PHYSICAL REVIEW D **72,** 074016 (2005)

In search of lonely top quarks at the Fermilab Tevatron

Matthew T. Bowen, Stephen D. Ellis, and Matthew J. Strassler

Department of Physics, University of Washington, P.O. Box 351560, Seattle, Washington 98195, USA

(Received 21 April 2005; published 14 October 2005)

Single top-quark production, via weak-interaction processes, is an important test of the standard model, potentially sensitive to new physics. However, it is becoming known that this measurement is much more challenging at the Tevatron than originally expected. We reexamine this process and suggest new methods, using shape variables, that can supplement the methods that have been discussed previously. In particular, by focusing on correlations and asymmetries, we can reduce backgrounds substantially without low acceptance for the signal. Our method also allows for a self-consistency check on the modeling of the backgrounds. However, at the present time, serious systematic problems remain, especially concerning the background from W-plus-jets; these must be studied further by experimentalists and theorists alike.

DOI: 10.1103/PhysRevD.72.074016 PACS numbers: 14.65.Ha, 13.85.Ni, 13.85.Qk, 13.87.Ce

I. INTRODUCTION

The electroweak production of single top quarks is an important standard model process which the Tevatron is guaranteed to be able to study. This reaction, which has been investigated previously [1,2], is interesting both because it provides a direct measurement of the V_{tb} Cabibbo-Kobayashi-Maskawa element and because it is sensitive to deviations of top-quark physics from standard model predictions, in particular, through their effects on the top-bottom-W vertex [3]. Limits on single-top production from run I at the Tevatron have been published [4,5], and the first run II limits have appeared [6,7]. In this article we discuss methods which we hope will improve the significance of the measurement by using more information encoded in the shape of the signal, in a way that will be less sensitive to systematic errors than a simple counting experiment. However, we also show that the size of the W-plus-jets background, and the difficulty of predicting it accurately, represent a serious obstacle.

Single-top production is a very unusual process. At Fermilab energies, the "tb" production of a top quark and bottom antiquark by an s-channel W boson, as can occur through the diagram in Fig. 1(a), has a lower cross section than "tbq" t-channel W boson of a t, $\bar{b}$, and an extra light-quark jet near the beam axis, as occurs through diagrams such as that in Fig. 1(b).[1] The tbq process has a distinct shape, both because of its unusual initial state, the hard light-quark jet (which has large p_T and large pseudorapidity) in the final state, and correlations between this jet and the lepton from the top decay. In this paper, we will explore a method for using these features to help separate single top from its major backgrounds: $t\bar{t}$, W-plus-jets, and QCD events.

This separation using the shape of the event is essential, because a simple counting experiment is extremely difficult to carry out. Both our studies and recent data indicate that the size of the relevant W-plus-jets background is larger than anticipated a few years ago [1,2]. This is compounded by other issues, such as the 2% decrease in Fermilab's energy compared to expectations, and lower cross-section calculations for the signal [8–10]. We are concerned that systematic errors in the understanding of the background will plague a direct counting experiment at a level that will make any claims of discovery suspect. Thus, in our view, additional methods, independent of (but perhaps to be used in conjunction with) a counting experiment, are needed even for a discovery of the process, as well as for a precision measurement. We will argue that it is necessary, and possible, to reduce backgrounds further by using more information about the final states. (Note that a shape fit using a single variable was used in [4,6].)

However, even with these improvements, our results suggest that the measurement of single top remains challenging. We have found that there is no simple way, even with 3 fb^{-1} of integrated luminosity, to achieve both good statistics and a satisfactory signal-to-background ratio. The essential problem, compared to earlier and more optimistic studies [1], is that the W-plus-jets sample with a single

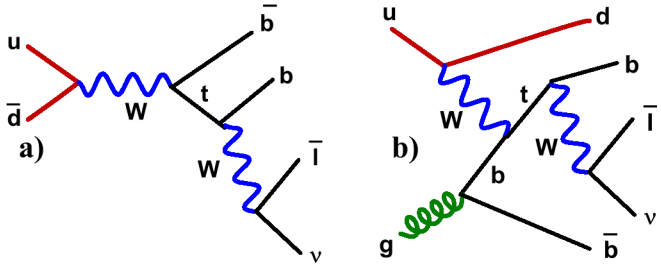

FIG. 1 (color online). Single top-quark production via (a) an s-channel W boson; (b) a t-channel W boson.

[1]There is, as always, some ambiguity as to whether the initial state contains a gluon which splits into a b and $\bar{b}$ as part of the scattering process or whether the initial state contains a b parton directly, i.e., the splitting process is part of the incoming wave function. Since the $\bar{b}$ jet is generally not used in the analysis below, a careful examination of this separation is not essential for us.

b-tagged jet is larger, less predictable, and more variegated than was understood a few years ago. We will discuss in some detail the difficulties with this background, and the need for a wide range of efforts to bring it under control.

The organization of this paper is as follows. In Sec. II we discuss the general structure of the events expected for both signal and background events. Section III includes a detailed discussion of how we have simulated both types of events, the cuts used to define the event samples, the expected (and observed) differences between signal and background event shapes, and new observables intended to highlight these differences. In Sec. IV we address the essential issues of uncertainties, both statistical and systematic, with a special focus on the difficulties inherent in understanding the background arising from W boson production accompanied by jets. In the final section we summarize our analysis and our conclusions.

II. THE BACKGROUNDS TO SINGLE-TOP PRODUCTION

From Fig. 1(b) one can see that a tbq event has a final partonic state consisting of at least the following: a charged lepton, a neutrino, b and $\bar{b}$ quarks, and a light quark. The bottom quark that comes directly from the initial gluon tends to have small transverse momentum, and is rarely in the pseudorapidity and p_T range necessary to be identified by b-tagging algorithms. Thus, in selecting t-channel signal events, one asks for (a) one or more b-tagged jets, (b) significant ''missing transverse energy'' (MET, the magnitude of the measured imbalance in the transverse momentum due to the undetected neutrino), (c) one and only one isolated $e^{\pm}$ or $\mu^{\pm}$, (d) at least one untagged jet. Typically, the highest-p_T untagged jet in a t-channel event is that from the light quark. Also, in typical events a t quark can be reconstructed from the tagged jet and a W, itself reconstructed from the charged lepton and the missing energy. The kinematics of the event tend to prefer a total visible scalar-summed transverse energy of order m_t.

The tb process has a b quark jet and a $\bar{b}$ jet, along with a lepton and a neutrino. In a significant number of events, one of the two b jets will not be tagged, so that the same criteria used for tbq—one b-tagged jet, missing energy, a charged lepton and an untagged jet—will have moderately high efficiency for this process as well. Since the tb process is smaller in cross section and considerably less distinctive in shape than tbq, and therefore harder to separate from background, it makes sense for us to optimize our approach for tbq. We will not, in this paper, discuss the measurement of the tb and tbq channels separately, as this will only be possible well after the initial measurements of the combined production process.

The main backgrounds to single-top production, which all can imitate the signature just described, are [1] (1) ''$t\bar{t}$,'' top-quark pair production, primarily from events where only one of the top quarks decays leptonically, but also from events with two leptonic decays; (2) ''QCD'' events with fake missing energy and with either a fake lepton or a lepton from a heavy-flavor decay; and, most problematic, (3) ''Wj^n'' events with a leptonically decaying W plus some number $n \geq 2$ of quark or gluon jets.

A. Top pair production

Top-quark pair production is a formidable background to all single-top channels. After both top quarks decay, there are two high-p_T b quarks, which typically give rise to at least one b-tagged jet, and two W bosons which can decay hadronically or leptonically. The signature for single top will be mimicked if only one W decays leptonically, or if both do and only one charged lepton is detected. The present measurement of the cross section for $t\bar{t}$ production has a large uncertainty, though this uncertainty is expected to drop to around 10% [11] by the end of run II. However, even this systematic uncertainty would prohibit observation of single top above the $t\bar{t}$ background, so a substantial amount of $t\bar{t}$ must be cut away. The contribution of $t\bar{t}$ with one hadronically decaying top is especially problematic, since often the leptonically decaying top quark can be reconstructed. On the other hand, this background can be reduced using the fact that these events tend to exhibit considerably more transverse energy than true single-top events, to have more jets, and to be more spherical. These kinematic handles are not present for events where both top quarks decay leptonically, but such events are suppressed by both the leptonic decay branching fraction and our requirement that only one $e^{\pm}$ or $\mu^{\pm}$ be observed in the detector. Furthermore, the accurate reconstruction of a leptonically decaying top quark is more difficult in such events.

B. QCD backgrounds

Pure QCD events can give rise to fake leptons and provide heavy-flavor jets (or jets which have no heavy flavor but are mistagged). As fluctuations in energy measurement can lead to a fake missing E_T signal, these QCD events form a background to single-top production. While the energetics tend to be lower than in signal events, and the invariant mass of the ''lepton,'' missing energy and b-tagged jet exhibits no peak at m_t, the number of events is so large that it is not obvious, without data (or a complete detector simulation), whether this represents a relevant background. Recent results from D0 [7] indicate that the number of QCD events entering their single-top samples (defined by somewhat different cuts than used here) is smaller than for other backgrounds, and is, in first approximation, comparable in size to the single-top signal. This conclusion depends in detail on cuts and on the flavor of the lepton (muons being more prevalent than electrons in their samples) and may differ for CDF. As we will see, the methods that we describe below provide substantial further reduction of this QCD background and the systematic

errors associated with determining it, to the point that it should not pose a serious issue. Consequently, we will largely disregard the QCD contribution to the sample, except for a discussion in Sec. III D as to when and why this is justified.

C. W-plus-jets

The W-plus-jets background is much more challenging. The Wj^n events potentially entering the sample consist of a real W boson decaying leptonically, and at least two other quarks or gluons in the final state. While the Wj^n events do tend to have smaller energetics than single top, and do not have a reconstructible t quark, the number of events is so large, and the energy resolution at Fermilab is sufficiently broad, that the Wj^n events form a large and problematic background to single-top production.

The difficulties involved in simulating Wj^n events have been discussed elsewhere [12,13] and we will not give a full review, but it is important to note that there are special problems for the sample with a single b tag that neither untagged nor double-tagged samples suffer from. In particular, many different partonic processes, with different shapes, contribute to the sample in a fashion which is difficult to predict accurately. We will discuss this in detail in Sec. IV D. As we will see, reducing the systematic error on the prediction and/or measurement of this process is essential for success.

III. SIMULATIONS, CUTS, SHAPES AND METHODS

A. Simulation and kinematic cuts

For the modeling of both signal and background, there are multiple tools available both for leading-order matrix elements and for showering and hadronization at leading order. In this study we use MADEVENT [14] to generate events, PYTHIA [15] to then simulate showering and hadronization (using the default value of the initial shower scale, $\sqrt{\hat{s}}$), and PGS [16] to act as a fast detector simulation. The single-top s- and t-channel cross sections have been normalized to 0.88 and 1.98 pb, respectively [8], and generated with factorization and renormalization scales $\mu = m_t = 175$ GeV. For the $t\bar{t}$ process, the cross section is normalized to 6.7 pb [17] and generated with factorization and renormalization scales also at $\mu = m_t = 175$ GeV. For the Wj^n channel, we have limited our simulations at the MADEVENT level to W-plus-two jets (henceforth Wjj), since we believe the uncertainties in our simulations of this process are as large as the contribution at the next order in α_s. We will discuss this point in more detail in Sec. IV. Samples for the Wjj channel were generated using a renormalization and factorization scale $\mu = M_W/2$, which is selected to give the correct normalization at next-to-leading order (NLO) [18]. We employ cuts at the MADEVENT level of $p_{Tj_1}, p_{Tj_2} > 10$ GeV (where p_{Tj_i} is the transverse momentum of the ith jet), with angular separation $\Delta R(j_1, j_2) > 0.4$, and $|\eta_j| < 4.0$.

TABLE I. Basic cuts for initial sample. Here p_T is the transverse momentum and η is the pseudorapidity.

Item	p_T	$\lvert\eta\rvert$
$\ell^{\pm}$	≥ 15 GeV	≤ 2
MET (ν)	≥ 15 GeV	$\cdots$
Jet (b tag)	≥ 20 GeV	≤ 2
Jet (no b tag)	≥ 20 GeV	≤ 3.5

The b tagging is simulated as follows. Jets containing a b quark (either perturbatively or produced during showering) are taken to be tagged with an efficiency of the form $0.5\tanh(p_T/36\ \text{GeV})$, where p_T is the transverse momentum of the jet. Jets containing a c quark (either perturbatively or produced during showering) are taken to be tagged with a rate of the form $0.15\tanh(p_T/42\ \text{GeV})$, while jets containing no heavy flavor are taken to be mistagged with a rate of the form $0.01\tanh(p_T/80\ \text{GeV})$.[2]

From the events generated in this fashion, we define our initial event sample to be those events that contain *one and only one* isolated charged lepton, missing energy, *one or more* b-tagged jets, and *one or more* untagged jets. These objects satisfy the "basic" cuts listed in Table I. The p_T constraints in this and later tables apply only to the highest-p_T b-tagged jet and the highest-p_T untagged jet. Additional cuts are necessary in order to bring backgrounds down to a reasonable level.

Given that Wj^n tends, compared to the signal, to have lower energy and fewer jets, and that $t\bar{t}$ tends to have higher energy and more jets, there are two natural choices of variables to cut on that take advantage of the overall kinematics of the events without appealing to event shapes. One possibility is to cut on the total transverse energy of the event; a second would be to cut on the number of high-p_T jets. The question of which of these (or whether a combination thereof) has the lowest theoretical uncertainties has not been resolved and we will not attempt to

[2]There is no agreed-upon convention for the light-quark/gluon mistagging function. The tanh form of the mistagging function is *not* what is used in the default PGS detector simulation, which instead is $-5.54 \times 10^{-5} + 1.66 \times 10^{-7}(p_T/1\ \text{GeV})^{2.507}$. Based on the fact that this function leads to large mistagging rates at very high p_T, in contradiction to measurements at CDF [19], we have changed this mistagging function within PGS to a tanh form. The overall mistagging rate that we use is larger than assumed in other papers, including [1] and the recent work of [9], where the size of Wj^n backgrounds is smaller as a result. The true form of the mistagging function is detector- and algorithm-dependent, and different mistagging functions do change the shape of certain distributions. Our results suggest that uncertainties in this function lead to important, but not dominant, systematic uncertainties in the Wj^n background.

TABLE II. Representative intermediate cuts.

Item	p_T	$\|\eta\|$
$\ell^{\pm}$	≥ 15 GeV	≤ 2
MET (ν)	≥ 15 GeV	$\cdots$
Jet (b tag)	≥ 20 GeV	≤ 2
Jet (no b tag)	≥ 20 GeV	≤ 3.5
	Min	Max
H_T	≥ 180 GeV	≤ 250 GeV
m_t	≥ 160 GeV	≤ 190 GeV

answer this very important question definitively here. For this study we have chosen to cut on the quantity

$$H_T = \sum_{\text{jets}} (p_T)_i + (p_T)_\ell + \not{E}_T,$$

where the sum is over all jets with $p_T > 20$ GeV and $|\eta| < 3.5$, $(p_T)_i$ is the magnitude of the transverse momentum of the ith jet, $(p_T)_\ell$ that of the lepton, and $\not{E}_T$ is the missing transverse energy in the event. Large values of H_T tend to favor $t\bar{t}$ final states, while lower values favor Wj^n final states, with the signal contribution peaking at intermediate values.

Furthermore, since the signal involves a t quark, we also impose a requirement that the invariant mass of the lepton, neutrino, and the leading tagged jet be approximately equal to the top-quark mass. In doing so, we must reconstruct the neutrino's momentum component $p_{\nu,z}$ along the beam axis, which has an ambiguity. We require that $(p_\ell + p_\nu)^2 = m_W^2$, and among the two solutions for $p_{\nu,z}$ we choose the solution with smallest absolute magnitude. (For complex solutions, only the real part is used.)

As suggested earlier, we find that such cuts cannot decrease the backgrounds to the point that they are comparable to the signal. Two choices of ''intermediate'' and ''hard'' cuts are indicated in Tables II and III. The resulting numbers of expected events for an integrated luminosity of 3 fb^{-1}, summing over $e^{\pm}$ and $\mu^{\pm}$ (and thus including both t and $\bar{t}$), can be seen in Table IV. We show the number of events which survive the basic cuts of Table I, the intermediate cuts of Table II, and the hard cuts of Table III. Consistent with [1], we find that while all of the cuts

TABLE III. Representative hard cuts.

Item	p_T	$\|\eta\|$
$\ell^{\pm}$	≥ 15 GeV	≤ 2
MET (ν)	≥ 15 GeV	$\cdots$
Jet (b tag)	≥ 60 GeV	≤ 2
Jet (no b tag)	≥ 30 GeV	≤ 3.5
	Min	Max
H_T	≥ 180 GeV	≤ 250 GeV
m_t	≥ 160 GeV	≤ 190 GeV

TABLE IV. Numbers of events for 3 fb^{-1} (summed over t and $\bar{t}$, e and μ channels) for the three sets of cuts in Tables I, II, and III.

Channel	Basic cuts	Intermediate cuts	Hard cuts
tbq	298	67	30
tb	145	27	13
$t\bar{t}$	2623	140	57
Wjj	6816	550	152
$(tbq + tb)/(t\bar{t} + Wjj)$	0.047	0.14	0.21

contribute to the background reduction, the Wj^n channel is reduced primarily by a combination of the stiffer p_T cuts and the m_t cut, while the $t\bar{t}$ background is affected primarily by the upper H_T cut. While the basic cuts reveal a signal-to-background ratio of approximately 1:21, this improves to 1:7.4 and 1:4.9 using the intermediate and hard cuts. This is, at best, disappointing.

The difference in the results between our study and that of [1] is striking, and requires an explanation.[3] We believe the main effects can be accounted for straightforwardly. First, in the present study we have the benefit of recent next-to-leading-order calculations [18], which increase the overall rate for Wj^n by of order 50% compared to that used in [1]. (In our leading-order computations, this is effectively obtained through our lower choice of renormalization scale.) Second, we find a much larger number of tagged jets in the Wj^n channel, because we include (by simulating parton showering) the fragmentation of leading-order partonic gluons into heavy-flavor jets at subleading orders. (This effect would appear already in a next-to-leading-order calculation, such as performed recently in [9].) Third, our more pessimistic estimate of energy resolution at the Fermilab detectors forces us to use a wider m_t window cut in order to have sufficient acceptance for the signal; this lets in more Wj^n background. Fourth, we use a more pessimistic b-tagging rate (50% vs 60%) and light-

[3]Because of our different cuts and somewhat different approaches, a detailed comparison between the results of the two analyses is difficult. To obtain some quantitative sense of the differences let us focus on the ratio of signal $(tbq + tb)$ to Wjj background, which is where the bulk of the difference arises and which provides an upper limit on the full signal-to-background ratio. For example, consider this ratio for the ''basic cuts'' results in Table IV yielding a ratio of 0.065. This is most usefully compared with the middle column (without parentheses) in Table 3 of [1], where the corresponding ratio is 0.24. The factor of nearly 4 difference results primarily from our much larger estimate of the Wjj contribution (larger by more than a factor of 3). Adding the m_t cut in [1] increases their value for the ratio to 0.77. However, with our much larger Wjj contribution, we can improve this ratio to only 0.28 when using the hard cuts that include a cut on m_t, i.e., the factor of 3 larger Wjj background is still there. The jet veto used in [1] (the last column in their Table 3) only reduces the background from $t\bar{t}$ (and also the signal), and does not help reduce the Wjj background.

quark-jet mistagging rate (1.0% vs 0.5%), a more pessimistic charm-to-bottom ratio in tagging (1:3.3 vs 1:4.0), and more realistic p_T distributions in tagging functions; these all hurt the signal-to-background ratio and the efficiency for the signal. Other small negative effects include a lower center of mass energy (1.96 TeV vs 2.00 TeV), and a lower cross section for the signal (due largely to a change in parton distribution functions). Note also that we have used $m_t = 175$ GeV, so our results may even be slightly optimistic in this regard.

The numbers in Table IV suggest that a further factor of 3–5 improvement in the signal-to-background ratio via more aggressive cuts will come at the cost of a factor of 5–10 reduction in the signal accompanied by a factor of 15–45 reduction in the background. The essential question then is whether a reduction by such a large factor can be achieved without large systematic and theoretical errors. From Table IV, one can see that systematic errors below about 10% in Wj^n are needed for a discovery. Unfortunately, the method for removing the background suggested in [1], namely, to use a jet veto to reduce $t\bar{t}$ to a small contribution, and do a sideband analysis on either side of the m_t window cut to remove Wj^n, is unlikely to work with such a large Wj^n background. This is illustrated in Fig. 2. With basic cuts, the Wj^n background (whose m_t distribution falls steeply and monotonically above 125 GeV) is very large, as shown in Fig. 2(a). A sideband analysis with a window centered around 175 GeV would be subject to large statistical errors. Meanwhile, the intermediate and hard cuts, shown in Figs. 2(b) and 2(c), tend to shape the Wj^n events such that the Wj^n background is no longer monotonic near and/or across the m_t window, making a sideband analysis problematic. Thus the shape of the m_t distribution of the Wj^n background after hard cuts must be predicted, with small errors. We will argue later that this is very difficult. Consequently, we doubt that a straightforward counting experiment can yield, on its own, a satisfactory measurement of the cross section for the production of single top.

B. Shape variables

Under the assumption that a counting experiment is insufficient, we turn to observables that (as in [4,6]) make use of other aspects of the signal. In particular, we will now explore variables that take advantage of the very special shape of single-top production compared to the background, and are less subject to, or less sensitive to, systematic errors.

As noted earlier, the dominant production process is tbq, in which there is a hard lepton, and also a hard untagged jet j with pseudorapidity $|\eta| \sim 1$–3 and p_T typically larger than 25 GeV. This strongly forward or backward jet is a distinctive signature which the backgrounds do not share. The b jet from the t decay tends to be produced centrally (low pseudorapidity) with high p_T, and is typically the

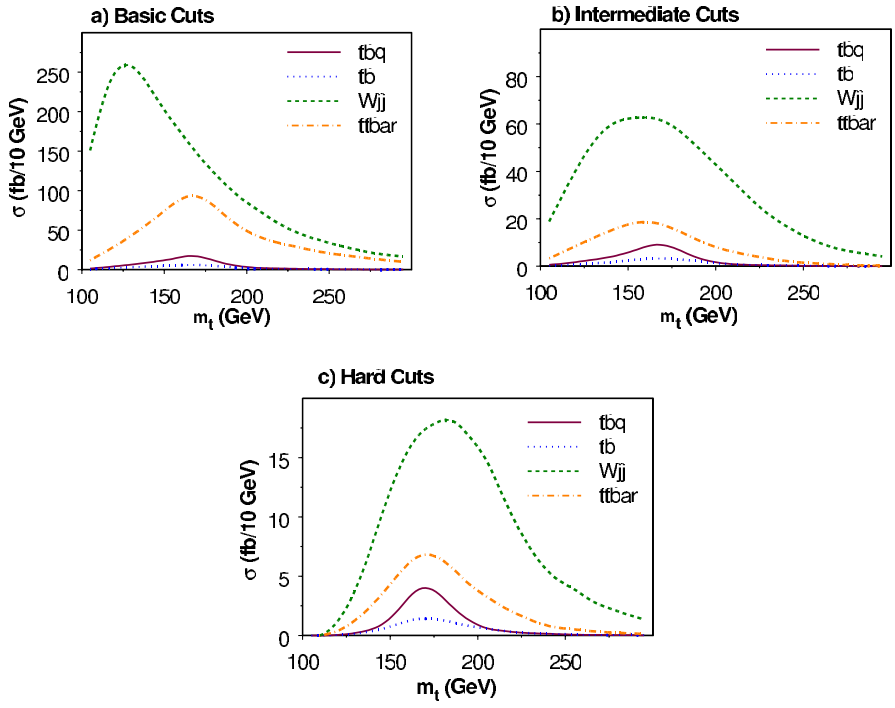

FIG. 2 (color online). Distribution of the reconstructed m_t, the invariant mass of the charged lepton, neutrino and highest-p_T b-tagged jet, for the backgrounds and signals. Additional details concerning the construction of this observable are given in the text. The three figures show the results for the (a) basic, (b) intermediate and (c) hard cuts given in Tables I, II, and III, with the cut on m_t omitted.

tagged jet (for which reason we only use the tagged jet in the t reconstruction). The other b jet tends to have low p_T; it is often unobserved, and is rarely tagged [1].

Importantly, the tbq process arises from an initial state light-quark or antiquark (typically a valence quark carrying moderate to large Bjorken x) scattering off a gluon (typically carrying lower x). This unusual initial state has important kinematic consequences. Because of this kinematic effect, the structure of the proton, and the details of the electroweak theory, the tbq signal has strong and distinctive correlations and asymmetries which we can use to separate it from the backgrounds.

First, the unusual kinematics and the structure of the proton combine in an interesting way. The creation of positively charged top quarks (and consequently positively charged leptons in the final state) in the $t\bar{b}q$ process requires either a u or $\bar{d}$ in the initial state, so that a W^+ can be emitted from the quark line. Since a reasonably large value of Bjorken x is necessary in order to produce the top quark, the required initial state is most often obtained from a valence quark striking a gluon. The most likely initial state uses a valence u quark from the p; the next most likely draws a valence $\bar{d}$ from the $\bar{p}$, and thereafter we must draw on the sea quarks from either p or $\bar{p}$. Since the quark usually has larger x than the gluon that it strikes, the $t\bar{b}q$ system is typically boosted in the direction of travel of the initial quark. Consequently, for t quarks, the $t\bar{b}q$ system is more likely to be boosted in the proton direction than in that of the antiproton, by a factor of roughly 2:1. The reverse is true for $\bar{t}$ production. Moreover, the light-quark jet in the final state, converted from the quark in the initial state by the emission of the W, tends to travel in the proton direction when a t is produced and in the antiproton direction when a $\bar{t}$ is produced.

Thus, because of the differences between the u and d parton distribution functions in the proton, and because of the quark-gluon initial state, large asymmetries under parity P and charge conjugation C result. These show up strongly in the differential cross section both for the top (and the positively charged lepton in its decays) and for the hard forward or backward light-quark jet.

Second, the momentum vectors of the lepton and the light-quark jet are correlated, as a result of both kinematics and spin polarization effects. The structure of the electroweak interactions ensures that the spin of the top quark tends to align with the momentum direction of the light-quark jet. Since the top decays before this spin information is lost, it is transferred to the momentum of the final-state charged lepton. In the t rest frame the lepton momentum and jet momentum tend to align [20]. The boost of the top quark relative to the lab frame, whose sign is also aligned with the momentum direction of the jet, further tends to push both lepton and jet into the same hemisphere.

These properties strongly distinguish the tbq process from its $t\bar{t}$ background. At tree level, $t\bar{t}$ is separately C- and P-invariant. Clearly this is true of the process $gg \to t\bar{t}$, since the initial state is P-invariant on average. It is also true of the process $q\bar{q} \to g \to t\bar{t}$, because the intermediate gluon state is a C eigenstate; this is the same as in $e^+e^- \to \mu^+\mu^-$, where there is no forward-backward asymmetry in the μ^+ distribution. The parity invariance is violated at the next order, due to radiative effects [21]; this is a few percent effect, both small and calculable. Moreover, there should be no strong correlation between the momenta of the lepton and the jets in the final state. In those $t\bar{t}$ events which have both low H_T and a reconstructible semileptonically decaying top quark, one or two jets from the hadronically decaying top tend to be lost or mismeasured. Meanwhile any high-p_T untagged jets whose momenta we might choose to compare with that of the lepton will also stem from the hadronically decaying top. The accidents which lead to the selection of any given jet as part of our analysis should largely wash out any correlation of its momentum with that of the lepton. Indeed, our simulations show that in $t\bar{t}$ the correlation between the lepton and the highest-p_T untagged jet (which we will use in our analysis below) is roughly a 10% effect.

Similar considerations apply, to a good approximation, to those QCD events which might pass our cuts. Parity asymmetries for these events are small. The charge of a fake lepton is unrelated to its momentum direction. Consequently, all distributions for events with fake leptons are invariant under flipping the sign of the lepton charge; as we will see, this implies that these events have P-invariant distributions for the observables we will choose. Isolated leptons from heavy flavor stem mainly from $c\bar{c}$ and $b\bar{b}$ events; these have similar C and P properties to $t\bar{t}$, so we expect small parity asymmetries. Meanwhile, a fake lepton and the highest-p_T untagged jet in the event should be essentially uncorrelated. However, if this jet contains heavy flavor, then a correlation can arise when the lepton observed in the event is from a wide-angle semileptonic decay of a heavy quark within the jet. The precise size of jet-lepton correlations from this source is unknown to us, and is detector- and cut-dependent. The small overall number of QCD events entering the single-top samples at D0 [7] suggests that QCD contributions to jet-lepton correlations are not a major issue for the single-top measurement, except possibly in the case of muons from heavy-flavor decays. We will return to this possible exception in Sec. III D.

The parity asymmetry in Wj^n unfortunately has the same sign as that of tbq, although it is less pronounced. The reasons for this are easy to see. As with a t quark, a W^+ is most likely to be produced moving in the proton direction, since it is most often produced in a $u\bar{d}$ event. This leads to a well-known asymmetry in its pseudorapidity. When produced in conjunction with two jets, the W^+ still tends to be boosted in the proton direction, since its initial state is most often ug or $u\bar{d}$. This leads to parity asymme-

tries which, though relatively small, are still quite large in absolute size compared to the signal.[4] Unfortunately, the size of the asymmetries and correlations in Wj^n appears to be very sensitive to assumptions, cuts, Monte Carlo parameters, and tagging, and will be a source of significant systematic error. We will return to this issue later.

C. Consequences of parity and correlations

In order to make the best use of these special properties of the signals and backgrounds, it is useful to consider these issues more formally. We next discuss the effect of C and P (non)invariance, and of lepton-jet pseudorapidity correlations, on two-dimensional distributions in pseudorapidity. The ''jet'' used throughout the analysis below is always *the highest-p_T untagged jet* in the event.

The $p\bar{p}$ initial state of the Tevatron is a CP eigenstate, and so all distributions of final-state particles are CP-invariant (to an excellent approximation, violated principally by the detector). This means that the differential cross section $d^2\sigma^+/d\eta_j d\eta_\ell$ with respect to the rapidities of the jet and the positively charged lepton, and the corresponding distribution $d^2\sigma^-/d\eta_j d\eta_\ell$ for processes with a *negatively* charged lepton, *must* be related by CP:

$$\frac{d^2\sigma^+}{d\eta_j d\eta_\ell}(\eta_j, \eta_\ell) = \frac{d^2\sigma^-}{d\eta_j d\eta_\ell}(-\eta_j, -\eta_\ell).$$

Consequently, we can combine data from positively and negatively charged leptons by defining a *lepton-charge-weighted pseudorapidity*, $\hat{\eta}_j = Q_\ell \eta_j$, $\hat{\eta}_\ell = Q_\ell \eta_\ell$, where Q_ℓ is the lepton charge. (The variable $\hat{\eta}_j$ was already introduced in [4,6].) For the remainder of this article, our entire discussion is based on the explicitly CP-invariant differential cross section

$$\frac{d^2\sigma}{d\hat{\eta}_j d\hat{\eta}_\ell}(\hat{\eta}_j, \hat{\eta}_\ell) \equiv \frac{d^2\sigma^+}{d\eta_j d\eta_\ell}(\eta_j = \hat{\eta}_j, \eta_\ell = \hat{\eta}_\ell) + \frac{d^2\sigma^-}{d\eta_j d\eta_\ell}(\eta_j = -\hat{\eta}_j, \eta_\ell = -\hat{\eta}_\ell).$$

However, the $p\bar{p}$ initial state is not an eigenstate of either C or P. Consequently, in general we expect parity-noninvariance

$$\frac{d^2\sigma^+}{d\eta_j d\eta_\ell}(\eta_j, \eta_\ell) \neq \frac{d^2\sigma^+}{d\eta_j d\eta_\ell}(-\eta_j, -\eta_\ell),$$

and similar noninvariance under charge conjugation

$$\frac{d^2\sigma^+}{d\eta_j d\eta_\ell}(\eta_j, \eta_\ell) \neq \frac{d^2\sigma^-}{d\eta_j d\eta_\ell}(\eta_j, \eta_\ell).$$

(Indeed these two statements are equivalent due to CP invariance.) In terms of the combined differential cross section, P (and C) noninvariance implies

$$\frac{d^2\sigma}{d\hat{\eta}_j d\hat{\eta}_\ell}(\hat{\eta}_j, \hat{\eta}_\ell) \neq \frac{d^2\sigma}{d\hat{\eta}_j d\hat{\eta}_\ell}(-\hat{\eta}_j, -\hat{\eta}_\ell).$$

Conversely, if we were to study a *parity-symmetric* process, such as the tree-level production of $t\bar{t}$, it would satisfy

$$\frac{d^2\sigma}{d\hat{\eta}_j d\hat{\eta}_\ell}(\hat{\eta}_j, \hat{\eta}_\ell) = \frac{d^2\sigma}{d\hat{\eta}_j d\hat{\eta}_\ell}(-\hat{\eta}_j, -\hat{\eta}_\ell)$$

(parity-even process).

Next, let us consider the effect of correlations. If the dynamics of a process is such that the jet and lepton directions are uncorrelated, then the differential cross section factorizes into a product of two distributions, one for the jet and one for the lepton:

$$\frac{d^2\sigma}{d\hat{\eta}_j d\hat{\eta}_\ell} = f(\hat{\eta}_j) g(\hat{\eta}_\ell) \quad \text{(jet and lepton uncorrelated)}.$$

Failure of this relation is proof of correlations. These might stem directly from correlations in the rest frame of the tbq system. However, even if the distributions in the rest frame are uncorrelated, they will be correlated in the lab frame, once they are convolved with a distribution of boosts of the rest frame.[5]

Finally, if the process is both uncorrelated *and* P-invariant, then both distributions must be even functions of their particle's pseudorapidity. In short

[4]Note these asymmetries in pseudorapidity, or equivalently angle, for fixed charge, are due to the Tevatron's proton-antiproton initial state. At the LHC, with a proton-proton initial state, the same effects will show up as *charge* asymmetries for fixed angle.

[5]For example, suppose $d^2\sigma/d\hat{\eta}_j d\hat{\eta}_\ell = f(\hat{\eta}_j) g(\hat{\eta}_\ell)$ in the tbq rest frame, with Gaussian lepton and jet distributions:

$$f(\hat{\eta}_j) \propto e^{-A_j \hat{\eta}_j^2}, \qquad g(\hat{\eta}_\ell) \propto e^{-A_\ell \hat{\eta}_\ell^2}.$$

Suppose further that the rest frame is boosted by an amount η_b with a probability which also has a Gaussian distribution

$$p(\eta_b) \propto e^{-B\eta_b^2}.$$

Then the observed distribution in the lab frame is

$$\frac{d\sigma}{d\hat{\eta}_j d\hat{\eta}_\ell} \propto \int_{-\infty}^{\infty} d\eta_b f(\hat{\eta}_j - \eta_b) g(\hat{\eta}_\ell - \eta_b) p(\eta_b) \propto e^{-[(B+A_\ell)A_j \hat{\eta}_j^2 + (B+A_j)A_\ell \hat{\eta}_\ell^2 - A_j A_\ell \hat{\eta}_j \hat{\eta}_\ell]/(A_j + A_\ell + B)}.$$

This is a correlated distribution; it cannot be written as $F(\eta_j)G(\eta_\ell)$, because of the cross term in the exponent. Note the correlation vanishes in the limit of a very narrow distribution of boosts, $B \to \infty$, even if the boost distribution is not centered at zero.

$$\frac{d^2\sigma}{d\hat{\eta}_j d\hat{\eta}_\ell} = f(\hat{\eta}_j) g(\hat{\eta}_\ell);$$
$$f(\hat{\eta}_j) = f(-\hat{\eta}_j);$$
$$g(\hat{\eta}_\ell) = g(-\hat{\eta}_\ell) \quad \text{(uncorrelated, parity even)},$$

with the fourway consequence

$$\frac{d^2\sigma}{d\hat{\eta}_j d\hat{\eta}_\ell}(\hat{\eta}_j, \hat{\eta}_\ell) = \frac{d^2\sigma}{d\hat{\eta}_j d\hat{\eta}_\ell}(-\hat{\eta}_j, -\hat{\eta}_\ell) = \frac{d^2\sigma}{d\hat{\eta}_j d\hat{\eta}_\ell}(\hat{\eta}_j, -\hat{\eta}_\ell) = \frac{d^2\sigma}{d\hat{\eta}_j d\hat{\eta}_\ell}(-\hat{\eta}_j, \hat{\eta}_\ell) \quad (1)$$

for such processes.[6] The $t\bar{t}$ process does indeed satisfy the relation (1) at the 90% level. We believe this continues to next-to-leading order: one-loop effects cause a 5% parity asymmetry in t and $\bar{t}$ production angles [21], which, when translated into jet and lepton pseudorapidities, is unlikely to violate parity by more than 10% (though this has not been simulated for our specific cuts). We believe that QCD contributions to the sample are similarly in good agreement with (1), except possibly for lepton-jet correlations in events with heavy flavor. The Wj^n background, with its moderate asymmetries and correlations, accords with (1) only to a very rough approximation. And as we have emphasized, the tbq signal strongly violates the relation (1).

D. Strategy and tactics

To illustrate the characteristic properties of the various channels, we will study a simulated event sample defined by the cuts in Table V. These cuts are more ''relaxed'' than those of Tables II and III, keeping a larger fraction of both the signal and the background, and yielding a sample which we believe is less sensitive to the systematic uncertainties stemming from the cuts. We will argue below that shape considerations will allow us to make some headway toward separating signal and background. The simulated differences in shape are summarized in the contour plots of Fig. 3, which give the distributions $(d^2\sigma/d\hat{\eta}_j d\hat{\eta}_\ell)$ of the various processes, plotted as functions on the $(\hat{\eta}_j, \hat{\eta}_\ell)$ plane. (Recall that we define the jet of relevance to be the highest-p_T untagged jet.)

Figure 3 illustrates the degree to which the various processes exhibit correlations and asymmetries. We label the four quadrants of the $(\hat{\eta}_j, \hat{\eta}_\ell)$ plane A, B, C, D as shown in Fig. 4. Positive jet-lepton correlations cause events to pile up in quadrants B and C, while parity asymmetries appear in the differences between quadrants B and C, and between quadrants A and D. The tbq signal in Fig. 3(b) shows clearly shows both effects. The Wj^n background, in Fig. 3(d), also has an asymmetric shape, though to a lesser relative degree. The tb process, Fig. 3(a), shows some correlation but no asymmetry, while Fig. 3(c) illustrates the uncorrelated and symmetric nature of (tree-level) $t\bar{t}$.

TABLE V. Relaxed cuts for analysis.

Item	p_T	$\lvert\eta\rvert$
$\ell^\pm$	≥ 15 GeV	≤ 2
MET (ν)	≥ 15 GeV	$\cdots$
Jet (b tag)	≥ 40 GeV	≤ 2
Jet (no b tag)	≥ 30 GeV	≤ 3.5
	Min	Max
H_T	None	≤ 300 GeV
m_t	≥ 155 GeV	≤ 200 GeV

As a quantitative measure of these statements, we consider the differential cross sections integrated separately over the four quadrants of the $(\hat{\eta}_j, \hat{\eta}_\ell)$ plane. For a given luminosity $\mathcal{L}$, the number of tbq events in the A quadrant is

$$\mathcal{L} \times \int_0^2 d\hat{\eta}_\ell \int_{-3.5}^0 d\hat{\eta}_j \frac{d^2\sigma^{tbq}}{d\hat{\eta}_j d\hat{\eta}_\ell},$$

that in the B quadrant is

$$\mathcal{L} \times \int_0^2 d\hat{\eta}_\ell \int_0^{3.5} d\hat{\eta}_j \frac{d^2\sigma^{tbq}}{d\hat{\eta}_j d\hat{\eta}_\ell},$$

and so forth. The resulting numbers of events with $\mathcal{L} = 3\ \text{fb}^{-1}$ of data appear in Table VI (summed over t and $\bar{t}$, e and μ). Statistical uncertainties in each bin are uncorrelated with other bins. Systematic uncertainties in these numbers, which are very substantial for Wj^n, deserve considerable discussion; since this table is merely intended for a general illustration, we defer this discussion until Sec. IV.

To capture quantitatively these differences in shape between signal and background, we suggest defining three orthogonal functions in the $(\hat{\eta}_j, \hat{\eta}_\ell)$ plane, based on the formal discussion of the previous section. For any differential cross section, whether a signal, a background or a combination, and whether simulated or measured experimentally, we may write it as a sum of three components

[6] Note the logic is not reversible; a distribution satisfying the condition (1) is not necessarily uncorrelated.

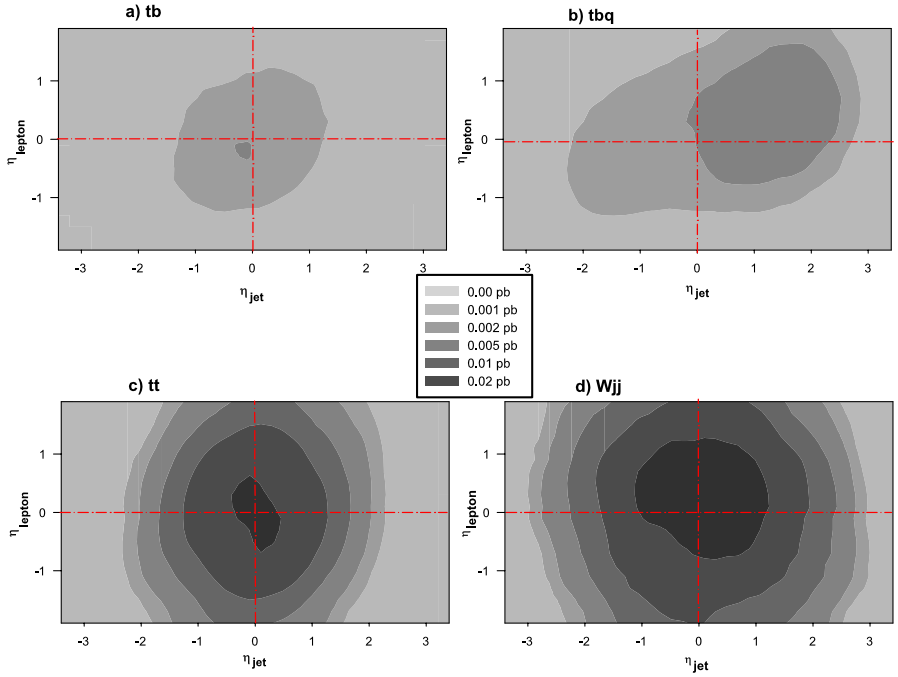

FIG. 3 (color online). Differential cross section ($d^2\sigma/d\hat{\eta}_j d\hat{\eta}_\ell$, summed over t and $\bar{t}$, e and μ) for the (a) tb channel (b) tbq channel, (c) $t\bar{t}$ channel, and (d) Wjj channel, after b tagging and the cuts of Table V.

$$\frac{d^2\sigma}{d\hat{\eta}_j d\hat{\eta}_\ell}(\hat{\eta}_j, \hat{\eta}_\ell) = \bar{F}(\hat{\eta}_j, \hat{\eta}_\ell) + F_+(\hat{\eta}_j, \hat{\eta}_\ell) + F_-(\hat{\eta}_j, \hat{\eta}_\ell), \tag{2}$$

where the components are of the form

$$\bar{F}(\hat{\eta}_j, \hat{\eta}_\ell) \equiv \frac{1}{4}\Big[\frac{d^2\sigma}{d\hat{\eta}_j d\hat{\eta}_\ell}(\hat{\eta}_j, \hat{\eta}_\ell) + \frac{d^2\sigma}{d\hat{\eta}_j d\hat{\eta}_\ell}(-\hat{\eta}_j, -\hat{\eta}_\ell) + \frac{d^2\sigma}{d\hat{\eta}_j d\hat{\eta}_\ell}(\hat{\eta}_j, -\hat{\eta}_\ell) + \frac{d^2\sigma}{d\hat{\eta}_j d\hat{\eta}_\ell}(-\hat{\eta}_j, \hat{\eta}_\ell)\Big] \tag{3}$$

$$F_+(\hat{\eta}_j, \hat{\eta}_\ell) \equiv \frac{1}{4}\Big[\frac{d^2\sigma}{d\hat{\eta}_j d\hat{\eta}_\ell}(\hat{\eta}_j, \hat{\eta}_\ell) + \frac{d^2\sigma}{d\hat{\eta}_j d\hat{\eta}_\ell}(-\hat{\eta}_j, -\hat{\eta}_\ell) - \frac{d^2\sigma}{d\hat{\eta}_j d\hat{\eta}_\ell}(\hat{\eta}_j, -\hat{\eta}_\ell) - \frac{d^2\sigma}{d\hat{\eta}_j d\hat{\eta}_\ell}(-\hat{\eta}_j, \hat{\eta}_\ell)\Big] \tag{4}$$

$$F_-(\hat{\eta}_j, \hat{\eta}_\ell) \equiv \frac{1}{2}\Big[\frac{d^2\sigma}{d\hat{\eta}_j d\hat{\eta}_\ell}(\hat{\eta}_j, \hat{\eta}_\ell) - \frac{d^2\sigma}{d\hat{\eta}_j d\hat{\eta}_\ell}(-\hat{\eta}_j, -\hat{\eta}_\ell)\Big]. \tag{5}$$

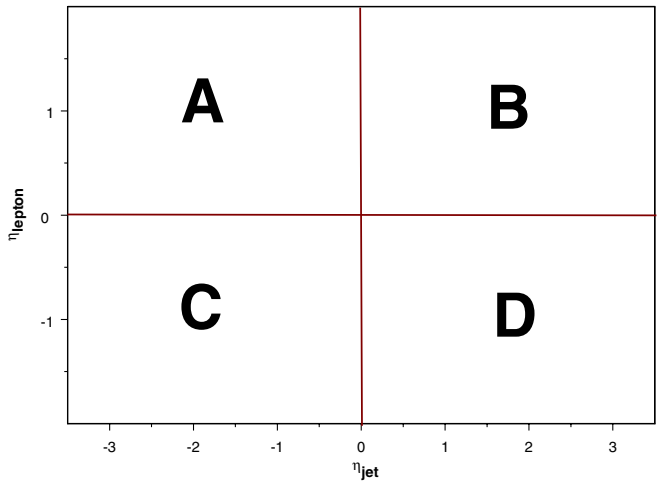

FIG. 4 (color online). The four quadrants of the $(\hat{\eta}_j, \hat{\eta}_\ell)$ plane.

Let us comment on some properties of these functions. First, they are explicit functions, not abstract devices: they can be directly constructed from any finite data set, simulated or measured. Second, they are orthogonal in the sense that

TABLE VI. Numbers of events for 3 fb^{-1} in the four quadrants for various channels (summed over t and $\bar{t}$, e and μ). See Fig. 4 for definition and labels of the quadrants.

Channel	A	B	C	D
tb	8	11	12	8
tbq	14	41	18	22
$t\bar{t}$	105	109	106	105
Wjj	204	207	159	180

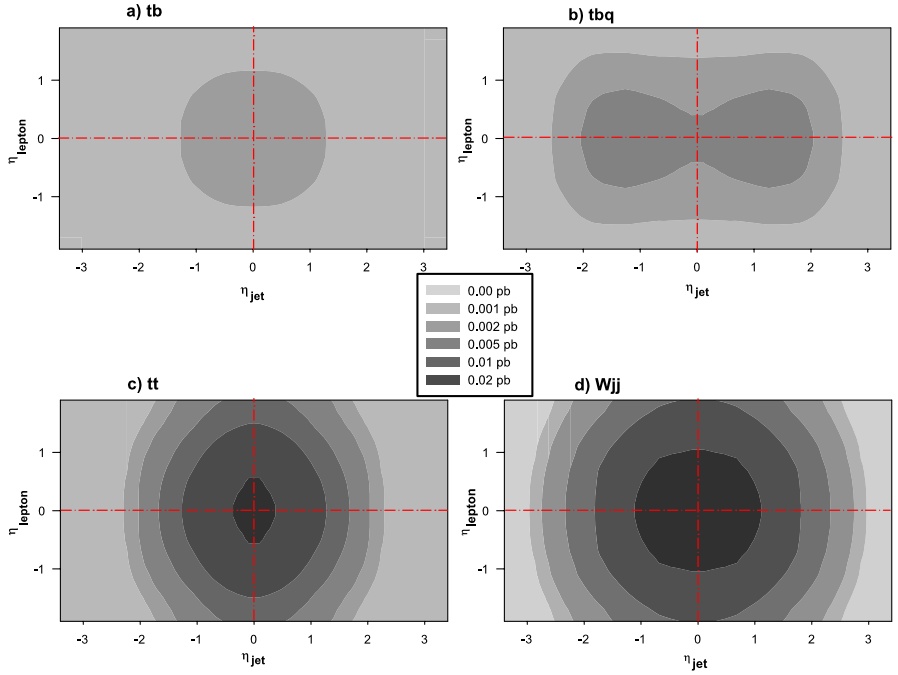

FIG. 5 (color online). Contour plots for the function $\bar{F}(\hat{\eta}_j, \hat{\eta}_\ell)$ for (a) tb, (b) tbq, (c) $t\bar{t}$, and (d) Wjj channels (summed over t and $\bar{t}$, e and μ).

$$\int_{-a}^{a} d\hat{\eta}_j \int_{-b}^{b} d\hat{\eta}_\ell \, \bar{F} \, F_{\pm} = 0;$$
$$\int_{-a}^{a} d\hat{\eta}_j \int_{-b}^{b} d\hat{\eta}_\ell F_{+} F_{-} = 0;$$

(where a and b are arbitrary positive numbers); indeed this orthogonality applies in any symmetrically defined region of the $(\hat{\eta}_j, \hat{\eta}_\ell)$ plane. Third, the functions provide important physical information about the symmetry properties of the differential cross section. $\bar{F}$ and F_+ are parity-even while F_- is parity-odd; thus $F_- = 0$ (within statistics) for any parity-even distribution. Meanwhile, because of Eq. (1), F_+ will also vanish if the distribution is parity-even *and* the leptons and jets are uncorrelated. Fourth, by construction, these functions have special symmetries under reflection in the $(\hat{\eta}_j, \hat{\eta}_\ell)$ plane, as will be obvious in the figures below. $\bar{F}$ and F_+ have fourway symmetry; in both cases, it is sufficient to know the function in any one quadrant to know it in all four quadrants. Meanwhile F_- has twoway symmetry; quadrants A and D are related, as are B and C, but quadrants A and B are independent and must be determined separately.[7]

In the problem at hand, the fact that the signal has strong asymmetries and correlations, while the backgrounds do not, is very usefully characterized using these functions. In particular, we expect, based on the properties we have discussed above, that

$$\bar{F}^{t\bar{t}} \gg |F_{\pm}^{t\bar{t}}|; \qquad \bar{F}^{Wj^n} > |F_{\pm}^{Wj^n}|; \qquad \bar{F}^{tbq} \sim |F_{\pm}^{tbq}|.$$

Our simulations further suggest that one can find cuts that are feasible at the Tevatron such that the backgrounds are still very large but only in $\bar{F}$, with

$$\bar{F}^{t\bar{t}} \sim \bar{F}^{Wj^n} \gg \bar{F}^{tbq},$$

while the signal has a much larger role to play in the other functions:

$$|F_{+}^{t\bar{t}}| \sim |F_{+}^{Wj^n}| \sim |F_{+}^{tbq}|, \qquad |F_{-}^{t\bar{t}}| \ll |F_{-}^{Wj^n}| \sim |F_{-}^{tbq}|,$$

especially away from the center of the $(\hat{\eta}_j, \hat{\eta}_\ell)$ plane. (The tb signal is smaller than the tbq signal for all quantities, but especially for F_+, by a factor of about 3, and for F_-, by a factor of order 10.)

These claims are illustrated in Figs. 5–7, where the functions $\bar{F}$, F_+, F_- are shown, for tb, tbq, $t\bar{t}$, and Wj^n. (The reader should note that the scale for the contours in Figs. 6 and 7 differs from that used in Figs. 3 and 5; this is because of the smaller dynamic range in the $F_\pm$ distributions.) The symmetry properties of the three functions discussed earlier are clearly evident.

To make this comparison more concrete, the integrals of the 3 functions over quadrants A and B, for an integrated luminosity $\mathcal{L} = 3$ fb^{-1}, are presented in Table VII; this table can be constructed from Table VI. The definitions of the entries in the table are

[7] In short, F_- contains twice as much information as F_+ and $\bar{F}$. In principle one could further separate F_- into two orthogonal functions, but this turns out not to be particularly useful.

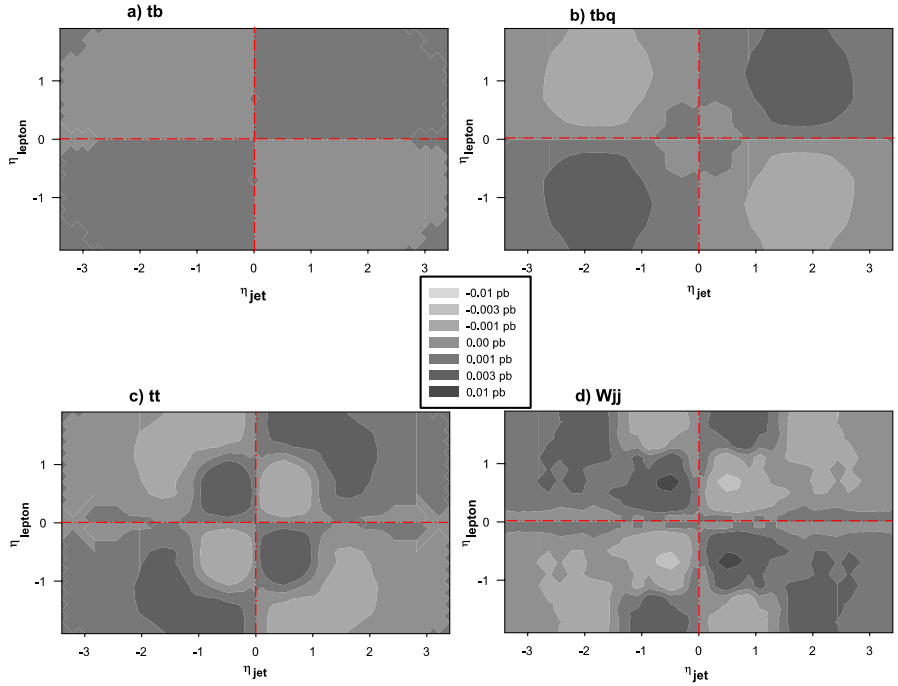

FIG. 6 (color online). Contour plots for the function $F_+(\hat{\eta}_j, \hat{\eta}_\ell)$ for (a) tb, (b) tbq, (c) $t\bar{t}$, and (d) Wjj channels (summed over t and $\bar{t}$, e and μ).

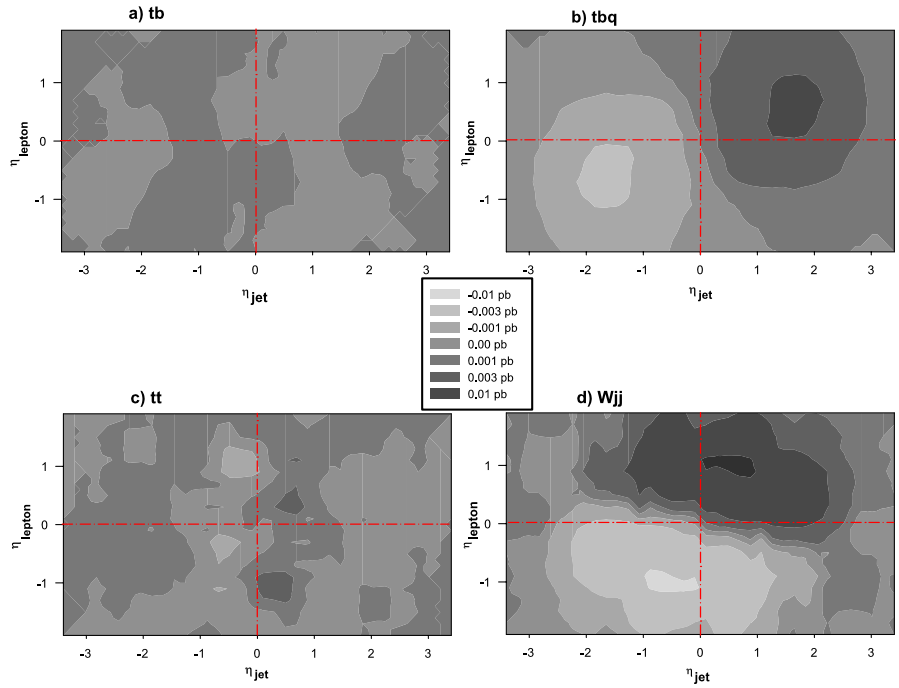

FIG. 7 (color online). Contour plots for the function $F_-(\hat{\eta}_j, \hat{\eta}_\ell)$ for (a) tb, (b) tbq, (c) $t\bar{t}$, and (d) Wjj channels (summed over t and $\bar{t}$, e and μ).

$$\overline{F}_{\mathrm{A}} \equiv \mathcal{L} \times \int_{\mathrm{A}} d\hat{\eta}_j d\hat{\eta}_\ell \overline{F} = \overline{F}_{\mathrm{B}} = \overline{F}_{\mathrm{C}} = \overline{F}_{\mathrm{D}}, \qquad F_{+,\mathrm{A}} \equiv \mathcal{L} \times \int_{\mathrm{A}} d\hat{\eta}_j d\hat{\eta}_\ell F_+ = -F_{+,\mathrm{B}} = -F_{+,\mathrm{C}} = F_{+,\mathrm{D}},$$

$$F_{-,\mathrm{A}} \equiv \mathcal{L} \times \int_{\mathrm{A}} d\hat{\eta}_j d\hat{\eta}_\ell F_- = -F_{-,\mathrm{D}}, \qquad F_{-,\mathrm{B}} \equiv \mathcal{L} \times \int_{\mathrm{B}} d\hat{\eta}_j d\hat{\eta}_\ell F_- = -F_{-,\mathrm{C}}.$$

TABLE VII. Numbers of events for 3 fb^{-1} in $\overline{F}$, F_+ and F_-, integrated over quadrants A and B, for the various channels (summed over t and $\bar{t}$, e and μ). Errors shown are statistical only. See text for further interpretation.

Channel	$\overline{F}_{\rm A} = \overline{F}_{\rm B}$	$F_{+,{\rm B}} = -F_{+,{\rm A}}$	$F_{-,{\rm A}}$	$F_{-,{\rm B}}$
tb	9.8 ± 1.6	1.4 ± 1.6	0.0 ± 2.1	-0.3 ± 2.4
tbq	23.8 ± 2.4	5.6 ± 2.4	-3.7 ± 3.0	11.6 ± 3.8
$t\bar{t}$	106.1 ± 5.2	1.2 ± 5.2	-0.2 ± 7.2	1.3 ± 7.3
Wjj	187.7 ± 6.9	-4.6 ± 6.9	11.8 ± 9.8	23.8 ± 9.6

The table also shows the statistical errors in these quantities. Since this table is intended only to emphasize qualitative points, we postpone discussion of systematic errors and next-to-leading-order corrections until we outline a more sophisticated approach with better statistical errors.

As we argued before, we are justified in disregarding QCD events. The number of QCD events entering the sample is small [7],

$$\bar{F}^{t\bar{t}}, \bar{F}^{Wj^n} \gg \bar{F}^{\rm QCD} \sim \bar{F}^{tbq}, \bar{F}^{tb}.$$

For the D0 detector, the number of events in the electron channel is smaller than the number in the muon channel. We expect lepton-jet correlations and parity asymmetries to be extremely small for fake leptons; for isolated leptons from heavy flavor, parity asymmetries are very small at tree level, and small at higher orders, while lepton-jet correlations might be a bit larger. Thus

$$\bar{F}^{\rm QCD} > |F_+^{\rm QCD}| > |F_-^{\rm QCD}|,$$

even for isolated leptons. This leads us to expect

$$|F_\pm^{tbq}| \gg |F_\pm^{\rm QCD}|,$$

except possibly for F_+ in the muon channel. The dominant source for F_+ will be from muons emitted at large angles during $b\bar{b}$ and $c\bar{c}$ events. The kinematics of these events can be studied using a double-tagged sample. We believe, therefore, that even if the contribution of QCD muon events to F_+ is large enough to be a concern, its size and shape can be determined from the data.

E. Lessons and caveats

We can now extract some important lessons from Table VII. Before doing so, we should comment on its limitations.

First, our study is done entirely using tree-level short-distance matrix elements; only the normalizations are at next-to-leading order. We must therefore emphasize that *the figures and tables in this paper are meant for illustration only.* Next-to-leading-order corrections to the matrix elements will change the shapes of the background, in ways which contribute nontrivially to Table VII. For example, we expect the above-mentioned next-to-leading-order effects on $t\bar{t}$ to affect our estimates of $F_\pm^{t\bar{t}}$, by something of order 10% of $\bar{F}^{t\bar{t}}$, which is not negligible. Such changes are large enough that they must be calculated and/or measured, but are small enough that they do not invalidate the *methodology* we are outlining here.

Second, it is clear from the table that statistical errors from the background are large in F_+ and F_-, comparable to the signal. While this looks discouraging, it applies for the distributions integrated over the entire A or B quadrant. A quick examination of Fig. 3 shows that the situation is not quite as bad as it appears, if one considers the region away from the center of the $(\hat{\eta}_j, \hat{\eta}_\ell)$ plane, where the backgrounds are much smaller and the signals are still quite large. We will discuss this in much more detail in the next section, where we will do a more sophisticated analysis, but for the moment we simply note that the size of the statistical errors is misleadingly large in the above table. Still, we will see that the situation with statistics remains unsatisfactory.

With these caveats in mind, we return to Table VII, on the basis of which we suggest the following general approach. One should first construct, for both the data and the Monte Carlo simulation output, the $\bar{F}$, F_+ and F_- functions. Using these functions, as well as information obtained from other measurements, one can systematically test one's understanding of each process. The $\bar{F}$ function allows a measurement of the sum of the backgrounds without much contamination from signal. We assume that the separation of Wj^n from $t\bar{t}$ can be obtained using the fact that the $t\bar{t}$ process can be measured and predicted with reasonable accuracy, using other data samples and Monte Carlo simulation. One can then cross-check one's understanding of the shape of the Wj^n background using the part of the F_- distribution (located roughly in quadrant A) where the signal is negligible. Finally, one can attempt to measure the signal from F_+ and from a different part (located largely in quadrant B) of the F_- distribution.[8] Effects on F_+ from QCD events with isolated leptons can be cross-checked by study of double-tagged events, and by comparing F_+ for electrons against F_+ for muons.

We now proceed to refine this approach, and to estimate the associated uncertainties.

IV. UNCERTAINTIES AND OPTIMIZATION

Our goal in this section is to show that the measurement of the signal using F_+ and F_- is potentially feasible, though difficult. We begin this section with an overview that, using Figs. 8 and 9 as guides, lays out our main points. We then turn to more detailed consideration of statistical

[8]Note that the study in [6] measures half this information; it is able to measure $\bar{F}$ and the difference of quadrants A and B in F_-.

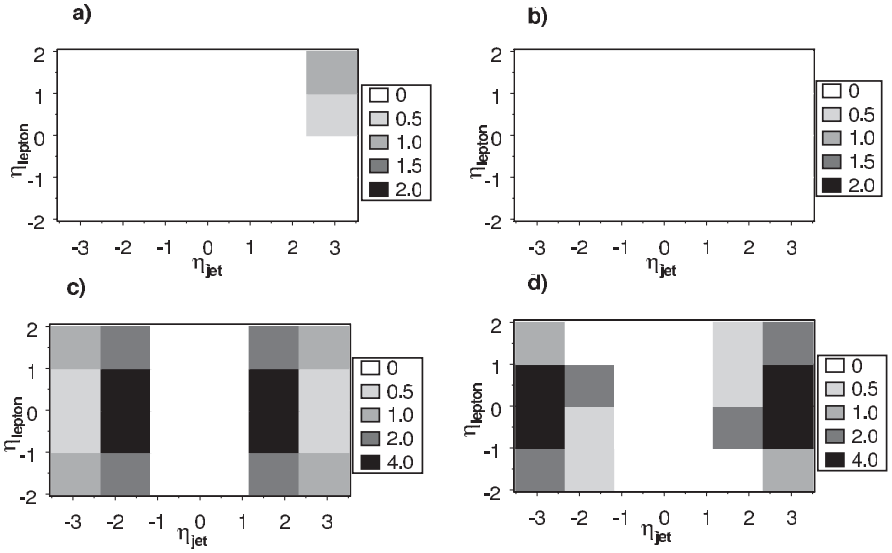

FIG. 8. Signal-to-background ratio S/B across the $(\hat{\eta}_j, \hat{\eta}_\ell)$ plane, showing, within each 1.16×1.0 bin in $\hat{\eta}_j \times \hat{\eta}_\ell$, the ratio of signal to background for (a) $\int d^2\sigma/d\hat{\eta}_j d\hat{\eta}_\ell$, (b) $\bar{F}$, (c) $|F_+|$ and (d) $|F_-|$.

and systematic uncertainties, especially those associated with W-plus-jets.

A. Overview

Figure 8 shows the signal-to-background ratio for the differential cross sections and for the three orthogonal functions that we have defined. Figure 8(b) shows the ratio

$$\frac{\bar{F}^{tbq} + \bar{F}^{tb}}{\bar{F}^{t\bar{t}} + \bar{F}^{Wjj}}$$

according to our simulations. Nowhere in the plane is the signal-to-background ratio of order unity, implying that sensitivity to even small systematic errors is severe. Instead, $\bar{F}$ should be viewed as insensitive to the signal, and therefore mainly useful in helping determine of the size of the backgrounds.

Figures 8(c) and 8(d) show the analogous ratios

$$\left| \frac{F_+^{tbq} + F_+^{tb}}{F_+^{t\bar{t}} + F_+^{Wjj}} \right|$$

and

$$\left| \frac{F_-^{tbq} + F_-^{tb}}{F_-^{t\bar{t}} + F_-^{Wjj}} \right|.$$

(We use absolute values for these two functions, as both numerator and denominator can be negative.) For both F_+ and F_-, the signal-to-background ratio reaches unity in some parts of the plane, implying that, for sufficiently high statistics, *the signal can be measured in these two observables even if the background has relatively large systematic errors.* In both cases, the dark regions are the best ones for the measurement; the other regions should be cut away.

In Fig. 9 the ratio of signal to the square-root of background-plus-signal is plotted, for an integrated luminosity of 3 fb^{-1}. Figure 9(b) shows

$$\frac{\bar{F}^{tbq} + \bar{F}^{tb}}{\frac{1}{2}(\bar{F}^{t\bar{t}} + \bar{F}^{Wjj} + \bar{F}^{tbq} + \bar{F}^{tb})^{1/2}}$$

while Figs. 9(c) and 9(d) show

$$\frac{|F_+^{tbq} + F_+^{tb}|}{\frac{1}{2}(\bar{F}^{t\bar{t}} + \bar{F}^{Wjj} + \bar{F}^{tbq} + \bar{F}^{tb})^{1/2}}$$

and

$$\frac{|F_-^{tbq} + F_-^{tb}|}{\frac{1}{\sqrt{2}}(\bar{F}^{t\bar{t}} + \bar{F}^{Wjj} + \bar{F}^{tbq} + \bar{F}^{tb} + F_+^{t\bar{t}} + F_+^{Wjj} + F_+^{tbq} + F_+^{tb})^{1/2}}.$$

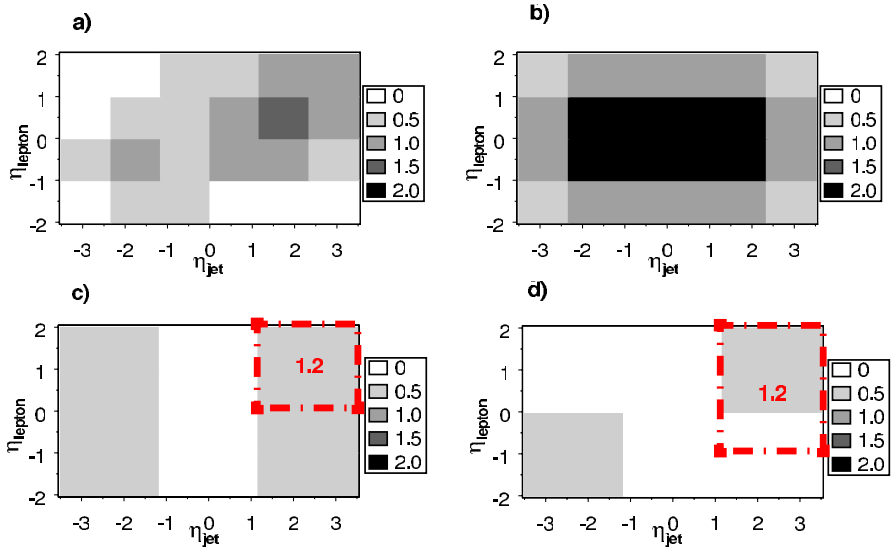

FIG. 9 (color online). Statistical significance $S/\sqrt{B+S}$, for an integrated luminosity of 3 fb^{-1}, across the $(\hat{\eta}_j, \hat{\eta}_\ell)$ plane. Within each 1.16×1.0 bin in $\hat{\eta}_j \times \hat{\eta}_\ell$, the number of events in (a) $\int d^2\sigma/d\hat{\eta}_j d\hat{\eta}_\ell$, (b) $\bar{F}$, (c) $|F_+|$ and (d) $|F_-|$, divided by the appropriate square-root of the number of events in background plus signal (see text for formulas). For each box outlined in dotted-dashed lines, the accompanying number indicates the improved value of $S/\sqrt{B+S}$ that results when the bins inside the box are combined.

(Note that, by construction, $\bar{F} > F_+$ for any distribution, so the expression under the square-root is always positive.) The form of the denominators (and the factors of $1/2$ and $1/\sqrt{2}$) follow from ordinary Gaussian statistics for each bin, using the definitions of $\bar{F}$ and $F_\pm$, Eqs. (3) and (4), as linear combinations of statistically independent bins.[9] The possibility of aggregating bins into larger regions, in which these ratios are of order 1.2, is illustrated using the dotted-dashed outlined bins in the case of F_+ and F_-.

Figures 8 and 9 show that *there is nonzero overlap, for both F_+ and F_-, between the region where the statistics is best and the region where sensitivity to systematic errors is small.* Moreover, we learn that *the two measurements should use data predominantly from certain "windows" in the $(\hat{\eta}_j, \hat{\eta}_\ell)$ plane.* The region where the jet has small $|\hat{\eta}_j|$ has severe problems with statistical background and sensitivity to systematics. The F_+ measurement should be made, roughly speaking, by using all data for $\hat{\eta}_j > 1$ and any $\hat{\eta}_\ell$ (recall that the regions with either or both $\hat{\eta}$ negative are redundant in the case of F_+). Meanwhile the F_- measurement of the signal should be made, roughly, by cutting away all but the region $\hat{\eta}_j > 1$ and $\hat{\eta}_\ell > -1$ (the mirror-symmetric region $\hat{\eta}_j < -1$ and $\hat{\eta}_\ell < 1$ being redundant in this case).[10]

Let us also call attention to the region $\hat{\eta}_j < 1$ and $\hat{\eta}_\ell > 1$ for F_- (ignoring the redundant mirror-symmetric region), away from the center of the plane but also outside the above-mentioned F_- window. Here, the signal is very small. This region is useful for a check of the modeling of the Wj^n background. In the standard model, the contribution of F_- should be consistent with the Wj^n background, with little $t\bar{t}$ and QCD background, and no measurable signal. Verifying that this is the case is both a cross-check on the analysis and a test of the standard model, so the measurement of F_- in this region is also very important.

B. Statistical errors

To explore the size of statistical errors, we have carried out two unsophisticated likelihood analyses, one optimistic, one overly pessimistic. In both cases, we assume 3 fb^{-1} of integrated luminosity. We assume that the shapes of the

[9]For instance, F_-, defined in Eq. (5), satisfies

$$\Delta F_- = \frac{1}{2}\left[\left\{\Delta\left(\frac{d^2\sigma}{d\hat{\eta}_j d\hat{\eta}_\ell}(\hat{\eta}_j, \hat{\eta}_\ell)\right)\right\}^2 + \left\{\Delta\left(\frac{d^2\sigma}{d\hat{\eta}_j d\hat{\eta}_\ell}(-\hat{\eta}_j, -\hat{\eta}_\ell)\right)\right\}^2\right]^{1/2}.$$

Since the two terms in the right-hand side are uncorrelated,

$$\Delta F_- = \frac{1}{2}\left[\left(\frac{d^2\sigma}{d\hat{\eta}_j d\hat{\eta}_\ell}(\hat{\eta}_j, \hat{\eta}_\ell)\right) + \left(\frac{d^2\sigma}{d\hat{\eta}_j d\hat{\eta}_\ell}(-\hat{\eta}_j, -\hat{\eta}_\ell)\right)\right]^{1/2} = \frac{1}{\sqrt{2}}(\bar{F} + F_+)^{1/2},$$

where we used Eqs. (3) and (4). This explains the formula used in Fig. 9(d).

[10]We do not mean to imply that we are specifying the precise form of these windows. Their shapes must be optimized for each particular analysis, based on the associated backgrounds, acceptances, detector issues, and integrated luminosity, as well as improved simulations. Also, no window cut need actually be used; a fit or neural network will naturally weight the data inside the windows heavily, while downweighting the regions outside the windows.

backgrounds and signals are known, and we fit for their normalizations. This is not quite appropriate, since some information about the normalization of the backgrounds will be known from other sources, and since there are significant uncertainties in the shape functions, especially for Wj^n, but we believe the measure we obtain is not far from the truth, and that the correct lesson can be extracted.

In the first likelihood analysis, somewhat optimistic, it is assumed that the shape of the distribution, in the $(\hat{\eta}_j, \hat{\eta}_\ell)$ plane, of the sum of all backgrounds, $t\bar{t}$ plus Wj^n (plus QCD) is known. Similarly, it is assumed that the shape of the signal, tbq plus tb, is known. Maximizing the likelihood as a function of the relative normalization of the signal over background, we find that the measurement of signal over background can be made to a precision of about 40%.

For the second likelihood analysis, it is assumed that the shape of $t\bar{t}$ (plus QCD) is known, the shape of Wj^n is separately known, and, as before, the shape of the signal tbq plus tb is known. We then fit for the two relative normalizations, and find that the measurement of the signal over the background can be made to a precision of about 50%. Meanwhile the measurement of the relative normalizations of the two backgrounds can be performed with an uncertainty of order 15%.

Both of these analyses show that statistical uncertainties are large even for $\mathcal{L} = 3\ \mathrm{fb}^{-1}$, probably comparable to or larger than the theoretical uncertainties in the shapes of the distributions for the signals and backgrounds. To the extent that normalizations of the backgrounds can be pinned down using other information, the situation may be slightly better than this estimate suggests.

Intuitively, from the figures, the region in the center of the $(\hat{\eta}_j, \hat{\eta}_\ell)$ plane plays a large role in fixing the backgrounds, while the large-$\hat{\eta}_j$ region of quadrant B plays a large role in fixing the signal. Indeed, as noted earlier, the statistical significance of the measurement in the signal comes dominantly from the regions outlined in dotted-dashed lines in Fig. 9. Both measurements of F_+ and F_- in these windows have $S/\sqrt{B+S}$ of order 1.2, reasonably consistent with the above likelihood analyses, which were applied to the full distributions over the entire plane.

C. Systematic errors: General comments

We now turn our focus to systematic uncertainties, which mainly stem from an inability to predict and simulate the backgrounds. Our use of the F_- and F_+ functions, we will argue, helps us reduce the sensitivity of the measurement to systematic uncertainties. However, the systematic errors on the backgrounds are very large at present, and must still be reduced.

Earlier in this section, using Fig. 8, we discussed the signal-to-background ratios for the various component functions. We noted that Fig. 8(b) indicates that $\bar{F}$ has a poor signal-to-background ratio. Although it has been argued that systematic uncertainties in predicting $\bar{F}$, using a combination of Monte Carlo simulations and data, will not be large, even a 15% systematic uncertainty is already enough to make $\bar{F}$ problematic for measuring the signal, no matter how good the statistics. Instead, it is better to treat $\bar{F}$ as a measurement which, along with other inputs, is used to help fix the normalizations of the backgrounds.[11]

Figures 8(c) and 8(d) show that the situation with $F_\pm$ is much better, outside of the central region of the plane where the background peaks. F_+ may receive some additional contributions which we have neglected. There is a QCD contribution, which we have argued is probably small, especially in the electron channel. Detector effects and correlations from cuts can also contribute to F_+. This means the signal-to-background ratio in the figure may be an overestimate.[12]

Meanwhile, F_-, the P-odd C-odd observable, has the feature that many potential sources of systematic uncertainty largely cancel. We expect QCD backgrounds, correlations from cuts, and effects of detector cracks or damage to be small in this variable, since they are largely P-even and/or C-even. Many uncertainties in simulating $t\bar{t}$ and even, to a degree, Wj^n will have substantial cancellations. This makes this variable especially compelling.

Eventually, the contribution of $t\bar{t}$ to $F_\pm$ should not have large systematic uncertainties. It is important to keep in mind that there are non-negligible effects, including a parity asymmetry [21], that appear at one loop. Next-to-leading-order calculations of the distributions in the $(\hat{\eta}_j, \hat{\eta}_\ell)$ plane are needed. These can be computed with small errors and should allow a precise background subtraction of $t\bar{t}$. At present, however, their absence causes a substantial uncertainty in this subtraction.

For both F_+ and F_-, the most serious systematic problems stem from the uncertainties in the shape of Wj^n. For this reason, we will now present an extended discussion of this background.

D. Systematic errors: Predicting W-plus-jets

One might hope that systematic errors stemming from the Wj^n background could be greatly reduced, for the variables $F_\pm$, using a sideband subtraction in the variable m_t. This approach, analogous to the method suggested in [1] for use in a counting experiment, would allow a background subtraction without much theoretical input.

[11] In the counting experiment discussed in Sec. III A, one essentially uses $\bar{F}$ to make the measurement, though with much tighter cuts than the ''relaxed cuts'' employed here.

[12] Moreover, it appears in the figure to be somewhat better than we actually expect it to be, perhaps by as much as a factor of 2 in any given bin. In our simulations there is some accidental cancellation between the Wj^n and $t\bar{t}$ contributions to F_+; see Fig. 6. Unfortunately our knowledge of these two backgrounds is too poor to be certain that this cancellation is robust, and moreover statistical fluctuations may also ruin the cancellation.

However, any such attempt will run into the same issues that cause this method to fail for a counting experiment. In Fig. 2, we showed that the application of aggressive cuts to bring Wj^n under statistical control simultaneously makes a sideband analysis problematic by deforming the shape of the Wj^n background, leaving it nonmonotonic across the region needed for the sideband analysis. To do a subtraction therefore requires that the shape of the background, after cuts, to be accurately *predicted,* using a combination of theory, Monte Carlo and data. Unfortunately, prediction of any aspects of Wj^n, especially with one or more b-tagged jets, is very difficult indeed. As we will now discuss in detail, we believe that Monte Carlo results for the Wj^n sample with b tags cannot currently be trusted at the level that is likely to be needed.

The shape of the Wj^n background is plagued with the usual concerns about the inability of PYTHIA or HERWIG [22] to reliably generate the correct pattern of additional radiated jets; for recent discussion, see [23]. This situation will improve over time as a consistent set of NLO tools becomes available [24]. But to this and other typical problems, which are known to be issues in many processes, we must add some others which are specific to the sample with one (and only one) b-tagged jet. (Note the event samples used in our single-top analysis above include events with one *or more* tagged jets; however, the problems detailed in this section are somewhat less severe for samples with more than one tag.)

A striking feature of this sample is that only a moderate fraction of Wj^n events with a single-tagged jet (and passing our cuts) actually have a short-distance b-quark parton present in the hard-scattering process; those that do are mainly $Wb\bar{b}$. Instead, the single b-tag sample is composed of many different contributions. This situation is made explicit in Table VIII, which shows the cross section for various Wjj channels, before and after the requirement of a single tag, and Table IX, which lists the relative contributions from the various Wjj channels to the single-tagged sample. As with all our simulations, these contributions were calculated using MADGRAPH [14] to evaluate the parton cross sections, PYTHIA [15] to provide showering and hadronization, and PGS [16] as the detector simulation of jet identification and tagging. The total Wj^n cross section (before tagging) was normalized to the one-loop result [18]. The basic cuts of Table I were used; note that an untagged jet was also required in the event. The labels Wxy in the first column indicate the perturbative final state at tree level (as evaluated in MADGRAPH); here q represents a light quark or antiquark ($u, d, s, \bar{u}, \bar{d}, \bar{s}$). In Table VIII, the cross section for each channel, before and after the requirement of a single-tagged jet, is shown; also shown is the ratio of after tagging to before tagging (i.e., the fraction of each channel containing a single-tagged jet). In Table IX, the three central columns divide the tagged events by whether the jet that was tagged in the event contained a bottom hadron, a charm hadron, or no heavy flavor. For example, the entry in the b-jet column of the Wqg row indicates that 11% of the entire Wj^n single-tag sample arises from Wqg events ($q = u, d, s, \bar{u}, \bar{d}, \bar{s}$) in which the tagged jet contains bottom hadrons, mainly due to the splitting $g \to b\bar{b}$ in the parton showering. In determining the entries in Table IX from those in Table VIII, we have attempted to correct for any double counting. For example, we have subtracted the contribution to the b-jet Wgg entry

TABLE VIII. The cross sections for Wj^n events with at least two jets, before and after tagging of one of the jets. Each row refers to the combination of processes with the tree-level final state shown in the leftmost column. The Wcq and Wcg entries include both c and $\bar{c}$ quarks; the symbol q stands for $u, d, s, \bar{u}, \bar{d}, \bar{s}$. After showering, hadronization and jet identification, and application of the cuts in Table I, the resulting cross section is indicated in the next column. The cross section corresponding to a single-tagged jet is shown in the third column, while the last column shows the ratio of the previous two. Simulation uncertainties in these numbers are of order $\pm 1\%$–3%, except for the Wqg and Wqq channels, with uncertainties of order $\pm 4\%$ and $\pm 8\%$, respectively. The (much larger) systematic uncertainties are discussed in the text.

Wjj channel	σ (before tags) [fb]	σ (after tags) [fb]	Fraction tagged
Wqq	16470	192	1%
Wqg	32000	732	2%
Wgg	14760	484	3%
Wcq	3200	318	10%
Wcg	2240	238	11%
$Wc\bar{c}$	600	104	17%
$Wb\bar{b}$	496	224	45%
Total	69766	2291	3%

TABLE IX. A budget of the sample of Wj^n events with a single-tagged jet (and containing at least one untagged jet), constructed as in the previous table. Each entry shows the fraction of the sample containing a single-tagged jet that was generated from the tree-level process labeling the row (as in Table VIII) and in which the tagged jet was of the class labeling the column. The Wcq and Wcg entries include both c and $\bar{c}$ quarks; the symbol q stands for $u, d, s, \bar{u}, \bar{d}, \bar{s}$. The entries in this table are subject to an additive uncertainty of $\pm 1\%$–2%. The (much larger) systematic uncertainties are discussed in the text.

Wjj channel	b-jet	c-jet	Non-b/c-jet	Total
Wqq	2%	1%	6%	9%
Wqg	11%	8%	14%	33%
Wgg	7%	5%	5%	17%
Wcq	0%	14%	1%	15%
Wcg	1%	10%	0%	11%
$Wc\bar{c}$	0%	5%	0%	5%
$Wb\bar{b}$	10%	0%	0%	10%
Total	31%	43%	26%	100%

(from $g \to b\bar{b}$ in the parton shower) arising from events with kinematic configurations already present in the b-jet $Wb\overline{b}$ entry. This leads to small corrections, of order the stated $\pm 2\%$ uncertainty, in several of the entries.

The central feature of these tables is the multitude of contributions to the tagged sample, all of similar magnitude. As clearly visible in Table VIII, the large parton-level cross sections for the light-quark and gluon processes are reduced by a low tagging fraction, while the much smaller tree-level heavy-quark cross sections are subject to much larger tagging rates. Consequently, within the single-tag sample, all such partonic processes end up contributing at roughly at the same level, as is clear in the ''after tagging'' column of Table VIII, and in the ''total'' column of Table IX. The breakdown of the resulting single-tag sample by the fraction of the sample for which the tagged jet is a bottom jet, a charm jet, or a jet without heavy flavor is visible in Table IX; as the numbers in the bottom row indicate, all contributions are again of the same order.

The simulations were performed with statistics sufficient to ensure such that each entry in Table IX is subject to an *additive* statistical uncertainty of order 2%–3%. However, the systematic uncertainties in the simulations are much larger, due to a host of important physical and technical issues. Since these systematic uncertainties are the most important obstacle to an accurate background estimate, we now discuss them in detail.

Consider first the uncertainties in the basic event simulation. A precision simulation of the Wjj sample cannot, at present, be carried out. The parton-level theoretical computation of the differential cross section for W-plus-two or more high-p_T jets has been advanced in recent years: Wjj has been calculated to next-to-leading order, and $Wjjj$ is known at leading order [25,26]. While the program ''MCFM'' [9,18] can provide accurate next-to-leading-order parton-level cross sections, no event generator valid at next-to-leading order currently exists. Consequently, there is at present no possibility of simulating this background without significant theoretical uncertainties, though this situation will improve with the advent of a next-to-leading-order Monte Carlo program MC@NLO [24] for this process.

But even when a next-to-leading-order event generator becomes available, there are serious questions concerning showering algorithms that must be addressed. The main problem is associated with the splitting of gluons to heavy quarks (or more precisely, with tuning PYTHIA or HERWIG to simulate correctly the process in which a partonic gluon generates one or more jets containing charm or bottom mesons). As illustrated in Table IX, a substantial fraction of the single-tagged sample, of order 33% in our simulations, originates from this process. It is not known with confidence how often gluons at short distance lead to jets with charm or bottom hadrons, or how often this process leads to two jets rather than one. Studies on this issue that compare data from LEP [27] and Tevatron [28] with QCD expectations [29] and with PYTHIA and HERWIG suggest that there is no serious disagreement. However, this conclusion rests on the substantial uncertainties (of order 30%) in both the experimental and the theoretical results. Generally the perturbative predictions (including resummed logarithms) and the Monte Carlo results (with default parameters) tend to underestimate the observed rates of heavy-flavor production in parton showers. Also, the rough agreement speaks mainly to overall rates of heavy-flavor production, not to charm-to-bottom ratios or kinematic distributions. Given the large sensitivity of the single-tag sample to these effects, it appears that a sizable uncertainty in both the normalization and shape of the Wj^n background arises from this source.

A related issue is the role of uncertainties in parton distribution functions. The single-tag sample receives significant contributions from events with charm in the final state that arise from initial states containing nonvalence partons. Examples include processes such as $u\bar{s} \to W^+ u\bar{c}$, $g\bar{d} \to W^+ g\bar{c}$, and $\bar{u}\,\bar{d} \to W^+ \bar{u}\,\bar{c}$. These processes have varying shapes and rates, and depend on the poorly determined parton distribution functions.

Even if one could precisely simulate these events, there is still the issue of determining the efficiencies for tagging of jets with bottom or charm quarks, and for mistagging of jets with neither. This must be done as a function of p_T and pseudorapidity. As indicated in Table IX, each of these tagging processes plays a comparable role in determining the Wjj sample. Uncertainties in tagging functions thus lead to uncertainties in the shape of Wj^n which cannot be ignored. An especially serious issue is that the ratio of c to b tagging rates is currently extracted not from data but from Monte Carlo programs, which, among other problems, are dependent upon the correct modeling of gluons splitting to heavy flavor.

Clearly, it would be beneficial to decrease the mistagging and charm-tagging efficiencies *relative* to the efficiency for tagging of bottom jets. This would improve the ratio of the signal to the Wj^n background, reducing sensitivity to systematic errors, and also would directly suppress some of the main sources of uncertainty in the prediction of Wj^n. However, tuning the b-tagging algorithm to improve purity of the sample generally comes at the cost of a small reduction in the b-tagging efficiency, which in turn decreases the signal and increases statistical errors. The right balance between these competing issues is detector- and luminosity-dependent and must be left to the experimental collaborations.

Another related concern is the subtle linkage between b tagging, jet definitions, and gluon splitting. When a gluon splits to two heavy quarks, the probabilities to obtain one tagged jet, two tagged jets, or one tagged and one untagged jet depend upon all three issues. This may mean that the separation of the samples with one tag versus two tags is

unstable and not well predicted by theory. It is for this reason that we have chosen to consider samples with one or more tagged jets. Alternatively, one might require strong angular separation between tagged jets in order to retain predictivity.

A further issue involves the potentially large sensitivity of the shape and normalization of the multichannel Wj^n background to the cuts used to bring backgrounds under control. In particular, both the H_T and m_t cuts used in the current analysis and in recent experimental papers [6] reshape the Wj^n background. While it has been shown [1] that jet vetoes are effective at substantially reducing the size of $t\bar{t}$ backgrounds, we would argue against the use of this approach. The systematic error which jet vetoes introduce into the prediction of the Wj^n background has not been quantified, but we expect it is prohibitively large. We are especially concerned about the requirement of two-and-only-two jets employed in [6]. We believe this will make the prediction of Wj^n unreliable, both because of problems with QCD corrections, and because of failures to correctly model the rate at which gluons in parton-level processes lead to zero, one or two jets containing heavy flavor. There is at present no consensus as to the safest method for reducing $t\bar{t}$, or, for that matter, Wj^n. We would like to argue strongly that this is a very important problem, which our methods simultaneously mitigate and highlight. On the one hand, neither the F_+ nor the F_- distribution is strongly sensitive to the overall size of $t\bar{t}$. Consequently, the use of a strong jet veto or harsh H_T cut is unnecessary, and indeed unjustified to the extent systematic errors in Wj^n become larger as a result. Our approach would instead prefer a method which cuts $t\bar{t}$ less severely, and in a safer fashion, such that theoretical errors in predicting Wj^n remain small. The H_T cut that we use is, we believe, safer, being a cut on a more inclusive variable; whereas the splitting of one jet into two, due to a fluctuation or a changed jet algorithm, affects a jet veto in a dangerous way, this is not so for an H_T cut. But this safety is only relative; there are problems with jets moving above or below the minimum p_T required for a jet to be included in the variable H_T that we defined. In any case, our method requires less focus on reducing the size of $t\bar{t}$ and more focus on keeping Wj^n as small and as predictable as possible.

The presence of so many sources of uncertainty (some of which are experimental, others theoretical, and still others a mixture) makes it very difficult to estimate precisely the current systematic uncertainty in the shape of the single-tag Wj^n sample. Presumably this uncertainty is larger than 20% and less than 100%. Our best guess for these uncertainties are of order 50%, based on our own experience with adjusting the parameters in tagging rates, gluon splitting to heavy quarks, and so on; but this value is not to be taken very seriously. In any case, the important issue is how large these uncertainties will be in a few years when the measurement is actually performed, and this is even more difficult to ascertain. For this reason, we have not attempted to be more quantitative, limiting ourselves instead to a list of action items where improvements are necessary.

The large uncertainties are the main reason that we have only simulated Wjj, neglecting $Wjjj$. Our studies of $Wjjj$ using MADGRAPH, PYTHIA and PGS indicate that this is justified. Rather than give detailed results backing this claim, we now present a simple-minded argument in its favor. If the shape of $Wjjj$ in the $(\hat{\eta}_j, \hat{\eta}_\ell)$ plane is roughly the same as in Wjj, our results are essentially unaffected (since the overall normalization of Wj^n will largely be extracted from data). But suppose instead that $Wjjj$ has a larger asymmetry, which would increase the background to the single-top signal. As indicated in Table VII, the asymmetry between quadrants B and C for Wjj is of order 13%; let us imagine that $Wjjj$ has an asymmetry which is twice as large, of order 26%. Suppose in addition that the cross section for $Wjjj$ with one or more tagged jets is of order $1/3$ of the total cross section for Wj^n with $n \geq 2$. (Naive estimates based on leading-order calculations [26] would suggest something closer to $1/5$.) With these assumptions, the resulting asymmetry of the $Wjj + Wjjj$ sample is increased by $4/3$, to about 17%. Thus, even in this situation, the $Wjjj$ contribution to the asymmetry variables is considerably smaller than our estimate for the current uncertainty in the Wjj asymmetry. In short, any apparent benefit from including $Wjjj$ in the present study would be illusory. Of course, for the actual measurement, this is not so; every effort must be made to reduce the errors on Wj^n to the point that a precise calculation of $Wjjj$ is required.

The various uncertainties also imply that any method used to extract and model the Wj^n background will have to pass many cross-checks. One important consistency check can be carried out by calibrating Monte Carlo simulations using both Wj^n and the process Z-plus-jets ("Zj^n"), where the Z decays leptonically. The processes Wj^n and Zj^n do not have the same shapes, rates, and heavy-flavor content, so one cannot directly take ratios of distributions or even of overall cross sections. However, the overall kinematics of Zj^n is similar to Wj^n, and is similarly sensitive to all of the above-mentioned issues. For these reasons, we expect that matching a Monte Carlo to the rates and shapes of the Zj^n and Wj^n distributions from data, with zero, one or two tagged jets, will significantly reduce systematic uncertainties. Such matching will require adjusting parameters which affect the splitting of gluons to heavy flavor, and adjusting tagging functions.[13] It would also be very helpful if direct measurements of the bottom-content, charm-content, and non-b/non-c content of the various single-

[13]We believe there should be enough Z-plus-jets data, with 3 fb^{-1} of integrated luminosity, for this analysis to be carried out.

and double-tagged Wj^n and Zj^n samples could be carried out, even with low precision and confidence. This would allow direct tests of numerous Monte Carlo predictions that at present have very large uncertainties. Studies of the Zb to Zj ratio have been performed [30], but the challenging study of Zc separately should also be considered as statistics improve.

In summary, determining the single-tagged Wj^n cross section will require a carefully crafted combination of theory, data, and theory-optimized Monte Carlo.

V. SUMMARY

A counting experiment for discovery of electroweak single-top production appears very challenging. In an effort to improve the situation, we have explored the possibility of using the distinctive shape of this process to separate it from background. We use as observables the pseudorapidity η_j of the leading-p_T untagged jet and the pseudorapidity η_ℓ of the charged lepton, weighted by the charge of the lepton Q_ℓ. (One of these variables was already used in [4,6].) Considering the distributions of signal and background in the $(\hat{\eta}_j, \hat{\eta}_\ell)$ plane (where $\hat{\eta}_j = Q_\ell \eta_j$ and $\hat{\eta}_\ell = Q_\ell \eta_\ell$), we have noted that the distributions for $t\bar{t}$ and for QCD are largely symmetric, while that of the signal is not; the Wj^n (W-plus-jets) background is intermediate between them. Constructing functions $\bar{F}, F_\pm$, defined in Eqs. (3)–(5), which have various symmetry properties, we have shown that the statistical and systematic errors in the functions $F_\pm$, which are orthogonal to the function $\bar{F}$ that would be used in a counting experiment, can be brought close to reasonable size without using extreme cuts.

Here is a summary of key ingredients that went into this analysis, as well as a list of elements which we did not account for, and a few of our crucial assumptions:

(i) All cross sections ($tb, tbq, t\bar{t}, Wbb, Wcc, Wcq, Wcg, Wqq, Wqg, Wgg$) were calculated at tree level.

(ii) These were then normalized to theoretical calculations at next-to-leading order. For Wj^n, only the sum of all channels was normalized in this way.

(iii) All processes were run through PYTHIA, to simulate showering, and through the detector simulation PGS.

(iv) We imposed hard p_T cuts on the leading tagged and untagged jets, and required that a top quark be reconstructible from the lepton, tagged jet, and missing energy. We applied an H_T cut aimed at reducing $t\bar{t}$; we did not use a jet veto.

(v) p_T dependence of tagging fractions was accounted for, with the maximal tagging rates at high p_T for $b, c, q/g$ being taken as 50%, 15%, 1%. The details of the analysis are sensitive to these numbers, as well as to the rate for hard gluons at leading order to evolve into jets containing heavy flavor.

(vi) We did not attempt to simulate QCD events. Instead, we argued QCD effects are (in theory) sufficiently symmetric in shape and (from D0 data) sufficiently small in rate that their contributions to all useful observables can be neglected (with the possible exception, for D0, of F_+ in the muon channel).

(vii) We discussed important shape effects on $t\bar{t}$ at next-to-leading order [21], which we expect to be about 10% or less, but did not simulate them.

It should be noted that we have not optimized our cuts to improve efficiencies and reduce systematics in a rigorous way. Indeed, it would not be too useful to do so, since the optimization is a moving target, depending on integrated luminosity, on tagging rates and other detector details, on next-to-leading-order shapes, and on Monte Carlo assumptions. We believe, therefore, that our results could be improved upon through such an optimization, though this is only worth doing in a concrete analysis. We also have not explored whether other methods of reconstructing m_t might be more effective, or whether variables other than H_T might be better as far as both statistics and systematics. This is certainly something that should be done as the integrated luminosity increases.

Moreover, there is a natural extension of our method which we did not consider, but which should be explored if the integrated luminosity becomes sufficiently high. Our observable F_+ focused on lepton-jet pseudorapidity correlations, but as we pointed out, these correlations can have two sources: *inherent* correlations in the rest frame of the tbq system (or of the top quark itself) and correlations which are *induced* by the boosting of these frames into the lab frame. Both of these effects are present in the signal, and they add coherently to give a large contribution to F_+. One could imagine measuring the two effects separately. This could potentially allow for even greater separation of signal from background.

We conclude with a summary of and comments upon what we see as the main lessons of our analysis.

(1) Our method largely removes $t\bar{t}$ and QCD events from the observables $F_\pm$, making extreme cuts on $t\bar{t}$ unnecessary, and focusing attention on Wj^n as the main background. While the statistical fluctuations from $t\bar{t}$ are still important, they are of less concern than systematic errors on Wj^n, since the former scale as the square-root of the $t\bar{t}$ rate, while the latter scale linearly with the Wj^n rate. Moreover, the $t\bar{t}$ background is much more safely calculated and simulated, and will be, in the end, easier to remove. In our approach to single top, *one should not cut hard on $t\bar{t}$ if doing so causes the systematic uncertainties in Wj^n to increase substantially.* In particular, this argues against the use of a severe jet veto, in which events with more than two observed jets are discarded.

(2) The method that we have introduced requires that the shapes of $t\bar{t}$ and Wj^n be properly modeled, but it also provides for cross-checks. The function $\bar{F}$ measures the

total background, and, assuming $t\bar{t}$ can be determined using other samples, this allows a measurement of the total Wj^n background. Meanwhile, there is a region in the $(\hat{\eta}_j, \hat{\eta}_\ell)$ plane where the F_- function gets small contributions from both signal and from $t\bar{t}$. This region allows a check of whether the shape of Wj^n has been correctly understood, as well as being interpretable as a worthwhile test of the standard model itself.

(3) The thorniest problem obstructing the measurement of single top is understanding the shape of the Wj^n background. This is a large and irreducible background which must be subtracted from the signal, even in the context of the methods we proposed here. This subtraction could be done directly, using our cuts, but this requires some prediction of the shape in the $(\hat{\eta}_j, \hat{\eta}_\ell)$ plane. Alternatively, a sideband analysis around the m_t window cut could be applied to $F_\pm$, in appropriate pseudorapidity windows, but this too requires prediction of the effect of cuts on the distribution of Wj^n in m_t and pseudorapidity.

While theory, Monte Carlo and data all can, and must, assist with these subtractions, many different types of uncertainties plague the sample with a single b tag (and therefore the sample with one or more b tags), making it unclear how to bring all the available resources together. We believe that a dedicated study, examining the rates, shapes, and flavor content (especially of bottom versus charm) of both Wj^n and Zj^n, with zero, one and two tagged jets, will be necessary. This will require a blend of multiple measurements, theoretically precise predictions, and careful tuning and cross-checking of Monte Carlo simulations. Since this issue affects many other measurements, including the Higgs search and numerous beyond-the-standard-model searches, we view this as a very high priority.

(4) There are very few paths toward reducing the Wj^n background relative to the signal. One clear need is to decrease the mistagging rate and charm-tagging efficiency while maintaining or increasing the b-tagging efficiency; this would both improve signal to background and reduce some of the uncertainties that make it difficult to model the background. Another important step would be taken if the resolution in reconstructing the top-quark mass from the b, lepton and missing energy could be improved. This would allow a narrowing of the m_t window cut, which would both reduce the size of the Wj^n background and reduce experimental and theoretical uncertainties in any sideband analysis. Beyond this, one would need to consider more radical ideas, such as finding methods which could, on average, differentiate bottom jets from charm jets, bottom jets from antibottom jets, and/or jets formed by short-distance heavy quarks from jets formed by gluons that have split into roughly collinear heavy-quark pairs.

In conclusion, our method for extracting single top confirms that one can use the distinctive shape of the signal to reduce backgrounds more effectively than in a pure counting experiment. However, we also find that backgrounds are much worse than was once thought. Improvements in (mis)tagging rates and in the understanding thereof, careful modeling of W-plus-jets cross-checked against both theory and data, and more theoretically trustworthy techniques for cutting away backgrounds will all be necessary for a robust measurement of single-top production. The single-tag W-plus-jets background, in particular, represents a challenge that the whole community must meet head-on.

ACKNOWLEDGMENTS

We thank our colleagues G. Watts, A. Garcia-Bellido, T. Gadfort, A. Haas, H. Lubatti, and T. Burnett for many useful conversations and direct assistance. We also thank K. Ellis, T. Junk, S. Mrenna, T. Stelzer, Z. Sullivan and E. Thomson for extended conversations and for their thoughtful comments and criticisms. This work was supported by U.S. Department of Energy Grants No. DE-FG02-96ER40956 and No. DOE-FG02-95ER40893, and by the Alfred P. Sloan Foundation.

[1] T. Stelzer, Z. Sullivan, and S. Willenbrock, Phys. Rev. D **58**, 094021 (1998).

[2] For other previous studies of single top-quark production see S. S. Willenbrock and D. A. Dicus, Phys. Rev. D **34**, 155 (1986); C. P. Yuan, Phys. Rev. D **41**, 42 (1990); S. Cortese and R. Petronzio, Phys. Lett. B **253**, 494 (1991); R. K. Ellis and S. J. Parke, Phys. Rev. D **46**, 3785 (1992); D. O. Carlson and C. P. Yuan, Phys. Lett. B **306**, 386 (1993); T. Stelzer and S. Willenbrock, Phys. Lett. B **357**, 125 (1995); Future Electroweak Physics at the Fermilab Tevatron: Report of the TeV-2000 Study Group, edited by D. Amidei and R. Brock (Fermilab Report No. FERMILAB-Pub-96/082, 1996); A. P. Heinson, A. S. Belyaev, and E. E. Boos, Phys. Rev. D **56**, 3114 (1997).

[3] T. Tait and C.-P. Yuan, Phys. Rev. D **63**, 014018 (2001).

[4] D. Acosta *et al.* (CDF Collaboration), Phys. Rev. D **69**, 052003 (2004); **65**, 091102 (2002).

[5] B. Abbot *et al.* (D0 Collaboration), Phys. Rev. D **63**, 031101 (2001); V. M. Abazov *et al.* (D0 Collaboration), Phys. Lett. B **517**, 282 (2001).

[6] D. Acosta *et al.* (CDF Collaboration), Phys. Rev. D **71**, 012005 (2005).

[7] See ''Search for single top production'' at http://www-d0.fnal.gov/Run2Physics/WWW/results/prelim/TOP/T09/T09.pdf.

[8] B. W. Harris *et al.*, Phys. Rev. D **66**, 054024 (2002).

[9] J. Campbell, R. K. Ellis, and F. Tramontano, Phys. Rev. D **70**, 094012 (2004).
[10] Q.-H. Cao, R. Schwienhorst, and C.-P. Yuan, Phys. Rev. D **71**, 054023 (2005); Q.-H. Cao and C.-P. Yuan, Phys. Rev. D **71**, 054022 (2005).
[11] D. Acosta *et al.* (CDF Collaboration), Phys. Rev. Lett. **93**, 142001 (2004).
[12] J. Campbell and J. Huston, hep-ph/0405276.
[13] S. Mrenna and P. Richardson, J. High Energy Phys. 05 (2004) 040.
[14] F. Maltoni and T. Stelzer, J. High Energy Phys. 02 (2003) 027.
[15] T. Sjostrand, P. Eden, C. Friberg, L. Lonnblad, G. Miu, S. Mrenna, and E. Norrbin, Comput. Phys. Commun. **135**, 238 (2001). Version 6.217 was used in this analysis.
[16] J. S. Conway *et al.*, hep-ph/0010338; see also http://www.physics.ucdavis.edu/~conway/research/software/pgs/pgs.html.
[17] M. Cacciari *et al.*, J. High Energy Phys. 04 (2004) 068.
[18] J. M. Campbell and R. K. Ellis, Phys. Rev. D **62**, 114012 (2000); **65**, 113007 (2002).
[19] See http://www-cdf.fnal.gov/physics/new/top/public/btag/ for relevant graphs.
[20] See, for example, G. Mahlon, hep-ph/0011349; hep-ph/9811219; Gregory Mahlon and Stephen Parke, Phys. Lett. B **476**, 323 (2000); Phys. Rev. D **55**, 7249 (1997).
[21] J. H. Kuhn and G. Rodrigo, Phys. Rev. D **59**, 054017 (1999); see also M. Fischer *et al.*, Phys. Rev. D **65**, 054036 (2002).
[22] G. Marchesini *et al.*, Comput. Phys. Commun. **67**, 465 (1992); HERWIG 6.5, G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour, and B. R. Webber, J. High Energy Phys. 01 (2001) 010; hep-ph/0210213.
[23] Z. Sullivan, Phys. Rev. D **70**, 114012 (2004).
[24] S. Frixione and B. R. Webber, J. High Energy Phys. 06 (2002) 029; S. Frixione, P. Nason, and B. R. Webber, J. High Energy Phys. 08 (2003) 007; S. Frixione and B. R. Webber, hep-ph/0506182.
[25] F. A. Berends, H. Kuijf, B. Tausk, and W. T. Giele, Nucl. Phys. **B357**, 32 (1991).
[26] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, J. High Energy Phys. 07 (2003) 001.
[27] A. Ballestrero *et al.*, hep-ph/0006259; see also K. Abe *et al.* (SLD Collaboration), Phys. Lett. B **507**, 61 (2001); R. Barate *et al.* (ALEPH Collaboration), Phys. Lett. B **434**, 437 (1998); P. Abreu *et al.* (DELPHI Collaboration), Phys. Lett. B **401**, 163 (1997); G. Abbiendi *et al.* (OPAL Collaboration), Eur. Phys. J. C **18**, 447 (2001).
[28] D. Acosta *et al.* (CDF Collaboration), Phys. Rev. D **69**, 072004 (2004).
[29] See, for example, M. H. Seymour, Nucl. Phys. **B436**, 163 (1995); D. J. Miller and M. H. Seymour, Phys. Lett. B **435**, 213 (1998); M. L. Mangano and P. Nason, Phys. Lett. B **285**, 160 (1992); M. L. Mangano, Nucl. Phys. **B405**, 536 (1993).
[30] See "Measurement of the Cross Section Ratio of $Z + b/Z + j$" at http://www-d0.fnal.gov/Run2Physics/WWW/results/HIGGS/higgs.htm.