Search for top in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV by constrained kinematic fitting

by

Myungyun Pang

An Abstract of

A Dissertation Submitted to the

Graduate Faculty in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Iowa State University
Ames, Iowa
1994

Search for top in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.8 TeV by constrained kinematic fitting

Myungyun Pang

Major Professor: J. M. Hauptman
Iowa State University

A search for the top quark has been pursued in $p\bar{p}$ collision at $\sqrt{s}$ = 1.8 TeV with the DØ detector. This documentation describes a method introduced to extract the mass of the top quark and the cross section limits on $t\bar{t}$ production. Also a preliminary result from the DØ experiment is described.

Search for top in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV by constrained kinematic fitting

by

Myungyun Pang

A Dissertation Submitted to the

Graduate Faculty in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Department: Physics and Astronomy
Major: High Energy Physics

Approved:

In Charge of Major Work

For the Major Department

For the Graduate College

Members of the Committee:

Iowa State University
Ames, Iowa
1994

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

Working in the field of Experimental High Energy Physics has given me tremendous opportunities to meet and work with a lot of people. I've learned a great deal about the field, collaboration, and science in general by just looking at my colleagues doing their jobs. Some of the things that they've shown, which I admired and appreciated very much, were the creative and critical thinking as scientists, leadership, perfectionism, organizational skills, and scientific honesty. For this, I would like to thank a few people with whom I worked or to whose work I was exposed. I would like to thank Howard Gordon, Ed Oltman, Tom Trippe, Peter Grudberg, John Hauptman, Al Clark, Mark Strovink, Rich Partridge, Chip Stewart, Tom Ferbel, Paul Grannis, Hugh Montgomery, Harrison Prosper, and a few others in DØ experiment.

I would like to thank Harrison Prosper for several very instructive discussions through which I learned a lot of useful ideas in statistical data analysis. Also, I thank Mark Strovink, Tom Ferbel, and Rich Partridge for their efforts in organizing the mass fitting group in DØ.

I would like to give the most special thanks to John M. Hauptman, my thesis adviser. The opportunities that he offered me were more than I could take. My scientific ideas were always encouraged to be pursued while I was strongly advised with his globally viewed and experienced viewpoint. This was incredibly beneficial

to me in terms of developing myself as an independent and critical thinker. I thank him for being an easy and patient partner in numerous discussions we had for the last 4 or 5 years. I give him many more thanks for his consistent willingness to help.

I would like to thank Iowa State University α HEP group for its active support for my research. I would like to thank Chip Stewart and John Hauptman for reviewing my thesis and for their criticism.

Lastly, I would like to thank my parents for being supportive and proud of my work for the last six years.

CHAPTER 1. THE SIGNIFICANCE OF TOP DISCOVERY AND PRECISION MEASUREMENT OF IT

Why should there be top Quark?

Historically, two of the most fundamental scientific questions have been "what are the constituents of matters?" and "how do they interact with each other?" These questions have been answered by chemists, atomic physicists, nuclear physicists, and now by particle physicists. This historical progression of science in the understanding of matter and its interactions led us to a more universal, unified, and inclusive description of our physical world at the smallest scales. Not so long ago, our understanding reached to questions such as 'what are nucleons made of?' and 'how do the constituents of nucleons interact among themselves?'. To answer these questions, particle physicists have performed experiments to test a series of models and theories during the past several decades.

In the 1960s, significant amounts of data on baryon and meson resonances were taken. Regularities and patterns were observed suggesting that there is a higher level of symmetry. This led to the static quark model in which the pattern could be accounted for in terms of three quark constituents (u, d, and s). The discovery of $\psi(c\bar{c})$ and $\Upsilon(b\bar{b})$ in the 1970s added two more quarks to our understanding of constituents. Over many of these experiences, a model was established from some experimental

results and put into a theoretical framework that has survived a significant amount of testing (some of which is yet to be made). This is what we call 'The Standard Model'.

In the Standard Model, the most fundamental particles are three generations of leptons

$$\begin{pmatrix} e \\ \nu_e \end{pmatrix} \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix} \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix} \tag{1.1}$$

and three generations of quarks.

$$\begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix} \tag{1.2}$$

Also, our understanding of the most fundamental interaction as of now (some of them are parts of the Standard Model) suggest four different forces, electromagnetic, weak, strong and gravitational forces, which are mediated by gauge bosons in the framework of a gauge theory. The summary of these forces is shown in Table 1.1.

Table 1.1: Fundamental forces and gauge bosons.

Force	Boson name	Symbol	Charge	Spin	Mass
Strong	gluon	g	0	1	0 GeV
Electromagnetic	photon	γ	0	1	0 GeV
Weak	W	$W^{\pm}$	± 1	1	80 GeV
	Z	Z^0	0	1	91 GeV
Gravitational	graviton	G	0	2	0 GeV

As mentioned above, five of the six quarks expected from the Standard Model have been found. Why did we ever expect a sixth one? One example that predicts the top quark is from the forward-backward symmetry in the process $e^+e^- \longrightarrow b\bar{b}$. In an e^+e^- collider, there are two contributions to $b\bar{b}$ production as shown in Figure 1.1. The angular distribution is symmetric with respect to 90° from either contribution,

Figure 1.1: Feynman diagram contributing to $b\bar{b}$ production.

but when both are present they can interfere, and an asymmetric distribution results. Measurement of this asymmetry provides the relative contributions to $b\bar{b}$ production and this gives the coupling of the b to the Z. Since the coupling of the b to the Z is proportional to $(T_3^b + \frac{1}{3}sin^2\theta_w)$, we can determine T_3^b from the coupling. The LEP experiments measure T_3^b to be $-0.49^{+0.046}_{-0.022}$ indicating that b is in a weak isospin doublet and there should exist its partner t.

Another indirect prediction of the top quark is from the b-quark decay in the Standard Model. b-quark decay occurs through quark mixing, and the allowed vertices are $b \longrightarrow c + W^-$ and $b \longrightarrow u + W^-$. They are proportional to the elements V_{bc} and V_{bu} of the Cabibbo-Kobayashi-Maskawa (CKM) matrix. But this picture requires that b be in a doublet, so t is required by the Standard Model.

Beyond the discovery of top quark

The Collider Detector of Fermilab (CDF) submitted a long paper on evidence for top quark production in $\bar{p}p$ collision at $\sqrt{s} = 1.8$ TeV [2], yet they still didn't claim the discovery. This experiment, DØ, has accumulated similar, but statistically weaker, evidence. If the top quark is there, it will be found one way or the other as both experiments accumulate more data. The next step is to accurately measure the top mass. This probably is as important as the discovery in the sense that it fixes one of the few unknown parameters which are fundamental in the Standard Model.

The Standard Model has a few parameters which are essential to describe the electroweak interaction. They are

1. The fine structure constant ($\alpha = \frac{1}{137.036}$) determined from the quantum Hall effect;
2. The Fermi constant ($G_f = 1.16639 \times 10^{-5} GeV^{-2}$) determined from the muon lifetime through

$$\tau^{-1} = \frac{G_f^2 m_\mu^5}{192\pi^2}(1 - 8\frac{m_e^2}{m_\mu^2})(1 + \frac{3}{5}\frac{m_\mu^2}{M_W^2})[1 - O(\alpha)]; \qquad (1.3)$$

3. $sin^2\theta_w$ determined from neutral current process, the W and Z masses, and Z-pole observables;
4. CKM mixing angles;
5. Fermion masses;
6. M_H, the Higgs mass.

Under the assumption that the Standard model is correct, fixing the top mass pro-

vides useful information in determining other important parameters in the Standard Model. For example, the value of $sin^2(\theta_W)$ which is dependent on the renormalization scheme, can be expressed as $sin^2(\theta_W)(M_Z) = C(M_t, M_H)$ relating M_t with M_H in the modified minimal subtraction scheme ($\overline{ms}$). Therefore, measuring the top mass accurately would narrow down the search region for higgs boson.

In a more global point of view, we are at the point where we have this beautiful Standard Model, which describes the real world correctly, but is not complete. Therefore we are facing two possibilities. One is that we show that this model is not complete. The search is then on to discover the missing elements. On the other hand, if we complete this model, we can use this model as a confirmed block of knowledge to build a larger picture that unifies all the different forces.

Before Einstein's special relativity, people thought that the electric and magnetic forces were two different forces. These two forces were combined within the framework of special relativity and called the electromagnetic force. Now, we have the Standard Model that describes the electromagnetic interaction as well as the weak interaction. But there is a fundamental difference between these two cases. The SU(2) (weak interaction)$\times$ U(1) (electromagnetic interaction) gauge group is a product of two disconnected sets of gauge transformations: the SU(2) group with coupling constant g and the U(1) group with strength $g\prime$. Therefore, these two couplings are not related by the theory but experimentally measured as $\frac{g\prime}{g} = tan(\theta_W)$ whereas in the previous case, we have only one coupling $g\prime$ for both electric and magnetic forces. Only if the SU(2) and U(1) gauge transformations are embedded into a larger transformation G, can g and $g\prime$ be related by gauge theory. Including the color gauge group SU(3), the

unified group would be represented as

$$G \supset SU(3) \times SU(2) \times U(1) \tag{1.4}$$

Once the gauge group G has been found, all the interactions (except gravitational force) would be described by a Grand Unified gauge Theory (GUT) with a single coupling G! This unification is pictured in Figure. 1.2. Georgi and Glashow have shown that the smallest such unified group gauge transformation is the group SU(5). This model requires new colored superheavy gauge bosons (X, Y) which mediate interactions which turn quarks into leptons. The estimation of the proton lifetime comes from the argument of long muon lifetime which is a direct result of the large mass of the W. The muon lifetime is approximately $\frac{M_W^4}{m_\mu^5}$ and by the same analogy proton lifetime would be $\frac{M_X^4}{m_p^5}$ where $M_X = 10^{15} GeV$ is the mass of the X boson. The estimated proton lifetime of 10^{30} years is lower than the experimental limit of 10^{33} years.

Some of these ideas may establish a solid foundation and some of them will be wrong when we test the Standard Model, which will allow us to step forward in a more focused way when we ask the same questions "what are the constituents of matter?" and "How do they interact with each other?" at a more fundamental level.

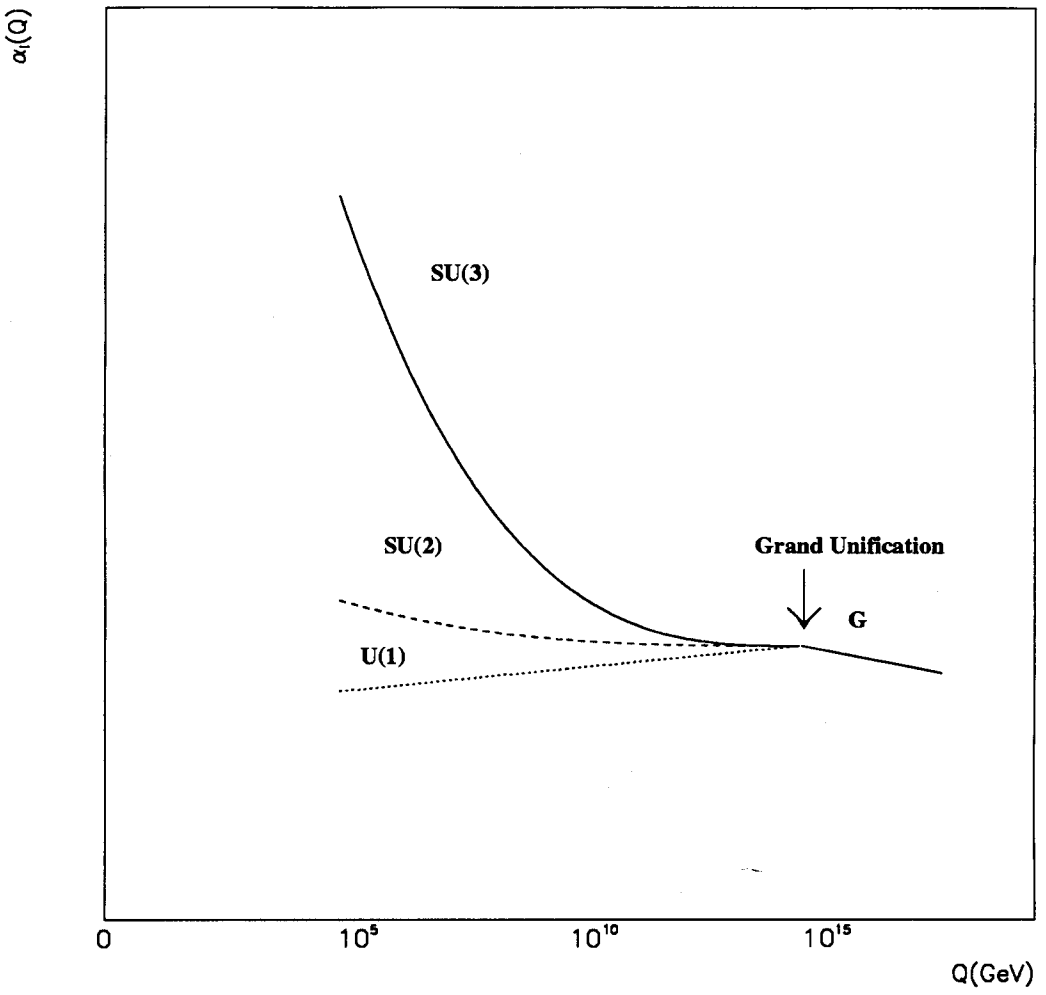

Figure 1.2: A speculative grand unification of SU(3), and electroweak (SU(2)×U(1)) interactions at very short distance $\frac{1}{Q} = \frac{1}{10^{15} GeV}$.

CHAPTER 2. OVERALL VIEW OF DØ DETECTOR

The DØ detector used in run 1a (1992-1993) can be described as a combination of three major parts; central tracking, calorimeter, and muon chambers. Central tracking provides information on the presence of tracks at particular coordinates, dE/dx of the tracks, accurate positions of the tracks and vertex position from which the directions of jets and electrons are determined, etc.

The calorimeter is finely segmented ($\Delta\eta = \Delta\varphi = 0.1$ for both EM and hadronic calorimeter, and $\Delta\eta = \Delta\varphi = 0.05$ in third layer of EM calorimeter where the shower profile is maximum). The calorimeter can identify electromagnetic showers by analyzing the longitudinal and transverse shower shape, and can measure the energies and the positions of electrons as well as jets. The fact that the calorimeter is hermetic and thick allows us to measure the $\not\!\!E_t$ well by requiring transverse momentum balance.

One of the three layers of the muon chamber is within the muon toroid just outside the hadronic calorimeter and two other layers outside the toroid are well separated to provide a long lever arm ($\geq 1m$) to yield good direction measurement after the bend in the muon toroid magnet. By measuring muon tracks before and after the magnetic field, we can reconstruct the bending angle and thus the momentum of the muon. Figure 2.1 shows the overall view of the DØ detector. To summarize

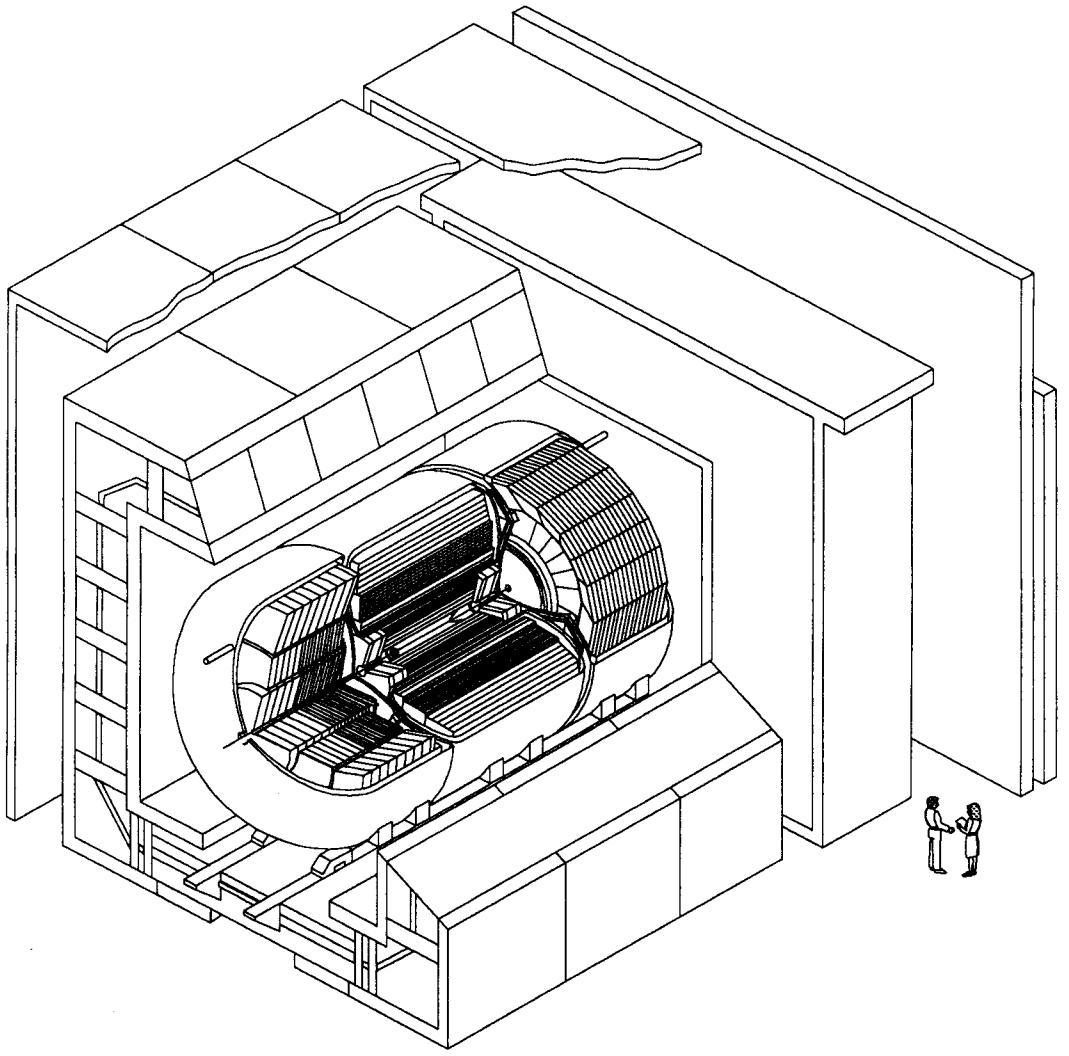

Figure 2.1: Overall view of DØ detector.

the strengths and weaknesses of the DØ detector:

Strengths

1. The calorimeter is hermetic so that $\not\!\!E_t$ can be well measured.
2. The calorimeter is thick so that all the electrons and jets are contained, and also the punch-through rate is low.
3. The calorimeter is finely segmented so that good position measurements as well as sophisticated cluster shape analysis for particle ID are possible.
4. The iron muon toroid allows muon momentum measurement.
5. Muon coverage is large.

Weaknesses

1. Central tracking suffers from high charged track multiplicity mostly from low momentum tracks due to not having a magnetic field in central tracking region.
2. Muon momentum resolution completely relies on its measurements after its passage through the calorimeter, and low momentum muons can suffer from the multiple scattering in the calorimeter.
3. No absolute z position measurement to which other detectors can be calibrated is made in central tracking.

DØ measures the energies of jets and electrons purely from the calorimeter. The energy resolution of jets and electrons is, for the most part, intrinsic to the

calorimetry. Therefore, it would be worthwhile to mention what contributes to the energy resolution. DØ is a sampling calorimeter and its energy resolution is often expressed as follows.

$$\frac{\sigma_E^2}{E^2} = C^2 + \frac{S^2}{E} + \frac{N^2}{E^2} \tag{2.1}$$

The first term (C) is a constant term and it's mainly due to the fluctuation of electromagnetic shower fraction over the total. Therefore, it's very small for electrons but relatively big for jets. N is the noise term which includes electronic noise, background radiation, and especially for DØ, uranium noise. These noise terms are independent of particle energy. S is the stochastic sampling term. This is due to the statistical nature of the shower development. In the showering process, the number of charged particles produced is roughly proportional to the energy of the incident particle. If we assume that each charged particle deposits the same amount of ionization on the average, the calorimeter response will follow Poisson statistics. Actual values of these constants for DØ will be discussed in later chapters when this information is needed for fitting.

CHAPTER 3. FUNCTIONAL DESCRIPTION OF DØ DETECTOR

In this chapter, I will describe some of the most fundamental measurements upon which our physics analysis is based. Some of the techniques are unique to DØ [1] and will be described rather qualitatively. I will focus mostly on the measurement techniques which are relevant to my analysis of top search in the $e + jets$ channel.

Z vertex measurement

Typically the interaction point along z axis (parallel to the beam) has a RMS spread of about a foot. Quantities such as E_t of electrons or jets can be calculated only if we know the z vertex position, since all that the calorimeter measures is the energy and the location of the shower, but not the direction. The direction information can be completed by knowing where the origin of the energy deposited in the calorimeter along z, namely the z vertex position. Therefore, it's one of the most fundamental measurements to reconstruct the four vectors of physical objects such as jets and electrons and even muons. This measurement is done by Central Drift Chamber (CDC). The CDC as a part of central tracking system is shown in Figure 3.1. When a charged track goes through 4 layers of CDC, there can be at most $7 \times 4 = 28$ hits of which we measure the x, y positions. x, y positions come from the drift time and the location of the sense wire. The z positions come from time

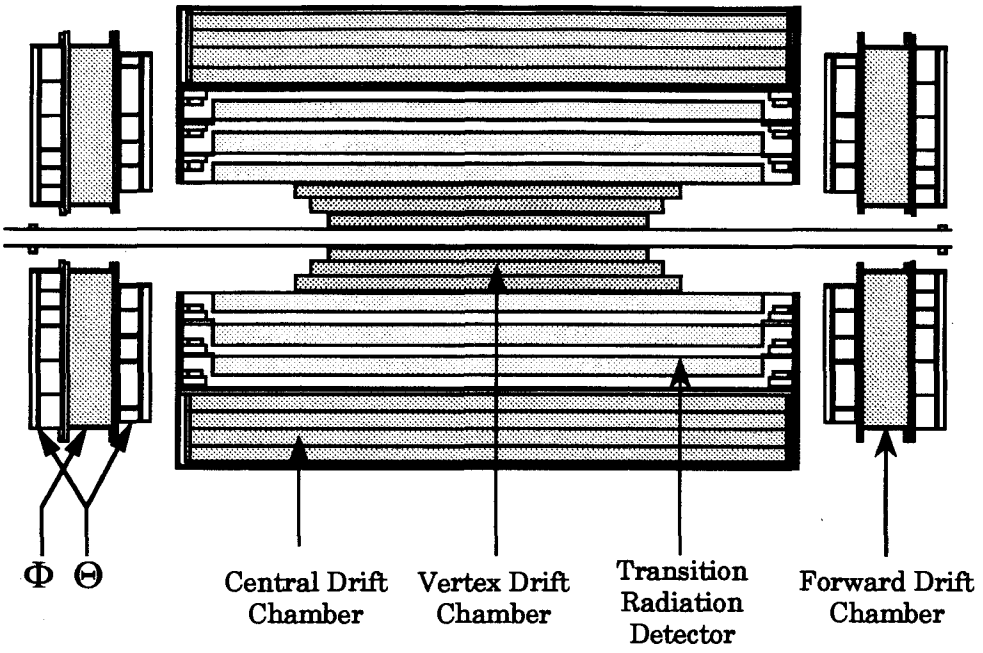

Figure 3.1: $X - Z$ view of central tracking system.

differences between two signals collected from both ends of the delay lines induced by the nearest anode (sense) wires. These 3 dimensional hits in space produced by a track are used to reconstruct a track by pattern recognition software.

For each event, CDC tracks with small $x - y$ impact parameters are chosen (to eliminate multiple scattered low momentum tracks) and they are projected to $x = y = 0$ in $r - z$ plane. These projected z positions are histogrammed to find the z position of the interaction. The distribution of the z interaction points at DØ is shown in Figure 3.2. The accuracy of this event-by-event measurement of the z interaction point is order of 1-2 cm.

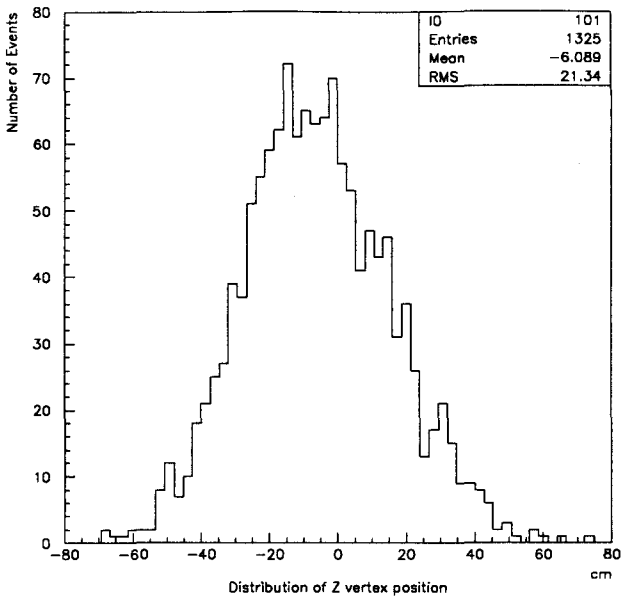

Figure 3.2: Distribution of Z vertex in $W \longrightarrow e\nu$ events.

$X - Y$ vertex measurement

The RMS spread of the interaction points in $x - y$ plane is much smaller than the one in z direction and is of order $50\mu m$. Our capability of measuring the tracks in transverse space with respect to the beam direction is much superior to that of measuring the z position. The systematic movement of the interaction point within a run (which is typically a few hours long) was measured to be also small ($< 50\mu m$). Again, we want to know the x, y interaction point to calculate the momentum vectors of the physical tracks we measure. However, our capability of measuring x, y position accurately is so much better than that of measuring just track direction that we can even think of looking for a displaced vertex in semi-leptonic b decay. To do this, we need to measure the $x - y$ positions of the primary interaction points (IP).

For the measurement of IP, we use Vertex Drift Chamber (VTX) and CDC

together. The reason why we need CDC is because we need the θ and Z_{CoG} (CoG = center of gravity) information of CDC track to correct the twist in layer 0 segment of the VTX track. (This can not be corrected by VTX information only due to poor z measurement of VTX chamber as of now.) After the twist correction on VTX tracks, we project VTX tracks to either x axis or y axis (centered at nominal IP) depending on their φ angle to get histograms of x, y positions. We determine the average interaction point (IP) for each run. The reason we calculate the IP for a whole run is because our measurement of IP from a single event has a larger error than the true spread of the IPs (about $50\mu m$) and also the movements of the IPs during a run are smaller than our measurement error from individual events. Using many events to calculate the beam position, we can calculate the average beam position to $50\mu m$ accuracy. Figure 3.3 shows the movement of the $x - y$ vertex position throughout the run 1a and the distribution of impact parameters calculated from these run dependent beam positions. The tracks used were reconstructed using both CDC and VTX tracks combined.

Electron identification

A lot of interesting physics events have leptons in them. For these events, by requiring a lepton in an event, we can usually eliminate a significant number of background events. To do so, however, it's essential that we identify leptons efficiently and accurately.

Electron identification starts with electromagnetic cluster finding in the EM calorimeter. The idea of this cluster finding method is to find seed towers above threshold and look at the next nearest neighboring towers to determine whether to

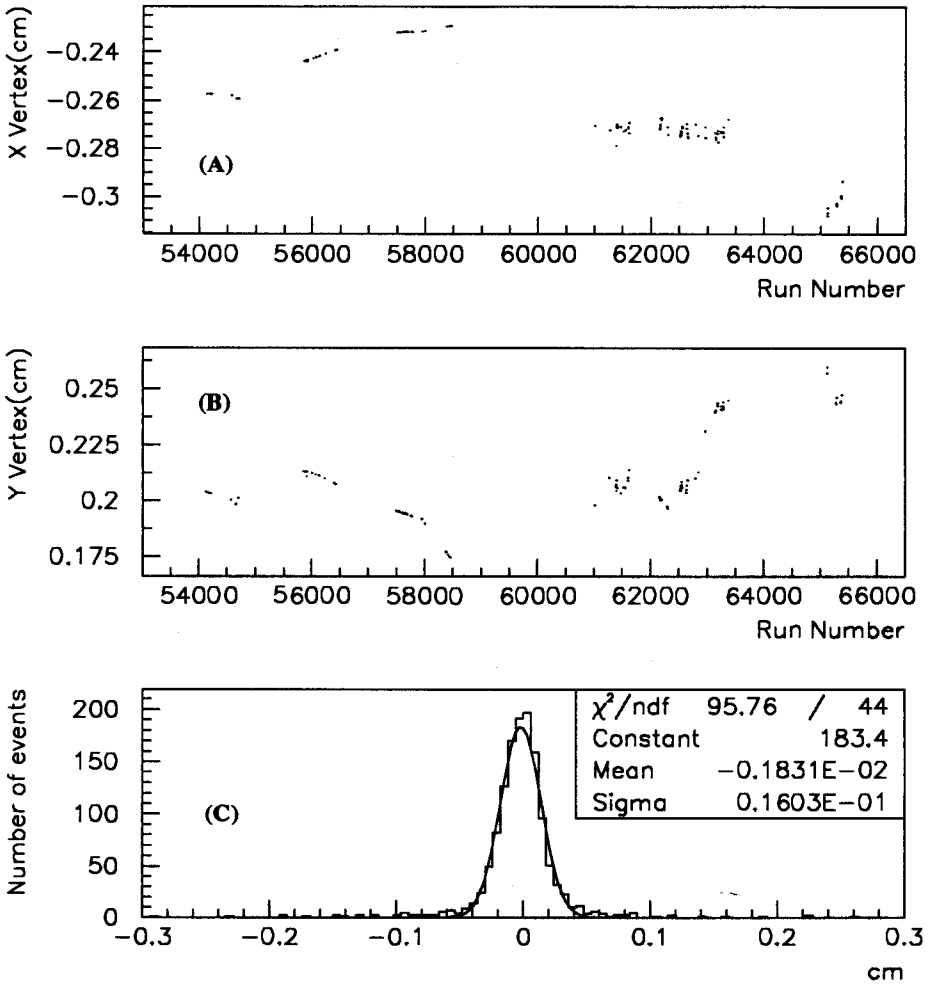

Figure 3.3: Mean x interaction point vs run number (A), mean y interaction point vs run number (B), and the impact parameter distribution of high P_t electrons from W decay calculated using the mean interaction points shown in (A) & (B).

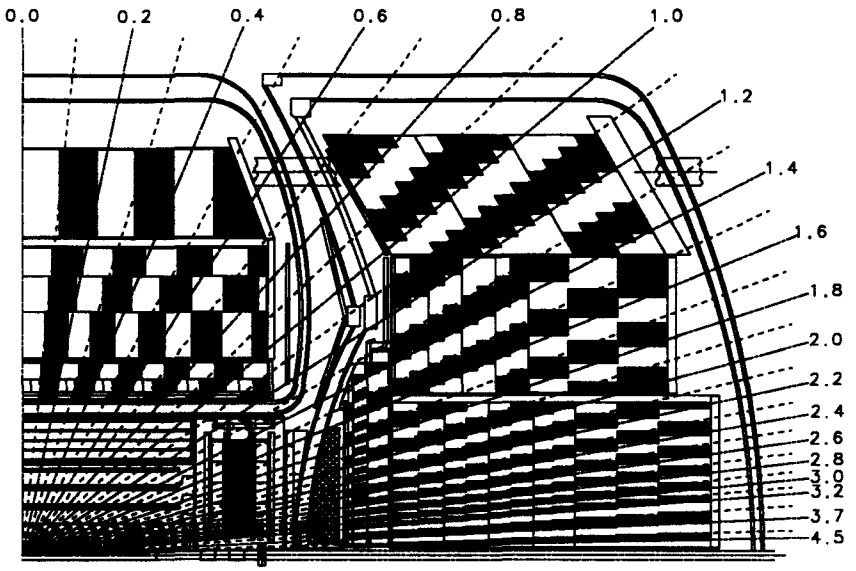

Figure 3.4: Calorimeter tower structure in η.

combine them with the seed towers or not based on some criteria. Calorimeter tower structure in η is shown in Figure 3.4. After cluster finding is done, a shape analysis for the cluster is done to see whether the shower shape of this cluster resembles a modeled electron. One of the method being used is the H-matrix method, using the inverse of the covariance matrix trained (calculated) from Monte Carlo electrons. The input to this matrix is general enough to describe the complete transverse and longitudinal shape of the cluster with its segmentation. Applying this trained H-matrix to a measured cluster provides a χ^2_H that is related to the probability that this cluster is from a real electron. The χ^2_H is defined as follows.

$$\chi^2_H = \sum_{i,j} (x_i^k - \langle x_i \rangle) H_{ij} (x_j^k - \langle x_j \rangle) \tag{3.1}$$

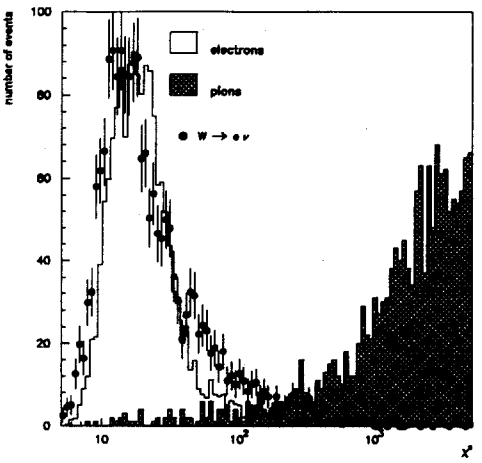

Figure 3.5: χ^2_H distribution for test beam electrons (*unshaded*), test beam pions (*shaded*), and electrons from W's (*dots*).

where H^{-1} is the covariance matrix,

$$H^{-1} = M_{i,j} = \frac{1}{N} \sum_{n=1}^{N} (x_i^n - \langle x_i \rangle)(x_j^n - \langle x_j \rangle) \tag{3.2}$$

and x_i^n is the observable i for electron n in the training sample.

Figure 3.5 shows the discriminating power of the χ^2_H when applied to the test beam data [1] and electrons from W decay. Along this line of cluster shape analysis for electromagnetic shower, a lot of effort and progress has been made on applications of Neural Network (NN) which can take nonlinear correlations among measured parameters into account, whereas the H-matrix method only takes linear

[1] Well controlled beam of electrons and pions in fixed target experimental area at Fermilab.

correlations into account.

For isolated electrons (electrons from W decay for example), we require an isolation cut. The isolation of a cluster is measured as follows.

$$\frac{E_{tot}(\Delta R < 0.4) - E_{em}(\Delta R < 0.2)}{E_{em}(\Delta R < 0.2)} \tag{3.3}$$

We also require an EM fraction defined as

$$\frac{E_{em}}{E_{em} + E_{FH1}} \tag{3.4}$$

to be above a certain value to assure it's a real electromagnetic shower. Finally, to distinguish electrons from photons, we require the track matching significance to be smaller than a certain value where the track matching significance is defined as

$$\frac{Cluster\ \ Position - Track\ \ Position}{\sqrt{\sigma^2_{track} + \sigma^2_{cluster}}} \tag{3.5}$$

For events with high P_t electrons from Ws, we can purify the event sample indirectly by requiring large missing E_t since the leptonic W decay produces a high P_t neutrino which doesn't interact with the detector, thereby leaving a huge P_t imbalance.

Missing E_t ($\not\!\!E_t$) measurement

As described in the previous chapter, one of the strong aspects of the DØ detector is that it's hermetic. The calorimeter covers almost the whole solid angle. Therefore, we can strictly impose the transverse momentum balance constraint to an event. The measurement is made by summing up transverse vector components of all the calorimeter energy cells (plus muon momenta, if there are muons in the

event) assuming that all the energy in a cell is deposited at the center of the cell (thus momentum = energy). After we sum up all the energies (or momenta) if it doesn't add up to zero, then this is due to at least one of the following reasons.

1. The fluctuation of energy measurement at each cell (and/or error on momentum measurement of muon), or
2. Missing neutrino, or
3. Particles hit dead material, or
4. Particles go through the beam pipe.

Usually $\not\!\!E_t$ of our interest is from the second contribution above. In practice, it's hard to separate the second contribution from the other contributions.

When there is no missing neutrino, the $\not\!\!E_t$ resolution is shown in Figure 3.6. In hard scattering the energy measured in each cell is either from parton energy (leptons, hadronized parton, etc) or from underlying events. I will try to separate these two contributions later in the fitting chapter.

Jet measurement

A jet is an ill-defined physical object, especially at low energy. A jet refers to a bunch of particles produced in the hadronization process of a quark or a gluon. It usually appears as a cluster of energy in the calorimeter. The measurement of a jet begins by identifying such a cluster. Usually we look for a calorimeter cell above certain E_t, and from there we have several different methods of further confirming the presence of jet and measuring the quantities of interest, such as the four vector of the jet.

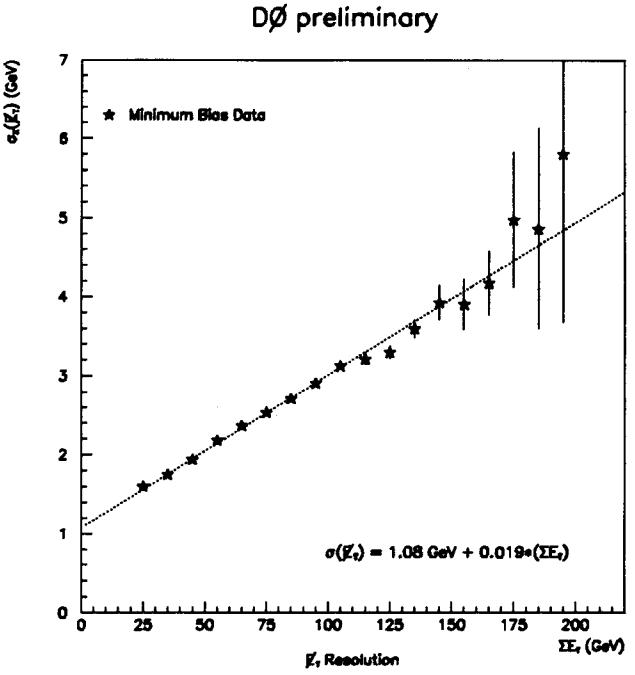

Figure 3.6: Calorimeter $\not\!\!E_t$ resolution function for the DØ detector for minimum bias data.

The Cone Algorithm (CA) is the most commonly used algorithm in DØ. This method draws a circular boundary in $\eta - \varphi$ space around the cluster of energy and measures the energy inside the cone by summing up the vector component of calorimeter cell energies within the cone. There are different cone sizes commonly used. They are 0.3, 0.5 and 0.7 in $\Delta R(= \sqrt{(\Delta\varphi^2 + \Delta\eta^2)})$. The advantage of this method would be that it's simple. On the other hand, the disadvantage would be that it doesn't take advantage of the cluster shape information but just decides whether to include a tower in the boundary of a fixed shape.

Nearest Neighbor Algorithm (NNA) is similar to electron cluster algorithm. It looks for a seed tower and grows the cluster based on more sophisticated information compared to the cone algorithm. In a situation where there are a lot of jets in an event so that the merging of jets is very likely, this method performs better in splitting the two jets that the cone algorithm could have merged because it not only looks at where the tower is but also its energy relative to the neighboring towers. Unfortunately, this method hasn't been getting much attention in DØ. But the optimization is being worked on and the test of its performance is under progress.

CHAPTER 4. DATA ACQUISITION

Six proton and anti-proton bunches circulate around the Tevatron ring of radius 1 km. This gives bunch crossing time of 3.5 μs. At $L = 10^{30} cm^{-2} s^{-1}$ there are on average 0.3 interactions per bunch crossing. Each interaction is filtered through three layers of triggers before it is written to tape for offline analysis. The three layers are called Level 0 trigger, Level 1 trigger, and Level 2 trigger.

Level 0 trigger is from hodoscopes of scintillators mounted on the front surfaces of the end calorimeters. It registers the presence of inelastic collisions and serves as the primary luminosity monitor for the experiment. Its efficiency of detecting inelastic collision is $\geq 99\%$.

Level 1 trigger involves three different detectors, calorimeter, muon chamber, and the Transition Radiation Detector (TRD). One of the important capabilities of Level 1 trigger is that it makes its trigger decision within the bunch crossing of 3.5 μs. Therefore it doesn't introduce any deadtime. The information available at this stage of Level 1 decision is

1. The number of electromagnetic (EM) and total (EM+Hadronic) trigger towers ($\Delta\eta = \Delta\varphi = 0.2$) above a preset E_t threshold.
2. The scaler sum of all E_t in the detector.
3. The $\not\!\!E_t$.

4. The number of muons in the various regions, etc.

Based on this information it performs various (32 allowable Level 1 triggers) vetos and also provides prescaling of triggers too copious to pass on without rate reduction. The typical rate that passes the Level 1 trigger is about 200 H_z.

Level 2 trigger is performed by 50 Level 2 processor nodes. It serves its purpose as an event builder as well as a more sophisticated filter to select events, reducing the input rate of about 100 H_z down to 2 H_z. Upon the arrival of the raw data, it does a preliminary reconstruction of the events and decides whether they pass at least one of the 128 filters set up based on different physics interests.

The events that pass all the triggers including the Level 2 are sent to the host computers to be written to tapes. Some of the triggers set up for very interesting physics topics (especially for top search) are processed directly by host computers to provide reconstructed information right away. This stream of data is called the Express Line.

CHAPTER 5. SIGNAL AND BACKGROUND

Signal production

At the Tevatron, the top quark will be mainly produced through $t\bar{t}$ pair creation. The lowest order Feynman diagrams are shown in Figure 5.1. As we include the higher order contributions, the $q\bar{q}$ channel is not affected significantly, whereas the gg channel needs 70% or more correction. The relative contribution of the gg and $q\bar{q}$ is such that the $q\bar{q}$ contribution keeps getting larger than the gg channel as the top quark mass increases. At M_t of 150 GeV, the gg contribution is around 20% and decreases down to 10% at around 200 GeV [3].

Including all these contributions, Table 5.1 shows the theoretical prediction of cross section of $t\bar{t}$ production as a function of the top mass [3].

Decay modes and backgrounds

When M_t is greater than the mass of W plus the mass of the b quark, the branching ratio of $t \longrightarrow W+b$ is almost 100%. Therefore the subsequent decay modes are determined by how the Ws in $t\bar{t}$ events decay. Table 5.2 shows the branching ratios of $t\bar{t}$ events.

The background depends on the channel (decay mode). For example, if one is looking at $t\bar{t} \longrightarrow$ *all jets*, the dominant background is QCD jet production. Since

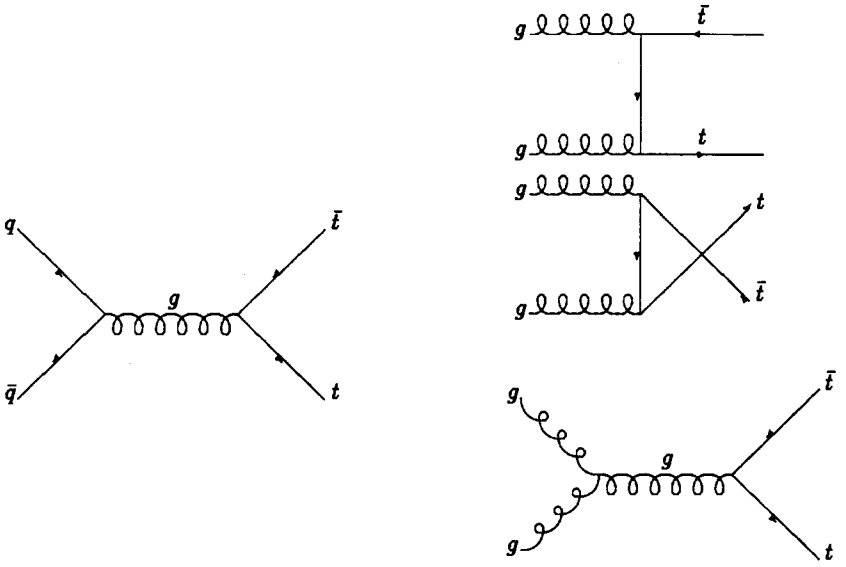

Figure 5.1: Lowest order $t\bar{t}$ production.

I am considering $t\bar{t} \longrightarrow e\nu + j_1 + j_2 + b\bar{b}$ channel, my background will have a high P_t isolated electron and large missing E_t. The type of events that satisfy these conditions are $W + jets$ events which are produced by the diagram shown in Figure 5.2.

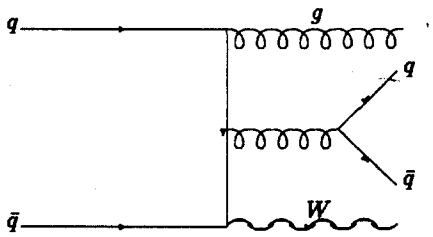

Figure 5.2: An example of background process to $e + jets$ channel.

Table 5.1: $t\bar{t}$ cross section ranges.

m_{top}	σ(pb), Lower	σ(pb), Central	σ(pb), Upper
90	148.00	180.00	259.00
100	86.30	102.00	141.00
110	52.70	61.60	81.40
120	33.70	38.90	49.70
130	22.30	25.40	31.20
140	15.10	16.90	20.50
150	10.50	11.70	13.80
160	7.41	8.16	9.53
170	5.32	5.83	6.68
180	3.86	4.21	4.78
190	2.83	3.06	3.44
200	2.09	2.26	2.52

Table 5.2: $t\bar{t}$ decay branching ratios.

Decay mode	Branching ratio
$t\bar{t} \longrightarrow (q\bar{q}b)(q\bar{q}\bar{b})$	36/81
$t\bar{t} \longrightarrow (q\bar{q}b)(e\nu\bar{b})$	12/81
$t\bar{t} \longrightarrow (q\bar{q}b)(\mu\nu\bar{b})$	12/81
$t\bar{t} \longrightarrow (q\bar{q}b)(\tau\nu\bar{b})$	12/81
$t\bar{t} \longrightarrow (e\nu b)(\mu\nu\bar{b})$	2/81
$t\bar{t} \longrightarrow (e\nu b)(\tau\nu\bar{b})$	2/81
$t\bar{t} \longrightarrow (\mu\nu b)(\tau\nu\bar{b})$	2/81
$t\bar{t} \longrightarrow (e\nu b)(e\nu\bar{b})$	1/81
$t\bar{t} \longrightarrow (\mu\nu b)(\mu\nu\bar{b})$	1/81
$t\bar{t} \longrightarrow (\tau\nu b)(\tau\nu\bar{b})$	1/81

CHAPTER 6. FITTING THE $e + jets$ CHANNEL (MC)

Introduction

Fitting of $t\bar{t}$ events that contain poorly measured objects such as jets and even unmeasured objects such as missing neutrino raises the question on whether we can reasonably measure the top mass for these events. Two of the the main purposes of this chapter are, first, to estimate the performance of how well this mass fitting technique works, and second, to separate various problems, investigate each of them to understand what problems are significant and what are trivial. These would give a direction on where we should spend our efforts to improve, and how we should utilize the result of this analysis.

Method

The $t\bar{t}$ decay hypothesis predicted by the Standard Model requires each top (t or $\bar{t}$) to decay into $W + b$ (or $\bar{b}$) where W is real (on mass shell) if the top mass is greater than the mass of W. The W from t or $\bar{t}$ decay subsequently decays into a lepton-antilepton or a quark-antiquark pair. The particular channel being studied here is when one of the Ws from $t\bar{t}$ decays into an electron and an anti-electron neutrino and the other W decays into two jets as shown in Figure 6.1. Under this assumption of $t\bar{t}$ decay, one can expect the following constraints being satisfied for

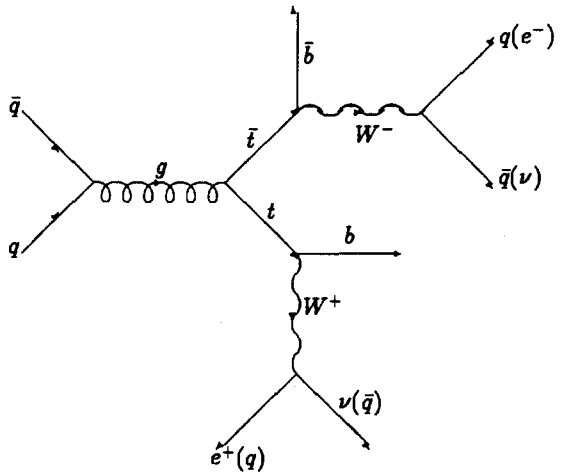

Figure 6.1: Feynman diagram for $t\bar{t}$ production and decay.

such events.

$$f_1 = \sum P_x = 0 \ (\textit{including neutrino}) \tag{6.1}$$

$$f_2 = \sum P_y = 0 \ (\textit{including neutrino}) \tag{6.2}$$

$$f_3 = M_{j_1 j_2} - M_W = 0 \tag{6.3}$$

$$f_4 = M_{e\nu} - M_W = 0 \tag{6.4}$$

$$f_5 = M_{e\nu b} - M_t = 0 \tag{6.5}$$

$$f_6 = M_{j_1 j_2 \bar{b}} - M_{\bar{t}} = 0 \tag{6.6}$$

Since we don't know the top mass, constraints 5 and 6 reduce to one constraint which is

$$f_5 = M_{j_1 j_2 \bar{b}} - M_{e\nu b} = 0 \tag{6.7}$$

For such $t\bar{t}$ events, measurements will be made for all the jets and electrons with certain efficiencies and resolutions; however, the momentum of the neutrinos will not

be measured directly since they don't interact in the detector. Therefore, a neutrino leaves three unmeasured quantities which are P_x, P_y, and P_z of the neutrino. There are several measured quantities, three unmeasured quantities and five constraint equations. Measured quantities don't introduce any unknowns to the system since they are measured. Each unmeasured quantity introduces one unknown to the system.

If the number of unmeasured quantities is the same as the number of independent constraints, one can always find a solution for the unmeasured quantities without varying the measured quantities. But if the number of the constraints exceed the number of unmeasured quantities, one has an over-constrained system where the measured quantities must vary from their measured values to satisfy all the constraints. The basic idea of the fitting is to find a set of numbers (fitting parameters) corresponding to the measured and unmeasured quantities which satisfy all the constraints and has minimum variations from their measured values (minimum χ^2). This χ^2 will be small for those events that meet the hypothesis of $t\bar{t}$ decay described above. Therefore, this χ^2 of the fit will tell us how well a particular combination of an event fits the hypothesis of top decay. And also, at this minimum χ^2, a certain value will be assigned to the invariant mass of electron, neutrino, and the b jet combined, which is the best estimate of the top mass if the combination is correct (i.e., the jets and electron are assigned to correct partons) within the $t\bar{t}$ event.

Measurements

In $t\bar{t}$ events with e+ jets decay mode, there will be at least 4 jets (sometimes more than 4 because of Final-State Radiation [1]), a neutrino, and an electron from

[1]FSR — The gluon jet radiated from quarks that decayed from t or $\bar{t}$

$t\bar{t}$. Also there will be some Initial-State Radiation (ISR) [2] jets and underlying event.

For jets and electrons, the four momenta are measured. $\not\!\!E_t$ is calculated from the vector sum of all the energies in calorimeter cells. These energies in the calorimeter cells are mainly from jets and electrons from hard collisions but some of them are due to the underlying minimum bias event. $\not\!\!E_t$, by definition, is the negative of the transverse vector sum of all the calorimeter cells. If the vector sum of energy cells is different from the vector sum of the reconstructed objects (jets and electrons), it's due to the fact that there are some residual energies which are not included in jets or electron. The fitting will have to know about this difference so that it wouldn't ignore the energies not being part of the reconstructed objects in its attempt to balance the transverse momentum. For this reason, I introduce a fictitious jet whose transverse vector momentum is the difference between the vector sum E_t of energy cells and the vector sum of reconstructed objects. I call this a baby jet. Strictly speaking, the baby jet is the measured quantity, not the $\not\!\!E_t$. But, in practice, it's equivalent to say that the $\not\!\!E_t$ is a measured quantity and the baby jet is not.

Fitting parameters

One consideration in determination of the fitting parameters was to minimize the correlation between the chosen parameters, that is, I choose uncorrelated parameters. For example, the momenta of jets and electron are directly correlated with the $\not\!\!E_t$. Therefore, I consider $\not\!\!E_t$ not as my measured quantity. Instead I introduce the baby jet as my measured quantity for which the correlation with jet/electron momenta is not as direct. Jet and electron parameters which are allowed to vary in the fitting are the magnitude of the momentum, azimuthal angle, polar angle, and the invariant

[2] ISR — Any gluon radiation that's not originated from t or $\bar{t}$

mass ($4 \times 5 = 20$), also P_x, P_y of ISRs and the baby jet ($2 \times 2 = 4$). The main reason why I don't combine ISRs and the baby jet is because I want to have a more clear understanding of the error assignment on them. Underlying events will have pretty much constant error or will depend on scalar sum E_t (only from underlying events), which will be small in general. The errors on P_x P_y of ISRs will depend on their relative orientation which we can calculate.

Considering the question of whether we should consider the mass of a jet as real measurement or an artifact of the process through which we detect the object, we can think of a couple of cases when jets would have masses. The first cases would be due to spread of the shower in the detector when a cluster of particles interact with material. The second case would be when a jet radiates gluons and that makes the jet broader. In this case, it will have rather large mass and it should be included in its calculation of energy as if the jet was a massive object. Since the effect of the first case is small compared to the second case, the decision was made to consider the jets as massive objects. However, it's pointed out later that the jet mass doesn't affect the kinematic fit a lot quantitatively.

Lagrange Multiplier and linear algebra

Let's define our variables as follows [14].

m = Measured value of well-measured variable

m^* = Measured value of badly-measured variable

G = Inverse square error matrix for m

$$G_{ij}^{-1} = \overline{\delta m_i \delta m_j} \tag{6.8}$$

G^* = Inverse square error matrix for m^*

$$G^{*-1}_{ij} = \overline{\delta m^*_i \delta m^*_j} \tag{6.9}$$

x = Fitted value of well-measured variable

x^* = Fitted value of badly-measured variable

f = Constraint equations

Here, the badly measured variables correspond to the P_x, P_y, P_z of neutrino to which we assign infinite errors, and therefore $G^* = 0$.

The χ^2 that we are minimizing is defined as follows.

$$\chi^2 = (x^* - m^*)^T G^* (x^* - m^*) + (x - m)^T G (x - m) \tag{6.10}$$

Introducing the Lagrange Multiplier, λ, we define

$$M = 2f^T\lambda + \chi^2 = 2f^T\lambda + c^{*T} G^* c^* + c^T G c \tag{6.11}$$

where $c = x - m$, $c^* = x^* - m^*$. We want to minimize M with respect to λ, x^*, and x. Thus we have

$$0 = \frac{1}{2}\frac{\partial M}{\partial \lambda} = f(x^*, x) \tag{6.12}$$

$$0 = \frac{1}{2}\frac{\partial M}{\partial x^*} = B^*\lambda + G^* c^* \tag{6.13}$$

$$0 = \frac{1}{2}\frac{\partial M}{\partial x} = B\lambda + Gc \tag{6.14}$$

where B is the matrix of derivatives of the constraints with respect to the fitting parameters. Solving the equations above is not an easy problem since the constraint equations $f_i(x^*, x)$ are not linear. However, one might be able to linearize these equations by expanding them and taking the leading terms under the assumption

that the constraints are reasonably linear within a given interval. Expanding $f_i(x^*, x)$ with respect to a trial solution $(\bar{x^*}, \bar{x})$ and taking the first order terms, we get

$$0 = f_i(x^*, x) = f(\bar{x^*}, \bar{x}) + B^{*T}(x^* - \bar{x^*}) + B^T(x - \bar{x}) = f + B^{*T}(c^* - \bar{c^*}) + B^T(c - \bar{c}) \tag{6.15}$$

or

$$B^{*T}c^* + B^T c = r \tag{6.16}$$

where it will be convenient to define

$$r = B^{*T}\bar{c^*} + B^T\bar{c} - f \tag{6.17}$$

All derivatives are evaluated at the point $(\bar{x^*}, \bar{x})$. Solving Eq. (6.14), I have

$$c = -G^{-1}B\lambda \tag{6.18}$$

Substituting Eq. (6.18) to Eq. (6.16), I have

$$-H\lambda + B^{*T}c^* = r \tag{6.19}$$

where

$$H = B^T G^{-1} B \tag{6.20}$$

Now combining all the equations, we get

$$\begin{pmatrix} -H & B^{*T} \\ B^* & G^* \end{pmatrix} \begin{pmatrix} \lambda \\ c^* \end{pmatrix} = \begin{pmatrix} r \\ 0 \end{pmatrix} \tag{6.21}$$

I can find λ and c^* by solving this equation. Then I calculate c by using Eq. (6.18). Since we have not solved the problem exactly but only in a linear approximation, we must check the new values of x^* and x to see if they satisfy the original equations. If they do not, we may use these values as new estimates $(\bar{x^*}, \bar{x})$ and repeat the procedure until convergence.

Combinatorial background

Given a $t\bar{t} \longrightarrow e + jets$ event which has two jets from a W, two jets from b and $\bar{b}$ and one electron, the fact that we don't know the parton identifications of the four jets introduces 12 ways of combining the jets to make an independent hypothesis, that is

$$\frac{P_4^4}{2} = 12 \tag{6.22}$$

Likewise, when there are 5, 6, 7 jets in an event, the numbers of independent combinations are 60, 180, and 420 respectively.

Out of these many combinations, only one is correct and is expected to give the right mass for the top and a small χ^2 from the fit. The question is "what fraction of the time would it give the smallest χ^2 of all combinations in the event?". The rest of the combinations are considered as background (combinatorial background) since there is no justification that these combinations should give correct mass or small χ^2 even if the event really is a $t\bar{t}$ event if the combination is completely wrong. However, it's worthwhile mentioning that a combination can be partially correct. An example would be the case when all the three jets from t or $\bar{t}$ are grouped correctly but b or $\bar{b}$ jet assignment is wrong. In this case, of course, the fit would preferentially give the correct t mass.

Since these wrong combinations don't fit the $t\bar{t}$ decay hypothesis, they would give larger χ^2. On the other hand, since there are so many wrong combinations, the chance that at least one of the wrong combinations giving better χ^2 than the correct one might be high. My goal is to quantify these various aspects.

Performance with Isajet MC and the effect of resolution

The level of combinatorial background depends on the number of combinations for each event, and also on the detector measurement resolution. As the resolution degrades, the parameter space (momentum, angles, mass space in this case) that gives wrong combination a χ^2 below that value for correct combinations enlarges, and more wrong combinations will give better χ^2 than the correct one. How much they merge at a certain resolution will tell us how serious the combinatorial background will be at that resolution. As an example, if the resolution is very good, selecting the correct combination based on χ^2 will guarantee a high efficiency, whereas when the resolution is poor selecting a combination based on χ^2 wouldn't necessarily guarantee a correct combination. In the latter case, the probability of selecting correct combination will asymptotically approach

$$\frac{1}{\#\ of\ combinations} \tag{6.23}$$

which means that χ^2 is not providing any useful information.

The purpose of this study is to see what our best results can be as a function of resolution in the absence of other systematic problems, so that we can set up an upper limit on what we can achieve.

ISAJET Monte Carlo generator was used to test the effect of resolution in selecting the correct combination in $t\bar{t}$ events at a top mass of 160 GeV. ISAJET simulates Final-State Radiation, but for simplicity, they were merged with the original partons that radiated them. So, I get one electron, 4 jets from $t\bar{t}$ and a few Initial-State Radiated jets. ISRs are correctly identified and, therefore, didn't introduce further combinatorial background. The momenta of those 4 jets and the

electron were smeared with a given resolution (% error$\times\sqrt{P}$) and the same errors were assigned in the fitting. Four vectors of ISRs were summed up and smearing was done on the momentum with a resolution corresponding to the sum of the jet energies. In summary, errors used are

$$\sigma_{P_{ISR}} = \frac{\%error_{jet}}{100\%}\sqrt{\sum E_{ISR}} \tag{6.24}$$

$$\sigma_{P_{jet}} = \frac{\%error_{jet}}{100\%}\sqrt{\sum E_{jet}} \tag{6.25}$$

$$\sigma_{P_e} = \frac{\%error_e}{100\%}\sqrt{\sum E_e} \tag{6.26}$$

$$\sigma_{\theta_e} = \sigma_{\varphi_e} = 0.005rad \tag{6.27}$$

$$\sigma_{\theta_{jet}} = \sigma_{\varphi_{jet}} = 0.05rad \tag{6.28}$$

Also the same errors were used for fitting.

Because ISAJET doesn't conserve momentum exactly at parton level for technical reasons, a baby jet was added to account for the momentum imbalance with resolution of 5 GeV.

The number of possible combinations in this case was 12 since I only loop over one solution of P_z^ν that minimizes the P_z of W. It was shown from Isajet MC that this choice of P_z^ν gives the correct solution 75% of the time. If the solution is imaginary, I changed the magnitude of the $\not\!\!E_t$ in both directions (increase/decrease keeping the direction the same) until the solution becomes real. It was shown, however, that

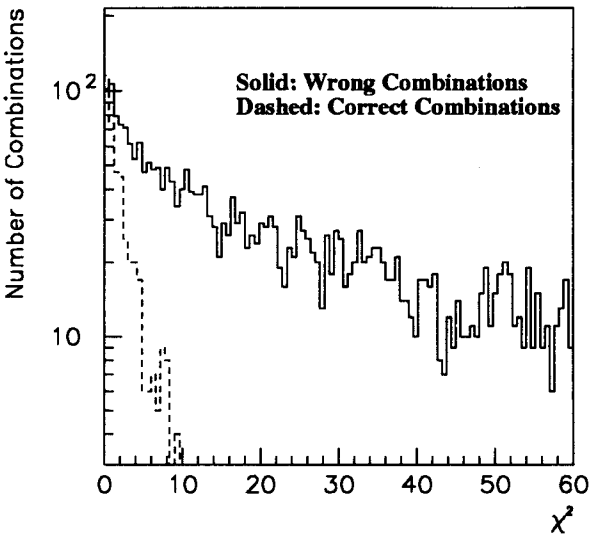

Figure 6.2: χ^2 of the correct combinations (*dashed*) and for the wrong combinations (*solid*).

looping over both P_z^ν solutions performs equivalently since the two combinations corresponding to the two P_z^ν solutions are not really independent, thus the additional contribution does not add combinatorial background. Figure 6.2 shows the difference in χ^2 distributions between correct combinations and the rest of the combinations. As one might notice, it's more likely that a correct combination will have the smallest χ^2. On the other hand, the correct combination is only a small fraction of the total. Figure 6.3 shows the effect of the jet resolution on fitted mass distribution when I choose the smallest χ^2 combinations. Table 6.1 shows more detailed aspects of what happens when the resolution degrades. One can notice the decrease in the

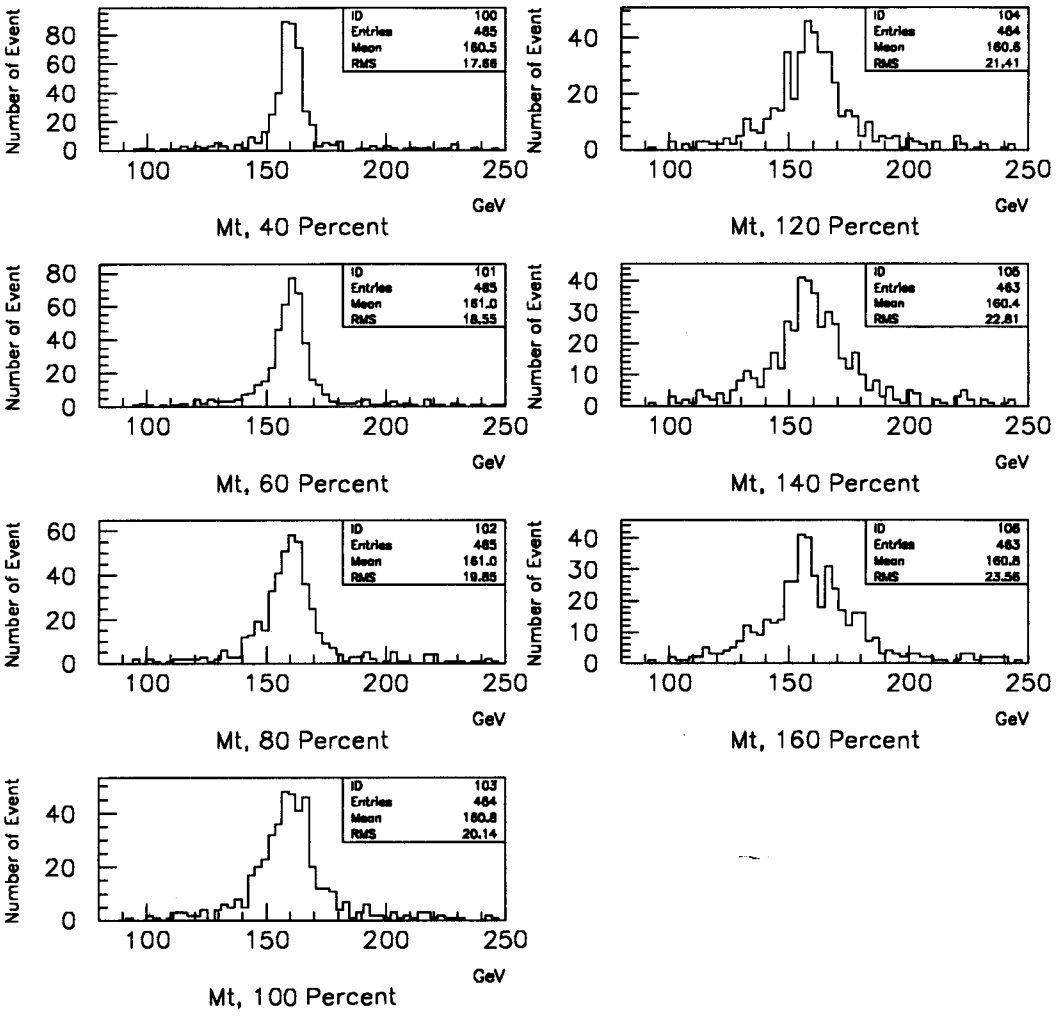

Figure 6.3: Mass resolution function of 160 GeV top from the smallest χ^2 combinations at various jet resolutions.

Table 6.1: Resolution vs. efficiencies. Fit attempt: Number of events fitted. Fit succeed: Number of events with at least one combination satisfying constraints. # Corr. 1st: Number of events in which the best χ^2 gives correct combination. σ_m 1st: Width of the mass distribution for '# Corr. 1st'. Eff. 1st: Correct selection efficiency for '# Corr. 1st'. # Corr. 2nd: Number of events in which the second best χ^2 gives correct combination. σ_m 2nd: Width of the mass distribution for '# Corr. 2nd'.

Jet Resol.	40%	60%	80%	100%	120%	140%	160%
Elec. Resol.	15%	15%	15%	15%	15%	15%	15%
Fit attempt	465	465	465	465	465	465	465
Fit succeed	465	465	465	464	464	463	463
# Corr. 1st	289	271	234	225	197	184	174
σ_m 1st	5.03	5.77	6.52	7.49	8.79	9.96	11.29
Eff. 1st	.622	.583	.500	.484	.424	.396	.374
# Corr. 2nd	106	106	133	97	103	95	85
σ_m 2nd	4.84	6.21	7.27	8.45	9.98	11.39	12.94
Eff. 2nd	.228	.228	.286	.209	.222	.204	.183

efficiencies of selecting the correct combinations based on χ^2 information as the jet resolution degrades. At the same time the resolution of the correct combination also degrades slowly. These efficiencies at given resolutions set the upper limits on how well we can select the correct combination out of 12 combinations and the resulting mass resolution.

These efficiencies give us the combinatorial background probabilities for a given $t\bar{t}$ event assuming that the probability of one wrong combination giving the best χ^2 is the same for all wrong combinations. For example, the probability of at least one wrong combination out of 11 giving better χ^2 than the correct one at jet resolution 80% is $\frac{(465-234)}{465} = 0.500 = (1-(1-x)^{11})$ where x is the probability that one wrong combination gives better χ^2 than the correct one. This probability is

$$x = 0.061 \tag{6.29}$$

If the number of combination increases by a factor of two (24, 1 correct, 23 wrong) and if I assume that all the wrong combinations are independent, the probability that at least one wrong combination gives better χ^2 than the correct one becomes (by binomial statistics)

$$\sum_{1}^{23} \frac{23!}{r!(23-r)!} x^r (1-x)^{23-r} \tag{6.30}$$

or simply

$$1. - (1. - x)^{23} = 0.765 \tag{6.31}$$

since $(1-x)^{23}$ is the probability that none of the 23 wrong combinations gives better χ^2 than the correct combination.

Table 6.2: Efficiency vs. number of combinations

Number of wrong combinations	Prob. that at least one wrong comb. gives better χ^2 than the correct
11	0.50
23	0.77
30	0.85
35	0.89
40	0.92
50	0.96
60	0.98
120	0.999

Likewise, the results for various numbers of wrong combinations are shown in Table 6.2 For large combinatorial background, fitting isn't capable of separating the correct combination from wrong combinations. This is understandable since the discriminating capability of the fitting is fixed as the number of wrong combinations increases.

If we can tag one b jet in $e+4jets$ event, then there are 6 ways of combining the jets instead of 12. Furthermore, as we get more information and reduce the number of combinations, we obtain the results shown in Table 6.3.

As a result of this study, we learn two things. One is that we are not utilizing all the information by choosing the best χ^2 combination since the second best χ^2 combinations also contain large number of correct combinations which give a peak at the correct mass (Table 6.1). The other is that the efficiency of selecting the correct combination is a very sensitive function of the number of possible combinations. Therefore, we can improve the efficiency by reducing the number of combinations (via b tagging, for example)

Table 6.3: Efficiency vs. number of combinations

Number of wrong combinations	Prob. that at least one wrong comb. gives better χ^2 than the correct
2	0.12
4	0.22
5	0.24

Other Isajet studies

Systematic errors on jet energy scale.

The fitted invariant mass of both the t and $\bar{t}$ is affected by the measurements of jets and electrons as well as the constraints. Therefore, an overestimate or underestimate of jet energy can certainly affect the top mass. For example, if we overestimate the energy of all the jets in an event and consider the three jets from t or $\bar{t}$, then two of three jets (corresponding to qq from W) will be scaled down to satisfy the W mass constraint, but the other jets (corresponding to the b or $\bar{b}$ along with the b jet

on leptonic side) will boost up the mass of top on each side, maintaining the t mass equality constraint. A more systematic study of this effect was done using smeared Isajet MC where we intentionally scaled up or down the jet energy. The result is shown in Figure 6.4, that the mass of the top scales linearly with a scale offset of the jet energy with the coefficient of

$$t\ mass\ shift \approx \frac{8.1 GeV}{10\%\ Offset} \tag{6.32}$$

Wrong assignment of errors

The correct assignment of errors has significance in making the χ^2 meaningful. However, χ^2 is used only to select the most probable correct combination. This means that only the relative differences in χ^2 among different combinations are important, not the absolute values. If an event consists of just jets, for example, whatever error I assign doesn't matter as long as I assign errors consistently to all the objects because it only scales up and down the χ^2s of all the combinations, but does not change the order.

In the $e + jets$ channel, the relative error assignment to jets and electrons has to be correct. Mis-assigned errors may increase the probability of promoting one of the 11 wrong combinations to have the best χ^2.

Detector simulation and sources of inefficiencies

Electron reconstruction efficiency

Electron reconstruction efficiency plays a little different role than the jet reconstruction efficiency in the sense that we require one electron, and based on whether

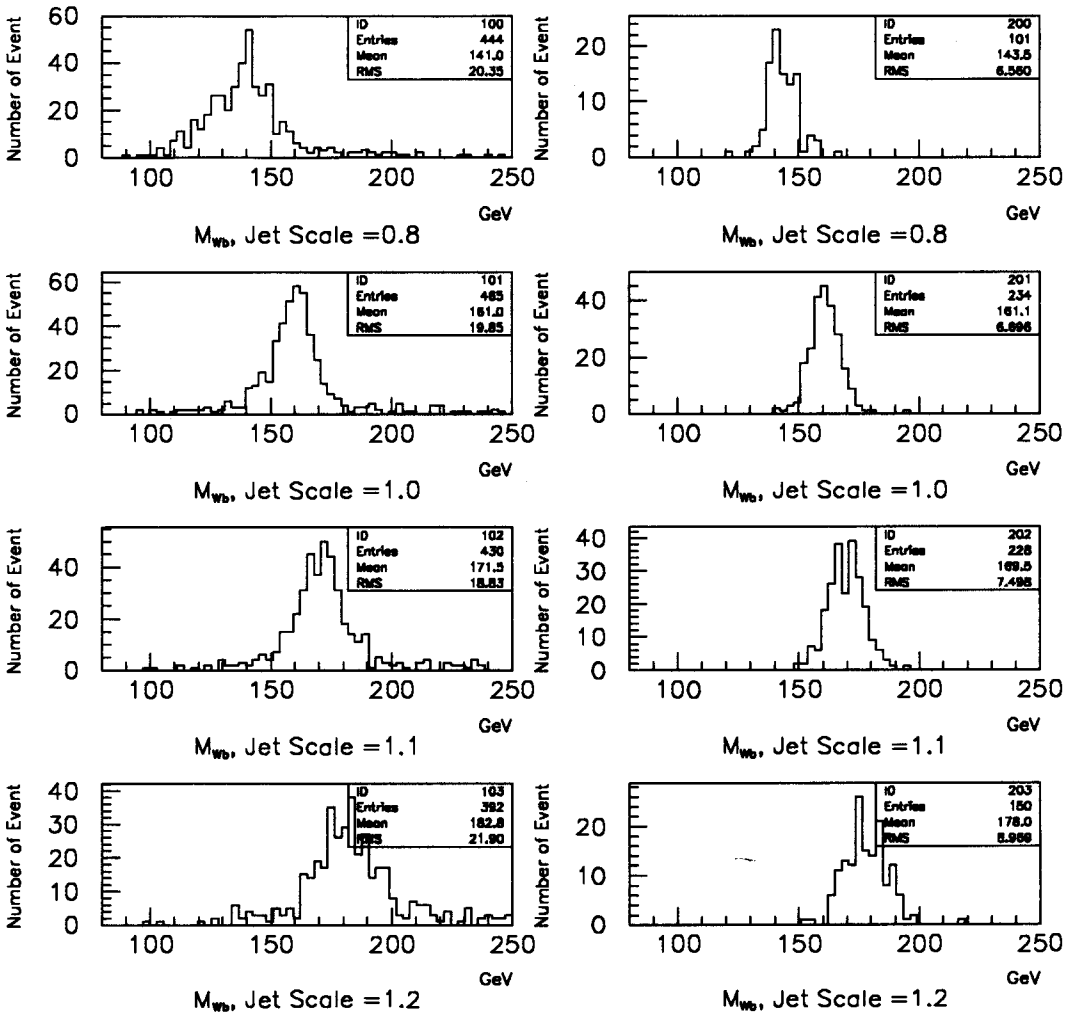

Figure 6.4: Fitted Wb mass with scale offsets in jet energy. (160 GeV top events at jet resolution 80%)

there was an electron or not, we accept or reject the event. In that sense, electron reconstruction efficiency is directly proportional to the event selection efficiency, and knowing this number correctly for Collider Data is important in the cross section measurement at the end of this chapter. However, once the electron is reconstructed and the event is selected, this efficiency doesn't affect the fitting.

Jet reconstruction efficiency

The fact that 4 jets are needed to fit an event is a rather stringent requirement, even when the efficiency for individual jet reconstruction is high. This is because we require all 4 jets from $t\bar{t}$ decay for the correct combination to exist. Since jet reconstruction efficiency is a function of jet transverse energy, the efficiency also depends on the top mass produced in $t\bar{t}$ events.

The purpose of this study is to estimate the fraction of $t\bar{t} \longrightarrow e + jets$ events that have all the jets from the $t\bar{t}$ reconstructed. ISAJET MC was used to generate $t\bar{t}$ events. These events were passed through DØ detector simulation (GEANT) and reconstructed (V11.17) using a couple of different jet reconstruction algorithms.

Each parton from Isajet (ISAQ) is compared to the reconstructed jets and matched with the closest one in ΔR. These partons were divided into two groups, Final-State Radiated (FSR) jets and jets from $t\bar{t}$ (usually high P_t quark jets). Here, I calculate jet finding efficiencies for FSR and quark jets from $t\bar{t}$ where 'jet finding' is defined as finding a reconstructed jet within ΔR of 0.2 from the parton (quark or gluon) direction. The data sample used in the following cases is an ensemble of 700 160 GeV $t\bar{t} \longrightarrow e + anything$ with the following requirements

1. Only one lepton (=electron) from W (ISAJET)
2. $\not\!\!E_t$ and electron E_t greater than 15 GeV (ISAJET)

3. Reconstructed event must have one electron, no muon and at least 4 jets with the implicit jet E_t cut of 8 GeV imposed at reconstruction

The jet reconstruction efficiencies for 0.3 and 0.5 cone algorithms are shown in Tables 6.4 and 6.5.

Table 6.4: Jet reconstruction efficiencies for 0.3 cone.

Type of Jets	Efficiency
FSR	185/432 = 0.428
Other jets from $t\bar{t}$	1026/1168 = 0.878

Table 6.5: Jet reconstruction efficiencies for 0.5 cone.

Type of Jets	Efficiency
FSR	141/427 = 0.330
Other jets from $t\bar{t}$	985/1144 = 0.861

Let me define the 'efficiency of event reconstruction' to be '*the number of* $t\bar{t} \longrightarrow e + jets$ *events with all the jets from* $t\bar{t}$ *reconstructed (except FSR) divided by the total number of* $t\bar{t} \longrightarrow e + jets$ *events.* Then, in the case of 0.3 cone, it would be $0.878^4 = 0.59$. The key point here is that this event reconstruction efficiency increases very rapidly as the jet reconstruction efficiency increases.

Jet mass

Jet mass can arise from three sources. One is actual mass of the parton. Another is the transverse shower spread within the calorimeter. The third is due to gluon radiation. If we assume that the parton was massless, the scalar sum of the calorimeter cell energies corresponding to the jet should be considered as the momentum of the

jet. In this case, we are assuming that the jet is massless and the mass due to the shower spread is an artifact of shower development.

On the other hand, the $\not\!\!E_t$ is calculated as the vector sum of all the cells, instead of the vector sum of all the massless jets. This $\not\!\!E_t$ calculation, therefore, assumes that jets balance momentum when they have mass.

One needs a consistent set of assumptions. Assuming that jets are massless and summing up all cell energies of a jet at its center is inconsistent with $\not\!\!E_t$ calculated as a vector sum of cells. But, if I assume that the jet masses are real we then have consistency in both calculation of $\not\!\!E_t$ and jet momentum since both of them are calculated as vector sums over cells. In addition it is consistent with real physical processes such as gluon radiation which gives mass to a jet.

Now, the question is how much does the jet mass affect our constraints. As described earlier I have two different types of constraints. The first involves momentum balance constraints and the other type involves the mass constraints. Momentum balance constraints are not affected by mass as long as the momentum scales are correctly calibrated. It turns out that the mass constraints are not affected much either. We can think of a simple but general example of W decaying into two jets in its center of mass frame. The mass of the W is

$$M_W = (E+E)^2 - (P-P)^2 = 4E^2 = 4(P^2 + M_j^2) \tag{6.33}$$

where M_j is the mass of each jet.

When the mass is zero, I get $M_W = 2P$. When the mass is $\frac{1}{5}P$, $M_W = \sqrt{4(P^2 + \frac{1}{25}P^2)} = 2.04P$. So, a jet mass at the level of 20% of the momentum of a jet in the CM frame affects the mass of M_W by only 2%!

ISR/FSR (Initial-State Radiation/Final-State Radiation)

ISR and FSR introduce a couple of significant difficulties in mass fitting. When FSR occurs and it's not properly combined with the parent parton, the mass resolution will degrade since the mass constraints are not valid any more. At the same time, ISR and FSR increase the number of jets in an event, introducing a huge combinatorial background.

Mass resolution

To study the effect of FSR on fitted mass resolution, Isajet MC events were generated with FSR turned on. Then I picked up 4 jets from $t\bar{t}$ which are not gluons (therefore not FSRs) and used them in the fitting. When there is FSR in an event, some of the energy from $t\bar{t}$ system will be missed, and the fitted mass will be lower than the value with no FSR. Figure 6.5 shows this effect at jet energy resolution of $100\%/\sqrt{E}$. This is to be compared with the results in Table 6.1 where there is no FSR. Both the efficiency of selecting the correct combination and the mass resolution of the correct combination degrade quite significantly. (resolution: 7.5 GeV $\longrightarrow$ 15.3 GeV, efficiency: $\approx 0.5 \longrightarrow 0.32$)

Combinatorial looping method

When there are only 4 jets in an event, and without any b tagging, there are 12 ways of combining these four jets to make up a $t\bar{t}$ decay hypothesis. When there are more than 4 jets in an event, the number of combinations increases very fast as the number of jets increases as shown in Table 6.6.

As shown previously, as the number of combinations increases the probability of selecting the correct combination is overwhelmed by the large number of wrong combinations for a given detector resolution. Practically, fitting doesn't do very much

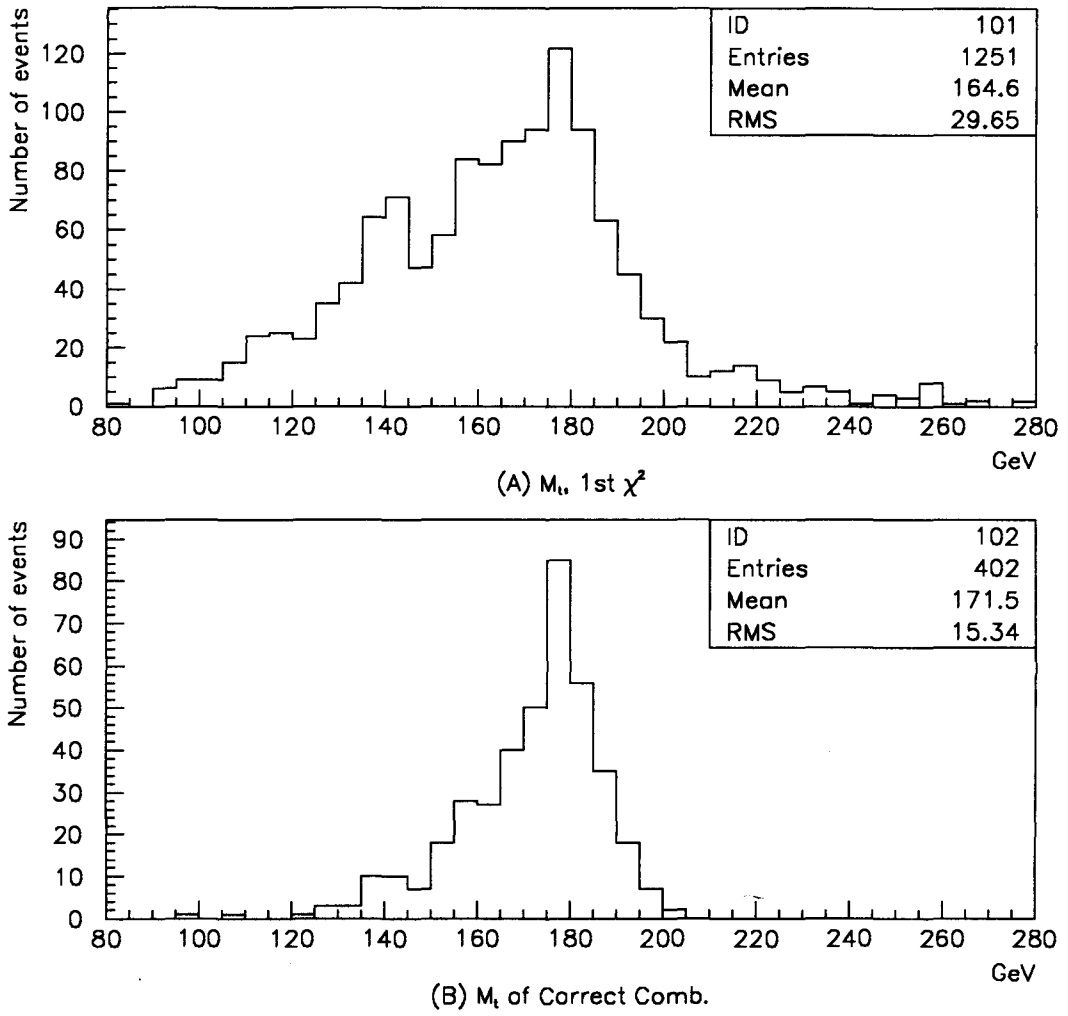

Figure 6.5: The effect of FSR. The fitted mass distribution from the smallest χ^2 combinations (A), and from the smallest χ^2 and also correct combinations (B). Jet energy resolution = $100\%/E^{1/2}$, generated top mass = 180 GeV.

Table 6.6: Jet multiplicity vs. number of combinations

Number of Jets	Number of Possible Combinations
4	4P4/2 = 12
5	5P4/2 = 60
6	6P4/2 = 180
7	7P4/2 = 420

in selecting the correct combination if the number of combinations is above 40 (10% efficiency).

Due to the ISR/FSR, there are extra jets in the $t\bar{t}$ events and it's not practical to loop over all the combinations when there are 5 or more jets. Fortunately, these ISR/FSR jets have some characteristics that help us distinguish them from jets from $t\bar{t}$ system such as low P_t, η, etc. Therefore, one way of handling this problem is to assign a probability to each jet for being a FSR/ISR, weight each combination accordingly, and use information from all the combinations.

An alternative is to come up with a way of looping over jets such that we wouldn't miss many correct combinations while keeping the size of combinatorial background manageable. When we order the jets within an event in transverse energy (E_t), the lowest E_t jet will have the highest probability of being an ISR or FSR, whereas the highest E_t jet will have a very small probability of being ISR/FSR. In this case it wouldn't make very much sense to loop over combinations that assume the highest E_t jet is ISR/FSR. Tables 6.7 and 6.8 show the results from two different looping methods: one looping over all the possible combinations, and the other one looping over the first 4 highest E_t jets assuming the rest are ISRs since they are low E_t jets. The sample used:

Number of events: 700 $t\bar{t} \longrightarrow$ e + anything, fully Geanted and recon-

structed.

Number of events fitted: 284 (4 or more jets, 1 electron)

Matching cut between Isajet in ΔR: 0.2

Jet reconstruction algorithm: 0.3 cone

Table 6.7: Looping over all combinations.

# of Jets	4	5	6	≥ 7
# of Events	101	99	45	45
# of Events with Correct Comb	53	52	25	26
# of Correct Comb. with Best χ^2	21	6	0	0
Efficiency	.40	.11	0	0

Table 6.8: Looping over the first 4 jets.

# of Jets	4	5	6	≥ 7
# of Events	101	99	43	45
# of Events with Correct Comb	53	52	24	26
# of Correct Comb. with Best χ^2	21	11	0	2
Efficiency	.40	.21	0	.08

One might think that looping over the first 5 jets will contain almost all the 4 jets from $t\bar{t}$ if there is a correct combination in that event. If this is the case, instead of looping over all possible 60 combinations one can think of looping over the first 4 jets (12 combinations), and replace the 4th jet with the 5th jet and loop again (12 combinations), and replace the 3rd jet with the 5th jet and loop again (12 combinations). This way, I am excluding the possibility that one of the first two

highest E_t jets is ISR or FSR, which is very unlikely. To see how well this looping method would perform, the same Geanted MC was used. There are 189 events out of 290 with more than 4 jets and 103 of which have correct combinations in them. 89 events out of the 103 event have all the 4 jets from $t\bar{t}$ decay within the first 5 jets. Table 6.9 shows how the 4 quark jets from $t\bar{t}$ distribute themselves among the first five jets in the 89 events.

Table 6.9: Configuration of the first 5 jets with correct combination within the first 5 jets.

Configuration	Number of Events
5th jet is ISR or FSR	38
4th jet is ISR or FSR	21
3rd jet is ISR or FSR	18
2nd jet is ISR or FSR	8
1st jet is ISR or FSR	4

This result indicates that the idea of considering the first 5 jets and looping over two or three groups of 4 jets is reasonable and consistent with the assumption that the last two or three jets can easily be ISR or FSR. Looping over the first four jets 1,2,3,4, then over jets 1,2,3,5 and 1,2,4,5, gives a total of $12 \times 3 = 36$ combinations. Thus, we keep 86% of the correct combinations out of 89 events that have correct combinations in them for 5 or more jets events. However, the inefficiency in selecting the correct combination due to other effects, such as the increase in the number of combination, should be taken into account at the same time.

Now, let's consider the overall efficiencies. There are two main inefficiencies in fitting the correct combination, as described in earlier sections. One is the inefficiency of selecting the correct combination. The other is the inefficiency of having the correct combination to begin with. The sources of the inefficiencies are due to

1. Jet energy resolution,
2. Combinatorial background,
3. Four jet reconstruction efficiency.

where 1. & 2. are related in the sense that one contributes to the other.

The overall efficiency of fitting would be written in the following form.

$$\epsilon = \epsilon_{reco} * \epsilon_{loop} * \epsilon_{\chi 2}(\sigma, n_{comb}) \tag{6.34}$$

where

ϵ: Given a 'e + 4 or more jets' event, the efficiency of selecting the correct combination using χ^2

ϵ_{reco}: The efficiency of reconstructing all 4 jets from $t\bar{t}$.

ϵ_{loop}: efficiency of keeping correct combination using a particular looping method given that the correct combination exists

$\epsilon_{\chi 2}$: The efficiency of picking up the correct combination out of n_{comb} number of combinations at a give energy resolution, σ, given that the correct combination exists.

If $\epsilon_{jet\ reco}$ is the efficiency to reconstruct one jet, then,

$$\epsilon_{reco} = (\epsilon_{jet\ reco})^4 \tag{6.35}$$

For example, for a jet reconstruction efficiency of 95%, from the information in Table 6.10 and 6.11,

$$\epsilon = 0.95^4 * 1. * 0.50 = 0.41 \tag{6.36}$$

Table 6.10: Efficiency of including the correct combination within the loop.

ϵ_{loop}	Cases
100 %	for 4 jets
77/103 = 74 %	for 5 or more jets (looping over $j_1j_2j_3j_4$, $j_1j_2j_3j_5$, $j_1j_2j_4j_5$)
38/103 = 37 %	for 5 or more jets (looping over first 4 jets)

Table 6.11: Efficiency of picking up the correct combination within N loops provided that there is correct combination within the N loops.

ϵ_{χ^2}	Cases
50 %	at n_{comb} = 12, resol=80%, 15%
23 %	at n_{comb} = 25, resol=80%, 15%
11 %	at n_{comb} = 35, resol=80%, 15%

for events with exactly 4 jets.

$$\epsilon = 0.95^4 * 0.74 * 0.11 = 0.066 \tag{6.37}$$

for 5 or more jets, and looping over $j_1j_2j_3j_4$, $j_1j_2j_3j_5$, $j_1j_2j_4j_5$.

$$\epsilon = 0.95^4 * 0.37 * 0.50 = 0.151 \tag{6.38}$$

for 5 or more jets, and looping over the first 4 jets. So, this clever idea doesn't do better than a simple loop over the first 4 jets.

The two most sensitive variables for the overall event efficiency are the jet reconstruction efficiency and the number of combinations. An increase in the jet reconstruction efficiency results in an increase in ϵ_{reco} as the 4th power. Reducing the number of combinations from 12 to 6 reduces the background by a factor of 2. By tagging b jets, one gets this down to 6 or even 3.

From the MC study described above, looping over the first 4 jets seems to be the most effective method without too much sophistication.

Jet energy scale correction (MC and data)

One of the most important quantities to calibrate is the jet energy scale, and the goal is to get the parton energy back from a reconstructed jet energy. Knowing this correctly at all energies will result in eliminating systematic biases in the mass distribution and improve the capability of selecting correct combinations as well as assigning correct masses to the correct combinations. The important aspect of this jet correction is that the corrections for MC and data have to be done in a coherent way so that we can model our data with MC.

DØ has developed it's own standard jet correction (CAFIX) for the study of QCD. This correction is intended to perform the following functions:

1. Given a parton energy, provide the correction to get the reconstructed energy. (It's not necessarily the other way around due the the 'resolution bias' [4])
2. Correct the calorimeter cell energy scale.
3. Correct for the out-of-cone energy due to showering in the calorimeter.

However, it's worthwhile mentioning that it does not correct for the out-of-cone energy due to the radiation. Again, the goals I want to achieve are

1. to get the parton energy back.
2. to establish an equivalent jet energy scale for both MC and data.

To achieve these goals, several assumptions are made which will be proved to be correct later, and other assumptions are made based on our best understanding of the physical process. The assumptions are

Assumption # 1. CAFIX corrects the calorimeter energy scale correctly. (This is shown by looking at the energy pull distributions from $Z \longrightarrow ee + 1jet$ data sample (see appendix) for both MC and data. Pulls on jet energy come out relatively unbiased as shown in Figure 6.6 in comparison with the systematic bias of the jet energies from the parton energies (P_t of Z in this case). This is because the calorimeter energies outside the jet cone are included to balance momentum.)

Assumption # 2. The jet energy bias after CAFIX is due to the out-of-cone radiation. (This is shown in Figure 6.7 by looking at the azimuthal angular (φ) correlation between the jet direction and the direction of the underlying event vector. The directions are the same indicating that the jet didn't include all the energies outside the cone. Also 0.3 cone seems to require significantly larger correction than 0.5 cone indicating the same effect of radiation loss)

Assumption # 3. The out-of-cone corrections are comparable for MC and data. (Figure 6.9 shows the comparison of data and the MC before and after radiative out-of-cone correction using $Z \longrightarrow ee + 1jet$ sample. One can see that the level of this correction for MC and data is comparable)

Assumption # 4. Loss of energy due to semileptonic decay of b quark jet in MC is a reasonable representation of the semileptonic b decay in data.

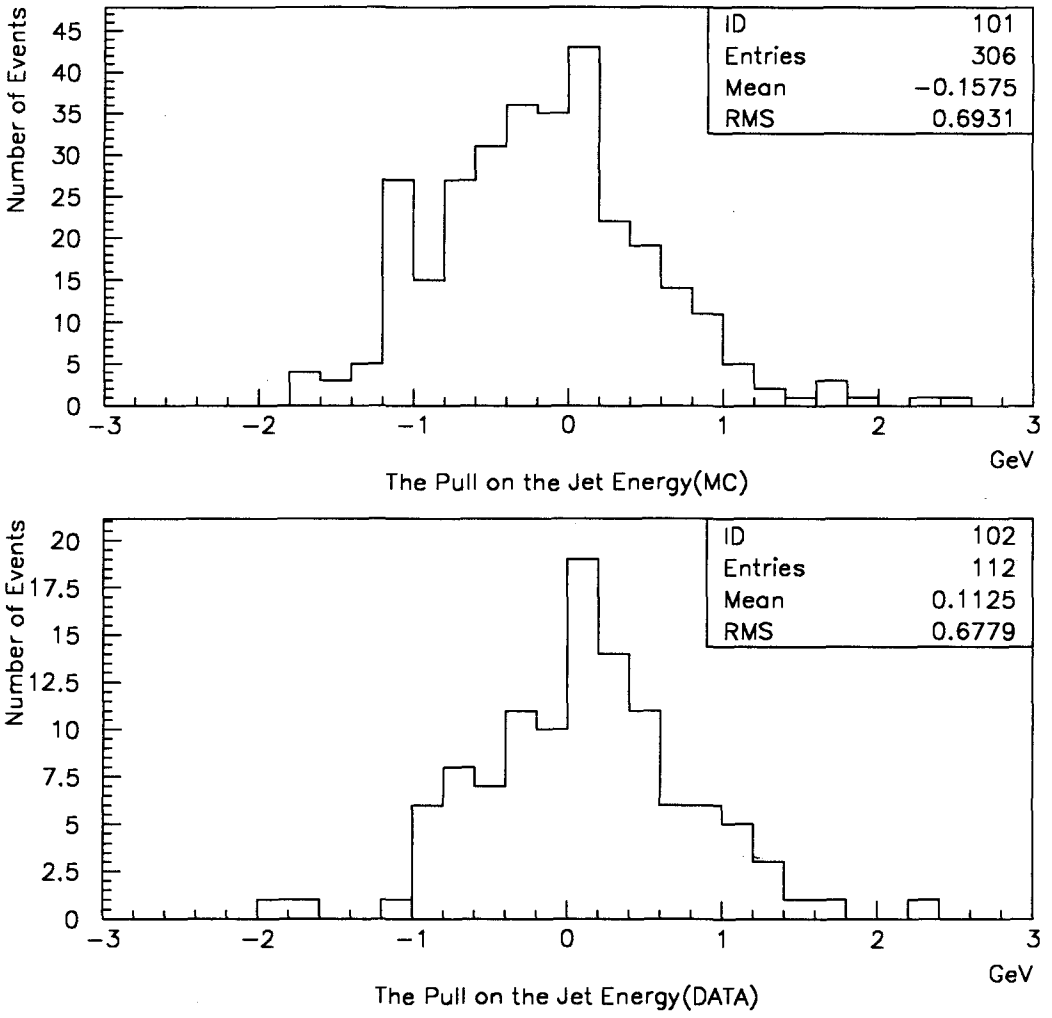

Figure 6.6: The pull quantities on jet energy for MC and DATA in $Z \longrightarrow ee + 1jet$. (After CAFIX only)

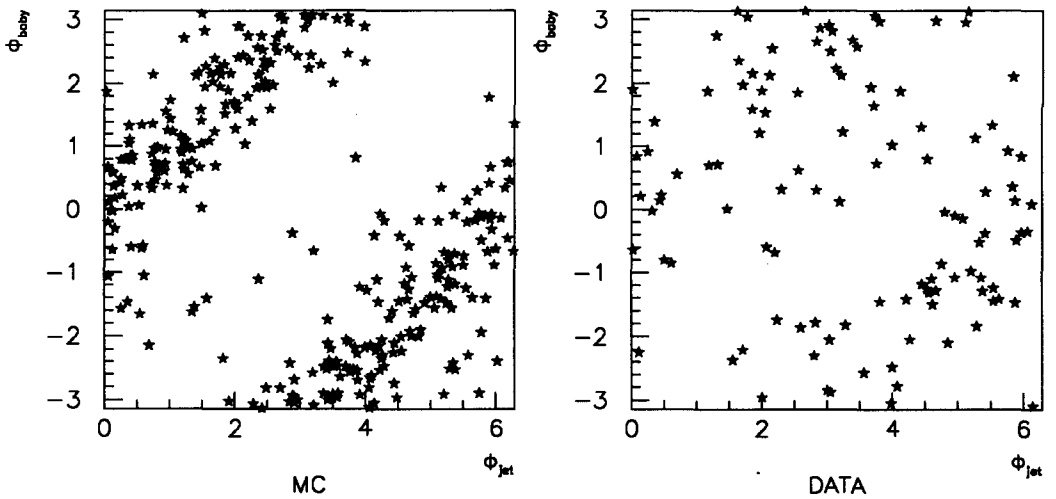

Figure 6.7: The φ of the underlying event vs. the φ of the jet.

Figure 6.8 shows the correlation between the reconstructed energy (0.3 cone) and the parton energy from MC. From this, we can extract the out-of-cone correction. (which can be used for both MC and data). Once we establish our confidence that the above assumptions are right, we can coherently apply our jet corrections as follows.

1. Apply CAFIX to both MC and data to correct the calorimeter energy scales.
2. Get the out-of-cone correction due to radiation on top of CAFIX to get the parton energy using MC.
3. Apply this out-of-cone correction to both MC and data.

(Do not correct $\not\!\!E_t$, as this correction is made since we are effectively adding the out-of-cone energy into the jet where we are changing the *baby jet* but not the $\not\!\!E_t$)

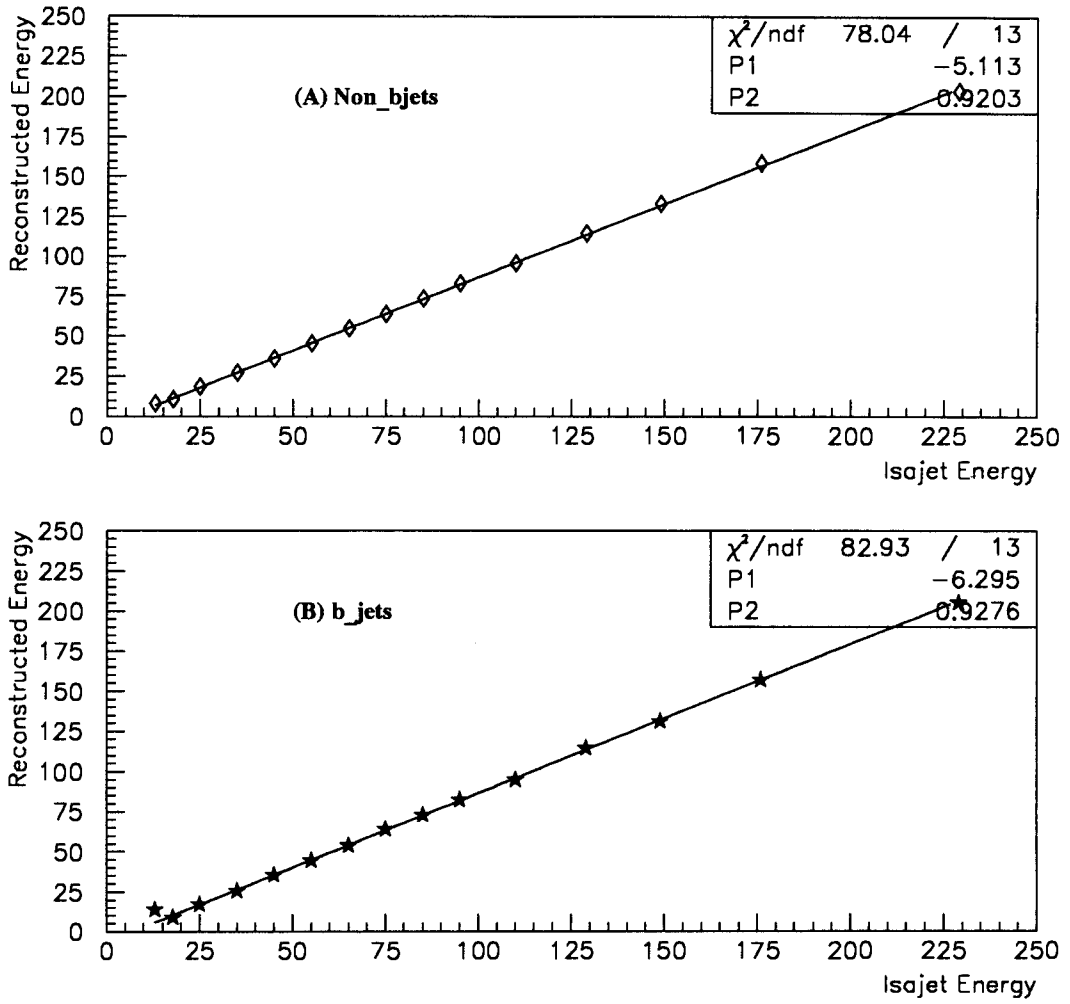

Figure 6.8: Reconstructed energy vs. the parton energy for non-b-jets (A) and b-jets (B)

For b-jets, we can add additional corrections based on the MC comparison between non-$b$jets and b jets. If this additional correction is due to $\not\!\!E_t$ introduced by neutrino, I correct the $\not\!\!E_t$ as I make this correction. One remark to be made here is that the $Z \longrightarrow ee + 1jet$ event sample is a test sample but not a calibration sample due to it's low statistics. Therefore, what is demonstrated here is that the overall calorimeter energy scale and the out-of-cone radiations are roughly the same for both MC and data. More detailed study at high statistics will have to be done using event samples such as direct photon data where we have more events.

Signal response vs. background response (MC)

The main objective here is to get the expected mass distributions from $t\bar{t}$ events (at any possible top mass) and also from background events so that when we get a mass distribution from collider data we can interpret our results based on what we expect from the MC.

After all the calibrations are done, a preselection of events was made, and I fit only those events that pass the cuts I will apply to the collider data in the following chapter. The mass distributions from top 140, 160, 180 GeV, and also from $W + jets$ background are shown in Figure 6.10. The mass distributions from $t\bar{t}$ events are fitted to a double gaussian and the $W + 4jets$ background mass distribution is fitted to the following functional form

$$Background\ mass\ probability\ density = p_1 e^{p_2 x + p_3 e^{p_4 (x - p_5)}} \qquad (6.39)$$

which is just an exponential that drops rapidly below a certain mass. It is important

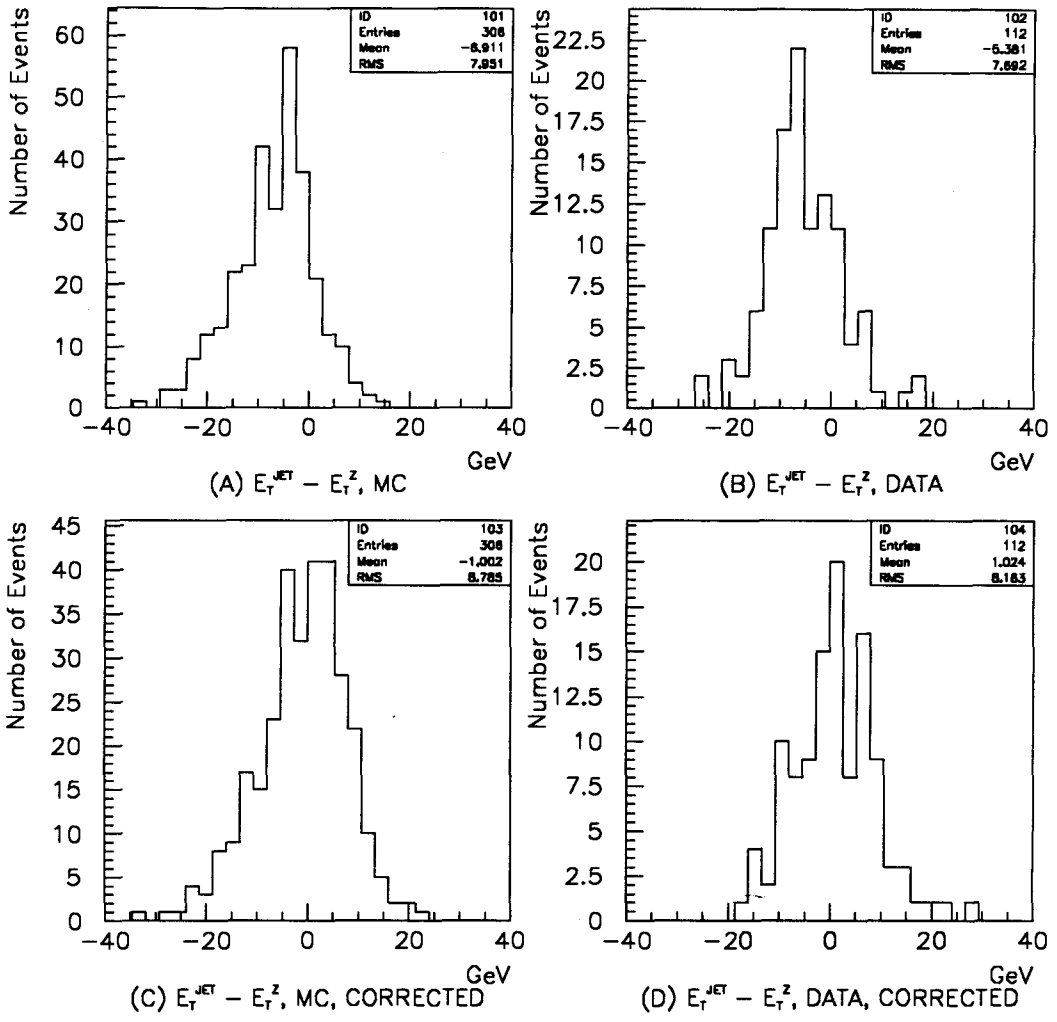

Figure 6.9: $E_t^{Jet} - E_t^Z$ shows how much energy we lose outside the jet cone. Plots are before radiative out-of-cone correction for MC (A) and data (B), and after the correction for MC (C) and data (D). Data and MC show good agreement.

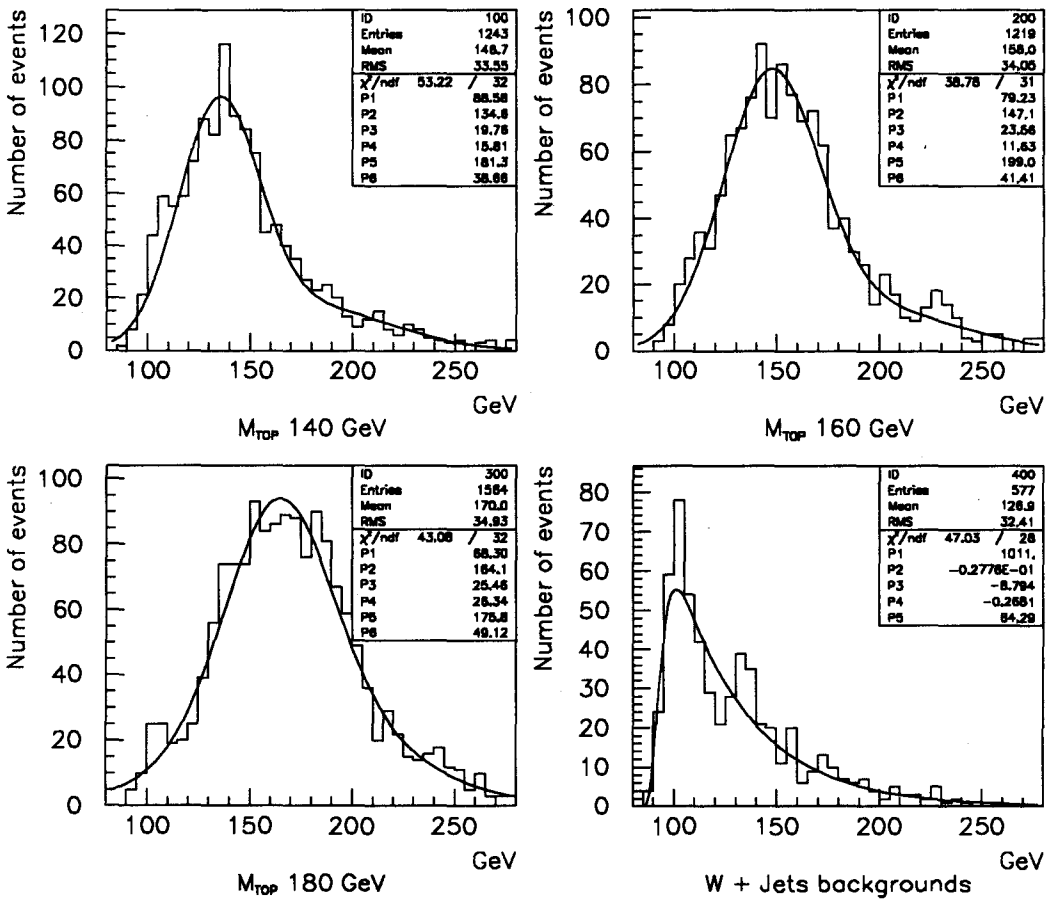

Figure 6.10: Fitted mass distribution of $t\bar{t}$ events (140 GeV, 160 GeV, and 180 GeV) and W + 4 or more jets events.

to note consistency in data selection criteria between data and MC to generate these modeled mass probability density functions.

One of the useful parameters in conventional analysis to discriminate background events from the signal events ($t\bar{t}$) is the variable H_t which is defined as follows.

$$H_t = \sum_{all\ jets} E_t^{jet} \tag{6.40}$$

This variable has a strong correlation with the fitted mass of an event. If our final events are selected after this cut, the mass distributions to be used in the likelihood analysis also have to be generated with this cut. Mass distributions after the cut $H_t > 140GeV$ is shown in Figure 6.11. As one might notice, the background mass distribution also has rather broad peak at high mass (around 150 GeV) after the H_t cut. But, one should also remember that the background level goes down significantly when this cut is made so that, overall, this cut might benefit us. However, this cut is not made in data selection because of its correlation with the fitted mass distribution. If making the H_t cut benefits us in data selection, it can also benefit us later in mass distribution because of this correlation even if I don't make H_t cut in data selection.

After these mass resolution functions are determined at three top masses, this function at any top mass is estimated by interpolating/extrapolating the parameters of these resolution functions so that I can get the likelihood as a continuous function of top mass. To do so, I interpolate/extropolate the following parameters.

1. Mean of the first Gaussian.
2. Sigma of the first Gaussian.
3. Mean of the second Gaussian.
4. Sigma of the second Gaussian.

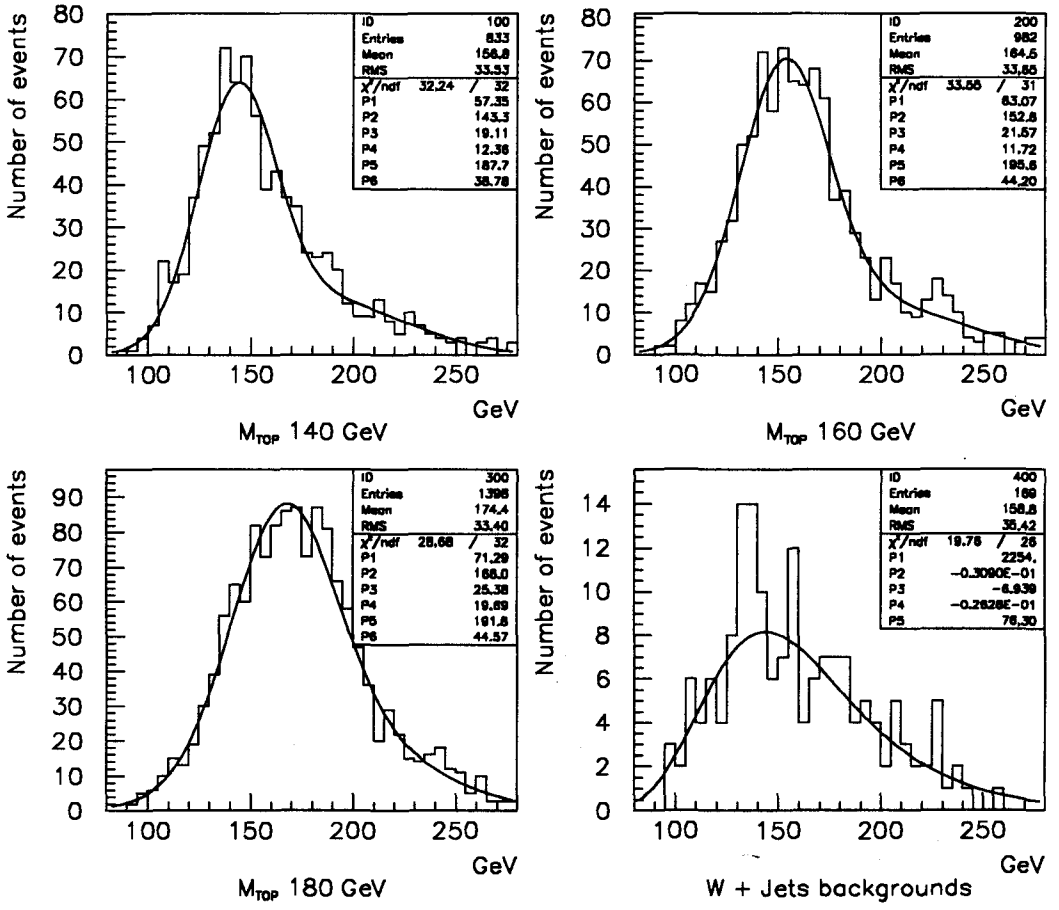

Figure 6.11: Fitted mass distribution of $t\bar{t}$ events (140 GeV, 160 GeV, and 180 GeV) and W + *4 or more jets* events after $H_t > 140GeV$ cut.

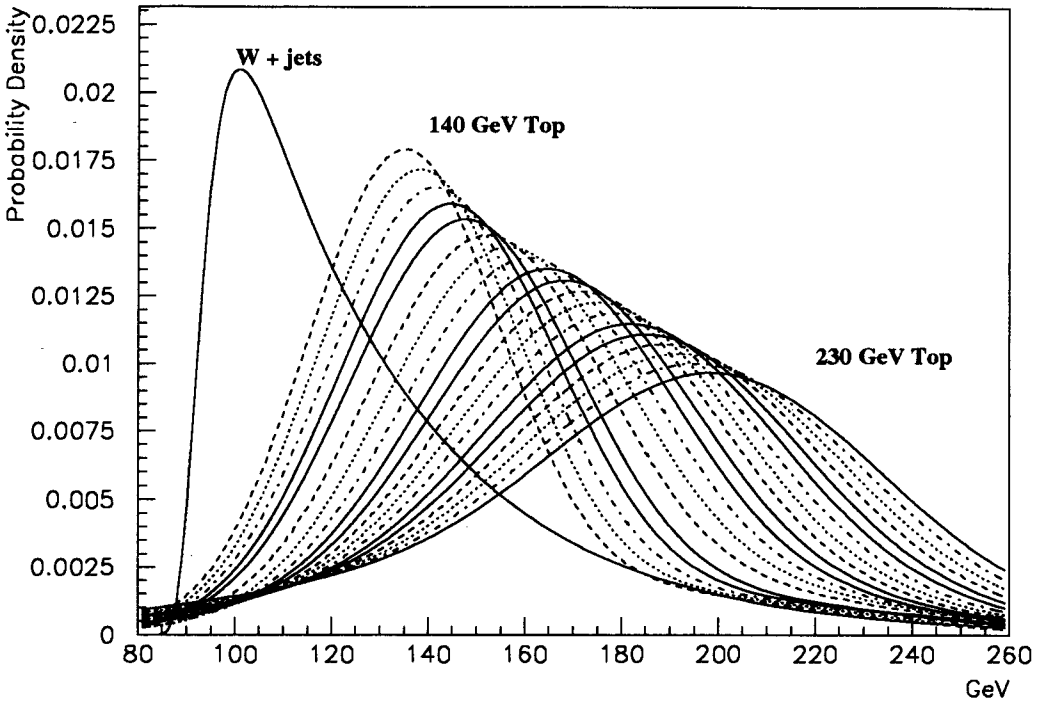

Figure 6.12: Interpolated and extrapolated mass probability density functions from 140 GeV top to 230 GeV top, and $W + jets$ background.

5. Ratio of the heights of the first and the second gaussians.

After interpolation, the normalized mass probability density functions are shown in Figure 6.12. These parameterized mass resolution functions will be used in the next chapter to determine top masses, to generate MC samples to test the method, and to estimate the errors on various fitted quantities.

Effect of b tagging on the result of fitting

In mass fitting, the only advantage of b tagging is that the combinatorial background is reduced by identifying a parton. One b tag reduces the number of combinations by factor of two in $lepton + 4jet$ events.

Since events with b tag contain more information about the parton ID, it can be advantageous if these events are treated separately using this extra information. But, first I have to make sure that this information really helps and if it does, how much.

Quantifying the improvements from b tagging can be measured in terms of the efficiency of selecting the correct combination. A summary of the results is shown in Table 6.12. But the ultimate test of the improvement is to see if the mass resolution function gets sharper. The improvement of mass resolution function from b tagging

Table 6.12: Comparison of efficiencies for events with a single b tag and without b tag.

Cases	Total # of events fitted	# of events with correct combination	Correct Combination selected
1 b tagged	1299	629	160
no b tagged	1375	658	116

is shown in Figure 6.13. The improvement is small, however, other effects such as resolution and FSR/ISR are dominant.

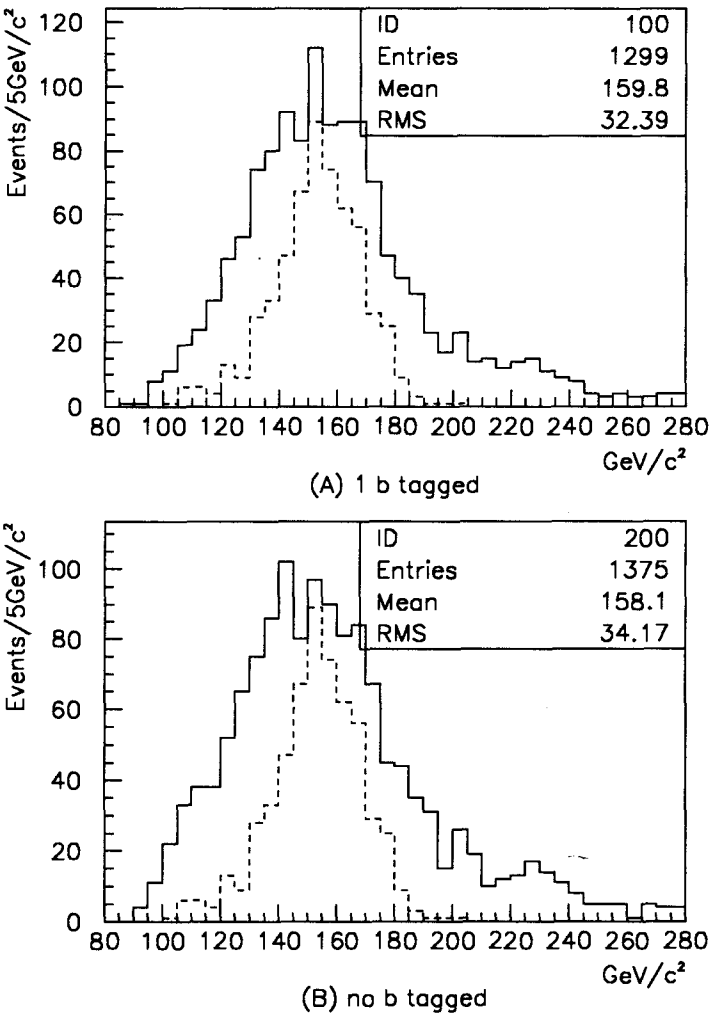

Figure 6.13: Fitted mass distribution with 1 b tagged (A), and without any b tagged (B). Dashed lines are for the correct combinations.

CHAPTER 7. FITTING THE $e + jets$ CHANNEL (COLLIDER DATA)

Two different modes of applications

The two-constraint (2C) kinematic fitting method described in chapter 6 can be used in two ways with slightly different philosophies. One way of using it would be to apply this fitting to a group of events selected with relatively loose cuts and search for a mass bump on a relatively smooth background. This might be convincing for the discovery of the top quark. But it requires high statistics. Even if one doesn't see a bump, one can calculate the expected number of signal events at a certain mass from the mass plot and the errors on the number of signal events at that mass. From this information at each mass point one can set a limit on the cross section for $t\bar{t}$ events as a function of top mass.

Another way of using this method would be to apply the fitting to a signal-enriched sample of events and estimate the mass from this small sample. This method for mass determination requires that a good portion of the event sample be signal.

I will be taking both approaches. First, I will apply this method to a group of events I selected with loose cuts to include as many signal events as possible. And later, I will apply this method to candidate events provided from the conventional analysis of DØ to extract the top mass and an estimated error.

A view of data selection efficiency

The data are accepted and written to tapes only when they pass three layers of triggers I described in chapter 4. Therefore, the trigger efficiency affects the final data selection efficiency at the very beginning. On top of the trigger efficiency is the offline selection cut efficiencies. However, these two efficiencies can not be simply multiplied to get the total efficiency after requiring a certain trigger and making a set of offline cuts. The reason is because we are not making selection cuts on the same variable in trigger level as we do in offline. The strategy for calculating the total efficiency, therefore, would be to make a set of offline cuts and require a certain set of trigger conditions at the same time, and at the end see how many events survive. This efficiency should be calculated from MC since we don't know which events are signal events and which events are not in real data. In case there is a systematic difference between MC efficiency and the collider data efficiency, one has to correct for this difference.

Each trigger condition is optimized to serve the best performance in selecting as many signal events as possible and reject as many background events as possible for a certain type of physics event. In the selection of the final sample of events, it's my freedom to choose which trigger conditions to require. On the other hand, it's my responsibility to optimize the signal-to-background ratio after the trigger requirements. It would be pointless to include a trigger which has very low efficiency for signal events and introduces large background events.

Data clean up

There are several instrumental backgrounds which are unique to the DØ detector. One of them is a group of events with calorimeter hot cells [1]. Events with hot cells have some unique features, such as no energy in neighboring cells, which can used to discriminate this type of event from others. DØ has developed a standard procedure called CLEAN_CAL_JUNK to get rid of these events at a high efficiency.

The other type is the class of events with proton showers in the calorimeter around the Main Ring. This happens because the Main Ring is running to accumulate $\bar{p}$ while the Tevatron is running. A lot of proton losses occur when they inject protons into the Main Ring and also when they ramp the energy of the Main Ring beam from 8 GeV to 120 GeV before hitting the target to produce $\bar{p}$. All the triggers veto this time interval of injection/transition (MRBS-LOSS) in level 1 and this contributes dead time. When Main Ring protons are in the vicinity of DØ, certain triggers veto this time interval (MICRO-BLANK) in level 1.

Some of the top triggers didn't apply this MICRO-BLANK veto and the elimination of events with Main Ring junk can be done offline, and also the corresponding correction has to be made in the integrated luminosity. The dead time due to this veto is approximately 0.08 [5]. Some runs didn't have any Main Ring activities. Therefore, this correction has to be made run-by-run.

In the following analysis, three types of potentially contaminated events are removed from the event sample. They are

1. Events taken during the MICRO-BLANK period.

[1] Repeatedly firing calorimeter cells, which are not due to energy loss of particles but due to HV short or leakage current, etc.

2. Events with hot cells.

3. Events in runs flagged as bad runs or special runs.

Integrated luminosity

Integrated luminosity depends on several things, such as

1. What triggers I select,

2. Are these triggers prescaled? If they are, what's the prescale ratio?

3. Did level 1 veto Main Ring activity (MICRO-BLANK)? If it did, what fraction is this to the total?

The second and the third questions should be asked for each run since the prescale ratio and the Main Ring condition could be different for different runs. In case I select the prescaled level 1 trigger and the prescale ratio varies within a store [2], it would be even more complicated to calculate the integrated luminosity for that integrated store. Fortunately, most of the triggers used for top search were not prescaled. Therefore, all we need to know is the integrated luminosity corresponding to the trigger bits selected for each run (and fractional loss due to MICRO-BLANK if the trigger didn't require the main ring veto)

Selecting non-prescaled triggers, the integrated luminosity after eliminating MICRO-BLANKed events for the run 1a (1992-1993) is [7]

$$\int Ldt = 13.5 \pm 1.6 \ pb^{-1}. \tag{7.1}$$

[2]The duration of the Tevatron beam

Sources of physics backgrounds and instrumental backgrounds

The main physics background to the $t\bar{t}$ signal events considered here is the production of single W events with extra jets, with the W decaying into $e + \nu$.

Another potentially important background comes from QCD events. This type of background is either from semileptonic heavy quark decay or a fluctuation of a jet into a false electron accompanied by $\not\!\!E_t$ from a fluctuation in jet energy. These events will have an electron and significant $\not\!\!E_t$ faking our e + jets signal.

Theoretical estimation of the $W + jets$ background events has a large uncertainty (40%) and therefore it's difficult to use this theoretical estimation of background to subtract from the experimentally observed number of events in any calculation of cross section limit. However, this uncertainty is eliminated if we use the data to estimate the background. What I want is the most accurate estimation of the number of background events at a jet multiplicity of 4 or higher. This can be done using a QCD scaling law [6]. The basic assumption is the QCD rule that the jet multiplicity distribution drops exponentially for $W + jets$ events. Therefore, with the number of events at jet multiplicities of 1, 2 and 3, we can predict the number of events at jet multiplicity of 4 or above. A complication arises if my background consists of two different processes, namely $W + jets$ and QCD because the scaling law doesn't hold for the QCD background events.That's because the probability that one of the jets in an event faking a high P_t electron from a W is a linear function of jet multiplicity. So, I either have to find out what the contamination from QCD at each jet multiplicity is, subtract them off from the data, estimate the number of background events only from $W + jets$, and apply scaling law to these $W + jets$ background events to estimate $W + jets$ background events at higher multiplicity (4 or more), or I have to get rid

of these QCD background events completely by making a tight cut (on $\not\!\!E_t$ in this case). A study shows that these QCD background events are negligible (less than 5% of the total) if I make a tight electron cut as well as a high $\not\!\!E_t$ cut at 30 GeV [8] [9].

Data streaming, electron definition, triggers

All the data taken from run 1A were filtered through what's called the RGE stream. The loose requirements for $e + jets$ channel in this stream are the following conditions

Electron: 1 (PELC/PPHO) $P_t > 12.$ GeV

$\not\!\!E_t$: $\not\!\!E_t > 12.$ GeV

The data sample that passes the above cut is the original sample. Clean up of potentially contaminated data is done afterward as described in the previous section.

Sometimes, data selection requires clear definitions of physical objects, which consist of a set of cuts on various quantities identifying the object. In the $e + jets$ channel, we have three distinct partons to identify; electrons, jets, and neutrinos ($\not\!\!E_t$). The tightness of the definition affects the signal efficiency as well as the background rejection. To require as many electrons from W decay as possible and to reject as many QCD background events as possible, optimization was done to decide on a tight definition of electron. This is accepted as DØ standard "tight election" definition as shown below.

1. em fraction ≥ 0.9
2. Isolation ≤ 0.1
3. Track matching significance ≤ 5.
4. $\chi^2_H \leq 100$.
5. Reject if $1.5 < dE/dx < 3.0$ in CDC

   Reject if $1.3 < dE/dx < 2.5$ in FDC

Since electrons from W decay tend to have higher P_t than QCD background, requiring a high P_t electron reduces electrons from QCD background significantly. QCD jets are mostly at high η where the cross section is bigger. The cut commonly used is $E_t^e \geq 20$ GeV.

The next step is to make trigger requirements. Even though it is important to make our trigger for $e + jets$ channel as efficient as possible, the reason why I make trigger requirement here is not only to make the signal-to-background ratio high but also to select the data corresponding to the calculated integrated luminosity. I require the triggers 'ELE_MAX' or 'ELE_JET' or 'ELE_JET_MAX' (or 'ELE_HIGH). None of these triggers are prescaled, and only a negligible fraction of ELE_HIGH was prescaled. The conditions of these triggers are shown below.

ELE_MAX: L2EM(1,20,EIS) & L2MS(20,0) meaning at least one electromagnetic cluster of E_t above 20 GeV with shape quality cut and isolation cut. It also requires level 2 $\not\!\!E_t$ greater than 20 GeV.

ELE_JET: L2EM(1,12,ELE) & JT(2,10,.3) & L2MS(10,0) meaning at least one electromagnetic cluster of E_t above 12 GeV with shape quality cut, at least 2 jets above E_t greater than 10 GeV with 0.3 cone algorithm. It also requires Level 2 $\not\!\!E_t$ greater

than 10 GeV.

ELE_JET_MAX: L2EM(1,12,ELE) & L2JT(2,16,.3) & L2MS(20,0) meaning at least one electromagnetic cluster of E_t above 12 GeV with shape quality cut, at least 2 jets above E_t greater than 16 GeV with 0.3 cone algorithm. It also requires Level 2 $\not\!\!E_t$ greater than 20 GeV.

Data selection I

Now, I have removed contaminated events, required triggers and also tight electrons with high P_t (>20 GeV) to ensure that the electron is from W decay. But, some ($\approx$14%) QCD background remains [11]. What I want to achieve in this section are

1. Include as many signal events which are fittable as possible.
2. Exclude almost all the QCD background so that I can use the scaling law (since I have only $W+jets$ background) and also I can introduce single type of kinematically well-known background in my likelihood analysis. (Neither MC nor real data (due to low statistics) can provide a reliable model for the mass resolution function of the QCD background)

This approach follows the first mode of application described at the beginning of this chapter, allowing as many signal events as possible and look for a bump in the mass distribution. To achieve the second goal, the most powerful way of eliminating the QCD background is to cut on $\not\!\!E_t$, since the $\not\!\!E_t$ in the QCD events is introduced by

a fluctuation in jet measurements, and since there is no real high P_t neutrino, the $\not\!\!E_t$s of the QCD events are small compared to those of $t\bar{t}$ events or $W + jets$ events. Studies show that $\not\!\!E_t$ cut of 30 GeV in addition to the tight electron ID requirements makes the QCD background negligible [8], [9], less than 5% of the total background to $t\bar{t}$ events.

Including the $\not\!\!E_t$ cut, the overall requirements for data selection are

1. The electron satisfies the Standard Tight Electron criteria.
2. $\not\!\!E_t > 30$ GeV.

There is no jet E_t cut other than the reconstruction threshold at 8 GeV since we want to include as many fittable signal events as possible.

Table 7.1: Number of events vs. jet multiplicity

# of Jets	Standard e	Estimated from +1, 2, 3 jets
$W + \geq 1$ Jet	1462	1464.5
$W + \geq 2$ Jets	313	307.4
$W + \geq 3$ Jets	62	64.5
$W + \geq 4$ Jets	19	13.5
$W + \geq 5$ Jets	5	2.8
$W + \geq 6$ Jets	3	0.6

Finally, the events are grouped with the same inclusive number of jets where jets are found by 0.3 ΔR cone algorithm with 8 GeV E_t trigger threshold. The resulting number of events *versus* the inclusive jet multiplicity is shown in Table 7.1. Figure 7.1 also shows the estimated number of events at each multiplicity from the fit to the first three data points ($\geq$1 jet, $\geq$2 jets, and $\geq$3 jets data). As shown in Table 7.1, the extrapolation of $W+ \geq 1jet$, $W+ \geq 2jets$, and $W+ \geq 3jets$ gives an estimated

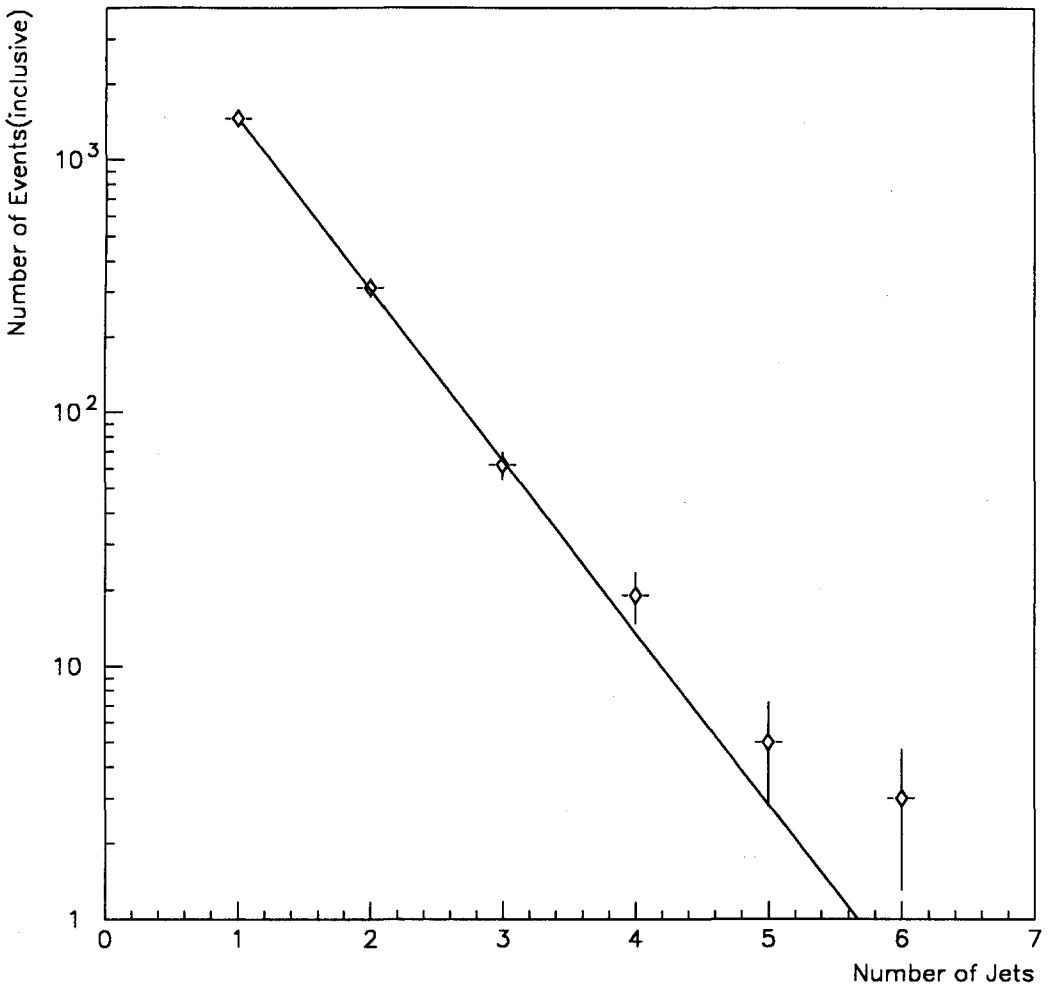

Figure 7.1: **Inclusive jet multiplicity distribution and a fit to the first three data points.**

background of

$$\approx 13.5 \ Events \tag{7.2}$$

for $W + 4$ *or more jets* events where I observe 19 events.

In the estimation of the number of $W + 4$ *or more jets* events, the assumption was that there are no $t\bar{t}$ events included in the three data sets 1 or more jets, 2 or more, and 3 or more jets. But if there are $t\bar{t}$ events in those data sets, the assumption will be wrong and the correct estimation of $W + jets$ background will have to be calculated after subtracting off the $t\bar{t}$ events. This can provide a systematic error on the estimation of the $W + jet$ background at higher multiplicity (4 or more). If I assume that the the top mass is around 160 GeV, the theoretical cross section is about $8.16pb$ and the branching ratio times the efficiency for the same cuts I make for the data (without the jet multiplicity cut) is about 0.0814. Therefore I expect the following number of $t\bar{t}$ events from $13.5pb^{-1}$ of data.

$$\sigma \times \epsilon \times Br. \int Ldt = 8.16pb \times 0.0814 \times 13.5pb^{-1} = 8.97 \ events \tag{7.3}$$

The expected multiplicity distribution for the 9 events is shown in Table 7.2. Table 7.2 also shows what the estimated number of $W + jets$ background with 4 or more jets would be if we subtract off the $t\bar{t}$ content from our data sample. As shown in Table 7.2, the expected $W + jets$ background with 4 or more jets is 11.8 events with $t\bar{t}$ subtraction. Therefore, within 100% error on the theoretical cross section of $t\bar{t}$, the systematic error on the estimated number of $W + jets$ background is

$$|13.5 - 11.8| = 1.7events \tag{7.4}$$

Table 7.2: Multiplicity of $t\bar{t}$ (160 GeV) events, $W + jets$ events after subtracting $t\bar{t}$ events, and estimated $W + jets$ events from a fit to the first three points.

# of Jets	$t\bar{t}$ events	Pure $W + jets$	Estimated $W + jets$ from fit
$W + \geq 1$ Jet	9.0	1462.-9.0=1453.0	1458.3
$W + \geq 2$ Jets	8.8	313.-8.8= 304.2	293.0
$W + \geq 3$ Jets	8.0	62.-8.0= 54.0	58.9
$W + \geq 4$ Jets	6.1	19.-6.1	11.8
$W + \geq 5$ Jets	3.4	5.-3.4	2.4
$W + \geq 6$ Jets	1.6	3.-1.6	0.5

Another important quantity I need is how efficiently the signal events pass my cuts. Later, this information will be used in the calculation of the $t\bar{t}$ cross section. For this study, I used unbiased (including all decay channels) Monte Carlo $t\bar{t}$ events put through detector simulation (shower library version [3]) for different top masses, 140 GeV, 160 GeV, 180 GeV. The efficiency (including the jet multiplicity requirement of 4 or more) times branching ratio is shown in Table 7.3.

Table 7.3: Efficiency times branching ratio of $t\bar{t}$ events with ≥ 4 jets (without jet E_t cut) and expected number of $t\bar{t}$ events from theoretical cross section.

Top mass	Eff. $\times$ Br.(%)	Expected # of events in 13.5 pb^{-1}
140 GeV	4.83	11.0
160 GeV	5.50	6.1
180 GeV	5.93	3.4

In summary, with all the cuts, I have

$$19.0 \quad \textit{candidate} \quad \textit{events} \tag{7.5}$$

with

$$13.5 \quad \textit{expected} \quad \textit{background} \quad \textit{events}. \tag{7.6}$$

[3] Fast MC smearing that replaces GEANT simulation

And the systematic error on the estimated background is 1.7 events.

Limit calculation

The main goal here is to set limits as a function of mass.

Now, we are provided with the following information.

1. Observed mass distribution from a set of selected events.
2. The estimated number of background events and the shape of the mass distribution from the MC background events.
3. The shape of the mass distribution from MC $t\bar{t}$ events at any given top mass.
4. The efficiency of the event selection cut at each given top mass. (We need this information when we calculate the cross section limit)
5. Integrated luminosity.

This information is a necessary and sufficient ingredient in interpreting the resulting mass distribution from the data. A powerful advantage of this mass analysis compared to the conventional analysis (which consist of data selection by cutting on parameters), is that the mass distribution is an additional handle in the interpretation of data. There can be various ways to utilize this information. A qualitative description of why a cross section limit as a function of top mass is interesting and can be more powerful when used with the mass information is discussed in the following subsection.

Advantage of σ limit as a function of top mass

To set a limit on cross section, one has to know the following information.

1. The number of events measured.
2. The number of expected background events.
3. The efficiency $\times$ branching ratio. (as a function of top mass)
4. The integrated luminosity.

The cross-section is determined by

$$\sigma = \frac{N_{observed} - N_{background}}{\epsilon \times Br. \int Ldt} \tag{7.7}$$

and the upper limit on the cross-section at 90 % confidence level, for example, is where the probability that the number of observed signal events would have been actually bigger than the number corresponding to this upper limit cross-section is only 10 %. This probability is governed by the Poisson nature of this statistical process. Since we have a mass distribution of these selected events, I can make the items 1 & 2 as a function of top mass. Here is one way of extracting the the number of observed events (or expected background events) at a particular top mass.

Let $B(x)$, $S(M_t, x)$ be the normalized mass distribution from background events and signal events respectively where S is not only a function of mass, x, but also of the assumed top mass M_t. When we have N background events and n signal events (at M_t), the total distribution will have the following shape

$$Total\ Distribution = NB(x) + nS(M_t, x) \tag{7.8}$$

If I make the correct assumption that the top mass is M_t, at that mass I can maximize the signal to background ratio by weighting each event with the probability

density function for signal, $W(M_t, x)$, which can simply be just $S(M_t, x)$, but to get the same number of observed signal events, the following condition has to be satisfied.

$$\int nS(M_t, x)W(M_t, x)dx = n \tag{7.9}$$

In other words, when there are n signal events, the integral should give n signal events as well. If we set $W(M_t, x) = G(M_t)S(M_t, x)$ for some function $G(M_t)$, then

$$\int nS(M_t, x)W(M_t, x)dx = G(M_t)\int n(S(M_t, x))^2 dx = n \tag{7.10}$$

$$G(M_t) = \frac{1}{\int (S(M_t, x))^2 dx} \tag{7.11}$$

$$W(M_t, x) = \frac{1}{\int (S(M_t, x'))^2 dx'} S(M_t, x) \tag{7.12}$$

Therefore, the number of observed events at this mass ($N_{inc}(M_t)$) is

$$N_{inc}(M_t) = \int W(M_t, x)(NB(x) + nS(M_t, x))dx \tag{7.13}$$

$$= N\int W(M_t, x)B(x)dx + n \tag{7.14}$$

where $N\int W(M_t, x)B(x)dx$ is the expected number of background events which is now much smaller than N. When we preserve the signal distribution within the mixed data sample, the number of integrated background events is suppressed outside this mass region resulting in a smaller number of estimated background events at this given mass.

This method will work if the statistics become infinite. But at low statistics, it doesn't take into account the statistical fluctuation of the number of the signal and the background events in conjunction with the shape of the background mass resolution function. But, it demonstrates how the mass information can be useful when added to the cross section calculation.

Likelihood method

An alternative way we can think of is to do some kind of fitting at each mass hypothesis and set a limit from the fitted error. One can think of a least squares fit. However, it wouldn't be appropriate because we are statistically limited, and the errors are not gaussian at low statistics. One reasonable method would be a maximum likelihood fit where the likelihood is defined as follows.

$$L = \frac{1}{\sqrt{2\pi}\sigma_b} e^{-\frac{(n_b - N_b)^2}{2\sigma_b^2}} \frac{e^{-(n_s+n_b)}(n_s+n_b)^N}{N!} \prod_{i=1}^{N} \frac{n_b f_b(m_i) + n_s f_s(m_i, M_{top}))}{(n_b + n_s)} \tag{7.15}$$

where n_s, n_b are the number of signal events and the number of background events, N is the number of fitted events on mass plot, N_b is the number of expected background events, σ_b is the systematic error on N_b, and $f_s(m_i, M_{top})$, and $f_b(m_i)$ are the mass probability density functions for signal and background, respectively.

The first term takes into account the systematic uncertainty on the estimated number of background events. The second term takes into account the Poisson fluctuation of the number of both signal and background events. Finally the last term is the likelihood of accommodating n_s signal events and n_b background events utilizing the mass shapes of the signal and background. n_s and n_b are the free parameters in the fitting.

The strategy is the following.

1. Fix the top mass.
2. Find n_s and n_b that maximize the likelihood.
3. From n_s, calculate the cross section at that mass.
4. From the error on n_s, calculate the limit.

5. Repeat at all masses.

To clarify step 4 of calculating error on n_s, let me go through a little digression of general probability arguments. n_s is one of the fitting parameters. In likelihood fitting from N measurements the error on a fitted parameter is given by the following.

$$-2ln(R) \longrightarrow \chi^2 \quad as \quad N \longrightarrow \infty \tag{7.16}$$

where

$$R = \frac{L(m,x)}{L_{max}(m,x_{max})} \tag{7.17}$$

and $L_{max}(m,x_{max})$ is the maximum likelihood with measurements m at fitted parameter of x_{max}.

This gives

$$ln(L) - ln(L_{max}) = -\frac{1}{2}\chi^2 = -\frac{1}{2}\frac{(x-\bar{x})^2}{\sigma_x^2} \tag{7.18}$$

and $\sigma_x = \frac{\sqrt{x^2-\bar{x}^2}}{\sqrt{2}}$ when $ln(L_{max}) - ln(L)$ is 1. This is how we usually estimate the error on x based on likelihood function $L(m,x)$. However this is true only if N is large, which certainly is not our case.

Another way of quoting this difficulty would be that when I find the n_s at the maximum likelihood with error on n_s (σ_{n_s}) calculated from the shape of the likelihood function, it still doesn't give the probability that the actual number of signal events was lower than $n_s + \sigma_{n_s}$ unless N is infinite. It only gives what value of n_s is most likely. But, when I set an upper limit on the cross section at a certain confidence level, I am basically quoting "the probability that the true cross section was actually lower than the limit is the value corresponding to that confidence level".

Therefore, to get the limit in this case of low statistics, I will have to rely on a different way of estimating the probability density function for cross section (or n_s), the probability density function of having n_s^{true} signal events at certain top mass M_t when the fitted number of signal events was n_s $(= P(n_s^{true}|n_s))$.

What I can do is to find the probability density function that I measure n_s when the true number of signal events is n_s^{true}. $(= P(n_s|n_s^{true}))$ This can be done by generating many ensembles of N events with average of n_s^{true} signal events at M_t in them, do the maximum likelihood fit to get the fitted n_s distribution from each ensemble. Then I can use Bayes Theorem to get $P(n_s^{true}|n_s)$ as shown below.

$$P(n_s^{true}|n_s) = \frac{P(n_s|n_s^{true})P(n_s^{true})}{\int P(n_s|n_s^{true})P(n_s^{true})dn_s^{true}} \tag{7.19}$$

Assuming that we don't have any *a priori* knowledge about the cross section $(P(n_s^{true}))$, we can just assume that it's flat function. Then we get

$$P(n_s^{true}|n_s) = P(n_s|n_s^{true}) \tag{7.20}$$

By integrating the tail of this probability density function from n_s^{limit} to infinity so that the percentage of the integrated area with respect to the total integrated area is 10%, for example, I can set a 90% confidence level limit on n_s from which I can also set a limit on cross section. This can be repeated at different masses.

A sample of 17 events out of the 19 candidate events from run 1a have fitted successfully. Their fitted mass distribution is shown in Figure 7.2. Assuming that both the signal events and background events have the same probability of failing the fit, my new number of candidate events is 17 and the estimated background is $13.5 \times \frac{17}{19} = 12.1$ events. As shown in Figure 7.2, I have estimated the number of signal events from maximum likelihood fitting at each mass. To get $P(n_s^{true}|n_s)$

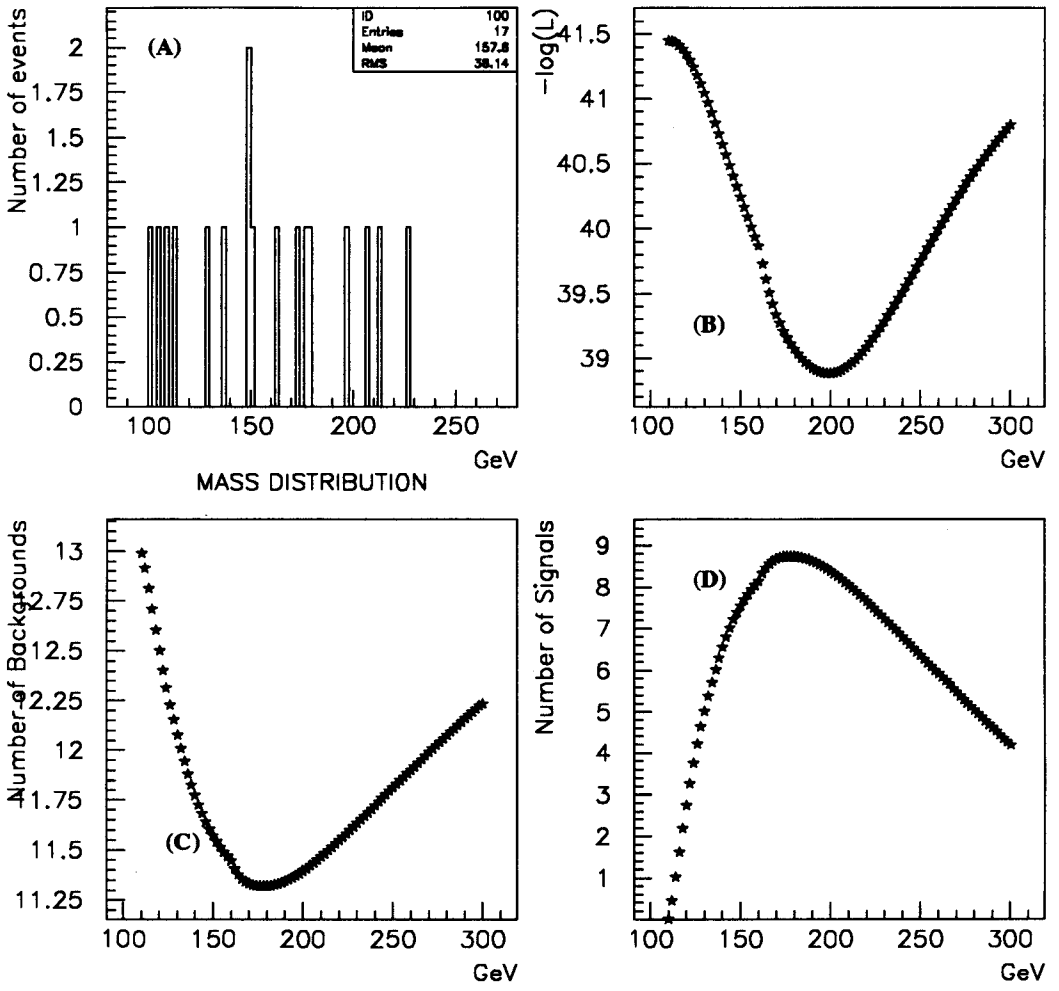

Figure 7.2: (A) Mass distribution of 17 candidate events. (B) -log(likelihood) vs top mass. (C) Fitted number of background events vs. top mass. (D) Fitted number of signal events vs. top mass.

which in my case is the same as $P(n_s|n_s^{true})$, I generate many ensembles of 17 events with average of n_s^{true} signal events in them, and get the distribution of fitted number of signal events (n_s). The systematic error on background estimation is 1.7 events. Figure 7.3 shows the average n_s and its 90% upper and lower limits vs. n_s^{true}. Figure 7.3 provides the 90% upper limit and lower limit on the true number of signal events when the fitted number of events n_s is given. The signal events in Figure 7.3 were generated at 160 GeV but it turned out that these limit curves almost don't depend on the top mass. These limits on the true number of signal events directly correspond to the cross section limits by

$$\sigma = \frac{n_s^{limit}}{\epsilon(M_{top}) \times Br. \times \int Ldt} \tag{7.21}$$

The resulting cross section limits from the fitted results shown in Figure 7.3 is shown in Figure 7.4.

Mass determination

Another measurement we can make using this technique is to extract the top mass. In this case, a strong assumption is that I have signal events in my data sample. If this assumption doesn't hold, there is no information from my data sample that can give any information about the mass of the top. Again, we are provided with the following information.

1. Mass probability density functions for both signal and background.
2. Mass distribution from the candidate events.
3. Estimated level of background.

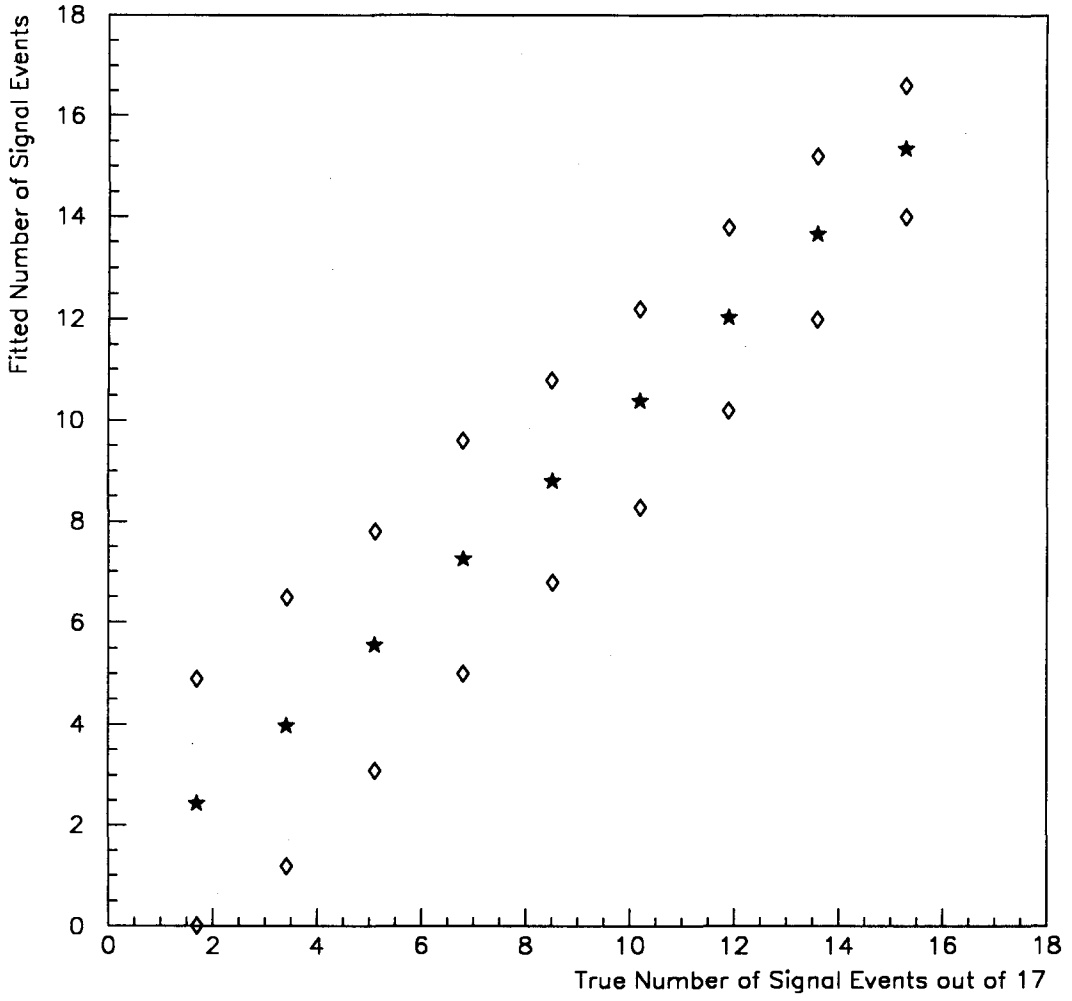

Figure 7.3: Average n_s from fitting (*stars*) and their limits at 90% CL (*diamonds*) vs the true number of signal events (n_s^{true}) from ensembles of 17 events. Signal events are generated at 160 GeV.

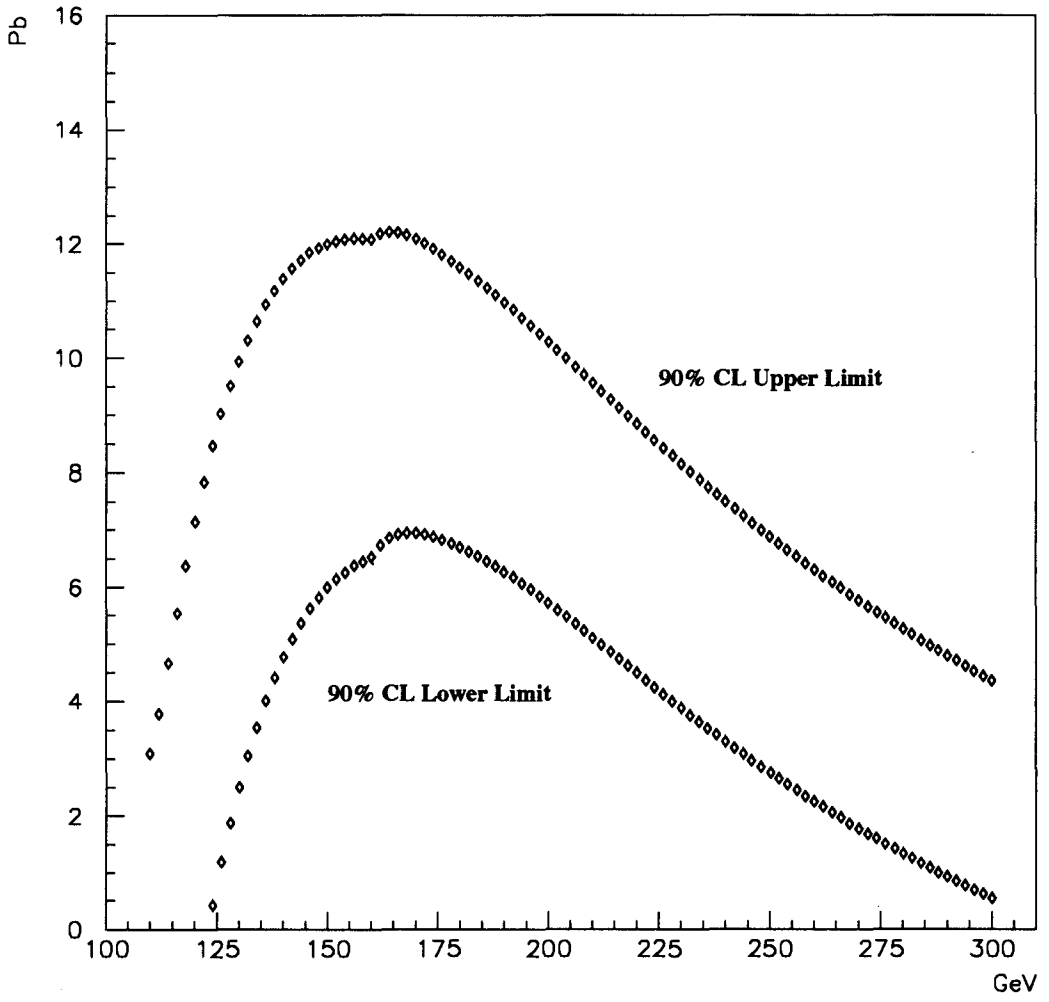

Figure 7.4: 90% CL upper and lower cross section limits as a function of top mass.

For this purpose also, we can use the same definition of likelihood as shown in Eq. (7.15). For an ensemble of events that provide a mass distribution, I can find the mass where the likelihood becomes maximum with some fitted value of n_s and n_b. When the top is generated at a certain mass, the probability that the likelihood becomes maximum at that mass will be roughly the largest. Therefore, we take this value as our best estimate of the top mass.

As described in the limit calculation, the error that maximum likelihood fit provides is not an accurate representation of the uncertainty on the fitted mass due to the fact that the number of events in the ensemble from which I calculate the fitted mass is not big enough to satisfy the following condition.

$$ln(L) - ln(L_{max}) = -\frac{1}{2}\chi^2 \tag{7.22}$$

An alternative way of estimating error would be to generate statistically independent ensembles of events (at certain mixture of signal and background) to get the fitted mass distribution from each ensemble. From the width of this distribution, I can estimate the error on the top mass. These ensembles of signal and background events are generated using a Monte Carlo method according to the mass resolution function of background and signal at different masses as shown in Figure (6.12). When we observe N_{ob} events and expect n_b background events. I generate N_{ob} events with binomial fluctuation of the number of background events with its average n_b.

To test the method, I consider various situations. First, when I have a large number of signal events with no background, Figure 7.5 shows how well this method performs in determining the mass of the top quark at high statistics. What one can notice from Figure 7.5 is that the uncertainty in mass is roughly proportional to $1/\sqrt{N}$ and also the fitted mass exactly matches the generated top mass.

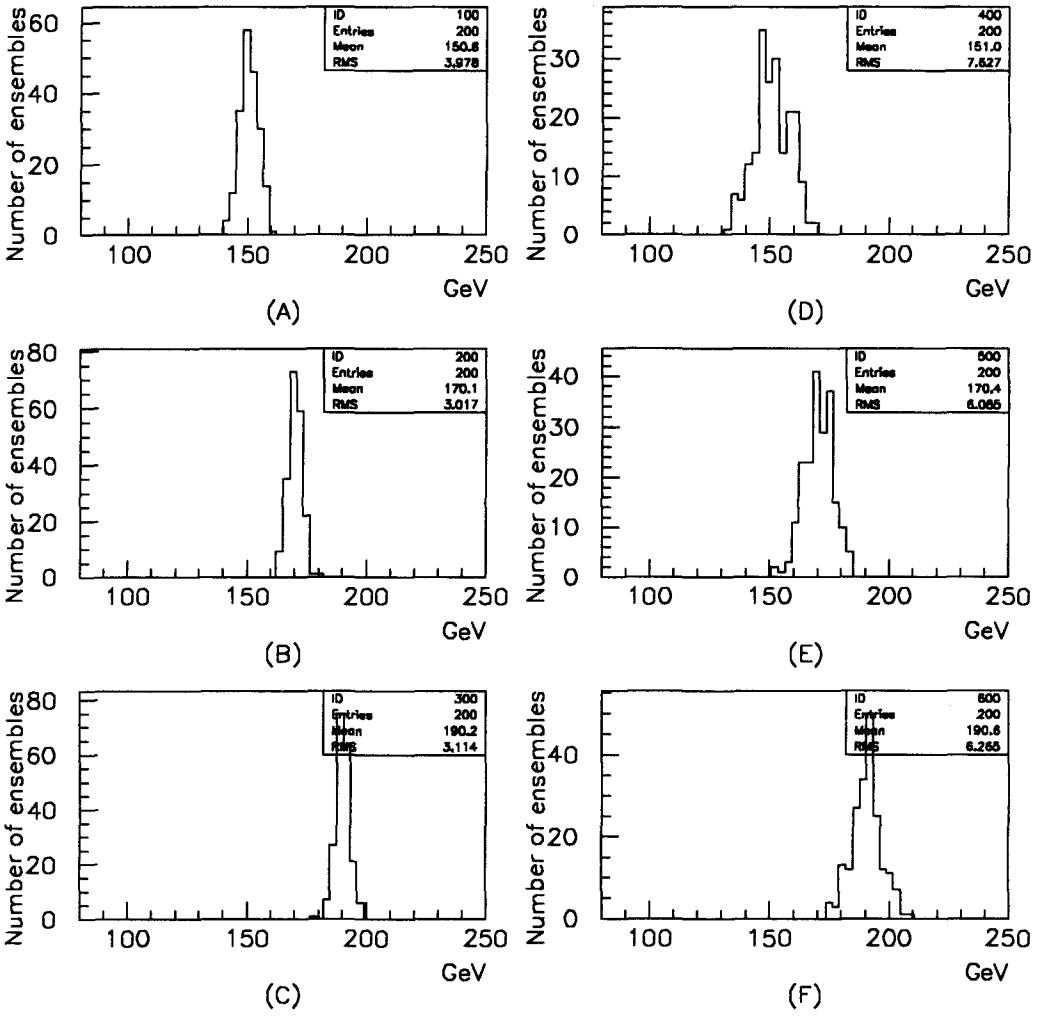

Figure 7.5: Mass determination from 200 ensembles of 200 signal events at M_t of (A) 150 GeV, (B) 170 GeV and (C) 190 GeV. Also from 200 ensemble of 50 signal events((D), (E), and (F))

Another example of its performance when I have rather poor statistics and small background contamination is shown in Figure 7.6. 1000 ensembles of 20 events with 10% background were generated at top masses from 120 GeV to 240 GeV. Figure 7.6 shows that the fitted mass is also linearly proportional to the generated mass but the proportional constant is not exactly 1.

As shown in Figure 7.2, when I apply this technique to the 17 candidate events with 12.1 estimated background events with 1.7 calculated background uncertainty, I get a top mass of 200 GeV. To estimate the uncertainty, I generated many ensembles of 17 events with 12.1 background events in them (with binomial fluctuation) at various top masses. Figure 7.7 shows the mapping between the fitted mass and the true mass as well as the error on the most probable true mass when the fitted mass is around 200 GeV. From Figure 7.2 and Figure 7.7, we can say that the measured top mass is 207 GeV and the statistical error on the mass is 28 GeV.

Data selection II

As mentioned earlier, an alternative approach is to use the mass fitting technique to get a signal-enriched sample of events and extract a mass from this sample. An advantage would be that most of the contribution in the likelihood fit comes from the signal events which carry the mass information and the fit doesn't have to consider various possible ways of accommodating the background content of the sample since it's small. For this, I adapted the DØ conventional $e + jets$ channel data selection cut. The differences between DØ standard cuts and my previous cuts are shown in Table 7.4 There are 8 events with jet multiplicity of 4 or more in the $e + jets$ channel.

The summary of this standard analysis is shown in Table 7.5 [10]. In background

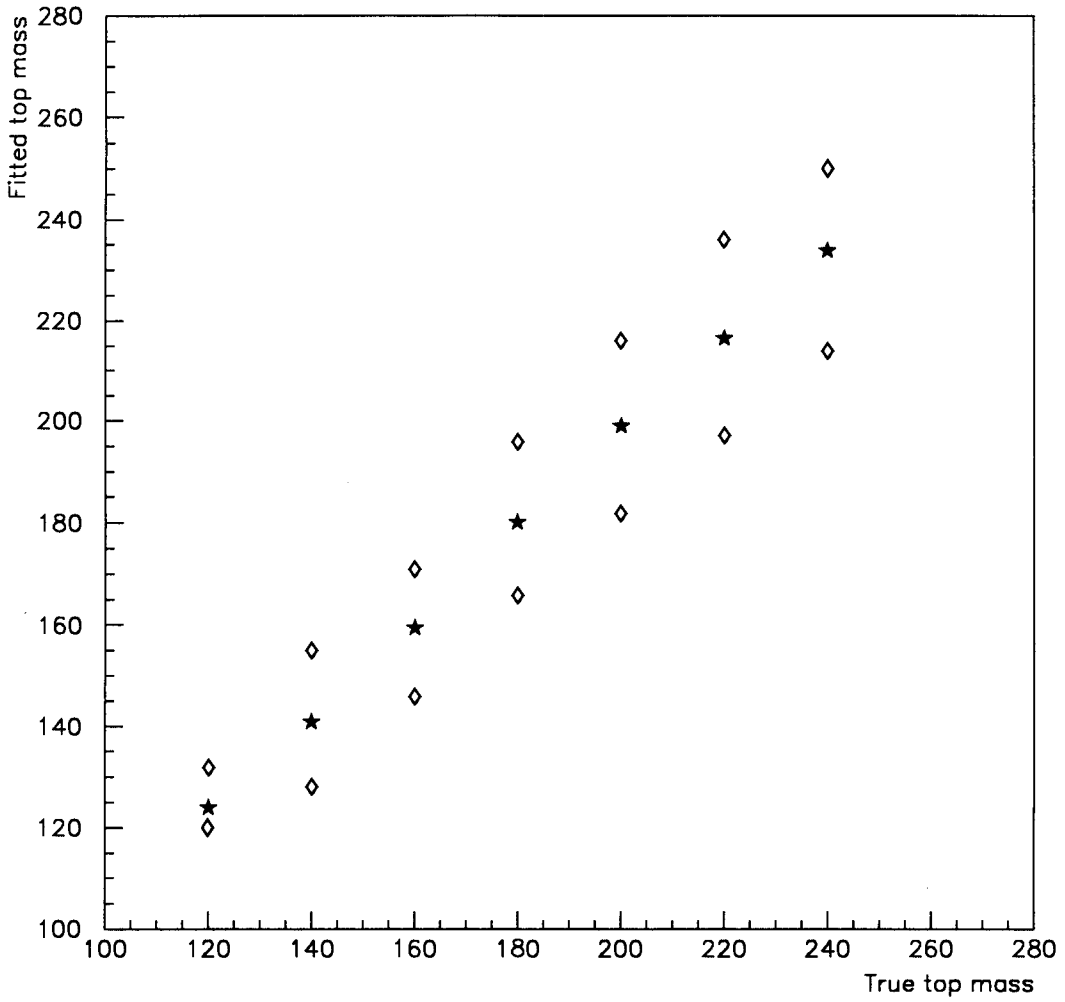

Figure 7.6: Average fitted mass (*stars*) and their 90% upper limit (*diamonds*) vs generated top mass. (20 signal events with 10% background)

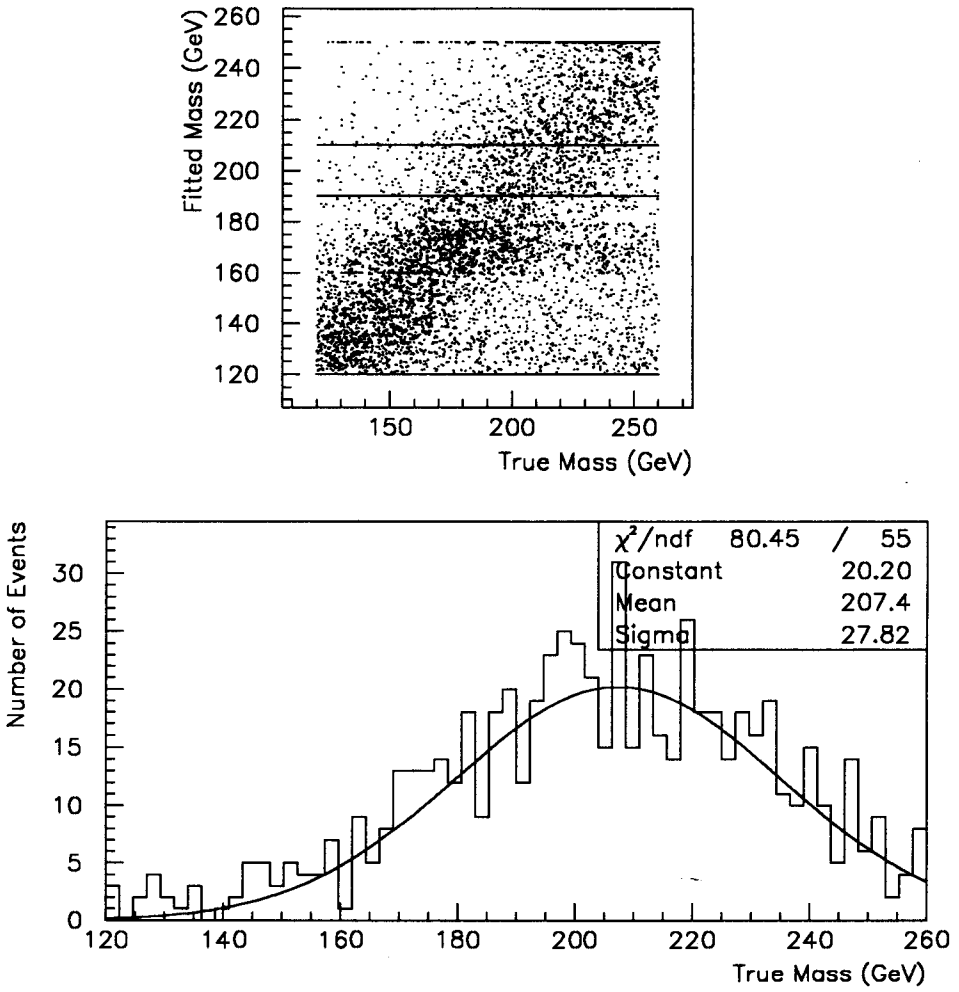

Figure 7.7: Maximum likelihood fitted mass vs. the true mass (*upper figure*), the distribution of the true masses when the fitted masses are within 200 $\pm$ 10 GeV (*lower figure*). Data selection I.

Table 7.4: The differences between DØ standard data selection cuts and cuts I used in the previous section.

Difference	DØ standard	Previous selection
Jet Def.	0.5	0.3
Jet E_t	15 GeV	8 GeV
$\not\!\!E_t$	25 GeV	30 GeV

Table 7.5: The summary of the standard e+ jets data selection.

Type of background	Estimated number in 8 events
QCD (Method I)	1.6
$W + jets$ (Method I)	4.8
QCD & $W + jets$ (Method II)	5.8

estimation, two different methods were used. The first method (Method I) is to use scaling law, and the other method (Method II) is to fit in aplanarity and H_t space.

Only 7 events out of 8 succeeded in mass fit. Therefore I scaled down the estimated background corresponding to the 7 events. I combine the QCD background and the $W+jets$ background and treat them as if all of them are from $W+jets$ since I have technical difficulty in estimating the mass resolution function from QCD background (due to low statistics) and also the QCD background is a small contribution. The estimated number of background events is

$$\approx 5.8 \times (7/8) = 5.08 \ \textit{events} \tag{7.23}$$

The systematic error on this number is calculated to be about 30% [12], which is 1.52 *events*. The fitted results are shown in Figure 7.8. The fitted top mass again come out to be about 200 GeV. This is not surprising since this sample of events is not independent from the sample in the previous section. 5 events out of the 7 are also included in the previous data set of 17 events. It's also true that the enrichment

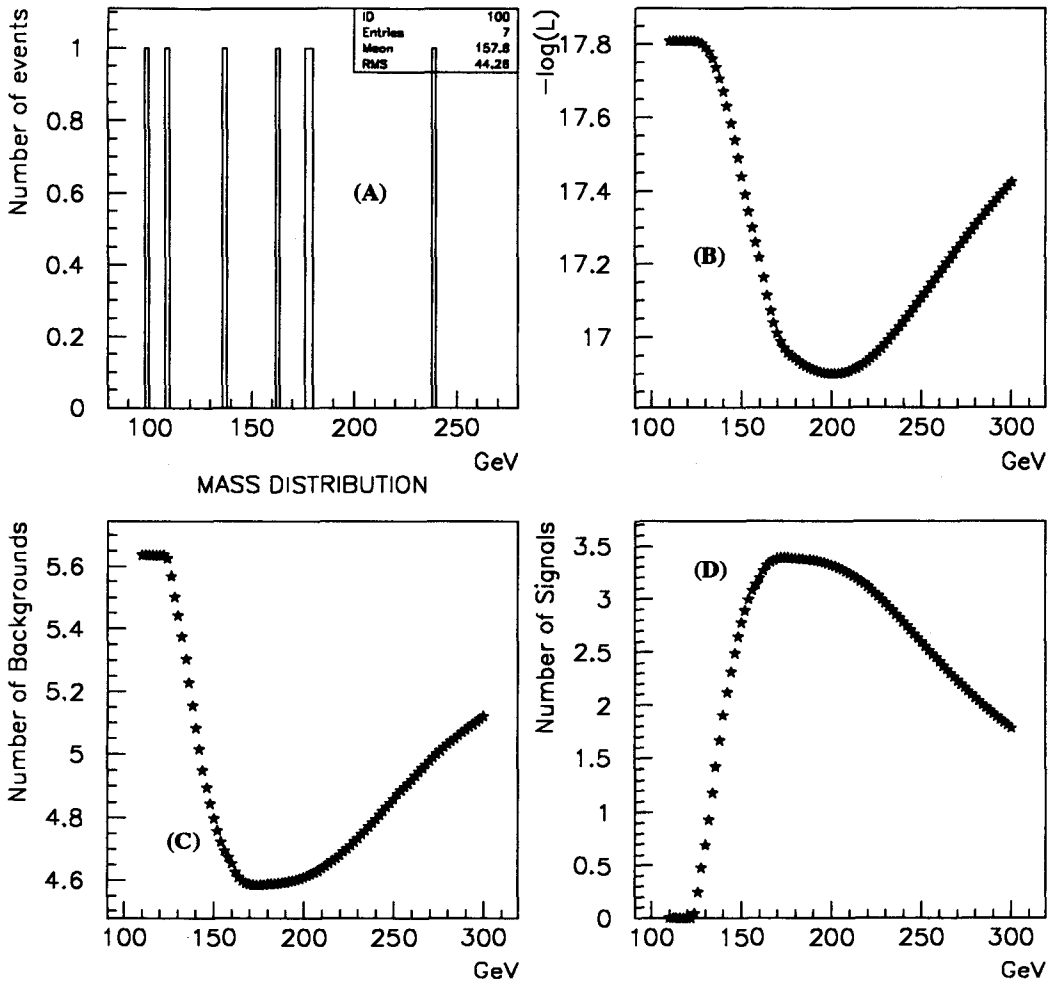

Figure 7.8: (A) Mass distribution of 7 candidate events. (B) -log(likelihood) vs top mass. (C) Fitted number of background events vs. top mass. (D) Fitted number of signal events vs. top mass.

of the signal wasn't noticeably better in this data selection to give better result in mass measurement.

To estimate the range of true top masses from this fitted mass of 200 GeV, I generated many ensembles of 7 events with 5.08 background events in them in average at various top masses. The results are shown in Figure 7.9 From Figure 7.9, I can conclude that the true top mass is 214 GeV with statistical error of 39 GeV.

Systematic error

The most noticeable contributions to the systematic error in mass determination are the following.

1. The systematic uncertainty of jet energy scale.
2. Systematic difference in the shape of the mass resolution function between MC and data.
3. Systematic shift introduced by the likelihood method itself.

The third one is already shown in Figure 7.7 and 7.9. Since this systematic behavior can be known from MC, it doesn't introduce any uncertainty. The first and the second contribution appear through the mass resolution functions. One interesting test would be to see the sensitivity of the systematic error on fitted mass due to our poor knowledge of the background mass resolution function. For this purpose, two different background mass probability density functions were used as shown in Figure 7.10. Here I consider the following two cases.

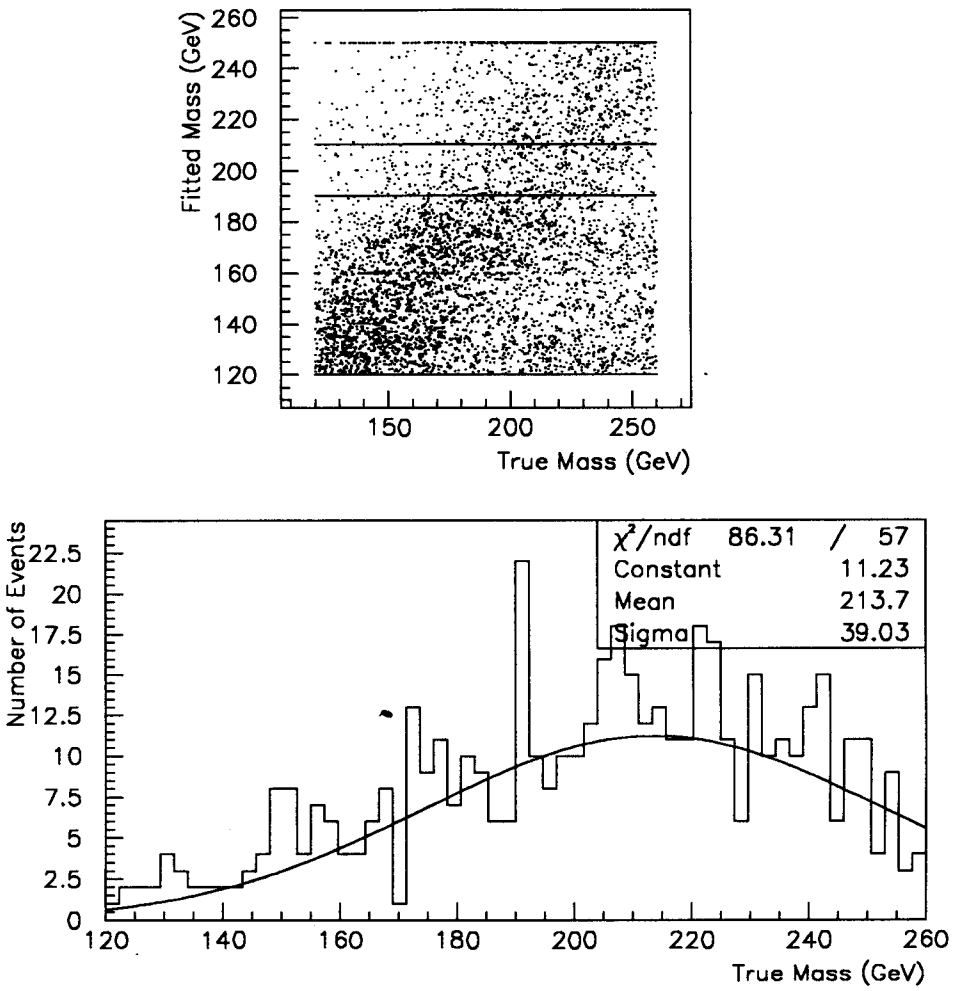

Figure 7.9: Maximum likelihood fitted mass vs. the true mass (*upper figure*), the distribution of the true masses when the fitted masses are within 200 $\pm$ 10 GeV (*lower figure*). Data selection II.

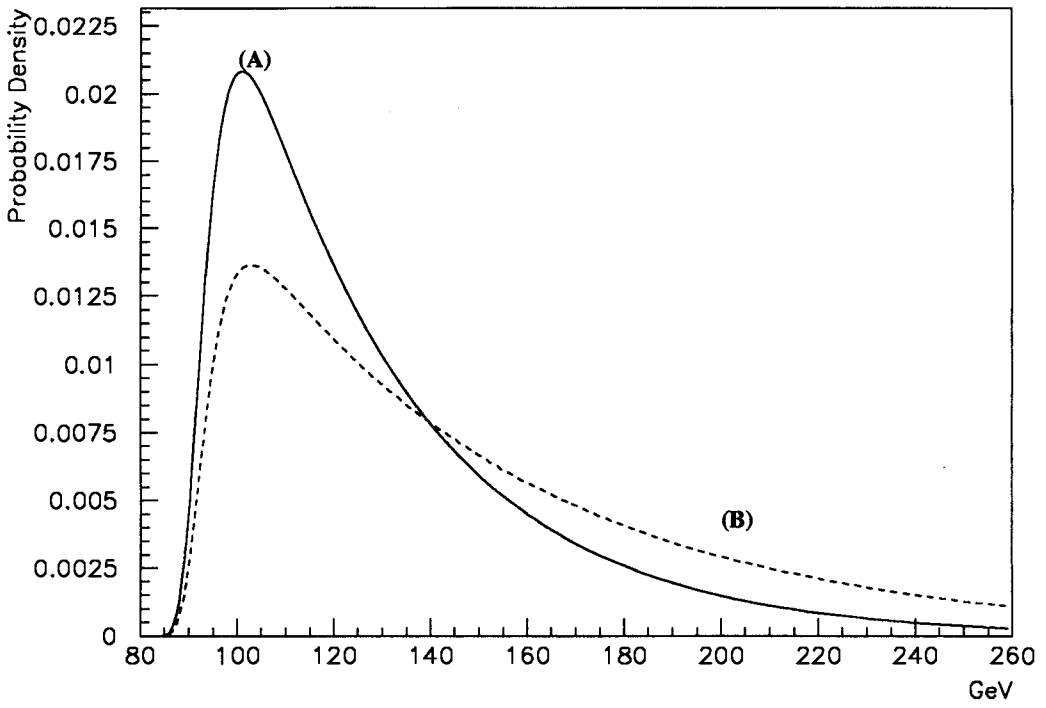

Figure 7.10: Two different background mass resolution functions. The distribution (B) has thicker tail than (A) by factor of 2 at around 200 GeV.

Case 1. Background mass distribution is generated by the probability density function (A) in Figure 7.10 and also the same function is used in likelihood fit.

Case 2. Background mass distribution is generated by the probability density function (B) in Figure 7.10 and the function (A) was used in likelihood fit.

The resulting differences in average fitted masses are shown in Figure 7.11. (Again, I used 17 event with 12.1 ± 1.7 background events in each ensemble) As shown in Figure 7.11, 50% uncertainty in the tail of background mass distribution introduces systematic error of about 30 GeV at this statistical level. (17 events with 12.1 ± 1.7 background events)

Conclusion

At present, when I am left with a handful of candidate events with all the systematic errors floating around, making a statement on whether I found the top or not seems rather religious than scientific. However, science, in my point of view, is about methodology of how we ask our question and how we attack our problems as much as it is about knowing the facts. Following this philosophy, my attempt throughout my research was to come up with a valid, consistent, and efficient method which can tell me not only about how much I do know but also about how much I do not know from the experimental observations in an objective way.

My objective has been to measure the cross section of $t\bar{t}$ events, and to measure the mass of the top quark. Considering that the determination of top mass is

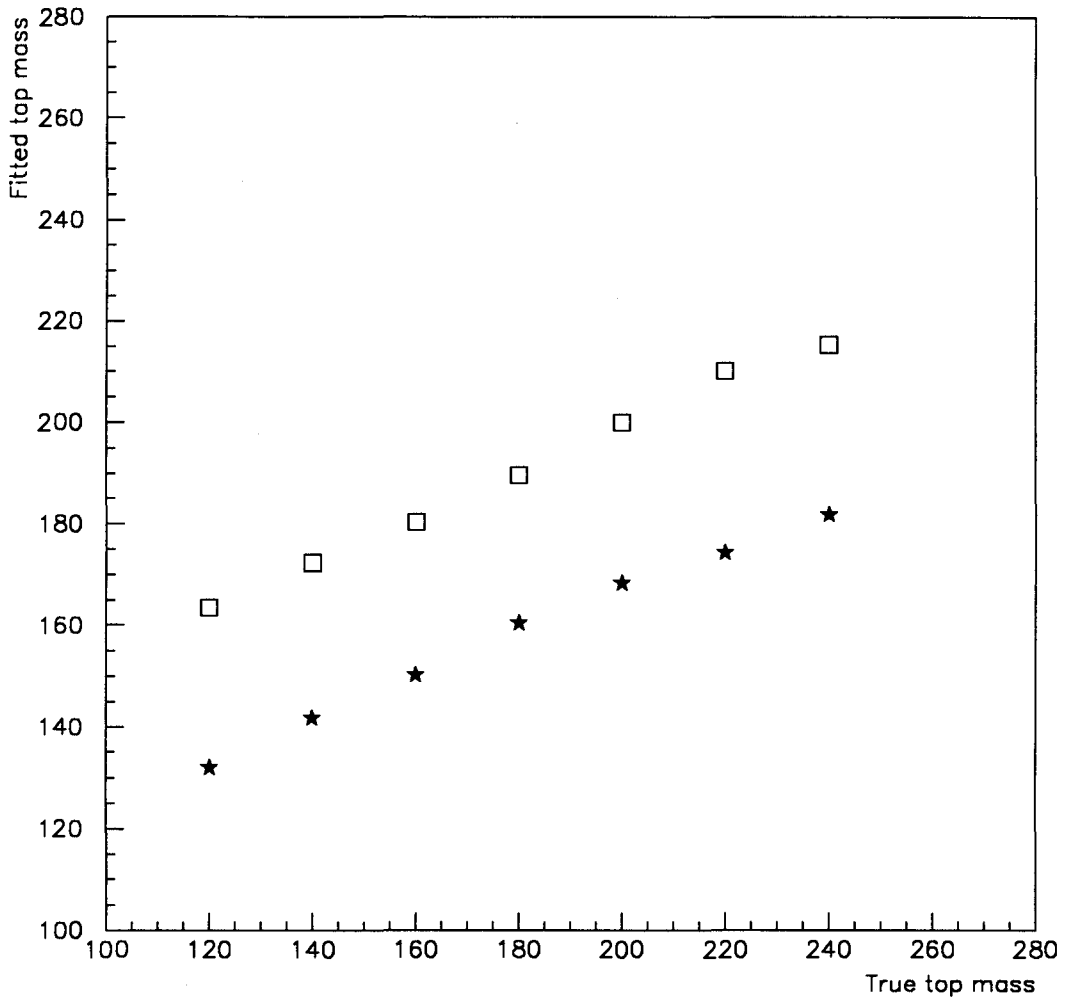

Figure 7.11: Average fitted mass vs. generated mass. When the background shape was correct (*stars*), and when the tail of the background mass distribution was underestimated by factor of 2 (*squares*).

meaningless if the cross section for $t\bar{t}$ production was zero, my analysis on top mass determination has been under the assumption that the cross section for $t\bar{t}$ production was non-zero even though this assumption wasn't strongly supported due to low statistics. Therefore, a reasonable interpretation of my analysis would be the following statement.

If the excess of events at high multiplicity (4 or more jets) in our data is due to $t\bar{t}$ events rather than a statistical fluctuation or a systematic effect, the measured top mass is 207 ± 27.8 (statistical) GeV/c^2.

It is a general tendency that our candidate events have higher H_t distribution than we expect from our modeled MC background events. Whether this is a real effect due to $t\bar{t}$ content of our data sample or not will have to be studied in detail, hopefully at high statistics, since it has significant effect on my analysis.

Depending on what data selection cuts we make, we will have different mass resolution functions, different number of candidate event, and different background estimation. This will result in a different statistical error on the measured top mass. Therefore, which selection cuts will provide the smallest error on the top mass will provide the justification for the data selection cuts, which also needs to be studied.

BIBLIOGRAPHY

[1] DØ collaboration, S. Abach *et al.*, Nucl. Instrum. Methods **A338**, 185 (1994)

[2] F.Abe, *et al.*, "Evidence for Top Quark Production in $p\bar{p}$ Collisions at $\sqrt{s}$ =1.8 TeV," Submitted to Phy. Rev. D April 22, 1994

[3] E. Laenen, J. Smith, W. L. van Neerven, "Top Quark Production Cross Section," Fermilab-Pub-93/270-T, August 1993

[4] Bob Kehoe, "Resolution Bias in Jet Response Measurement," DØ note #2052, February 8, 1994

[5] John M. Butler, "Main Ring Deadtime," DØ note #1682, February 28, 1993

[6] F. Berends, H. Kuijf, B. Tausk, and W. Giele, Nucl. Phys. **B357**, 32 (1991)

[7] S. Abachi, *et al.*, "Search for High Mass Top Quark Production in $p\bar{p}$ Collisions at $\sqrt{s}$ =1.8 TeV," Final version to be submitted to Phy. Rev. Letts. Oct. 1994

[8] Haowei Xu, Rich Partridge, "$W + jets + \mu$ tag Background study," DØ note #2206, July 14, 1994

[9] Sailesh Chopra, Rajendran Raja, "Estimation of the QCD background to $W \longrightarrow e\nu + jets$," DØ note #2098, April 14, 1994

[10] Dhiman Chakraborty, "A search for $t\bar{t} \longrightarrow electron + \not{E}_t + jets$ signature in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.8 TeV with the DØ detector," Ph.D dissertation, State University of New York, Stony Brook, NY (1994)

[11] Serban Protopopescu, "Search for top in lepton + jets final state with DØ detector," 27th International Conference on High Energy Physics, Glasgow, Scotland, July 1994

[12] Private conversation with Serban Protopopescu, DØ DAB, Fermi Lab. Oct. 1994

[13] Andy James Milder, "Dijet Angular Distributions at $\sqrt{s}$ = 1800 GeV Using the DØ Detector," Ph.D dissertation, The University of Arizona (1993)

[14] O. I. Dahl, T. B. Day, F. T. Solmitz, N. L. Gould, "SQUAW kinematic fitting program," LBL, Group A Programming note No. P-126, July 1968

APPENDIX FITTING OF $Z \longrightarrow ee$ DATA SAMPLE

The $Z \longrightarrow ee$ data sample is exceptionally clean in the sense that there are two very well measured electrons and there is no missing neutrino. In addition, we know what physical process is responsible for these events from the sharp Z resonance peak. Figure A.1 shows the data sample being used. This provides a unique situation where I can apply my constrained kinematic fitting method under the known correct assumption of $Z \longrightarrow ee$ decay. By looking at the pull quantities, I can make sure that the errors assigned are reasonable for various quantities such as η, φ, E of electrons and jets, and the baby jet P_x, P_y. The fitted parameters and the constraints for $Z + 0jet$ case are shown in Table A.1 .

In $Z + 0jet$ events, we have η, φ, and E of electron and P_x, P_y of the baby jet to which errors are assigned. What I want is to adjust the errors of these parameters until the pull quantities of these parameters are centered at zero and have σ of one. One question to be addressed is whether a set of resolutions that calibrates the pull quantities is unique. The answer to this question seems to be 'No'. For example, the θ resolution of electrons and the energy resolution of the electrons can have a different set of values still satisfying the pulls centered at zero with width of one. Therefore what one needs is a reasonable estimate of the set of errors to start with. If all the errors are exactly known except one, the error on that one quantity can be

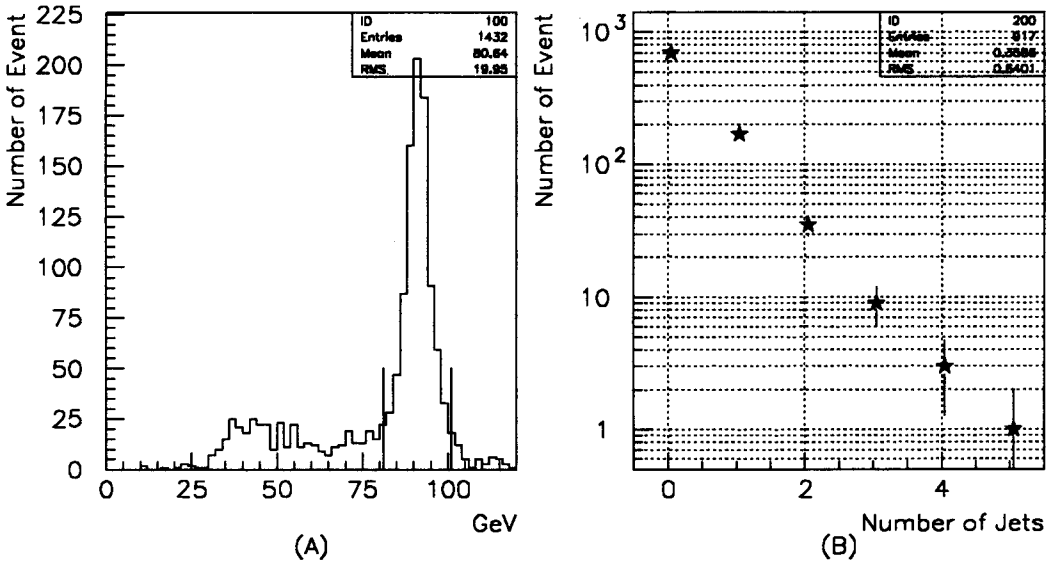

Figure A.1: (A) The two electron mass distribution after 'tight' electron cut on both electrons. (B) Jet multiplicity distribution within the Z mass band. (no jet E_t cut)

Table A.1: Fitted parameters and constraints in $Z + 0jet$ events.

Parameter 1	Energy of the electron 1
Parameter 2	Energy of the electron 2
Parameter 3	Phi of the electron 1
Parameter 4	Phi of the electron 2
Parameter 5	Theta of the electron 1
Parameter 6	Theta of the electron 2
Parameter 7	Px of the baby jet
Parameter 8	Py of the baby jet
Constraint # 1	Px momentum balance
Constraint # 2	Py momentum balance
Constraint # 3	Mass of the Z

determined by calibrating the pull quantity.

We know the errors of the electrons in $Z \longrightarrow ee$ events relatively well compared to the baby jet, and also the P_x and P_y of the baby jet are weakly coupled to the Z mass constraint so that changing the baby jet resolution doesn't affect the pulls on electron energy or θ, but affects the electron φ a lot. On the other hand, electron θ and electron energy are coupled strongly by the Z mass constraint, so changing the resolution of one affects the other significantly.

With the best estimate from results of independent study and calibration with the pull quantities, I come up with the following resolutions shown in Table A.2 and the pull quantities with this set of resolutions are shown in Figure A.2. Using these calibrated errors from $Z + 0jet$ data, we can expand our calibration of errors to jets in $Z + 1jet$ events. In this case, we have some more parameters to use with the same constraints. They are shown in Table A.3.

To ensure that the jet in the event is a recoiling jet of the Z, I made the following cuts.

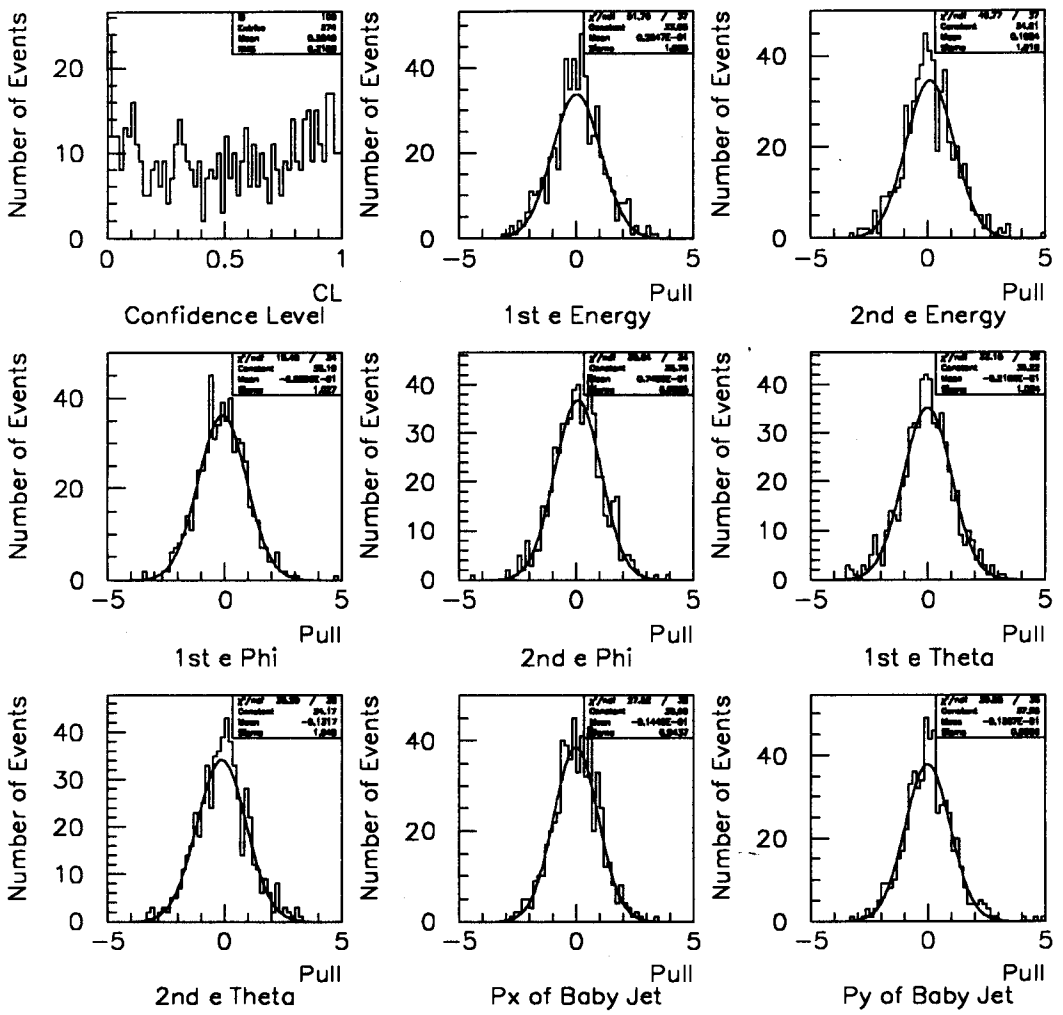

Figure A.2: Confidence level of the fit and pull quantities of the parameters with the best estimate of the errors on them.

Table A.2: Best estimate of electron and baby jet resolutions.

Description	Resolution
e Energy (Sampling Term)	15 %
e Energy (Constant Term)	2 %
e Energy (Noise term)	2.0 GeV
e θ Resolution	0.011 rad
e φ Resolution	0.003 rad
Baby Jet P_x, P_y Resolution	6. GeV

Table A.3: Fitted parameters and constraints in $Z + 1jet$ events.

Parameter 1	Energy of the electron 1
Parameter 2	Energy of the electron 2
Parameter 3	Energy of the jet
Parameter 4	Phi of the electron 1
Parameter 5	Phi of the electron 2
Parameter 6	Phi of the jet
Parameter 7	Theta of the electron 1
Parameter 8	Theta of the electron 2
Parameter 9	Theta of the jet
Parameter 10	Px of the baby jet
Parameter 11	Py of the baby jet
Constraint # 1	Px momentum balance
Constraint # 2	Py momentum balance
Constraint # 3	Mass of the Z

$E_t^z > 10$ GeV

$||\varphi^z - \varphi^{jet}| - \pi| < 0.5$

The jet correction on top of CAFIX is made based on the out-of-cone radiation correction as described in Chapter 6. The pull quantities from $Z+1jet$ events after the correction are shown in Figure A.3. It was true that the systematic imbalance of the Z and the recoiling jet was also shown consistently in the pull quantity distribution

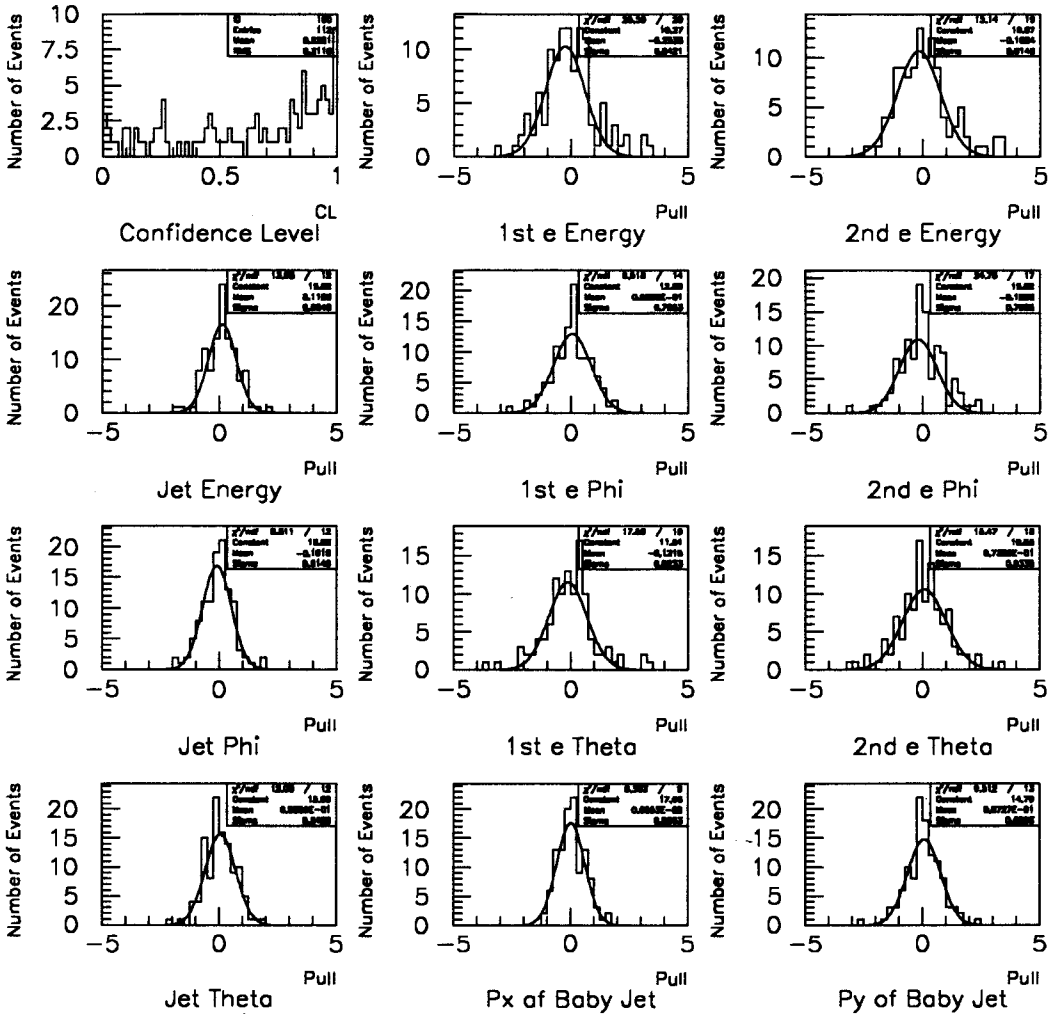

Figure A.3: Confidence level of the fit and pull quantities from $Z + 1Jet$ events.