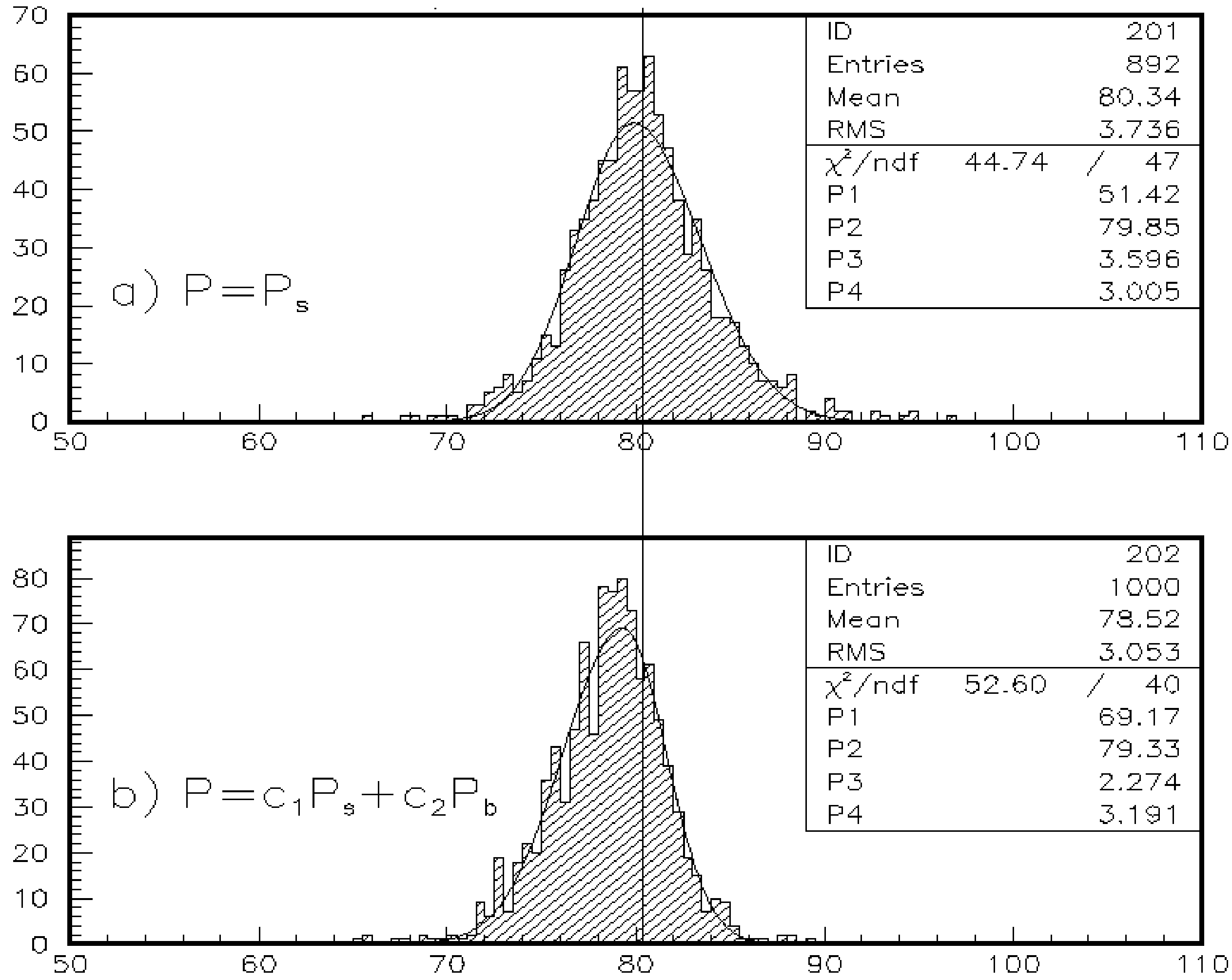
Number of events for signal and backg
(something to consider in the future)



Because of radiation, ~20% of the signal events look more like background than signal. When only events with good jet-parton matching are used, this is resolved (better treatment of higher order effects will help).

M_W, MC studies

We see that the signal events that are reconstructed as background are responsible for the 1.5 GeV shift in M_W.

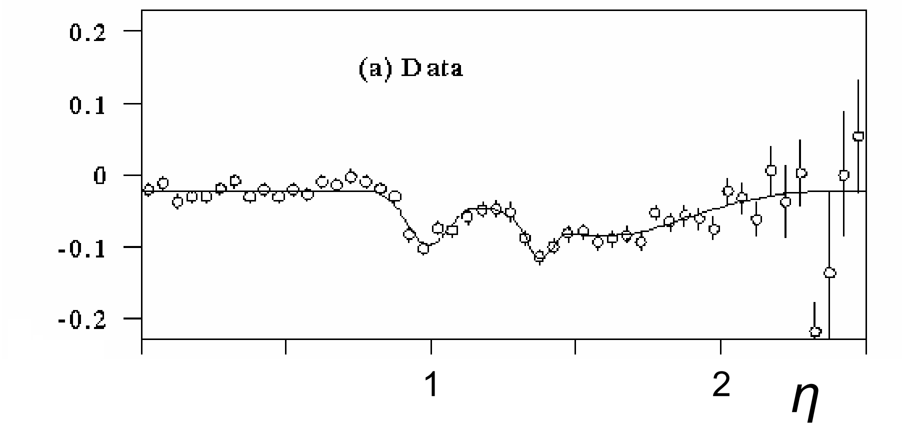


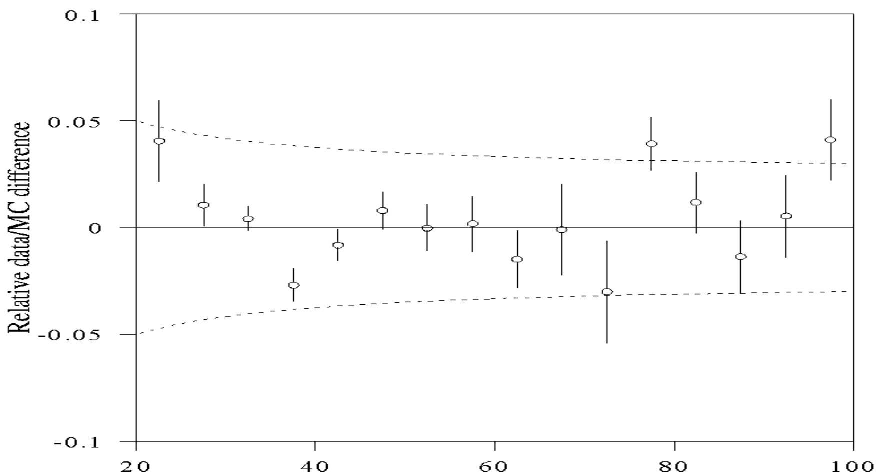
This will have to be solved if we want high precision top physics (~1 GeV). Higher order corrections will have to be included.

Jet Energy Scale (main systematic effect)

- We use a **Monte Carlo simulation of the detector to build the transfer function** (or the templates in our previous analysis).
- It is essential to **check that the energy scale in the MC simulation is representative of that in the detector**. This can be done for the electromagnetic showers using $Z \rightarrow e^+e^-$ decays. It is **not so easy to do for hadronic showers**.
- Our **γ+jet sample gives 2.5% uncertainty in JES.** In 1998 this translated to 4.0 GeV in M_t.

Systematic Uncertainty JES
(presented HCP2002)





E_t^γ (GeV)

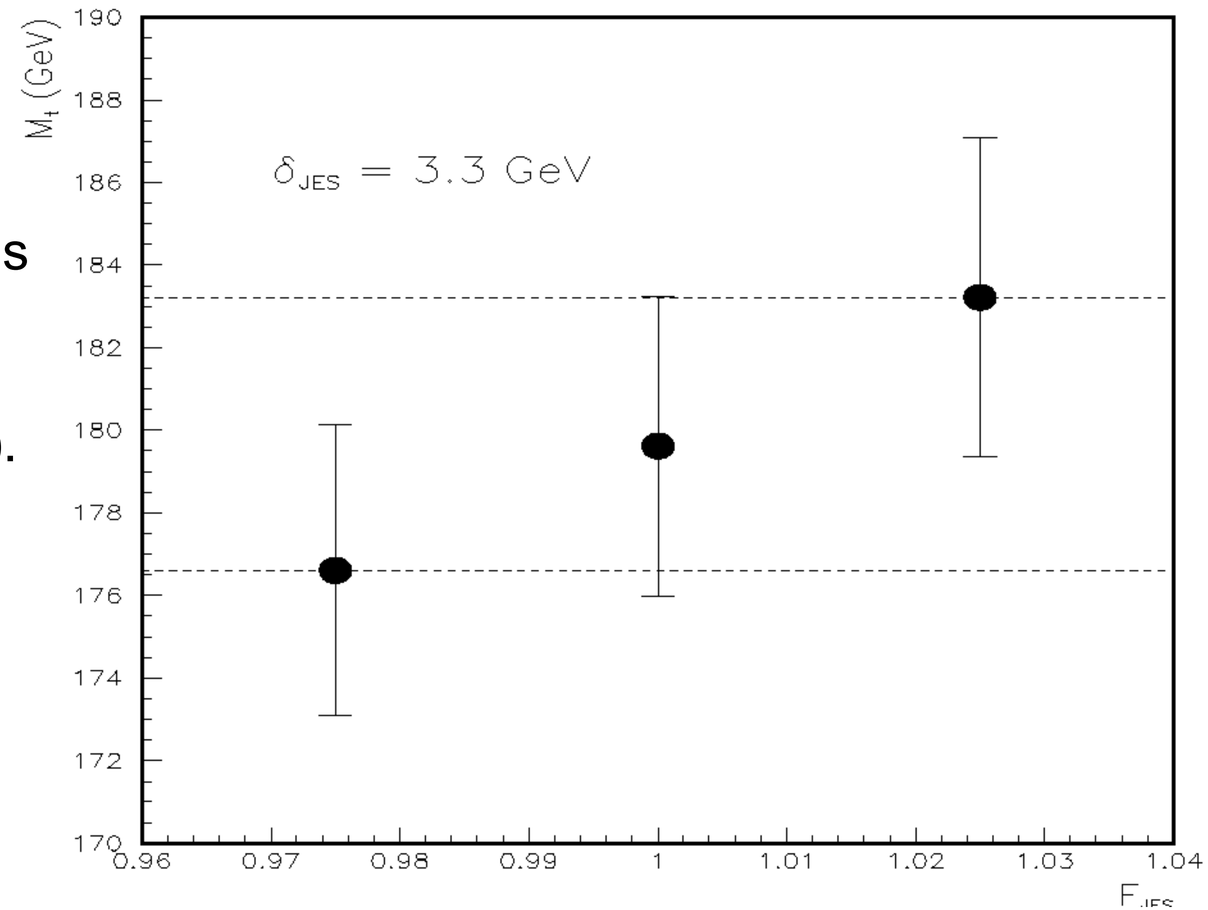
Using a sample of γ+jet [PRD 58 52001, (1998)] we got a function that matches the JES in MC and data with 2.5%+0.5 GeV uncertainty per jet.

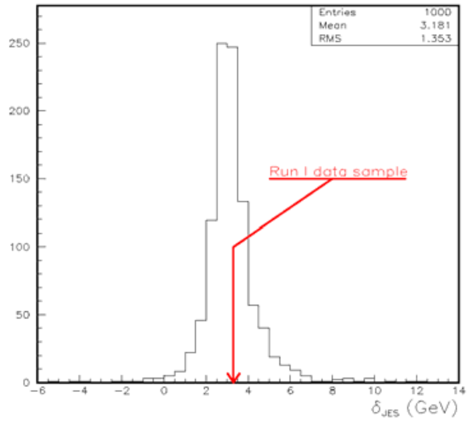
For the result presented in HCP2002 (analysis tools session), **we did the analysis of data with and without this correction**.
This was a **very conservative approach**, because we just wanted to demonstrate the statistical power of the analysis.

Systematic error due to JES

For the result presented here, we are using the previous calibration and its uncertainty for each jet: ±(2.5%+0.5 GeV).

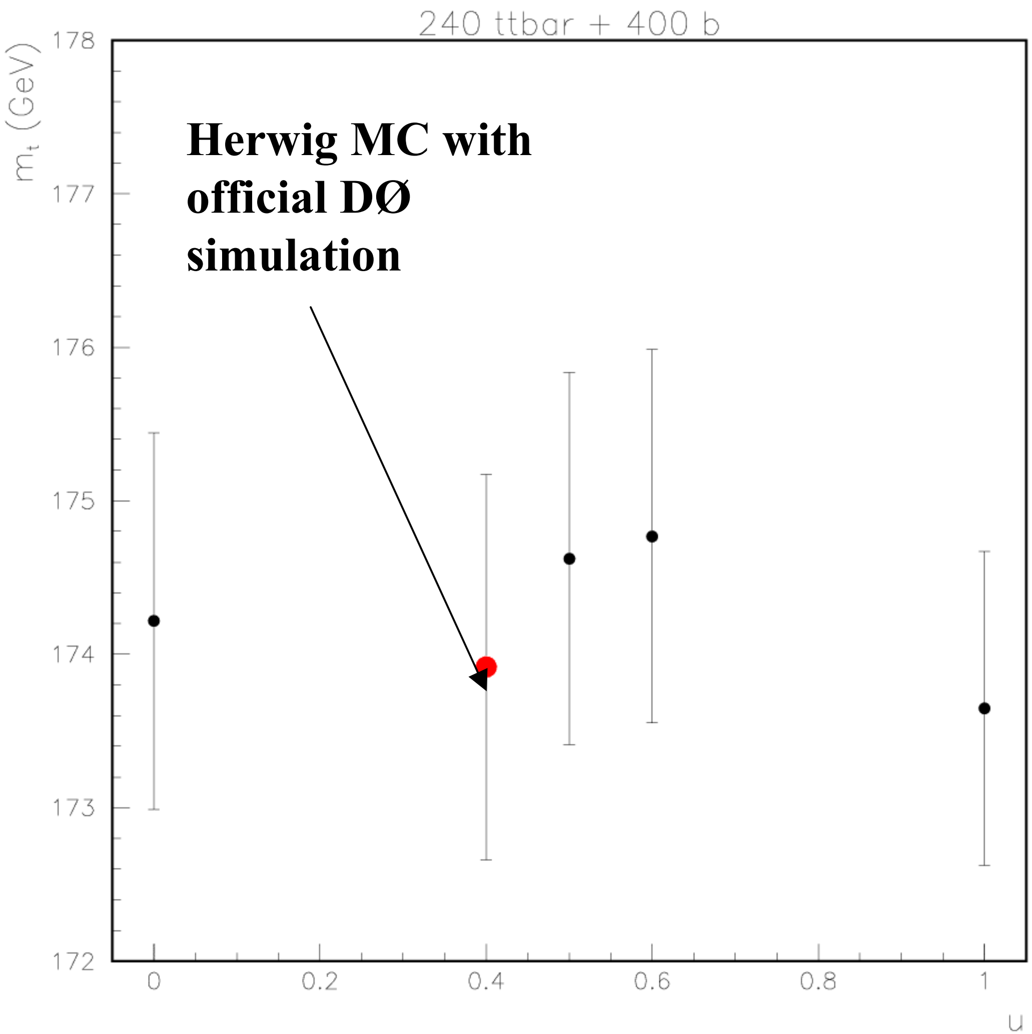
δ_{JES}=3.3 GeV





Consistent with MC expectations

Systematic uncertainty due to top-production model



u: fraction of events in the experiment where all the jets can be matched with partons from top quark decays. Increasing the fraction u, effectively turns on radiation and hadronization effects. The systematic uncertainty is:

$$\delta = 1.5 \text{ GeV}$$

(Each point corresponds to the maximum of a likelihood for a large event sample).

Total Uncertainty

I. Determined from MC studies with large event samples:

Signal model	1.5 GeV
Background model	1.0 GeV
Noise and multiple interactions **PRD 58 52001, (1998)**	1.3 GeV

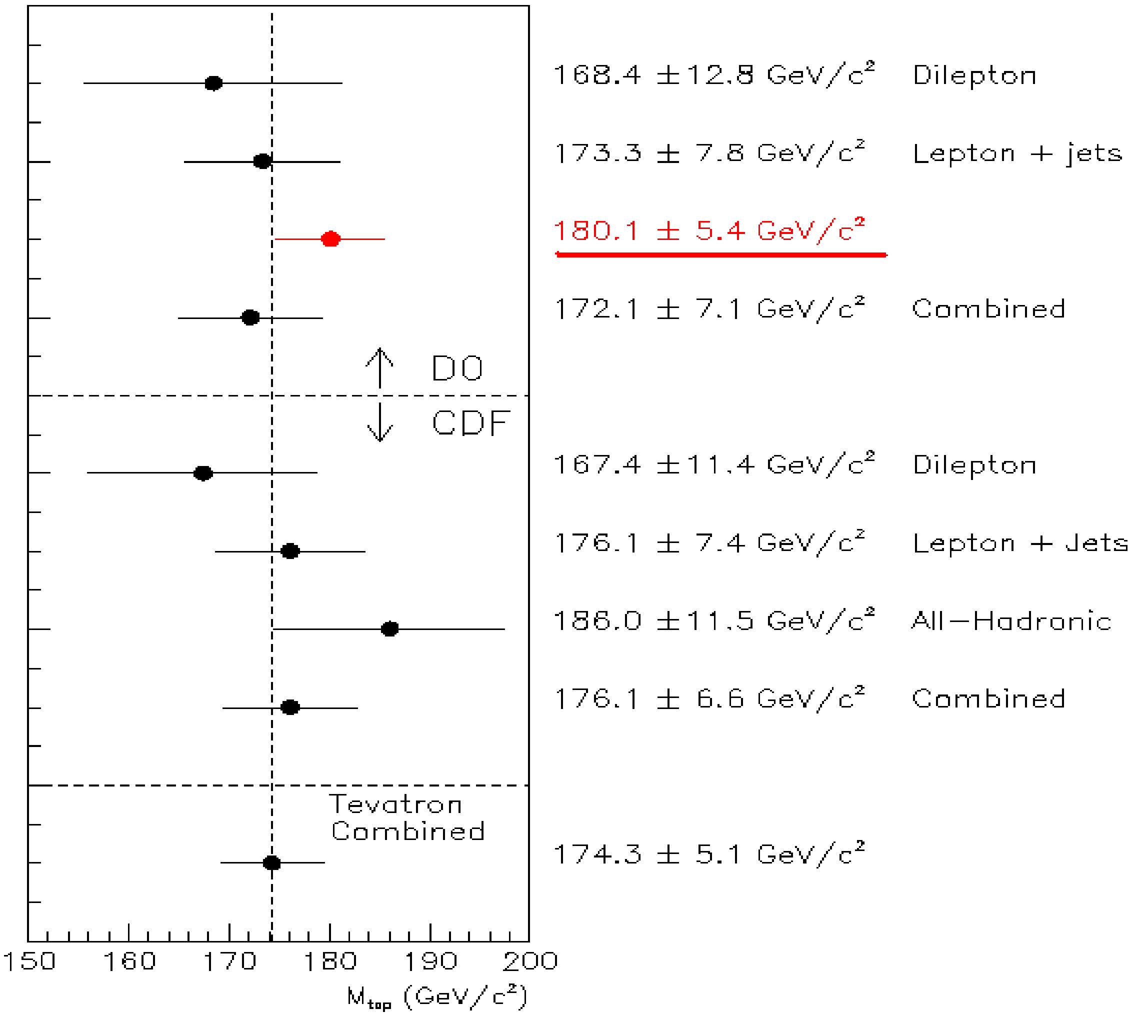
II. Determined from data:

Jet Energy Scale	3.3 GeV
Parton Distribution Function	0.2 GeV
Acceptance Correction	0.5 GeV

Total systematic: 4.0 GeV

$M_t = 180.1 \pm 5.4$ GeV (preliminary)

New [preliminary] Result



The relative error in this result is 3%, compare to 2.9% from the previous CDF and DØ combined average for all channels.

Conclusions

Using LO approximation (and parameterized showering) we calculated the event probabilities, and measured:

$$M_t = 180.1 \pm 3.6\ (\text{stat}) \pm 4.0\ (\text{syst})\ \text{GeV} \quad \text{preliminary}$$

Significant improvement to our previous analysis, is equivalent to 2.4 times more data:

1. Correct permutation is always considered (along with the other eleven)
2. All features of individual events are included, thereby well measured events contribute more information than poorly measured events.

To consider for the future:

- The possibility of checking the value of the *W* mass in the hadronic branch on the same events provides a **new handle on controlling the largest systematic error**, namely, the jet energy scale.
- A very general method (application to W boson helicity, Higgs searches, ….)