for including ELE_JET is its lower threshold on the E_T of the electron (the 20 GeV required by ELE_HIGH is too high for us at the trigger level).

Our studies indicate that in general, the $\not{E}_T$ in an event is correlated with the electromagnetic contents of the jets. However, once there is a jet that satisfies the EM requirements made by the above triggers, it is safe to assume that the correlation is not too different from the events that actually appear as backgrounds after all our selection cuts. In order to interpret (the bulk of) this sample as non-W background, we treat the electromagnetic cluster satisfying the triggers as an electron. Basically, this amounts to a loosening of the electron identification. The procedure is described in details in Subsection 6.2.1.

Chapter 6

Analysis

We shall describe the details of our analysis in this chapter. We take two independent approaches at the first stage to estimate the amount of signal in our data. The first approach assumes that the distribution of the number of events as a function of the number of jets has an exponential form for the background. The distribution has a different form for the signal, which is determined using the ISAJET Monte Carlo. Fitting the two forms to that observed in our data enables us to estimate the amount of signal in it. This approach does not attempt to study explicitly the correlations between the jets in an event. It requires Monte Carlo simulation mainly for the signal: conclusions about the background are based almost entirely on the data itself. The second approach does not make any assumption of the jet-multiplicity spectra in signal and background. Requiring the events to have at least 4 jets, it focuses on the correlation between the momenta of the jets and the overall energy scale of the event. This offers a remarkable discrimination between signal and backgrounds, the parton jets being more isotropically distributed in the former than in the latter. In this method, we rely on Monte Carlo for the distributions and the correlations of the jets in the W background. The study of the non-W background is still based entirely on our own data. The results from the two approaches agree very well within errors, which are relatively large. Finally, we pick a subspace of the phase space covered by the two previous approaches, calculate the background in it using the average of the two approaches and arrive at yet another estimate for the amount of signal in our data.

In the first section, we describe the event selection. The next two sections contain descriptions of the two analysis methods and the results obtained from them. In the section after that, we give the estimate based on the "purest" albeit "smallest" sample combining the background estimates from either approach. The last section contains some remarks on our analysis.

6.1 Event selection

6.1.1 Triggers

In order to receive further attention, an event has to satisfy at least one of the two Level 2 triggers described in Section 5.2. (ELE_HIGH and ELE_JET). ELE_HIGH requires neither any $\not{E}_T$ nor any jet activity in an event. It would thus have been ideal for all our studies but for the high cut on the E_T of the EM object. As mentioned before, we do not want to place offline cuts too close to the trigger requirements because then the trigger inefficiency becomes high. Making the offline requirement on the E_T of the electron much higher than 20 GeV costs us too much in efficiency. In fact, as we shall soon see, we want to place the offline cut at 20 GeV. Therefore, we include ELE_JET which has a looser requirement on the E_T of the EM object (see Section 5.2). Because ELE_JET requires one jet in addition to the EM object, we cannot study the "$W + 0$ *jet*" candidates. It also introduces some trigger bias for the "$W + 1$ *jet*" events since an $R = 0.5$ jet barely passing the $E_T > 15$ GeV cut that we want to make after corrections offline, will occasionally fail the $E_T > 10$ GeV cut at L2 where R = 0.3 cones are used and no corrections are applied. However, the effect of this bias is barely noticeable since the chance of having such an event which also fails ELE_HIGH is small. The cut on $\not{E}_T$ in ELE_JET does not affect us in a big way since our final offline cut is much tighter.

Note that in ELE_HIGH, $\mid \eta \mid$ of the EM cluster is limited only by the acceptance of the detector ($\mid \eta \mid < 3.5$). The tighter constraints on $\mid \eta \mid$ of the objects in ELE_JET are inconsequential to us since we make tighter cuts offline. According to TRIGSIM, the combined efficiency of these two triggers on those signal events that would have at least 4 jets after our final cuts (there we do not count the electron as a jet) is 0.98.

6.1.2 Offline cuts

For the offline event selection, we shall first give the motivation, and then the values of various cuts.

- The E_T spectra of the parton jets are expected to be stiffer in signal than in the backgrounds. We therefore require our jets to pass a minimum threshold on E_T:

$$E_T(jet) > 15 \text{ GeV} \tag{6.1}$$

- The p_T distribution of the leptons from real W's peaks at about $\frac{M_W}{2}$. Hence, the electron and the neutrino candidates are required to pass a minimum threshold on their E_T's:

$$E_T(e) > 20 \text{ GeV} \tag{6.2}$$

$$\not{E}_T > 25 \text{ GeV} \tag{6.3}$$

  The reason for cutting harder on $\not{E}_T$ is that unlike the electron for which we have the luxury of making so many other quality requirements, for the neutrino we have none.

- The jets from our signal have a lesser tendency to crowd near the beamline compared to the backgrounds. This prompts us to choose only those jets which have a small pseudorapidity:

$$\mid \eta(j) \mid < 2.0 \tag{6.4}$$

- A non-w background enters the stage when statistical fluctuations cause a non-electron object to be misidentified as an electron. Like the other jets, such a "fake" electron also tends to be in a forward direction more than the electrons from a real W. Therefore, we require the electrons to be central:

$$\mid \eta_D(e) \mid < 2.0 \tag{6.5}$$

where η_D refers to "detector η" i.e. the pseudorapidity calculated using $z = 0$ for the vertex. The reason for using this variable, instead of the "physics η" which uses the primary interaction vertex determined for the event is that the the probability of a photon converting into an electron-positron pair before reaching the tracking volume depends the amount of material upstream that it has had to traverse which, in turn dictates the probability of its misidentification as an electron. Thus, the cut on η_D is partly "fiducial" in nature.

6.2 Number of jets produced in association with a *W* (Method A)

Let $\sigma(W + nj)$ represent the cross-section for the process of production of n parton jets in association with a W. We define

$$\rho(n) \equiv \frac{\sigma(W + nj)}{\sigma(W + (n+1)j)} \tag{6.6}$$

$\rho(n)$ is a function of the strong coupling constant (α_s) and is determined by the threshold imposed on the E_T of the parton jets $(E_T^{\rm min}(j))$. This part of our analysis is based on the conjecture that $\rho(n)$ is independent of n for small values of n $(n \leq 5)$ for that part of the process that does not involve any resonant combination (for example, a t) among the jets [103]. Gradual departure is expected as the phase space available to an additional parton becomes more and more restricted at higher values of n. The conjecture is based on the current understanding of perturbative QCD. Our signal, in contrast, results in the production of $W + jets$ via a cascade decay process through high-mass resonances. Thus, the conjecture does not apply to our signal. In this part of our analysis, we attempt to exploit this difference between the signal and the background by studying how well our data lends itself to the conjecture. In particular, we look for an excess in the number of events that have 4 or more jets since that's where our signal is expected to have the optimum enhancement relative to the background. Henceforth, unless explicitly stated otherwise, the jet multiplicity of an event will always stand for the inclusive count. For example, N_m will stand for "the number of events with m or more jets" i.e. $S_l \subset S_k$ for $l \geq k$ where S_m stands for the m-jet sample. Note that barring departures at large n, equation 6.6 holds with the same $\rho(n)$ whether one uses inclusive counts or exclusive.

6.2.1 Estimation of non-W background in data

First, we estimate the amount of non-W background in each bin of jet multiplicity by the following procedure. Let us imagine a 2-dimensional space defined by the quality of the leptonic decay products of the W candidate (i.e. the electron and the neutrino). It is here that we expect to see the most pronounced separation between the W background and the non-W background. For the electron, we have numerous handles as discussed in the last two chapters while for the neutrino, we only have the missing transverse momentum. We label the axes as "electron quality" (which is the logical .AND. of all the individual electron-id requirements) and "neutrino quality" (just $\not{E}_T$). We have cuts separating "good" (g) from "bad" (b) in either direction, giving us four quadrants: gg, bg, bb, gb, where for example, an event passing the "neutrino quality" cut but failing to meet the the electron-id requirements would fall in the quadrant bg. For brevity, we label the quadrants numerically: $1 \Rightarrow gg$, $2 \Rightarrow bg$, $3 \Rightarrow bb$, $4 \Rightarrow gb$. We take the ELE_HIGH and ELE_JET triggers (described in Section 5.2), the same as our data set, which have rather loose requirements on electron-id and on $\not\!\!E_T$ (for the most part, ELE_HIGH is satisfied which has no cut on $\not\!\!E_T$). To avoid trigger bias, at least one jet is required offline. Let N_i be the number of events observed in the $i^{\rm th}$ quadrant and let N_i^W and N_i^Q be the contributions of the real and fake W's respectively to it:

$$N_i = N_i^W + N_i^Q \tag{6.7}$$

Our aim is to estimate N_1^W. We define

$$a_{i,j}^W \equiv \frac{N_i^W}{N_j^W} \tag{6.8}$$

so that equation 6.7 can be rewritten as

$$N_i = a_{i,1}^W \cdot N_1^W + N_i^Q \tag{6.9}$$

We have $N_1 = 1316$, $N_2 = 812$, $N_3 = 6026$, $N_4 = 1624$. The $a_{i,1}^W$'s are estimated using Monte Carlo, taking into account the fact that it overestimates the electron-identification efficiency by about 15%: $a_{2,1} = 0.35$, $a_{3,1} = 0.11$, $a_{4,1} = 0.31$. Thus, we have 4 equations for 5 unknowns (N_i^Q and N_1^W).

For the $5^{\rm th}$ equation, we need to make an assumption. We assume that having met the loose electron-id criteria of the trigger, the two variables, $\not\!\!E_T$

and electron quality, are independent for the non-W events in this sample. This assumption would be wrong in a minimum-bias sample since the $\not{E}_T$ of an event is sensitive to the electromagnetic content of its jets. But the trigger assures us that at least one of the jets (which may well be a legitimate electron) is highly electromagnetic, just what we have in our signal and in the W background. This gives us the 5^{th} equation we need:

$$\frac{N_1^Q}{N_2^Q} = \frac{N_4^Q}{N_3^Q} \tag{6.10}$$

We solve the equations iteratively starting with a guess for N_1^W. For the first guess $N_1^W = 0$, the procedure converges after 4 iterations, giving $N_1^W = 1223.6$. This lets us estimate the amount of non-W background in our sample:

$$N_1^Q = N_1 - N_1^W \tag{6.11}$$

The process is repeated for higher multiplicities. The third column of Table 6.1 lists the results (see the next subsection for the difference between the convention we are using at this point and that used in the table). Figure 6.1(a) shows our estimate of N_1^Q as a function of jet multiplicities for two thresholds on $E_T(j)$. The line is an exponential function drawn through the points of 2 and 3 jets The point observed at 1 jet falling slightly below the line agrees with our expectation due to the aforementioned trigger-bias. Figure 6.1(b) shows the distribution for what is left of our data after subtracting the estimated non-W background from it. This is then the sum of the W-background and the signal (to be estimated next). Here, the exponential fit is made using the points of 1, 2 and 3 jets. 1-jet is included since the effect of the trigger bias is

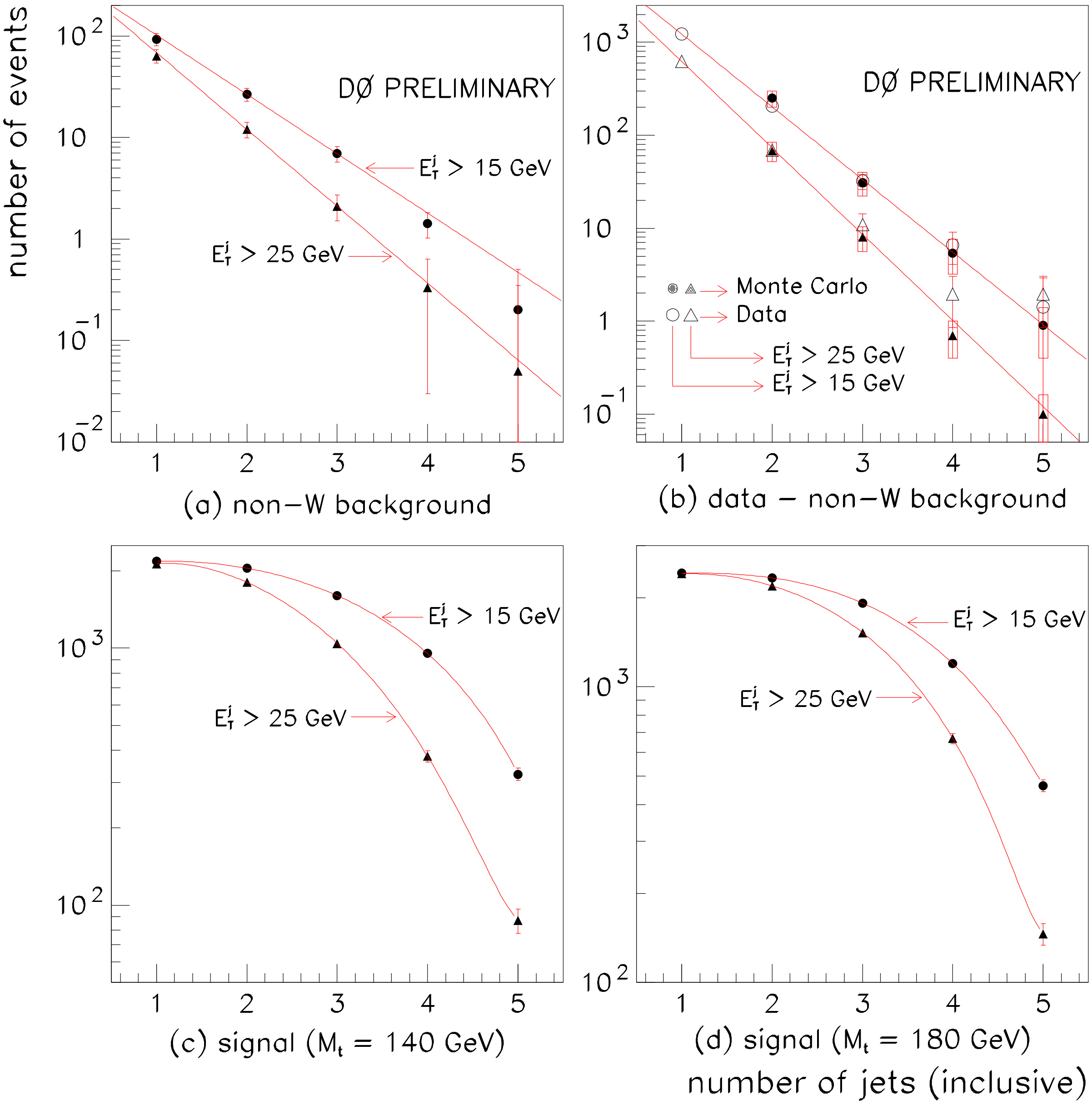


Figure 6.1: The jet multiplicity spectra for our backgrounds, signal and data: (a) non-W background; (b) our data after the subtraction of non-W background; (c) signal ($M_t = 140$ GeV) and (d) signal ($M_t = 180$ GeV).

m	N_m			
	data	non-W bkg.	W bkg.	$t\bar{t}$
1	1316	92.4 ± 12.8	1223.6 ± 74.1	6.2 ± 4.9
2	233	26.5 ± 3.8	189.5 ± 11.0	6.0 ± 4.8
3	39	6.9 ± 1.2	28.3 ± 4.0	4.8 ± 3.8
4	8	1.4 ± 0.4	4.2 ± 2.8	2.9 ± 2.8

Table 6.1: The inclusive jet multiplicities in $e + \not\!\!E_T$ events: data, estimated non-W background and the results of fitting for the W background and signal ($M_t = 160$ GeV)

expected to be negligible for events containing real W's. The slight apparent excess in data over the fit at the 4-jet point is not statistically significant, as we shall soon see. The large open markers in Figure 6.1(b) represent our estimate as derived here while the small closed markers represent the estimate obtained from the VECBOS Monte Carlo. The error bars apply to data and the error boxes to the Monte Carlo. We do not use the Monte Carlo distribution in this section of our analysis, yet we plot it here to strengthen the case for using it in the next section (Method B). We see that the Monte Carlo agrees very well with our data for either threshold. Note that no relative normalization was done between the Monte Carlo and data. This shows that the cross-sections predicted by VECBOS are quite accurate, although we do not use them anywhere in our analysis. Both Figures 6.1(a) and 6.1(b) correspond to $\int \mathcal{L}dt = 13.5\ \mathrm{pb}^{-1}$.

6.2.2 Estimation of signal in data

Since the rest of our analysis will be limited to the gg quadrant of the last subsection, we simplify our notations a bit by defining N_m as N_1^W for the sample S_m. Let N_{mt} stand for the contribution of $t\bar{t}$ events to N_m. Then, if we assume that Berend's scaling rule is valid for the $W + jets$ component of our sample, Equation 6.6 gives

$$N_m - N_{mt} = \frac{(N_{m-1} - \alpha_{(m-1),m} \cdot N_{mt})^2}{N_{m-2} - \alpha_{(m-2),m} \cdot N_{mt}} \tag{6.12}$$

where $\alpha_{k,l} \equiv \frac{N_{kt}}{N_{lt}}$ are calculated using the ISAJET $t\bar{t}$ Monte Carlo. Figures 6.1(c) and 6.1(d) show the distribution of N_{mt} as predicted by the ISAJET Monte Carlo for $M_t = 140$ GeV and $M_t = 180$ GeV respectively. The numbers are obtained using a sample of 30000 $t\bar{t}$ events in each case where all decay channels were allowed. Therefore, the Figures correspond respectively to about 175 and 700 times our data sample according to the cross-section curve shown in Figure 1.3. The fits shown are obtained using a 4th degree polynomial function. For $M_t = 160$ GeV, the relevant ratios are $\alpha_{2,4} = 2.076$, $\alpha_{3,4} = 1.664$.

Rephrasing Equation 6.12 for $m = 4$ we get

$$(N_2 + \alpha_{2,4} \cdot N_4 - 2\alpha_{3,4} \cdot N_3) \cdot N_{4t} + (\alpha_{3,4}^2 - \alpha_{2,4}) \cdot N_{4t}^2 = N_2 \cdot N_4 - N_3^2 \tag{6.13}$$

The term containing N_{4t}^2 can be safely ignored to get a simple solution for N_{4t}:

$$N_{4t} = \frac{N_2 \cdot N_4 - N_3^2}{N_2 + \alpha_{2,4} \cdot N_4 - 2\alpha_{3,4} \cdot N_3} \tag{6.14}$$

The third and the fourth columns of Table 6.1 summarizes the results thus obtained. Only the statistical errors are shown. We shall take the systematic errors into consideration when we convert the signal into an estimate for cross-section in the next chapter. Note that in general, the central values in the last three columns do not add up to the number in the first column since after extracting the signal from the estimated total of "real W events", the remainder is re-fitted using Equation 6.6 to estimate the W background (which, after all, is the key conjecture in this method). So, this method gives us

$$N_{4t}^{A} = 2.9 \pm 2.8 \qquad (6.15)$$

where the superscript is used as a tag for the method. i.e. we find 8 events where we expect 5.1 ± 2.8 from background processes.

6.3 Shape and energy-scale of W candidate events (Method B)

In the correlations between the momenta of the objects in an event, there are features that can be of use in differentiating signal and background. Much of such features remain unexploited in the technique of jet counting done in the last section. To this end, we seek two variables with minimal inter-correlation: one to measure the total amount of energy in the event ("energy-scale" of the event) and another to quantify its distribution among the individual objects as well as the spatial distribution of the objects ("shape" of the event). Due to the high mass of top, we expect enhancement of signal over background at a large energy-scale. Due to the masses of the intermediate resonances involved in the cascade decay, the total energy of the signal should be more equitably distributed among the individual jets compared to the background. For the spatial distribution, we would like to be able to separate a collinear topology from a non-collinear one. The physical motivation to seek a variable which quantifies the non-collinearity of jets in an event comes from the following observation. The background to our signal consists of QCD multijet events with or without a real W. In either case, the jets are produced primarily by quark bremsstrahlung or gluon-splitting. On an average, the jets in such events tend to have smaller p_T relative to each other than in the $t\bar{t}$ events where they are produced by cascade decays of high-mass resonance states. Thus, the jets in the background events can be expected to have smaller separation between them than those in the signal events.

Several variables have been developed and used by various collider experiments to describe the shape of an event in the past. The common denominator of such variables is that they depend only on the relative momenta of the objects and not at all on the overall energy-scale of an event. In particular, such variables were found to be very useful in studying the correlations between particles in the e^+e^- experiments. Even though in $p\bar{p}$ experiments such as ours, objects often consist of jets representing partons rather than particles, our studies reveal that several of these variables offer considerable discrimination between signal and background in the monoleptonic decay channels of $t\bar{t}$. We employ only one of these in this part of our analysis [1].

---

[1] Another promising variable is discussed in Reference [104]

6.3.1 The shape of an event: Aplanarity ($\mathcal{A}$)

For each event we define the *normalized momentum tensor* (M) [105]:

$$M_{ab} = \frac{\sum_i p_{ia} p_{ib}}{\sum_i p_i^2} \tag{6.16}$$

where $\vec{p_i}$ is the 3-momentum of the i^{th} object in the laboratory frame and a, b run over the three space directions. M is a symmetric matrix which can be diagonalized. We compute the eigenvalues Q_j and order them:

$$Q_1 \leq Q_2 \leq Q_3 \tag{6.17}$$

The following relations hold:

$$Q_1 + Q_2 + Q_3 = 1 \tag{6.18}$$

$$Q_1 \geq 0 \tag{6.19}$$

These eigenvalues can be subsequently used to quantify the shape of the event:

- for roughly spherical (ball-shaped) events, $Q_1 \approx Q_2 \approx Q_3$
- for coplanar (pancake-shaped) events, $Q_1 \ll Q_2$
- for collinear (cigar-shaped) events, $Q_2 \ll Q_3$

Particular linear combinations of the Q_j have been given suitable names. The *aplanarity ($\mathcal{A}$)* defined as

$$\mathcal{A} \equiv \frac{3}{2} Q_1 \tag{6.20}$$

is normalized to lie in the range $0 \leq \mathcal{A} \leq 0.5$. Clearly, the higher the minimum one requires of $\mathcal{A}$, the more spherical the event has to be in order to meet that requirement.

Analytically, by diagonalizing the momentum tensor and ordering its eigenvalues, we have picked the direction (given by the eigenvector $\hat{n}_1$ corresponding to Q_1) along which the the net second moment of the momenta of the objects, which can be interpreted as a measure of "activity" in an event, is the minimum. Q_1 thus quantifies the activity normal to the plane of maximum activity. In a highly collinear or coplanar event, most of the activity will be limited in the proximity of the plane normal to $\hat{n}_1$. Therefore, we can reject such events by requiring our candidates to have Q_1, and hence $\mathcal{A}$, above some threshold. In this analysis, we require the events to have at least 4 jets and calculate $\mathcal{A}$ using only the jets as measured in the laboratory frame (i.e. the electron and $\not{E}_T$ are not used). Figure 6.2 shows the distribution of $\mathcal{A}$ in our backgrounds and signal. The fake background is obtained using a sample which has electromagnetic objects that fail our quality criteria and have small $\not{E}_T$ (i.e. events with at least 4 jets in the quadrant 3 as described in Subsection 6.2.1) No relative normalization has been done. The dotted line shows the cut to be applied later. $\mathcal{A}$ is independent of the overall scale of the event but its distribution for signal depends on M_t. As a function of M_t, the shape of a $t\bar{t}$ event is expected to be most spherical when the momenta of the final state partons are likely to be relatively close in absolute value and have opening angles neither too small, nor too large. This situation is favored most in the region of $1.5M_W < M_t < 2M_W$. For $M_t < \sim 100$ GeV, a b-jet is likely to very

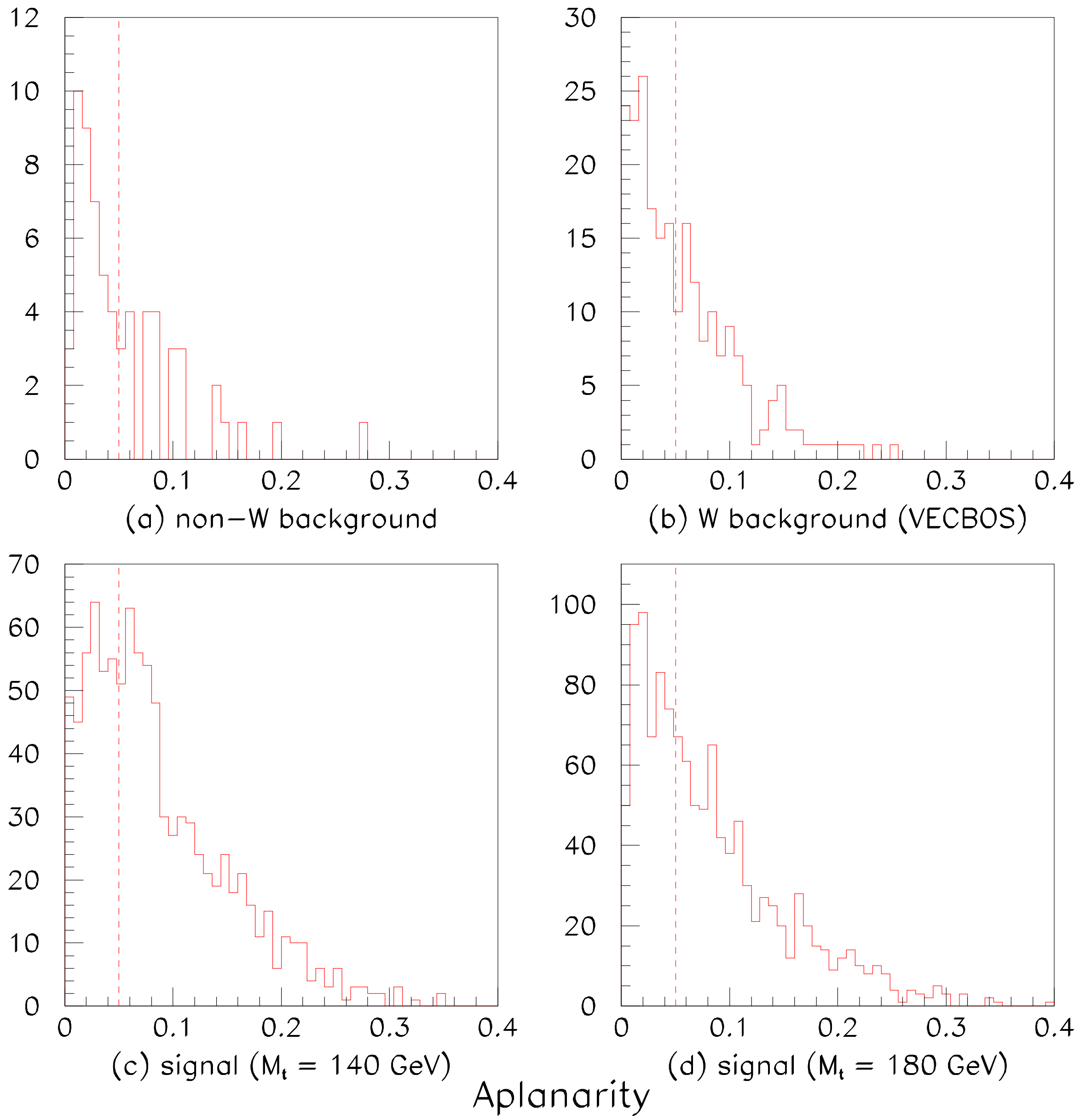


Figure 6.2: Aplanarity ($\mathcal{A}$) in our backgrounds and signal: (a) non-W background; (b) W background (VECBOS); (c) signal (ISAJET, M_t = 140 GeV) and (d) signal (ISAJET, M_t = 180 GeV).

weak and have a small opening angle with respect to its sister W. At the other extreme, for $M_t >\sim 200$ GeV, a b-jet is likely to be back-to-back with its sister W (again resulting in a small relative p_T) and the opening angle between jets from the W-decay is also likely to be small owing to the large p_T of the parent W. In either case, one would expect the events to be less spherical than when M_t is in the intermediate range. Indeed, the variable turns out to be most useful for ~ 100 GeV $< M_t < \sim 200$ GeV, precisely the region of M_t where our current interest is focused.

6.3.2 The size of an event: H_T

To quantify the size of an event, we use the scalar sum of the E_T of all the jets in an event:

$$H_T \equiv \sum_{j=1}^{m} E_T(j) \tag{6.21}$$

where j is the jet index. Figure 6.3 shows the distribution of H_T in our backgrounds and signal. No relative normalization has been done. The dotted line shows the cut to be applied later.

6.3.3 Analysis using $\mathcal{A}$ and H_T

Here, we require the events to have at least 4 jets. First, we test the validity of our expectation on Monte Carlo simulation of our signal and the W background. Figure 6.4 shows the distribution of the backgrounds, signal and data in the $\mathcal{A}$ *vs.* H_T plane. No relative normalization has been done. The dotted lines show the division of the plane into 4 quadrants. The W back-

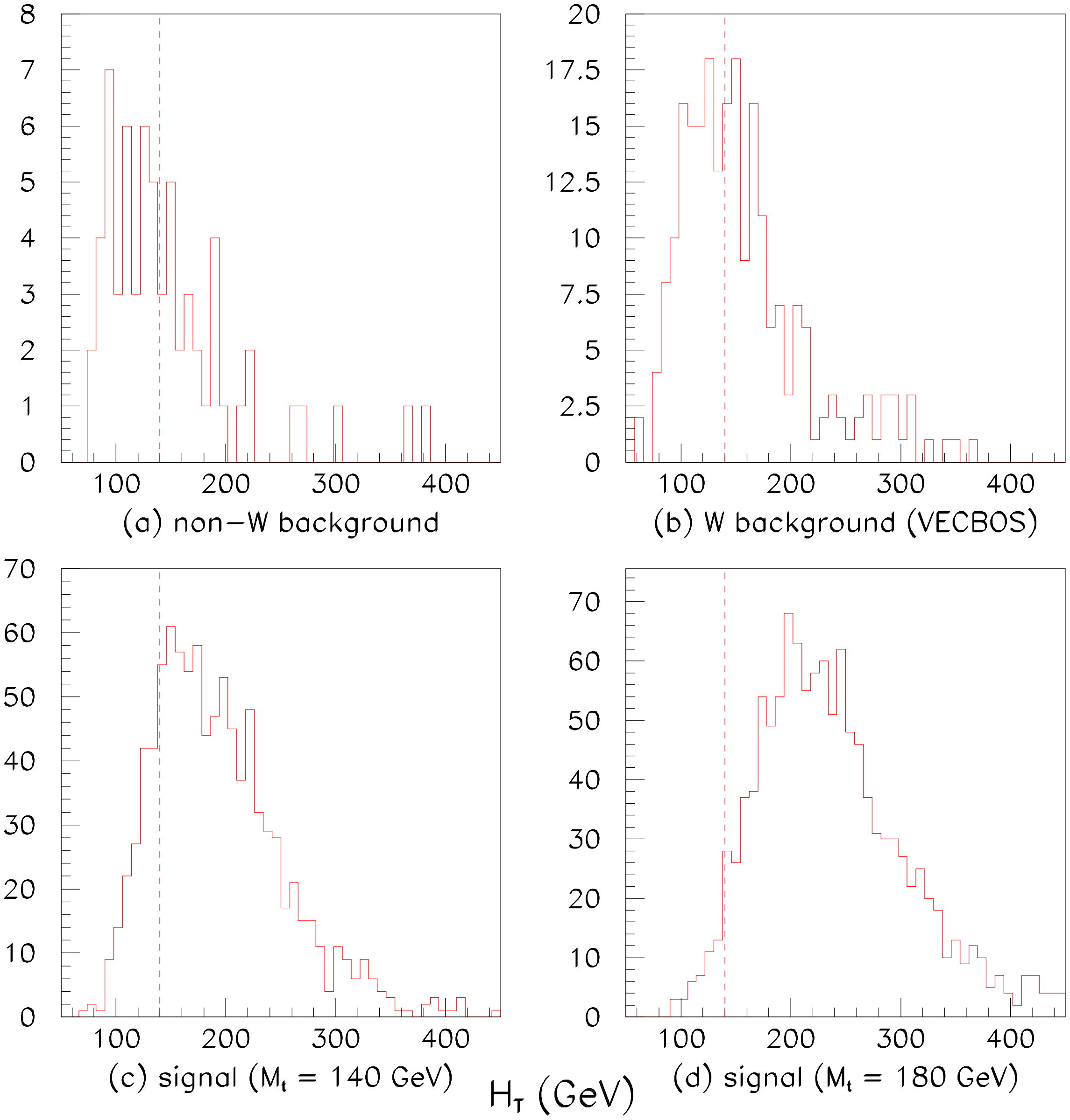


Figure 6.3: H_T in our backgrounds and signal: (a) non-W background; (b) W background (VECBOS); (c) signal (ISAJET, $M_t = 140$ GeV) and (d) signal (ISAJET, $M_t = 180$ GeV).

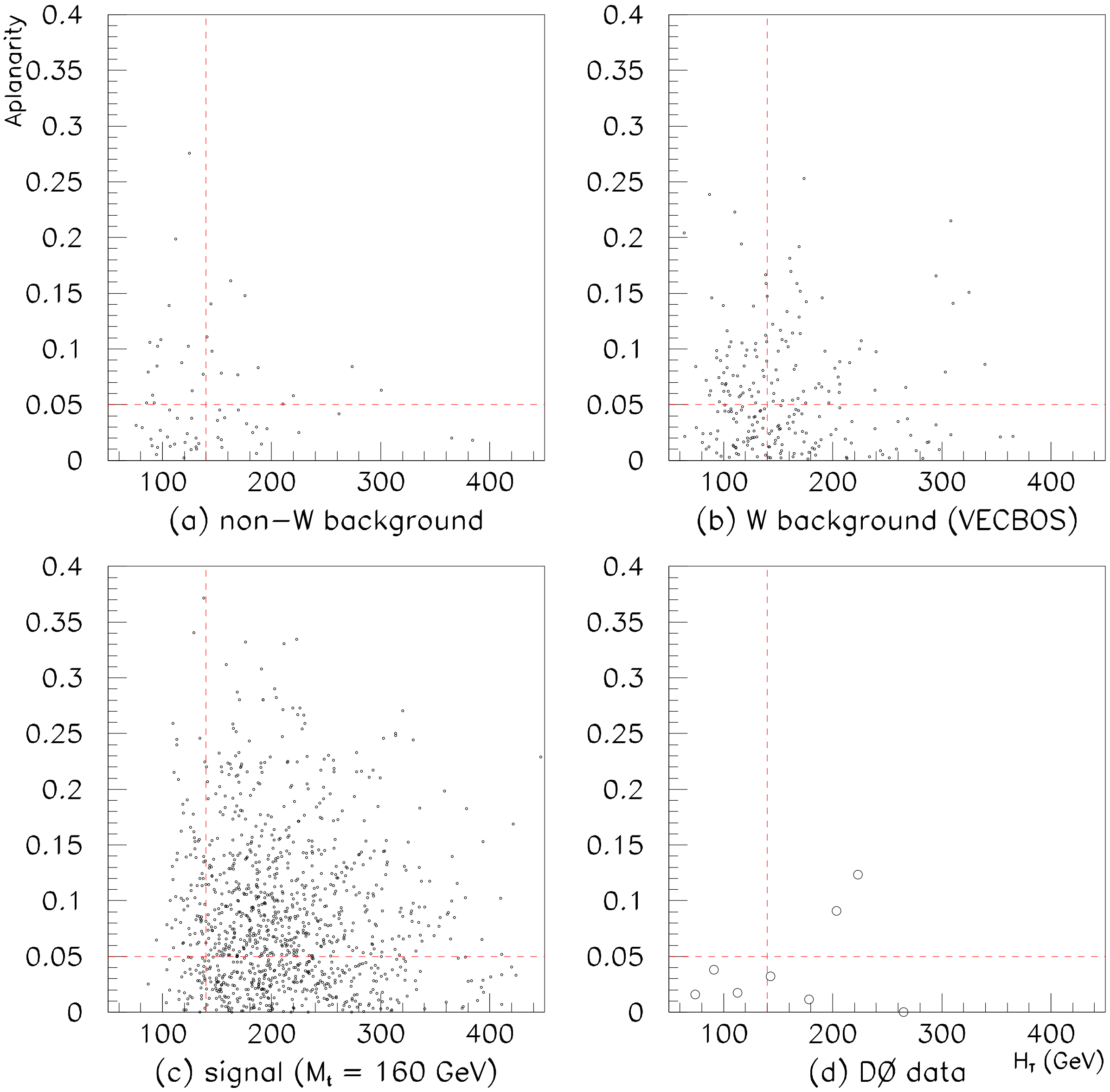
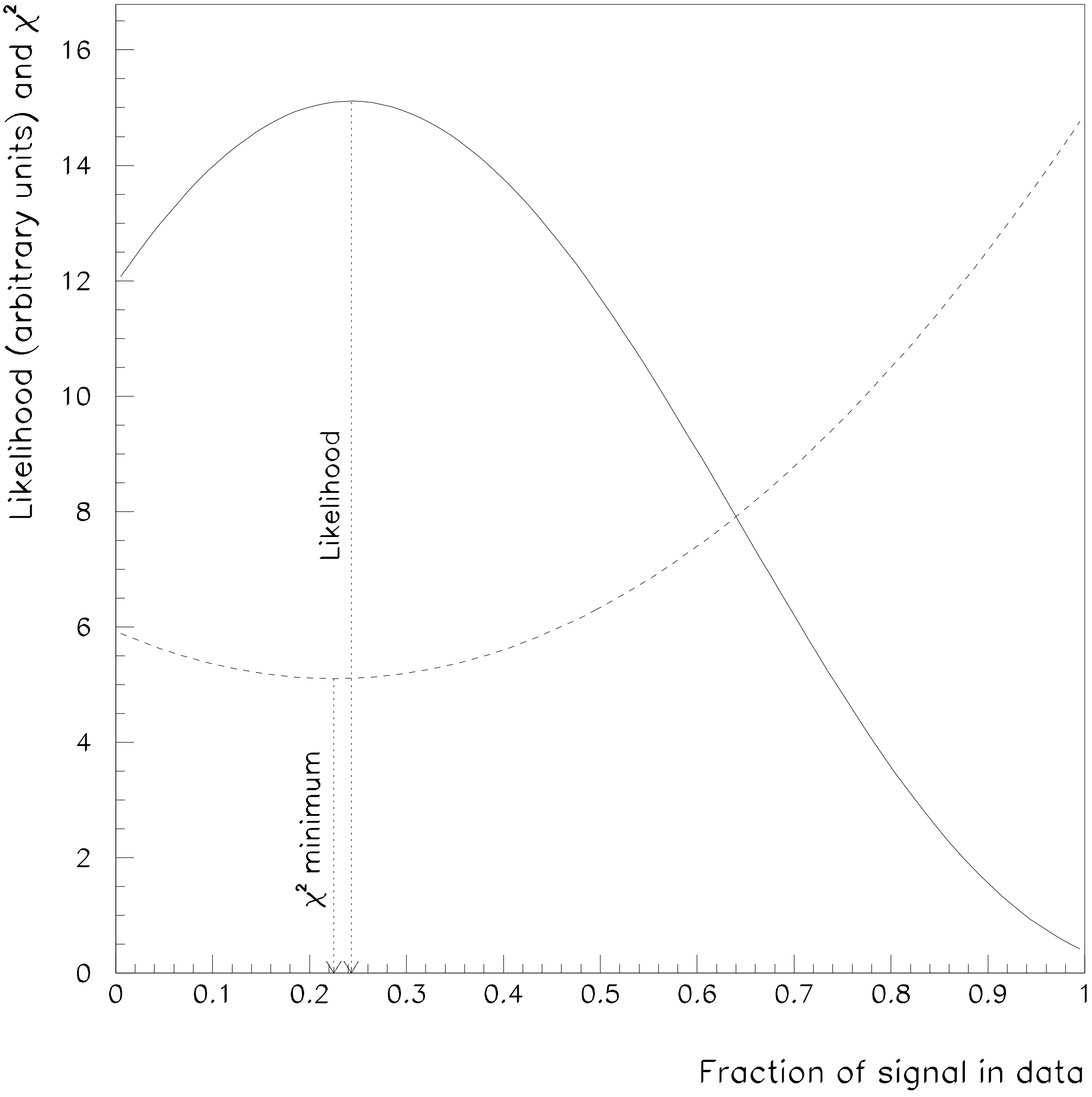


Figure 6.4: $\mathcal{A}$ *vs.* H_T in our backgrounds and signal: (a) non-W background; (b) W background (VECBOS); (c) signal (ISAJET, $M_t = 160$ GeV) and (d) data.

ground is modelled by the VECBOS Monte Carlo and non-W background by our data as described earlier. For the purpose of calculating $\mathcal{A}$ and H_T in the non-W sample, we take events with $\not\!\!E_T < 15$ GeV from the ELE_HIGH and ELE_JET triggers as described in Subsection 6.2.1. We treat any electromagnetic jet (a jet with EM fraction exceeding 0.9) failing our electron-id cuts as an electron, but only one at a time. That is to say, if an event has k such EM jets, it enters our calculation k times, each time a different jet being substituted for the electron.

First, with two variables ($\mathcal{A}$ and H_T) at hand, we can move the origin on the plane described by them (with no rotation of the axes) so that either background is split more or less evenly among the 4 quadrants. This puts the origin at $H_{T_0} = 140$ GeV, $\mathcal{A}_0 = 0.05$. The choice of origin is prompted by two considerations. Firstly, the pattern of split in the two types of backgrounds being nearly identical, the relative weight of them in data becomes inconsequential. Secondly, the uniform split of backgrounds among the four quadrants minimizes the errors in the fitting procedure described below. Table 6.2 summarizes the split of signal, background and data in the four quadrants.

Let f be the fraction of signal in our data. A maximum likelihood (ML) fit to the fractions in the four quadrants with respect to f then gives us an estimate for the value of f for which the mixture of simulated signal and backgrounds comes closest to our data [94]. In view of the small statistics, we perform a cross-check by the method of least squares (LS). Figure 6.5 shows the likelihood and the χ^2 of least squares as functions of the fraction of signal in our data. for our data, the two estimators give close results: $f_{_{ML}} = 0.243$


Likelihood (arbitrary units) and χ^2
Fraction of signal in data
Likelihood
χ^2 minimum
0
2
4
6
8
10
12
14
16
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1