Detecting very massive top quarks at the Fermilab Tevatron

H. Baer
Physics Department, Florida State University, Tallahassee, Florida 32306

V. Barger
Physics Department, University of Wisconsin, Madison, Wisconsin 53706

J. Ohnemus
Physics Department, Florida State University, Tallahassee, Florida 32306

R. J. N. Phillips
Rutherford Appleton Laboratory, Chilton, Didcot, Oxon, England
(Received 19 January 1990)

We consider the dilepton-dijet signal from very heavy $t\bar{t}$ pair production with semileptonic $t \to b\ell\nu$ decays via real W bosons in the standard model. We evaluate the principal backgrounds, arising from the direct production of W^+W^- plus two jets or Z plus two jets (with $Z \to \tau\bar{\tau} \to \ell\bar{\ell} + 4\nu$), and show that they can be cleanly separated by suitable cuts. We demonstrate how the top-quark mass m_t can be estimated via dynamical distributions or event reconstructions or event rate and calculate the statistical correlations between different mass estimators.

I. INTRODUCTION

The top quark t is an essential ingredient of the standard model (SM) and its mass is an important parameter. It has so far eluded discovery[1,2] and must therefore be rather heavy: the most stringent present limit is $m_t > 77$ GeV from the Collider Detector at Fermilab (CDF) experiment[1] assuming SM decay modes. On the other hand, the consistency of SM radiative corrections places an upper bound[3] variously estimated to be of order $m_t \lesssim 150$–250 GeV. The whole of this mass range can be explored in the next five years at the Fermilab Tevatron $p\bar{p}$ collider, given the planned luminosities of several hundred inverse picobarns. Heavy top searches at the Tevatron, based on single-lepton and dilepton signals from SM decays, have received much theoretical attention;[4-10] most emphasis has been on the lighter end of the m_t range where the principal backgrounds come from $b\bar{b}$ and W+jets production and are well known. As m_t increases, the single-lepton backgrounds become more serious and dilepton signals appear more promising.[5] In the present work we consider isolated dilepton-dijet signals from top quarks in the upper part of the mass range, where the most difficult background comes from W^+W^- + dijet production; the latter has hitherto only been estimated approximately;[5] but we now evaluate it explicitly. The other serious background here[5] comes from Z + dijet production, with $Z \to \tau\bar{\tau} \to \ell\bar{\ell} + 4\nu$; it too has not been evaluated explicitly until now. We show how these backgrounds can be removed by suitable cuts (along with other backgrounds). We also discuss the determination of m_t from a small number of eventual candidate events and propose a new method based on event reconstruction as an alternative to dynamical distributions or maximum-likelihood methods.[11] Altogether we study six different observables that can be used to estimate m_t; we evaluate their covariance matrix to exhibit their correlations, and discuss the possibilities for combining different m_t estimates.

II. DILEPTON-DIJET SIGNALS

We consider the hadroproduction of heavy $t\bar{t}$ pairs, followed by semileptonic decays via real W bosons:

$$p\bar{p} \to t\bar{t}X, \quad t \to bW^+ \to b\ell_1^+\nu_1, \quad \bar{t} \to \bar{b}W^- \to \bar{b}\ell_2^-\bar{\nu}_2 . \tag{1}$$

This gives typically two isolated high-p_T leptons $\ell_1^+\ell_2^-$ ($\ell_1\ell_2 = ee$, $e\mu$, μe, or $\mu\mu$) plus two high-p_T b jets, plus possible additional jets from QCD radiation from the initial and final states. We calculate the initial $t\bar{t}$ production as in Ref. 5, from $2 \to 3$ QCD subprocesses[12] with a cutoff at small $p_T(t\bar{t})$ chosen to give the correct total cross section[13] to order α_s^3; thus some initial and final-state QCD radiation is explicitly included. We calculate t and $\bar{t}$ decays in the usual spectator approximation, neglecting spin-correlation effects that are known to be small.[14] We coalesce partons with $\Delta R = [(\Delta\eta)^2 + (\Delta\phi)^2]^{1/2} \leq 0.7$ into jets and identify summed parton transverse momentum p_T with jet E_T; here $\eta = \ln\tan(\theta/2)$ is pseudorapidity, θ and ϕ are polar and azimuthal angles with respect to the beam axis. We find that the following set of cuts progressively optimizes the $t\bar{t}$ signal-to-background ratio

for very heavy m_t (of order 200 GeV):

$$p_T(l) > 15 \text{ GeV}, \quad E_T(j) > 15 \text{ GeV}, \quad \not{p}_T > 20 \text{ GeV},$$

$$|\eta(l)| \leq 3.0, \quad |\eta(j)| \leq 2.5, \tag{2a}$$

$$\sum_c E_T \leq 3 \text{ GeV in cone } \Delta R \leq 0.4 \text{ about lepton momentum},$$

$$30^\circ < \Delta\phi(l_1 l_2) < 150^\circ;$$

$$n_j \geq 2\,, \tag{2b}$$

$$m_T(\ell\bar{\ell}, \not{p}_T) > M_Z\,, \tag{2c}$$

$$E_T(j) > 30 \text{ GeV}\,. \tag{2d}$$

Here j denotes jet, $\not{p}_T$ denotes missing p_T, the third line defines lepton isolation ($\sum_c$ denotes summation of hadronic E_T within a cone), and the cut on azimuthal difference $\Delta\phi$ helps discriminate against $b\bar{b}$, $c\bar{c}$, and Drell-Yan backgrounds.[4,5,15] The cluster transverse mass $m_T(c, \not{p}_T)$ is defined by[16]

$$m_T^2(c, \not{p}_T) = [(m_c^2 + p_{cT}^2)^{1/2} + |\not{p}_T|]^2 - (\mathbf{p}_{cT} + \not{p}_T)^2, \tag{3}$$

and Eq. (2c) refers to the dilepton cluster $c = \ell_1\ell_2$.

The acceptance cuts in Eq. (2a) were designed[5] to separate out the $t\bar{t} \to WW$ dilepton signal for top masses $m_t > M_W$; in Ref. 5, distributions and tables which motivate these cuts and provide efficiency information may be found. It was found in Ref. 5 that the dilepton signal remains above backgrounds out to m_t of 200 GeV for jet multiplicity $n_j \geq 2$, which motivates Eq. (2b). The remaining significant backgrounds are due to $Z + n$ jet production with $Z \to \tau\bar{\tau} \to \ell\bar{\ell} + 4\nu$ decays and $WW + n$ jet production with leptonic decays of both W bosons. We find that the transverse-mass cut in Eq. (2c) eliminates the $Z \to \tau\bar{\tau}$ background at little cost to the top signal. The more restrictive jet E_T cut in Eq. (2d) can then be used to suppress the remaining WW background at slight cost to the signal event rate.

We compute production, decay, and acceptance effects by Monte Carlo methods, assuming Duke-Owens set 1 parton distributions[17] evolved up to $Q^2 = \hat{s}$, the subprocess c.m. energy squared, at an overall c.m. energy $\sqrt{s} = 2$ TeV corresponding to future Tevatron running. We include a Gaussian measurement uncertainty on jet E_T with standard deviation $0.8\sqrt{E_T\,(\text{GeV})}$; we impose a beam-pipe cut at $|\eta| = 4.2$, beyond which particles escape detection; we evaluate $\not{p}_T$ from the imbalance of the measured jet and lepton momenta, including an estimate of measurement errors on the underlying soft event.[5] In modeling lepton isolation, we do not explicitly hadronize our final-state partons. Instead, we examine whether or not final-state partons lie within a cone of $\Delta R = 0.4$ about the lepton direction. If they do, their E_T is summed and added to a contribution due to the underlying soft scattering event. The latter is represented by the modulus of a Gaussian random variable with mean 0 and standard deviation 1.8 GeV, following the UA1 results[18] for leptons produced in W decays, rescaled to our cone size, and the central multiplicity at $\sqrt{s} = 2$ TeV.

III. BACKGROUNDS

The best known dilepton background is from the Drell-Yan subprocess $q\bar{q} \to \gamma^*, Z \to \ell^+\ell^-$; this gives typically isolated unlike-sign dileptons, sometimes accompanied by QCD jets radiated from the quarks, but is strongly suppressed by our cuts on $\not{p}_T$ and $\Delta\phi$ (see Ref. 5). The next background is from $b\bar{b}$ (or $c\bar{c}$) hadroproduction followed by semileptonic decays, with possible additional QCD jets. The leptons from this source are typically nonisolated, however; they are also typically either parallel or back to back in the transverse plane;[15] hence they are essentially removed by our isolation and $\Delta\phi$ acceptance cuts[5] and the additional requirement of two high-p_T jets renders them even more innocuous. Another background comes from $p\bar{p} \to W + b\bar{b}$+jets. We have estimated the rate for this via a perturbative calculation of the $q\bar{Q} \to Wb\bar{b}$ subprocess. Performing the usual convolution with structure functions yields a cross section of $\sigma(p\bar{p} \to Wb\bar{b}) = 31$ pb, for $\sqrt{s} = 2$ TeV. Applying the cuts of (2a) yields a rate of ~ 2 fb; imposing the $n_j \geq 2$ requirement of (2b) reduces this to a negligible level.

A potentially more difficult background comes from direct W^+W^- production[19] with leptonic W decays; the leptonic content of such events closely resembles that of heavy $t\bar{t}$ events, but the jet content is expected to be rather different. Hitherto there have been no explicit calculations of W^+W^- + dijet production; in the present work we remedy this deficiency, using the new formulas of Ref. 20 for $WWjj$ production cross sections. We find that this background can be suppressed relative to the signal by requiring high enough E_T for the two hardest jets (see below). Another difficult background comes from $p\bar{p} \to \gamma^*, Z \to \tau\bar{\tau} \to \ell_1^+\ell_2^- + 4\nu$ with two QCD jets; a significant number of such indirect Drell-Yan events survive the $\not{p}_T$ and $\Delta\phi$ cuts.[5] However, in these events $m_T(\ell_1\ell_2, \not{p}_T)$ is bounded by M_Z, up to corrections due to energy mismeasurement and off-mass-shell Z production. The cluster transverse mass cut in Eq. (2c) is therefore expected to remove this background. The $\tau\bar{\tau}jj$ background has not been explicitly calculated before (only the inclusive $\tau\bar{\tau}$ production has been studied). We find it is possible to calculate

$$p\,\bar{p} \to Z\,j\,j \to \tau\,\bar{\tau}\,j\,j \to \nu_\tau\,\ell_2\,\bar{\nu}_{\ell_2}\,\bar{\nu}_\tau\,\bar{\ell}_1\,\nu_{\ell_1}\,jj \tag{4}$$

directly from the formulas of Ref. 21, which are for $p\bar{p} \to Zjj$, by simply replacing the Z polarization vector e^μ with the $Z \to \tau\bar{\tau} \to \nu_\tau\,\ell_2\,\bar{\nu}_{\ell_2}\,\bar{\nu}_\tau\,\bar{\ell}_1\,\nu_{\ell_1}$ decay current J^μ. Using the notation of Ref. 21 and using a particle's symbol to denote its four-momentum, the decay current is

$$J^\mu = D_Z(Z)\, D_\tau(\tau)\, D_{\bar\tau}(\bar\tau)\, \frac{e^2}{2\sin^2\theta_W}\, \sqrt{4\, E_{\nu_\tau}\, E_{\bar\nu_\tau}} \Bigg(\frac{e}{\tan 2\theta_W}\, \chi_-^\dagger(\nu_\tau)\, (\not{J}_{W-})_-\, (\not{\tau})_+\, (\sigma^\mu)_-\, (\not{\bar\tau})_+\, (\not{J}_{W+})_-\, \chi_-(\bar\nu_\tau) + m_\tau^2 e \tan\theta_W\, \chi_-^\dagger(\nu_\tau)\, (\not{J}_{W-})_-\, (\sigma^\mu)_+\, (\not{J}_{W+})_-\, \chi_-(\bar\nu_\tau) \Bigg) , \tag{5}$$

where θ_W is the weak mixing angle and $\chi_-(p)$ is a normalized helicity eigenspinor:

$$\chi_-(p) = [2|\mathbf{p}|(|\mathbf{p}| + p_z)]^{-1/2} \begin{pmatrix} -p_x + ip_y \\ |\mathbf{p}| + p_z \end{pmatrix} . \tag{6}$$

The symbol $(\sigma^\mu)_\pm$ denotes the 2×2 matrices

$$(\sigma^\mu)_\pm = (1, \pm\boldsymbol{\sigma}) , \tag{7}$$

where $\boldsymbol{\sigma}$ are the Pauli matrices and $(\not{a})_\pm$ denotes the contraction of a four-vector a_μ with the matrices in Eq. (7):

$$(\not{a})_\pm = a_\mu\, (\sigma^\mu)_\pm = a^0 \mp \mathbf{a}\cdot\boldsymbol{\sigma} . \tag{8}$$

The $\pm$ subscripts are chirality indices and $D_X(p)$ denotes a propagator

$$D_X(p) = (p^2 - m_X^2 + i m_X \Gamma_X)^{-1} . \tag{9}$$

Finally, $J_{W\pm}$ are the decay currents for virtual $W^+ \to \bar\ell_1\, \nu_{\ell_1}$ and $W^- \to \ell_2\, \bar\nu_{\ell_2}$:

$$\begin{aligned} J^\alpha_{W+} &= D_{W+}(W^+)\, \frac{e}{\sqrt{2}\sin\theta_W}\, \sqrt{4\, E_{\bar\ell_1}\, E_{\nu_{\ell_1}}}\, \chi_-^\dagger(\nu_{\ell_1})\, (\sigma^\alpha)_-\, \chi_-(\bar\ell_1) , \\ J^\beta_{W-} &= D_{W-}(W^-)\, \frac{e}{\sqrt{2}\sin\theta_W}\, \sqrt{4\, E_{\ell_2}\, E_{\bar\nu_{\ell_2}}}\, \chi_-^\dagger(\ell_2)\, (\sigma^\beta)_-\, \chi_-(\bar\nu_{\ell_2}) . \end{aligned} \tag{10}$$

The virtual-photon contribution to the decay current has been neglected and the narrow-width approximation has been used for the Z, τ, and $\bar\tau$ particles.

Figures 1 and 2 illustrate the suppression of these backgrounds. Figure 1 shows the $m_T(\ell\ell, \not{p}_T)$ distributions with the cuts of Eq. (2a) and (2b). It is clear that requiring $m_T(\ell\ell, \not{p}_T) > M_Z$ removes virtually all the $\tau\bar\tau jj$ background. Figure 2 shows the dependence of the $t\bar t$ signal and the remaining $WWjj$ background on the transverse energy $E_T(j_2)$ of the second jet (ordering jets with decreasing E_T) after applying the $m_T(\ell\ell, \not{p}_T)$ cut. It is clear that requiring $E_T(j) > 30$ GeV effectively suppresses this remaining background for the $m_t = 200$, 250 GeV cases; however this additional cut is

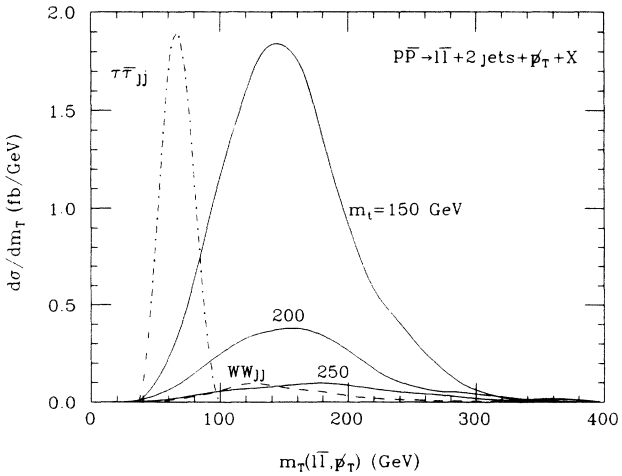

FIG. 1. Distributions of $t\bar t$ signal and $WWjj$, $\tau\bar\tau jj$ backgrounds vs the dilepton cluster transverse mass $m_T(\ell\bar\ell, \not{p}_T)$, with the acceptance cuts of Eqs. (2a) and (2b). Solid curves denote the $t\bar t$ signal, dashed and dotted-dashed lines denote the backgrounds from $WWjj$ and $\tau\bar\tau jj$, respectively.

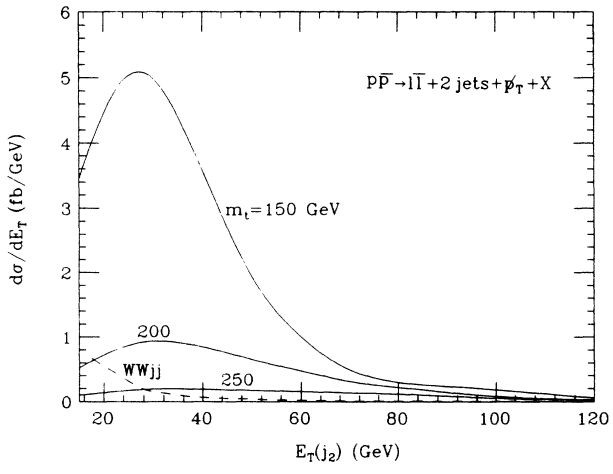

FIG. 2. Distributions of $t\bar t$ signal and $WWjj$ background vs $E_T(j_2)$, the transverse energy of the second hardest jet, with the acceptance cuts of Eqs. (2a), (2b), and (2c). Solid and dashed curves denote the signal and background, respectively.

not necessary or desirable for $m_t = 150$ GeV. This explains and justifies the additional cuts of Eqs. (2c) and (2d).

IV. RESULTS

Figure 3 shows the calculated net cross sections for dilepton-dijet events at $\sqrt{s} = 2$ TeV, after acceptance cuts, summing over all four lepton-flavor channels e^+e^-, $e^+\mu^-$, μ^+e^-, $\mu^+\mu^-$. We emphasize again, these are events with two isolated high-p_T leptons plus missing p_T plus two high-E_T jets; acceptance cuts are fully included but possible detector efficiency factors are not. The solid curves denote the $t\bar{t}$ signal; the dashed curves show the $WWjj$ background; in each case the lower curve refers to the full cuts of Eqs. (2a)–(2d), the upper curve shows the effect of removing the more stringent jet requirement (2d). Figures 2 and 3 show that the optimum $E_T(j)$ cut depends on m_t.

The lepton transverse momenta $p_T(\ell_i)$ reflect the energy release in t decay and offer a simple measure of m_t. We therefore choose as our first dynamical variable of interest (labeled x_1) the mean value

$$x_1 = \tfrac{1}{2}[p_T(\ell_1) + p_T(\ell_2)] . \tag{11}$$

Its statistical mean and standard deviation are shown in Table I for $m_t = 150$, 200, 250 GeV, with the full cuts of Eqs. (2a)–(2d).

If a lepton and jet are decay products of a common parent, their invariant mass gives a lower bound on the parent's rest mass. But there are two ways in which two leptons can be paired with two jets. For a $t\bar{t} \to (b\bar{\ell}\nu)(\bar{b}\ell\bar{\nu})$ signal event, one of these pairings is "correct," associating b with $\bar{\ell}$ and $\bar{b}$ with ℓ; in this case both the lepton-jet

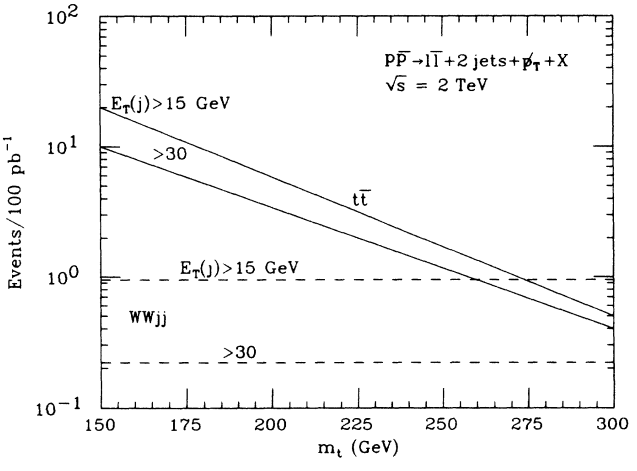

FIG. 3. Cross-section dependence on m_t at $\sqrt{s} = 2$ GeV, summing over all four lepton-flavor channels. Solid and dashed curves denote the $t\bar{t}$ signal and $WWjj$ background, respectively; in each case the upper curve is for the cuts (2a)–(2c) and the lower curve is for the cuts (2a)–(2d). The event rate shown takes full account of the acceptance cuts, but assumes ideal detector efficiency of 100%.

TABLE I. Calculated mean values μ_i and standard deviations σ_i of the six dynamical variables defined in Eqs. (11)–(16), for $m_t = 150$, 200, 250 GeV.

	$m_t = 150$		$m_t = 200$		$m_t = 250$	
	μ_i	σ_i	μ_i	σ_i	μ_i	σ_i
x_1	45	16	52	20	63	25
x_2	90	24	119	30	153	36
x_3	88	22	122	31	158	40
x_4	163	38	207	46	253	52
x_5	129	17	160	25	194	36
x_6	265	78	317	97	363	94

invariant masses are less than m_t and the maximum of the two is the best lower bound on m_t (modulo measurement errors and jet coalescence). Since in practice we do not know which pairing is correct, we must evaluate $\max[m(\ell j)]$ for both pairings; the best available lower bound on m_t is then the lower of these two values, which we denote $\min \max[m(\ell j)]$. Alternatively, the mean value of all four $m(\ell j)$ assignments is also sensitive to m_t. We choose these quantities as our next two dynamical variables of interest:

$$x_2 = \text{mean}\{m(\ell j)\} , \tag{12}$$

$$x_3 = \min \max[m(\ell j)] . \tag{13}$$

Their statistical means and standard deviations are given in Table I.

Figure 4(a) shows the predicted distributions of events versus $\min \max[m(\ell j)]$ for $m_t = 150$, 200, 250 GeV, with the acceptance cuts of Eq. (2). The distribution peaks well below m_t and offers a rather safe lower bound for m_t; in our calculations less than 1% of events have $\min \max[m(\ell j)] > m_t$ (due to simulated measurement errors). Note all curves are normalized to unity; the normalizations can be obtained from Fig. 3.

Another dynamical measure of the complete dilepton-dijet system, including the missing p_T, is the cluster transverse mass defined by Eq. (3) where we now take the dilepton + dijet cluster $c = \ell_1\ell_2 j_1 j_2$. This quantity is in fact the minimum invariant mass of the measured cluster and an unmeasured system of neutrinos with net transverse momentum $\not{p}_T$; for $t\bar{t}$ signal events it is therefore a useful lower bound on the $t\bar{t}$ invariant mass (modulo measurement errors and extraneous sources of jets and $\not{p}_T$), which in turn is expected to lie close to the threshold value $2m_t$ for very heavy top quarks. We therefore choose this quantity divided by 2 as our fourth variable:

$$x_4 = \tfrac{1}{2}\, m_T(\ell\ell jj, \not{p}_T) . \tag{14}$$

Its calculated mean and standard deviation are given in Table I. Figure 4(b) shows the predicted distribution of events versus $m_T(\ell\ell jj, \not{p}_T)$ for $m_t = 150$, 200, 250 GeV. We see that in practice the mean value of x_4 is close to m_t.

For processes with very small event samples,

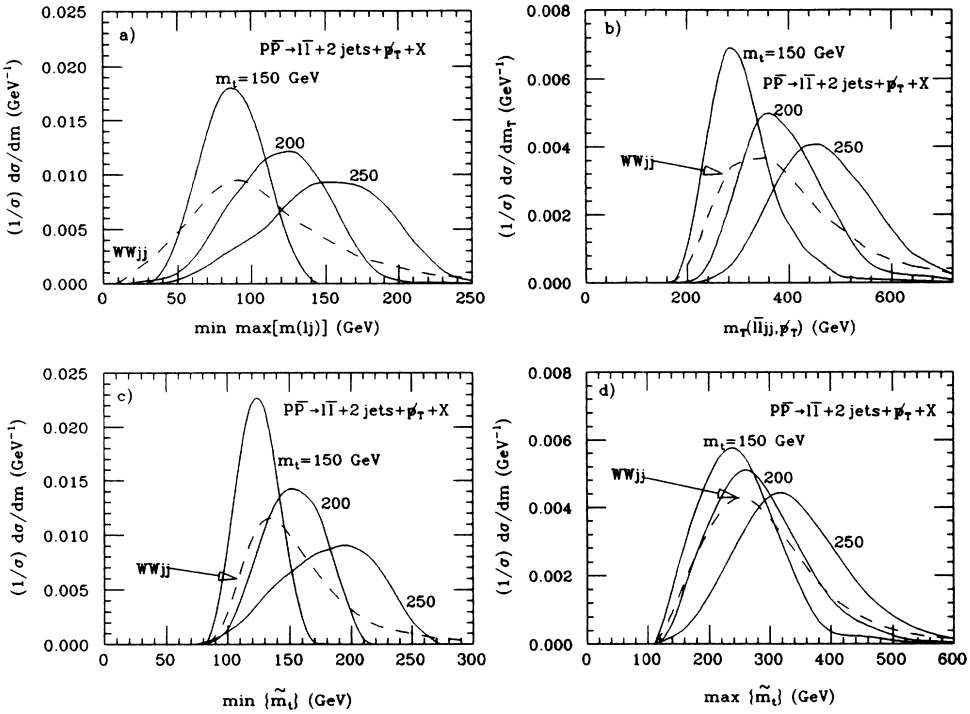

FIG. 4. Dependence of the $t\bar{t}$ dilepton + dijet signal, for $\sqrt{s} = 2$ GeV and the acceptance cuts of Eqs. (2a)–(2d), vs four of the variables of interest: (a) $\min\max[m(\ell j)]$, (b) $m_T(\ell\bar{\ell}jj, \not{p}_T)$, (c) $\min\{\tilde{m}_t\}$, and (d) $\max\{\tilde{m}_t\}$. Solid and dashed curves denote the $t\bar{t}$ signal and $WWjj$ background, respectively. All curves are normalized to unity; the background is small compared to the signal. The cross sections are 100 fb, 32 fb, 12 fb, and 2.3 fb for $m_t = 150$, 200, 250 GeV, and $WWjj$, respectively.

maximum-likelihood methods[11] appear attractive. Here one varies the unknown theoretical parameters to maximize the likelihood of the events, by reference to a theoretical formula for the multiply differential cross section (with respect to the various observed energies, angles, etc.). However, it is necessary to have a background-free sample and also a precise cross-section formula. In the present case these requirements are incompatible; to remove backgrounds we need strong cuts, after which the cross section can only be calculated by Monte Carlo methods, with which multiply differential cross sections are practically unattainable. An alternative approach that we adopt here is to inject the maximum of theoretical information by attempting event reconstruction.

For a heavy-$t\bar{t}$ event, we know that the two leptons come from $W \to \ell\nu$ decay and that the two neutrinos combine to give $\not{p}_T$ (ignoring other contributions). Hence if we first fix the transverse momenta ν_{1x} and ν_{1y} of ν_1 (associated by definition with ℓ_1^+), the transverse momenta ν_{2x} and ν_{2y} are given. There are then in general two solutions for the longitudinal momentum ν_{1z} such that $m(\ell_1\nu_1) = M_W = 80$ GeV; similarly there are two solutions for ν_{2z} such that $m(\ell_2\nu_2) = M_W$ (we can exclude unphysical solutions where ν_{1z} or ν_{2z} exceeds the beam momentum). This reconstruction necessarily neglects the finite width of W. There are then two possible ways to pair the two hardest jets with W_1 and W_2; i.e., 8 possible ways to reconstruct $t\bar{t} \to (b\ell_1^+\nu_1)(\bar{b}\ell_2\bar{\nu}_2)$ for each pair of ν_{1x}, ν_{1y} values. However, a credible assignment would give the same invariant mass to t and $\bar{t}$, an additional constraint. Hence there is ideally a family of credible reconstructions, with one continuous parameter and 8 branches. Realistically, to take account of measurement errors and Γ_W, we allow the two reconstructed m_t values to differ within a specified resolution chosen to be $\Delta m = 40$ GeV, and take their mean value. A small fraction of events fail to reconstruct and are rejected. In among these reconstructions is the correct reconstruction. The minimum and the maximum of all the reconstructed masses are therefore lower and upper bounds, respectively, on m_t (modulo measurement errors, etc.) and may hopefully lie close to m_t. We therefore adopt these quantities as our fifth and sixth variables to study:

$$x_5 = \min\{\tilde{m}_t\}\,, \tag{15}$$

$$x_6 = \max\{\tilde{m}_t\}\,, \tag{16}$$

introducing the symbol $\tilde{m}_t$ for reconstructed top mass. Their distributions are shown in Figs. 4(c) and 4(d) from which we see that they do indeed behave like lower and upper bounds on m_t (despite measurement errors only a small fraction of events violate these bounds). Note that the constraint $\tilde{m}_t > M_W$ cuts off the x_5 distribution at the lower end.

Table I lists the calculated mean values μ_i and stan-

dard deviations σ_i of these variables for $m_t = 150$, 200, 250 GeV. It is interesting to use the variables x_i to construct estimators x_i' for m_t, based on a linear approximation in a band of m_t; for example, in the range $150 \leq m_t \leq 200$ we can construct (in GeV units)

$$x_i' = 150 + 50[x_i - \mu_i(150)]/[\mu_i(200) - \mu_i(150)]. \quad (17)$$

These quantities have the correct mean values $\mu_i' = m_t$ at $m_t = 150$ and 200 GeV. We can directly compute their standard deviations from Table I and find $\sigma_1' = 123$, $\sigma_2' = 41$, $\sigma_3' = 34$, $\sigma_4' = 43$, $\sigma_5' = 29$, $\sigma_6' = 75$ in GeV units in the neighborhood of $m_t = 150$ GeV. This shows that x_3' and x_5' based on the variables $\min\max[m(\ell j)]$ and $\min\{\tilde{m}_t\}$ are the most accurate estimators of m_t here. For a sample of n events, the mean of the n values of x_i' will have statistical standard deviation $\sigma_i'/\sqrt{n}$ and approach a Gaussian distribution as n increases.

It is tempting to consider combining several estimators, but we must take proper account in that case of their correlations. We have computed their covariance matrix. It is more transparent in fact to illustrate the covariance of related variables

$$x_i'' = x_i'/\sigma_i', \quad (18)$$

normalized so that the variance of each variable is unity. We compute the covariance matrix $V_{ij}'' = \langle (x_i'' - \mu_i'')(x_j'' - \mu_j'') \rangle$ directly from a sample of Monte Carlo events (for each of which the values x_i are recorded and hence x_i', x_i'' can be generated). At $m_t = 150$ GeV, for example, we obtain

$$V'' = \begin{pmatrix} 1.00 & 0.67 & 0.44 & 0.50 & 0.18 & -0.19 \\ 0.67 & 1.00 & 0.68 & 0.75 & 0.42 & 0.26 \\ 0.44 & 0.68 & 1.00 & 0.32 & 0.55 & 0.05 \\ 0.50 & 0.75 & 0.32 & 1.00 & 0.45 & 0.28 \\ 0.18 & 0.42 & 0.55 & 0.45 & 1.00 & 0.15 \\ -0.19 & 0.26 & 0.05 & 0.28 & 0.15 & 1.00 \end{pmatrix}. \quad (19)$$

If two variables x_i'' and x_j'' are 100% correlated [i.e., $x_j'' = f(x_i'')$] then the off-diagonal element $V_{ij}'' = 1$ with our present normalization (to first order in a Taylor expansion of f about the mean). Equation (19) shows that the four most accurate estimators x_2'', x_3'', x_4'', x_5'' (or equivalently x_2', x_3', x_4', x_5') are quite strongly correlated. There is some advantage in combining them appropriately, but not very much.

To discuss combining estimators, we return to the x_i' which have common mean $\mu_i' = m_t$ and covariance matrix $V_{ij}' = \sigma_i'\sigma_j' V_{ij}''$ determined by Eqs. (11)–(19) and Table I. We seek the optimum linear combination defined by $x_0' = \sum_i a_i x_i'/(\sum_i a_i)$ with mean $\mu_0' = m_t$ and minimum variance $(\sigma_0')^2 = \sum_{ij} a_i a_j V_{ij}'$, minimized with respect to the coefficients a_i. For general V' the optimum prescription is $a_i = \det[V'(i)]$ where $V'(i)$ is obtained from V' by replacing all elements in row i by unity. With this prescription, we find that the optimum combination of the four best estimators $x_2' \cdots x_5'$ has $\sigma_0' = 28$ GeV; the optimum combination of all six estimators $x_1' \cdots x_6'$ has $\sigma_0' = 26$ GeV; the reduction in standard deviation is rather small.

A similar analysis near $m_t = 200$ GeV yields $\sigma_5' = 41$ GeV for the best single estimator and $\sigma_0' = 36$ GeV for the optimum linear combination of all six estimators.

Next there is the question of systematic uncertainties in the theory; this can be approached by studying several different Monte Carlo simulations, with various alternative treatments of parton fragmentation, structure functions, etc., that we do not attempt here. We expect these uncertainties to be smaller for the shapes of the dynamical distributions than for the overall event rate, which is sensitive to the choice of structure functions and the renormalization/factorization scale?[22] There are also systematic experimental errors that depend on the apparatus and are not discussed here.

Finally, the event rate itself is an approximate measure of m_t. It is statistically uncorrelated with the distribution shape variables x_i, but systematic errors can introduce biases, and the latter may be correlated. For a small event sample of say 10 events, there would be a $\pm 30\%$ statistical uncertainty; from Fig. 3 with $E_T(j) > 15$ GeV cuts, this is equivalent to about ± 10 GeV in m_t, comparable to the accuracy from the best estimators x_3' and x_5' with this number of events. Theoretical uncertainties in the $t\bar{t}$ production cross section[22] alone are of order $\pm 20\%$; decay and fragmentation introduce more, perhaps $\pm 30\%$ altogether, equivalent to a further ± 10 GeV in m_t. Our results above are all based on calculations at the parton level; they include the leading effects of initial- and final-state QCD radiation (through the $2 \to 3$ production processes) and also realistic estimates of many measurement uncertainties. They allow for jet overlaps and smearing, plus the possibility that the two hardest jets may not simply be identified with the b and $\bar{b}$ quarks in any given event. However, it will be desirable also to include multiple QCD radiation, explicit hadronization, and detailed detector simulation, before comparing with eventual experimental data. Meanwhile, our results should be regarded as indicative rather than definitive.

In an eventual data sample, it will be desirable also to investigate all other possible top-quark signals, such as those based on one lepton plus multijets. Since previous studies[4,5] have shown that single-lepton signals from very heavy top quarks have very severe background problems at the Tevatron, we have not pursued them here.

V. CONCLUSIONS

Our results lead to the following conclusions.

(i) The dilepton-dijet signal from very massive $t\bar{t}$ hadroproduction at the Tevatron can be separated cleanly from backgrounds, up to mass $m_t = 250$ GeV and beyond, by suitable cuts.

(ii) In particular the $\tau\bar{\tau}jj$ and $WWjj$ backgrounds, calculated explicitly for the first time in this paper, can be suppressed by $m_T(\ell\ell, \not{p}_T)$ and $E_T(j)$ cuts. At lower m_t values, a less severe $E_T(j)$ cut suffices.

(iii) A signal of 10 events at $\sqrt{s} = 2$ TeV requires an integrated luminosity of order 100 pb^{-1} for $m_t = 150$ GeV, rising to 2000 pb^{-1} for $m_t = 250$ GeV (with our recommended cuts and assuming 50% detector efficiency).

(iv) With a sample of 10 events, the event rate alone would allow m_t to be estimated approximately within ± 10 GeV statistically (with perhaps ± 10 GeV additional theoretical uncertainty) for $m_t \sim 150$–200 GeV.

(v) Additionally, various dynamical variables and quantities derived from event reconstructions can be used to estimate m_t. The most accurate estimates are those based on mean$\{m(\ell j)\}$, min max$[m(\ell j)]$, $m_T(\ell\ell jj, \not{p}_T)$, and min$\{\tilde{m}_t\}$. Individually, these estimators give m_t within ± 10 GeV statistically for a sample of 10 events near $m_t = 150$ GeV (± 14 GeV near $m_t = 200$ GeV).

(vi) The estimators in (v) are strongly correlated; combining several such estimators gives little statistical advantage. Comparing such estimators however provides some check on systematics. Also this approach to m_t is independent of the event rate (iv) and is expected to have smaller theoretical uncertainty.

(vii) The variables min max$[m(\ell j)]$ and min$\{\tilde{m}_t\}$ have direct interest also as lower bounds on m_t; max$\{\tilde{m}_t\}$ is an upper bound on m_t; $\frac{1}{2}m_T(\ell\ell jj, \not{p}_T)$ has mean value close to m_t.

[1] CDF Collaboration, F. Abe *et al.*, Phys. Rev. Lett. **64**, 142 (1990); **64**, 147 (1990); UA1 Collaboration, K. Eggert *et al.*, in *Proceedings of the XIV International Symposium on Lepton and Photon Interactions at High Energies*, Stanford, California, 1989, edited by M. Riordan (World Scientific, Singapore, 1990); UA2 Collaboration, L. DiLella *et al.*, *ibid.*

[2] Mark II Collaboration, G. Feldman *et al.*, in *Proceedings of the XIV International Symposium on Lepton and Photon Interactions at High Energies* (Ref. 1); AMY Collaboration, S. Igarashi *et al.*, Phys. Rev. Lett. **60**, 2359 (1988); TOPAZ Collaboration, I. Adachi *et al.*, *ibid.* **60**, 97 (1988); VENUS Collaboration, K. Abe *et al.*, Phys. Lett. B **207**, 355 (1988).

[3] U. Amaldi *et al.*, Phys. Rev. D **36**, 1385 (1987); G. Costa *et al.*, Nucl. Phys. **B297**, 244 (1988); J. Ellis and G. L. Fogli, Phys. Lett. B **213**, 526 (1988); **232**, 139 (1989); P. Langacker, Phys. Rev. Lett. **18**, 1920 (1989); F. Halzen and D. Morris, University of Wisconsin—Madison Report No. MAD/PH/517, 1989 (unpublished).

[4] H. Baer, V. Barger, H. Goldberg, and R. J. N. Phillips, Phys. Rev. D **37**, 3152 (1988); H. Baer, V. Barger, and R. J. N. Phillips, Phys. Lett. B **221**, 398 (1989); Phys. Rev. D **39**, 2809 (1989).

[5] H. Baer, V. Barger, and R. J. N. Phillips, Phys. Rev. D **39**, 3310 (1989).

[6] S. Gupta and D. P. Roy, Z. Phys. C **39**, 417 (1988).

[7] R. Kleiss, A. D. Martin, and W. J. Stirling, Z. Phys. C **39**, 393 (1988); W. T. Giele and W. J. Stirling, Leiden report, 1989 (unpublished).

[8] F. Halzen, C. S. Kim, and A. D. Martin, Mod. Phys. Lett. A **4**, 1531, 1989.

[9] J. L. Rosner, Phys. Rev. D **39**, 3297 (1989); **40**, 1701(E) (1990); W. Kwong, L. H. Orr, and J. L. Rosner, *ibid.* **40**, 1453 (1989).

[10] C. P. Yuan, Phys. Rev. D **41**, 42 (1990).

[11] See, e.g., K. Kondo, J. Phys. Soc. Jpn. **57**, 12 (1988).

[12] R. K. Ellis and J. C. Sexton, Nucl. Phys. **B282**, 642 (1987).

[13] P. Nason, S. Dawson, and R. K. Ellis, Nucl. Phys. **B303**, 607 (1988).

[14] V. Barger, J. Ohnemus, and R. J. N. Phillips, Int. J. Mod. Phys. A **4**, 617 (1989).

[15] V. Barger and R. J. N. Phillips, Phys. Lett. **143B**, 259 (1984).

[16] V. Barger, A. D. Martin, and R. J. N. Phillips, Phys. Lett. **125B**, 339 (1983).

[17] D. W. Duke and J. F. Owens, Phys. Rev. D **30**, 49 (1984).

[18] UA1 Collaboration, G. Arnison *et al.*, Lett. Nuovo Cimento **44**, 1 (1985).

[19] R. W. Brown and K. O. Mikaelian, Phys. Rev. D **19**, 922 (1979).

[20] V. Barger, T. Han, J. Ohnemus, and D. Zeppenfeld, Phys. Rev. D **41**, 2782 (1990).

[21] K. Hagiwara and D. Zeppenfeld, Nucl. Phys. **B313**, 560 (1989).

[22] G. Altarelli *et al.*, Nucl. Phys. **B308**, 724 (1989).