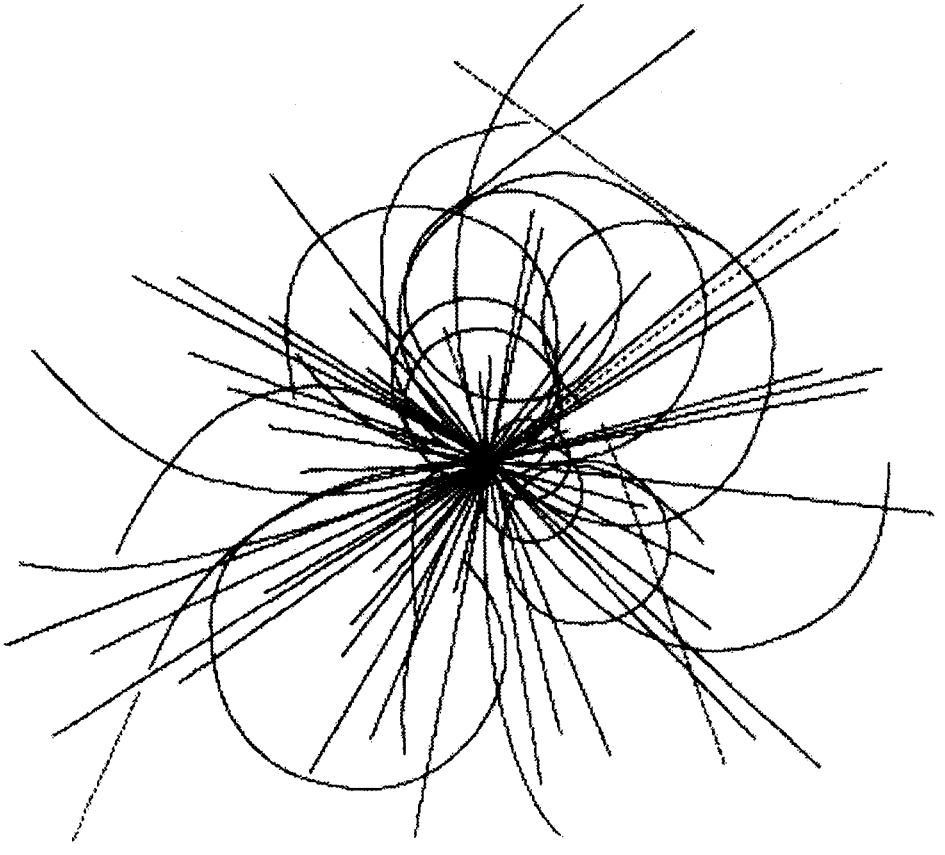
SSCL-Preprint-480

SSCL-Preprint-480
June 1993
Distribution Category: 414

M. Cobal
H. Grassmann
S. Leone

Exploiting the Single-Lepton Event Structure in the Search for the Top Quark



Superconducting Super Collider Laboratory

SSCL-Preprint-480

Exploiting the Single-Lepton Event Structure in the Search for the Top Quark*

M. Cobal and S. Leone

Univ. and INFN Pisa

and

H. Grassmann

Superconducting Super Collider Laboratory†
2550 Beckleymeade Avenue
Dallas, TX 75237

June 1993

* Submitted to *Il Nuovo Cimento*, May 1993.

† Operated by the Universities Research Association, Inc., for the U.S. Department of Energy under Contract No. DE-AC35-89ER40486.

Contents

1 Introduction **1**

2 Monte Carlo **2**
2.1 Generators 2
2.2 Normalization 3

3 Event Structure **4**
3.1 Angular Correlations between Proton Beam and Hadronic Activity 4
3.2 Reconstruction of $\cos\theta^*$ 4
3.3 Jet Multiplicities 5
3.4 Separation of Signal from Background 5

4 Estimate of the Systematic Errors **7**
4.1 Problems from the experimental jet definition 7
4.2 Higher Jet Multiplicities 8

5 Comments on Observability of a Top Signal **9**
5.1 Additional Variables as Checks 9
5.2 Estimate of the absolute expectation 10

6 Outlook **11**

7 Conclusion **11**

8 Acknowledgment **12**

Exploiting the Single-Lepton Event Structure in the Search for the Top Quark

M. Cobal[(1)], H. Grassmann[(2)], S. Leone[(1)]

[(1)] *Univ. and INFN Pisa,* [(2)] *SSCL*

Abstract

We discuss whether the event structure can be useful in searching for the top quark in the single-lepton channel at the Tevatron, where the main obstacle is a considerable background from QCD W+ multi-jet production. We restrict our discussion to the case of a Standard Model top quark of heavy mass, M_{top}>100 GeV.

1 Introduction

The most stringent lower limit on the mass of the top quark (hereinafter referred to as "top") comes from the Tevatron $p\bar{p}$ collider [1]. It is based on a search for events with two leptons. There was also a search for the top in single-lepton events, based on a study of the transverse mass of the charged lepton and the neutrino [2]. The limit from the di-lepton channel is 91 GeV and from the single lepton channel, 77 GeV.

The main background in the single-lepton channel is from QCD W+jet events [3]. However, the event structure of QCD W+jet events and W+jet events from $t\bar{t}$ decays should be different, because the matrix elements of these processes are different. Consequently, a large fraction of QCD jets tends to be emitted at small angles to the beams, while jets that are prongs of a heavy top decay tend to be more central. In this paper we investigate whether these differences are significant enough to allow a top search in the single-lepton channel. We will consider events with a W decaying leptonically (into an electron or a muon) and at least three jets.

A number of papers [4] have discussed how to reconstruct heavy $t\bar{t}$ decays, and have included the case of one-leptonic $t \rightarrow Wb \rightarrow l\nu b$ decay. However, no experimental paper has been based on this approach jet. We believe that one good reason for this is that a major problem, besides handling the signal, is overcoming the large and hardly controllable background.

Also, the above methods [4] are rather complex. The complexity of the procedures makes difficult an independent check of the results, which may not be reproducible.

Our approach is as follows:

- We emphasize background rejection without trying to reconstruct decay of the top.
- We aim at the highest possible simplicity [5], in the hope that the simplicity of the analysis will make it possible to study the influence of systematic errors quantitatively.

Of course, were a top signal established in a reliable way, one would profitably apply more complex analysis on the details of the signal, for example, with the techniques developed in Reference [4].

2 Monte Carlo

2.1 Generators

We made use of publicly available Monte Carlo programs, PAPAGENO for the top, and the new VECBOS program for W+multi-jet production. PAPAGENO was written by I. Hinchliff [6]; VECBOS was written by Berends et al. [7] and is now maintained and has been further improved by W. Giele [7],[8]. We use these programs because they are based on matrix elements and because they are sufficiently tested, having already been used in various studies [3],[9],[10]. PAPAGENO uses the correct matrix elements for lowest-order $t\bar{t}$ production, and VECBOS uses the correct lowest-order matrix elements for W+n jet production, $n = 1, 2, 3, 4$.

The W from QCD W+jets production decays leptonically into an electron or muon. One of the two Ws from $t\bar{t}$ production decays leptonically, the other hadronically.

We simulated Tevatron collider events with proton and antiproton beam energies of 900 GeV. We did not attempt to model any specific existing detector, but as in the case of the CDF, we required P_t (charged lepton)> 20 GeV, P_t (neutrino)> 20 GeV, and the pseudorapidity of the electron less than 1.2. In addition, we required transverse mass M_t (charged lepton + neutrino)> 40 GeV.[1] We decay the b quarks using a simple fragmentation routine with the appropriate branching ratios but without second-generation decays. We assume that neutrinos and muons escape the detector without depositing energy, while the remaining hadronic partons are assumed to show up in the detector as jets (and will often be called "jets" in the following). The spatial separation between jets in rapidity and phi, ΔR, is $\Delta R > 0.6$. Also, we require P_t (parton)> 10 GeV and |rapidity (parton)| < 3.5. These cuts avoid coming close to singularities in the QCD matrix elements.

We produced 10,000 PAPAGENO top events with M_{top} =100 GeV, 100,000 VECBOS W+3 jet events, and 50,000 W+4 jet events. This large number of VECBOS events was needed because in the version available to us at the time of this analysis, the final state phase space was populated randomly by VECBOS. The resulting weight distribution is broad and imposes the need to generate a large number of events in order to arrive at reasonable statistical fluctuations. New and improved sampling methods are presently being worked on [8].

2.2 Normalization

Our main goal is to study the event structure at the parton level and to estimate relative efficiency when comparing QCD W+jet and top W+jet events, taking into account systematic errors. We do not take into account energy measurement resolution, lepton identification, or trigger efficiencies, which may differ for different detectors. Therefore, we will not be able to predict absolute numbers of events for a certain integrated luminosity. For W+jet events, such a prediction would suffer from considerable systematic uncertainties due to the restriction to lowest-order matrix elements.

Only in Section 5 will we attempt a rough comparison between CDF data and a prediction from a top Monte Carlo.

[1] P_t is the momentum component perpendicular to the beam axis; the transverse mass is the invariant mass of the P_t vectors.

3 Event Structure

3.1 Angular Correlations between Proton Beam and Hadronic Activity

The most obvious difference that comes to mind between W+n jet events from QCD or from top production is the presence of the W mass peak in the di-jet invariant mass. However, it has been shown that this method by itself is not sufficiently powerful, given the small signal/background ratio and the experimental uncertainties [11].

An additional difference between top and QCD W+jet events should be that many jets in QCD W+jet events come from "initial state" gluon radiation or from scattering of initial state particles. Those jets will mostly be emitted in the forward direction, close to the beams. For various types of QCD processes this behavior has indeed been observed [12]. In contrast, the decay products of centrally produced heavy particles are expected to be less correlated with the beam direction. This difference is important because it is general in nature and is suited to distinguish between QCD-dominated background processes and decays of heavy particles of any kind. This is in no way a new idea; it has been used in a number of studies, for example, in the search for leptoquarks [13].

It was shown in these studies that a suitable parameter to monitor these effects is $\cos\theta^\star$, where $\theta^\star$ is the angle between the outgoing particle and incoming proton beam in the rest system of the event.

3.2 Reconstruction of $\cos\theta^\star$

We take as observables in an event the muon or the electron from the leptonic decay of the W, the transverse component of the neutrino, and the three jets highest in E_t(jet). The longitudinal component of the neutrino momentum cannot be reconstructed in an experiment. We compared the jet angular distributions in top and QCD events in three different rest frames:

1) the center of mass system, which is known in Monte Carlo events
2) the c.m.s. without the longitudinal neutrino momentum
3) the lab system.

We found that the analysis described in the following sections can be done efficiently in each of these systems. The signal/background separation is best in system (1), worse in (2), and worst in (3). But the differences are small. We will work in system (2) rather than in (3), because that seems to be a more common choice. (Compare Refs. [12],[13].)

3.3 Jet Multiplicities

In a real experiment top events may show up in samples with different jet multiplicities. There are several possible strategies to analyze them. For example, one might divide the QCD and top samples in various subsamples with precisely n jets, $n = 1, 2, 3 \ldots$, and examine each of them separately. We did not choose that approach because the jet multiplicity depends on the jet definition and reconstruction algorithms, as well as on the top mass.

Rather, we choose to discuss W events with three or more jets and proceed in the following way:

For top: The sample will contain a certain fraction of W+4 jet events. Using the information carried by the three jets highest in E_t(jet), we will not distinguish between events that have or do not have an additional fourth jet, and can study all events in an unified fashion.

For QCD events: We first investigate W+3 jet events, which are expected to dominate the W+n jet sample, $n \geq 3$. Next, we investigate by how much the additional amplitude for W+4 jet (partons) events can change the result obtained on W+3 jet events. We found that the picture does not change appreciably if this production amplitude is taken into account. This study was not continued further, because W+5 jet matrix element Monte Carlos are not available.

3.4 Separation of Signal from Background

We order the jets in E_t(jet), jet_1 being the highest one in E_t. We require $E_t(\text{jet}_3) > 15$ GeV because above an energy of this order the experimental jet reconstruction is fairly reliable in an actual experiment. As mentioned in Section 3.3, in the first instance we did not consider possible additional jets. The effects from additional jets will be discussed later.

We begin by investigating the two leading jets. Figure 1 shows the event distribution as a function of $\cos\theta^{max}$, which is the maximum value of $|\cos\theta^{\star}(\mathrm{jet}_1)|$ and $|\cos\theta^{\star}(\mathrm{jet}_2)|$, for both QCD W+3 jet events and top W+3 jet events ($M_{top} = 100$ GeV/c^2). The vertical scale is in arbitrary units. The relative normalization between QCD and top events will be explained below. We see that there is already a clear difference in shape between top and QCD events. As expected, the jets in QCD events tend to be emitted more in the forward direction.

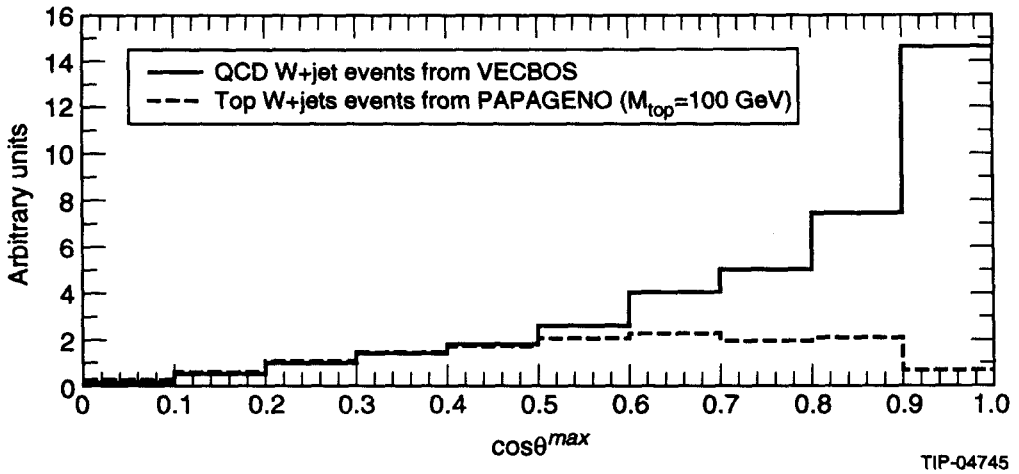


Figure 1: $\cos\theta^{max}$, which is the maximum of $|\cos\theta^{\star}(\mathrm{jet}_1)|$ and $|\cos\theta^{\star}(\mathrm{jet}_2)|$.

To continue with our analysis we apply a cut at $\cos\theta^{max}<0.8$. The signal/background ratio improves by a factor of 1.85. (For the W+3 jet sample the statistical errors on the numbers we quote are less than 10%.)

For the events that pass the $\cos\theta^{max}<0.8$ cut, we plot $|\cos\theta^{\star}(\mathrm{jet}_3)|$ in Figure 2. Again top events are more central. A cut $|\cos\theta^{\star}(\mathrm{jet}_3)| < 0.8$ further improves signal/background by an additional factor of 1.7. In Figures 1 and 2 the relative normalization between top and QCD events is done such that in Figure 2 at $|\cos\theta^{\star}(\mathrm{jet}_3)| < 0.8$ the number of top and QCD events is the same. After these cuts in $\cos\theta^{\star}$ we have reached a suppression of the QCD background by almost a factor of 4, while losing only 30% of the signal. The relative efficiencies of the cuts are summarized in Table 1.

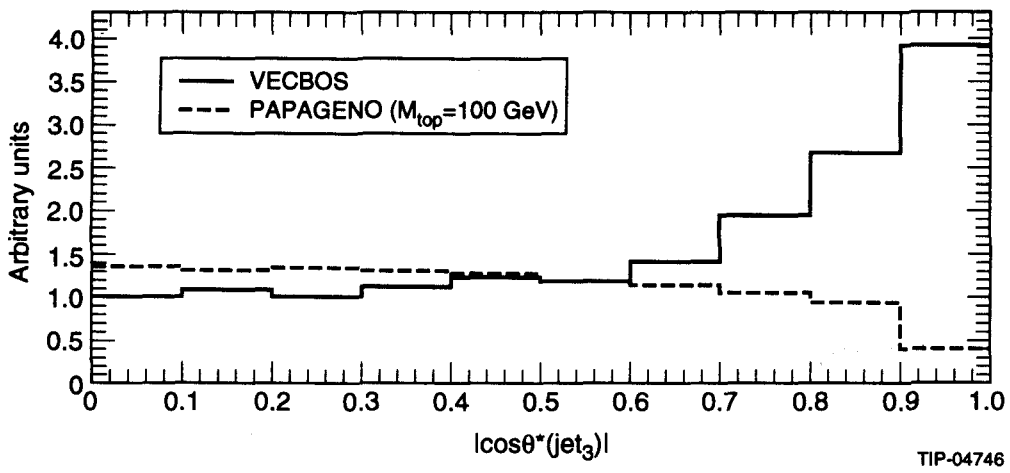


Figure 2: Event distribution versus $|\cos\theta^\star(\mathrm{jet}_3)|$. A cut $\cos\theta^{max}<0.8$ has been applied. VECBOS and PAPAGENO are normalized to each other in the region $|\cos\theta^\star(\mathrm{jet}_3)|<0.8$.

cut	QCD	top
$\cos\theta^{max}<0.8$	2.3	1.3
$\cos\theta^\star(\mathrm{jet}_3)<0.8$	1.7	1.1
total	3.9	1.4

Table 1 : Event reduction factors caused by $\cos\theta^\star$ cuts.

4 Estimate of the Systematic Errors

4.1 Problems from the experimental jet definition

$E_t(\mathrm{jet}_3)$ is a rather steeply falling spectrum for QCD W+jet events. Because $E_t(\mathrm{jet})$ and $\cos\theta^\star(\mathrm{jet})$ are correlated, the experimental jet energy resolution can have large effects and can modify the number of events accepted within given cuts. Due to the finite energy resolution of any experiment, a parton with $P_t(\mathrm{parton}) < 15$ GeV can be reconstructed as a jet with $E_t(\mathrm{jet}) > 15$ GeV.

We studied this problem by requiring again $E_t(\text{jet}_1, \text{jet}_2) > 15$ GeV, but changing the request on jet_3 to 10 GeV$<E_t(\text{jet}_3)<15$ GeV (instead of $E_t(\text{jet}_3)>15$ GeV). We find that this sample with low $E_t(\text{jet}_3)$ is reduced under the cut $|\cos\theta^\star(\text{all jets})|<0.8$ by a factor of 4.5. This compares favorably to the reduction factor of 3.9, which we got for events with $E_t(\text{jet}_3)>15$. The rejection power of the $\cos\theta^\star$ cuts in an actual experiment is thus larger than that derived from a parton level calculation.

For top events the effect from jet smearing will be small, because the $E_t(\text{jet}_3)$ distribution is harder for top events ($M_{top}>100$ GeV) than for QCD events.

4.2 Higher Jet Multiplicities

The existence of W+n jet events, $n \geq 4$, must be considered for two reasons.

First, we cannot restrict ourselves to an inclusive W+3 jet analysis, excluding events with higher jet multiplicities. Especially at higher top masses we would lose a large fraction of the signal.[2]

Second, W+n jet events would show up as W+(n–1) jet events in the experimental data samples if one of the jets is not reconstructed. We limited our investigation of this problem to W+4 jet events, because there is no W+5 jet matrix element Monte Carlo presently available.

We applied the $\cos\theta^\star$ cuts, as described in Section 3.3, on a sample of W+4 jet events. Again we used only the three leading jets; jet_4 is ignored. The $\cos\theta^\star$ cuts now result in a reduction factor of 3.7, with a 25% statistical uncertainty.[3] It follows that jet reconstruction problems and higher QCD jet multiplicities will not change the reduction factors given in Table 1 by very much. Of course, the precise numbers would have to be determined for an individual experiment using the appropriate detector simulation.

[2]One way to treat also events from a very high mass top as W events with precisely three jets would be to require very high E_t(jet) values. We consider this an interesting alternative that we will not address further in this paper.

[3]The production of W+4 jet events needs a great deal of CPU time; in addition, the event weight distribution gets broader. We have not yet been able to produce a sample large enough to ignore statistical uncertainties.

In summary: a matrix element Monte Carlo calculation predicts that by applying cuts in $\cos\theta^\star$ one can suppress QCD background events relative to top events at M_{top}=100 GeV by a factor of approximately 2.8. Were the top mass larger than this, the improvement must be expected to become larger as well.

5 Comments on Observability of a Top Signal

5.1 Additional Variables as Checks

One might begin a top search in an experimental W+jet data sample by just comparing the number of events with jets in the central and forward region, as discussed in Section 4. But the signal/background ratio would still be too high. In addition, since we have only lowest-order matrix elements available, significant systematic uncertainties are unavoidable in an estimate of this ratio.

However, once $\cos\theta^\star$ cuts have reduced the background, other variables can be used in addition to identify the signal. A strategy could be to subdivide the event sample into a background-enriched subsample and a signal-enriched subsample. For example, one might use the events at $\cos\theta^{max}<0.8$ and define as a signal-enriched sample the events with $|\cos\theta(\text{jet}_3)|< 0.8$, and as a background-enriched sample the events $|\cos\theta(\text{jet}_3)|> 0.8$. There will be variables that are distributed in a similar way for QCD events in both the signal- and background-enriched samples, while they would be different for top events. As an example, in Figure 3 we show (arbitrary normalization) the distribution in $M(\text{jet}_1+\text{jet}_2)$ for QCD W+3 jet events with the two leading jets at $|\cos\theta^\star(\text{jet})|< 0.8$. The solid line is from events with $|\cos\theta^\star(\text{jet}_3)| > 0.8$, and the dotted line is from $|\cos\theta^\star(\text{jet}_3)| < 0.8$. One observes that they are very similar.

One could use the background-enriched sample (solid line, Figure 3) to check the detector performance; if the detector is understood, one would know what di-jet mass spectrum to expect for the signal-enriched sample. Top events might show up as anomalies in this spectrum. A bump in $M(\text{jet}_1+\text{jet}_2)$ might be expected at about 80 GeV for moderate top masses. For higher top masses the leading jets would come more frequently from the

beauty quarks, and one may use different variables like $E_t(\text{jet}_{1,2,3})$. For very high masses ($\geq$ 140 GeV), one could exploit the distribution of the fourth jet.

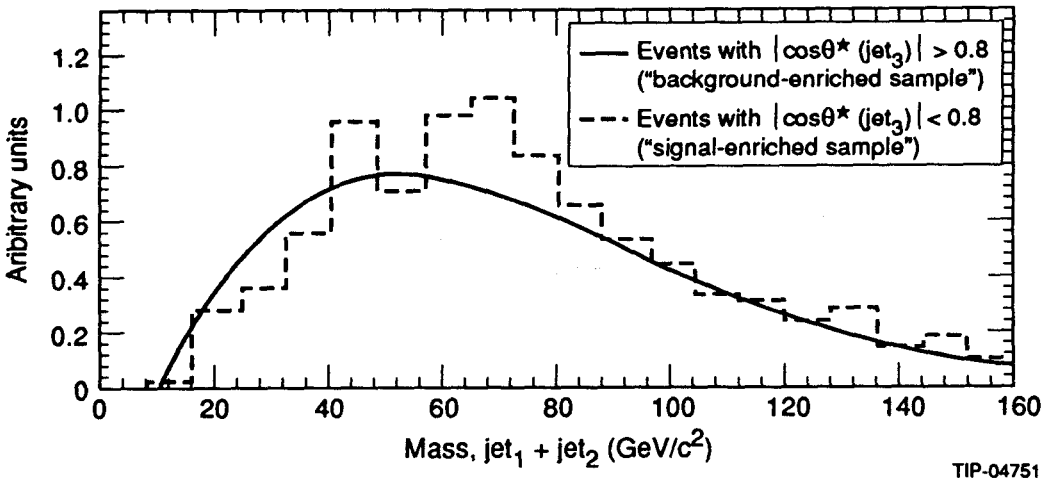


Figure 3: Invariant mass of the two leading jets, QCD W+jets events from VECBOS. A cut $\cos\theta^{max}$<0.8 has been applied.

5.2 Estimate of the absolute expectation

In an attempt to get a rough estimate of the expected absolute signal/ background ratio, we can make use of published data from CDF at the Tevatron collider.

In Reference [9] a sample of W+3 jet events from the Tevatron collider with loose cuts in $\cos\theta^\star$ is compared to the prediction from VECBOS. Tightening the $\cos\theta^\star$ cuts to 0.8 for all jets would reduce that sample to 10 data events (as seen in Reference [9]). Background from non-W events was not removed from that sample. Higher jet multiplicities were excluded. Removing the non-W background and including events with more than three jets should still result in O(10) events.[4]

We applied all the cuts explained in Reference [9] to our sample of PAPAGENO events. Assuming the lepton reconstruction efficiencies quoted in Reference [3] and the integrated luminosity quoted in Reference [9], PA-

[4]In Ref. [3] it is explained that most of the non-W background comes from photon conversions. They can be identified and removed on an event-to-event basis.

PAGENO predicts about 10 events from a 100-GeV top. If we assume that the W+3 jet data events of Reference [9] do not contain any top, this would mean that the $\cos\theta^\star$ cuts result in a signal/background ratio of O(1). If some of the data events are from top, then the signal/background ratio is correspondingly better. For higher top masses the top production cross section decreases, but up to masses of about 120 GeV this is approximately compensated by the increasing efficiency of the $E_t(\text{jet}_3)$ cut.

We understand that the above arguments are more qualitative than quantitative. Our goal was to show a possibility, not to predict a quantitative result. We believe that we demonstrated that such an analysis is feasible. On the other hand, the correct number of top events to be expected in an actual experiment will have to be based on a well-studied detector simulation.

6 Outlook

The opportunity to obtain information on the top quark from the single-lepton channel is attractive. The price one has to pay is some dependence on Monte Carlo simulations. The predicted $\cos\theta^\star$ distributions might turn out to be different, once higher-order processes are better taken into account. That is true for both QCD W+jet and top W+jet events. Improving the already sophisticated matrix element calculations will require time and effort. In the meantime, however, significant progress could come from a comparison between existing calculations, though they may not be perfect, and experimental data. Such a comparison could settle some of the questions that cannot be answered reliably by theoretical calculations, such as the true rejection power of the $\cos\theta^\star$ cuts.

7 Conclusion

With a few simple cuts one can reduce the W+jets QCD background (as simulated by VECBOS) to $t\bar{t}$ events (as simulated by PAPAGENO for M_{top} = 100 GeV) into one lepton plus jets by about a factor of 4, losing 30% of the top events.

This kind of analysis could be applied to existing data. A comparison with CDF data indicates that a signal/background ratio of O(1) could be achieved in the same way for m(top) > m(W).

It is also noteworthy that the procedure suggested here should be applicable in a similar way to a search for any new, heavy, centrally decaying particles.

8 Acknowledgment

This work would not have been possible without the help of the VECBOS Monte Carlo program, which has been further developed and is supported by W. Giele. We thank Prof. G. Bellettini for many useful discussions and for his careful reading of the manuscript.

References

[1] F. Abe et al. (CDF coll.), "A Lower Limit on the Top Quark Mass from Events with Two Leptons in $p\bar{p}$ Collisions at $\sqrt{s}$ = 1.8 TeV," Phys. Rev. Lett. 68, 447 (1992). A more stringent limit of 109 GeV/c^2 was announced by CDF at the 1993 Winter Conferences (C. Campagnari for the CDF Collaboration, "Search for New Particles at CDF," to appear with Proceedings of the 1993 Rencontree de Physique de la Vallee d'Aoste, La Thuile, March 8–13, 1993).

[2] F. Abe et al. (CDF coll.), "Top Quark Search in the Electron + Jets Channel in $p\bar{p}$ Collisions at $\sqrt{s}$ = 1.8 TeV," Phys. Rev. D 43, 664 (1991).

[3] F. Abe et al. (CDF coll.), "A Lower Limit on the Top Quark Mass from $p\bar{p}$ Collisions at $\sqrt{s}$ = 1.8 TeV," Phys. Rev. D 45, 3921 (1992).

[4] F. A. Berends, J. B. Tausk, W. T. Giele, "Top Search in Multijet Signals," Fermilab-Pub-92/196-T.

R. H. Dalitz and G. R. Goldstein, "Decay and polarization properties of the top quark," Phys. Rev. D45, 1531 (1992).

R. H. Dalitz and G. R. Goldstein, "The Analysis of top-antitop production and dilepton decay events and the top quark mass," University of Oxford, Dept. of Physics - Theoretical Physics, preprint (1992).

K. Kondo, J. Phys. Soc. Japan, 57, 4126 (1988); K. Kondo, J. Phys. Soc. Japan, 60, 836 (1991).

[5] "Io stimo piu' il trovar un vero benche' di cosa leggiera che'l disputar lungamente delle massime questioni senza conseguir verita' nessuna," G. Galilei. ("I prefer to find truth, even if simple, over discussing in length most sophisticated problems without arriving at any truth at all.") Inscription at the Physics Institute, Univ. Padova.

[6] I. Hinchliffe, private communication.

[7] F. A. Berends, W. T. Giele, H. Kuijf, B. Tausk, "On the production of a W and jets at Hadron Colliders," Fermilab-Pub-90/213-T, and

W.Giele, Ph.D. Thesis, Univ. Leiden.

[8] W. Giele, private communication.

[9] H. Grassmann, "Status Report from CDF," Presentation at the DESY Theory Workshop, Oct. 1991, Fermilab-Conf-92/105.

[10] W. T. Giele, T. Masuura, M. H. Seymour, B. R. Webber, "W Boson plus Multijets at Hadron Colliders: HERWIG Parton Showers versus Exact Matrix Elements," Fermilab-Conf-90/228-T.

[11] M. Shochet, "The Physics of Proton Antiproton Collision," Fermilab-Conf-91/341-E.

[12] F. Abe et al. (CDF Coll.), "Dijet Angular Distributions from $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV", Phys. Rev. Lett. 62, 3020 (1989), and

R. Harris, Proceedings Particles and Fields 91, Univ. of Vancouver, Vancouver, British Columbia, Canada.

[13] H. Grassmann, "A Limit on Leptoquarks from Missing Energy and from Muon Events at the $p\bar{p}$ Collider," Ph.D. Thesis, RWTH Aachen, 1988.