Top-Quark Pairs at High Invariant Mass: A Model-Independent Discriminator of New Physics at the Large Hadron Collider

Vernon Barger,[*] Tao Han,[†] and Devin G. E. Walker[‡]

Department of Physics, University of Wisconsin, Madison, Wisconsin 53706, USA
(Received 7 December 2006; published 22 January 2008)

We study top-quark pair production to probe new physics at the CERN Large Hadron Collider. We propose reconstruction methods for $t\bar{t}$ semileptonic events and use them to reconstruct the $t\bar{t}$ invariant mass. The angular distribution of top quarks in their c.m. frame can determine the spin and production subprocess for each new physics resonance. Forward-backward asymmetry and *CP*-odd variables can be constructed to further delineate the nature of new physics. We parametrize the new resonances with a few generic parameters and show high invariant mass top pair production may provide an early indicator for new physics beyond the standard model.

In the next few years, high energy physics will experience the excitement of major discoveries when the CERN Large Hadron Collider (LHC) opens up the unexplored TeV energy scale. Besides the long anticipated Higgs boson that is responsible for the mass generation in the highly successful standard model (SM), the quadratic sensitivity of the Higgs boson mass to radiative corrections indicate the existence of new physics associated with electroweak symmetry breaking naturally at the scale of order $4\pi v$. (Here $v \approx 246$ GeV is the Higgs field vacuum expectation value). Numerous extensions to the SM have been proposed to describe electroweak symmetry breaking. A sample of popular scenarios include: the minimal supersymmetric standard model (MSSM) [1], new strong dynamics [2–5], composite Higgs boson at the TeV scale [6], Little Higgs theory [7], and extra dimensions at the electroweak scale [8,9]. String-inspired extensions in the gauge sector associated with an extra $U(1)$ symmetry [10] also lead to striking signatures. It is therefore highly expected that many new signatures will become manifest at TeV energy scales that can be probed at the LHC.

The LHC will be a "top factory:" About 80×10^6 $t\bar{t}$ events will be generated by QCD production with an integrated luminosity of 100 fb^{-1}. Thus studying top-quark production can be fruitful. The fact that the top-quark mass is at the electroweak scale ($m_t \approx v/\sqrt{2}$) suggests that top-quark production may be sensitive to new physics near the TeV scale. Generically, if the new physics contributes to $t\bar{t}$ production as an s-channel resonance, we want to identify the signal as a bump [11] on the smoothly falling $t\bar{t}$ invariant mass distribution. Once we can reconstruct the $t\bar{t}$ c.m. frame, the integer spin ($J = 0, 1, 2$) of any resonances can be determined from the polar angular distribution of the top quark. An asymmetry of this distribution would probe the chiral structure of the couplings. It may be possible to explore the *CP* property of the couplings with the help of *CP*-odd kinematical variables constructed from the final state particle momenta. Moreover, the relative importance of gluon-gluon and quark-

antiquark subprocesses can be inferred from the spin and angular distributions. It is thus of fundamental importance to effectively reconstruct the $t\bar{t}$ invariant mass via their decay products.

We focus on the semileptonic decay mode, $t\bar{t} \to b j_1 j_2 \bar{b} \ell^- \bar{\nu}$ + c.c., where $\ell = e$ or μ. The purely hadronic decay mode of $t\bar{t}$ not only suffers from a much larger QCD background, but also loses the identification of t from $\bar{t}$. For the purely leptonic mode, with a small branching fraction of about 4/81, one cannot reconstruct the $t\bar{t}$ invariant mass with two missing neutrinos. The signal to search for is an isolated charged lepton plus missing energy ($\not{E}_T$), 2 b-jets plus 2 light jets. The branching ratio of the semileptonic to the hadronic channel is 2/3.

New reconstruction methods.—A primary focus of our study is to reliably reconstruct $t\bar{t}$ kinematics at high invariant mass on an event-by-event basis. The challenge is to reconstruct the momentum of the missing neutrino. The transverse momentum of the neutrino is identified with the observed $\not{E}_T$. The longitudinal momentum is subject to a twofold ambiguity from solving the kinematic quadratic equation.

Several top reconstruction methods have been used at the Tevatron [12]. There, however, the top quarks are produced near threshold and the kinematics of the subsequent decay products are very complicated. Since we are interested in new physics in the TeV region, demanding a high invariant mass for the $t\bar{t}$ events will tremendously simplify the kinematics, especially by distinguishing the b quark from $\bar{b}$. Throughout our study, we use a $2 \to 6$ partonic level Monte Carlo simulation that incorporates full spin correlations from production through decay [13]. We made a Pythia simulation, including gluon radiation and hadronization, that confirmed our basic results.

We impose a cluster transverse mass cut on the $t\bar{t}$ system

$$M_T = \sqrt{(p_b + p_{\bar{b}} + p_{j_1} + p_{j_2} + p_\ell)^2 + \not{E}_T^2} + \not{E}_T > 600 \, \text{GeV}.$$

We adopt kinematical cuts from the ATLAS and CMS [14] top studies. We smear the hadronic energy according to a

Gaussian error given by $\Delta E_j/E_j = 0.5/\sqrt{E_j/\mathrm{GeV}} \oplus 0.03$, and the lepton momentum by $\Delta p_T^\ell/p_T^\ell = 0.36(p_T^\ell/\mathrm{TeV}) \oplus 0.013/\sqrt{\sin\theta}$, where θ is the polar angle of the lepton with respect to the beam direction in the lab frame. We present two schemes to reconstruct semileptonic $t\bar{t}$ events and evaluate their efficacy.

(i) (M_W, m_t) *scheme.*—In this scheme, the key assumption is to take M_W and m_t as inputs for their on-shell production and decays. *Step I:* Demand $m_{l\nu}^2 = M_W^2$. The longitudinal momentum of the neutrino is formally expressed as

$$p_{\nu L} = \frac{1}{2p_{eT}^2}(Ap_{eL} \pm E_e\sqrt{A^2 - 4p_{eT}^2\slashed{E}_T^2}),$$

where $A = M_W^2 + 2\vec{p}_{eT} \cdot \vec{\slashed{E}}_T$. If $A^2 - 4p_{eT}^2\slashed{E}_T^2 \geq 0$, the value of $p_{\nu L}$ that best yields the known top mass via $m_{l\nu b}^2 = m_t^2$ is selected. This ideal situation may not always hold when taking into account the detector resolutions. For cases with no real solutions, we then proceed to the next step. *Step II:* To better recover the correct kinematics, we instead first reconstruct the top quark directly by demanding $m_{l\nu b}^2 = m_t^2$. The longitudinal momentum of the neutrino is expressed as

$$p_{\nu L} = A'p_{blL}/2(E_{bl}^2 - p_{blL}^2) \pm \frac{1}{2(E_{bl}^2 - p_{blL}^2)}(p_{blL}^2 A'^2$$
$$+ (E_{bl}^2 - p_{blL}^2)(A'^2 - 4E_{bl}^2\slashed{E}_T^2))^{1/2},$$

where $A' = m_t^2 - M_{bl}^2 + 2\vec{p}_{blT} \cdot \vec{\slashed{E}}_T$. The twofold ambiguity is broken by choosing the value that best reconstructs $M_W^2 = m_{l\nu}^2$. A plot of the top and W mass distributions is shown in Fig. 1(a). The solid histogram is from the procedure *Step I*, and the dashed histogram from *Step II*. With these two steps, there could still be some events that do not lead to a real solution. We thus discard them in our event collection. The discard rate is about 16%.

(ii) *Small θ_{bW} angle selection scheme*—(We have chosen to use θ_{bW} in our reconstruction instead of, e.g., $\theta_{l\nu}$, because the b quark is on average much more energetic (highly boosted) than the lepton.) This scheme reconstructs the $t\bar{t}$ system without relying on the top mass reconstruction, thus avoiding potentially large QCD corrections due to gluon radiations. Since we are interested in new physics in the TeV regime, the top quarks will be relativistic with a γ-factor of $(1 \text{ TeV})/2m_t \sim 3$. We thus expect that the decay products are fairly collimated along the top-quark moving direction. This is illustrated in Fig. 1(b) for the normalized opening-angle distribution between b and W^+ in the lab frame, where an increasing cluster transverse mass cut has been imposed for the dotted, dashed and solid curves for $M_T > 0$, 600, 1000 GeV, respectively. With the initial requirement $m_{l\nu}^2 = M_W^2$, the twofold ambiguity in θ_{bW} is resolved by choosing the smaller reconstructed angle. This scheme should work better at higher $t\bar{t}$ invariant masses.

Backgrounds to the $t\bar{t}$ signal.—The major backgrounds to our $t\bar{t}$ events include the processes $W + $ jets, $Z + $ jets, WW, WZ and ZZ. The ATLAS and CMS technical design reports [15] detail studies of the selection efficiencies for these background processes in comparison to a reconstructed $t\bar{t}$ semileptonic signal. The ATLAS (CMS) group found for an integrated luminosity of 10 fb^{-1} (1 fb^{-1}) a signal to background ratio of $S/B = 65$ ($S/B = 26$) [14]. Because of the expected high S/B ratio, our analysis is concentrated solely on the $t\bar{t}$ events without including the small background contamination. Our analysis does not include misidentification of faked leptons from jets in $t\bar{t}$ total hadronic decays.

Although the t ($\bar{t}$) is primarily identified by the charged lepton, ℓ^+ (ℓ^-), a concern is the matching of the b jet associated with top-quark decay. Both ATLAS and CMS studies [15] show a combination of kinematic fits, designed to properly reconstruct the W boson and the hadronically decaying top significantly reduces misidentification. Our cut on M_T helps significantly in this regard.

Search for new physics.—We wish to explore the new physics searches in a model-independent manner for $t\bar{t}$ semileptonic decays. We consider $t\bar{t}$ production via

$$gg \to \phi \to t\bar{t}, \quad q\bar{q} \to V \to t\bar{t}, \quad q\bar{q}, \quad gg \to \tilde{h} \to t\bar{t},$$

where ϕ, V and $\tilde{h}$ are the spin-0, spin-1, and spin-2 resonances. We characterize the effects on the invariant mass spectrum with three parameters: mass, total width, and the signal cross section normalization (ω^2). The normalization $\omega = 1$ defines our benchmark for the spin 0, 1
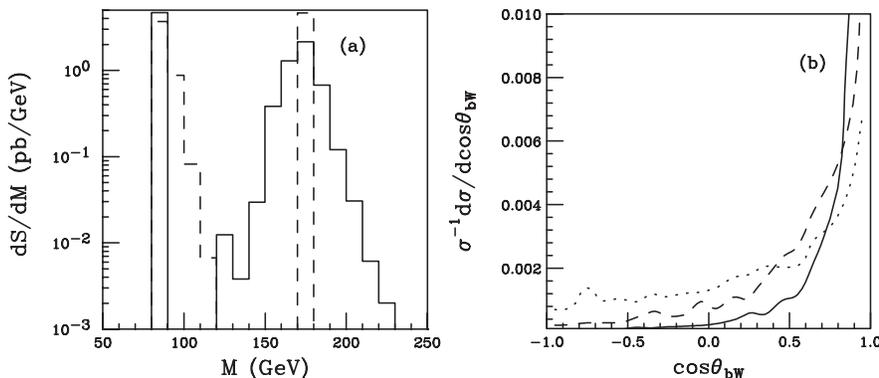

FIG. 1. (a) The W and top mass reconstructions from the (M_W, m_t) scheme, with the procedure Step I (solid) and Step II (dashed). (b) Differential cross section versus $\cos\theta_{bW}$ with no invariant mass cut (dotted), and with cuts of 600 GeV (dashed) and 1000 GeV (solid).
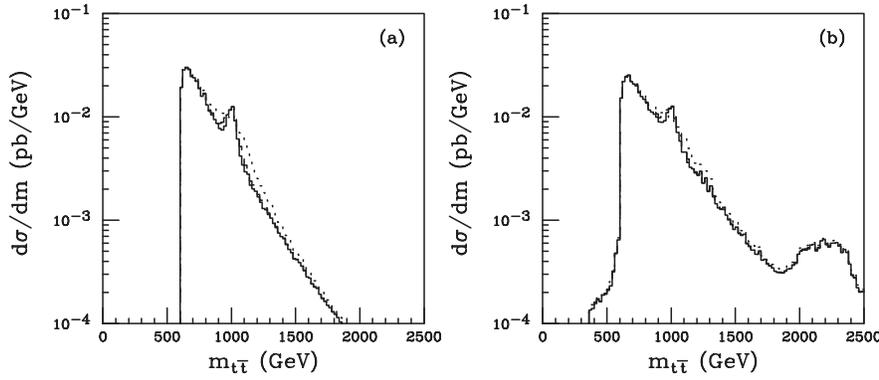
FIG. 2. $t\bar{t}$ invariant mass distributions reconstructed from (a) the (M_W, m_t) scheme, and (b) the small angle selection scheme. Both plots featured a 1 TeV resonance with a total width of 2% (solid), 5% (dashed), and 20% (dotted) of the resonance's mass.

and 2 resonances. They correspond to the SM-like Higgs boson, a Z' with electroweak coupling strength and left (L) or right (R) chiral couplings to SM fermions, and the Randall-Sundrum graviton $\tilde{h}$ with the couplings scaled as Λ^{-1} for $\tilde{h}q\bar{q}$, and $[\Lambda \ln(M_{pl}^*/\Lambda)]^{-1}$ for $\tilde{h}gg$, respectively. More precisely, we use the Feynman rules given in [16] and include the additional warp correction factors from [17]. Numerically, we take $\Lambda = 2$ TeV.

In Fig. 2 we show the reconstructed $t\bar{t}$ invariant mass distributions for the two reconstruction schemes. The SM $t\bar{t}$ total cross section is theoretically known beyond the leading order in QCD [18]. We thus expect to have a good control of this distribution even at high invariant masses. As for new physics, we include the contribution of a 1 TeV vector resonance for illustration, for $\omega_v = 1$, with total widths specified in the caption of Fig. 2. We note that a very high invariant mass tail exists for the $t\bar{t}$ invariant mass reconstructed via the small angle selection. This comes from the misreconstructed events in the low invariant mass region. When a large enough transverse mass cut is applied for a given resonance, the tail will not obfuscate the resonance signal. For example, if searching for 2 TeV resonance, a 800 GeV minimum cut will eliminate the tail for the mass region of interest.

We maximize the signal observability by isolating the resonance within an invariant mass window of ± 100 GeV, ± 30 GeV and ± 25 GeV for the scalar, vector, and graviton resonance, respectively. Given a resonance mass and total width, we can quantify how large ω needs to be for a 5σ discovery. With the number of events for a signal (S) and background (B), we require $S/\sqrt{B+S} > 5$. This translates to a bound $\omega^2 > (25 + 5\sqrt{25 + 4B})/2S_1$, where S_1 is the benchmark signal rate for $\omega = 1$. This is illustrated by Fig. 3 versus the mass for a scalar, vector, and graviton resonance for total widths of 20%, 5%, and 2% of its mass, respectively, for an integrated luminosity of 10 fb^{-1}.

It is of critical importance to reconstruct the c.m. frame of the resonant particle, where the fundamental properties of the particle can be best studied. In Fig. 4, we show the top-quark angular distribution, $\cos\theta^*$, with θ^* defined as the angle in the $t\bar{t}$ c.m. frame between the top-quark

momentum and the incident quark momentum, with latter determined by the longitudinal boost direction of the c.m. system. Although events in the forward and backward regions are suppressed due to the stringent kinematical cuts, we still see the impressive features of the d-function distributions [19] in Fig. 4(a): a flat distribution for a scalar resonance (dashed), d_{11}^1 distribution for the left or right chiral couplings of a vector (dotted), and $d_{1\pm1}^2$ from $q\bar{q}$ (solid), and $d_{2\pm1}^2$ from gg (dot-dashed) for a spin-2 resonance. To illustrate the statistical sensitivity for observing the characteristic distributions, we show in Fig. 4(b) the expected SM $t\bar{t}$ events (solid) with 1σ statistical error bars in each bin for a 10 fb^{-1} integrated luminosity, along with a 5σ signal of a chirally coupled vector summed with the $t\bar{t}$ background in the resonant region (dashed). Because of the large event sample, the statistical significance is evident in the central and forward region. The forward-backward asymmetry in $\cos\theta^*$ can thus be constructed to probe the
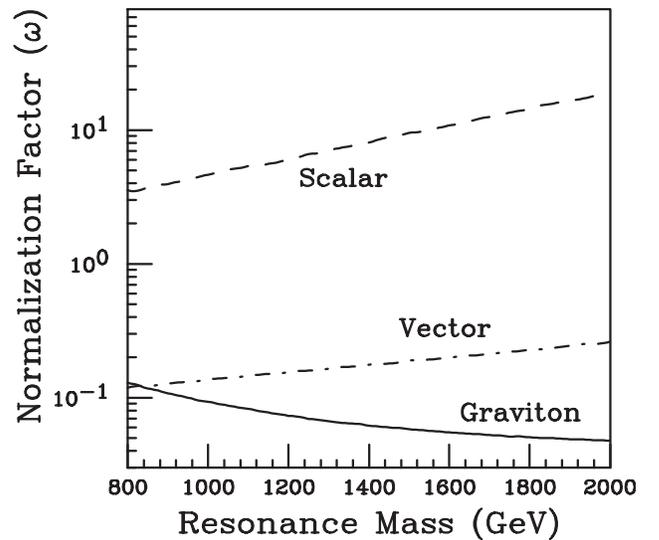

FIG. 3.   Normalization factor versus the resonance mass for the scalar (dashed) with a width-mass ratio of 20%, vector (dot-dashed) with 5%, and graviton (solid) 2%, respectively. The region above each curve represents values of ω that give 5σ or greater statistical significance with 10 fb^{-1} integrated luminosity.
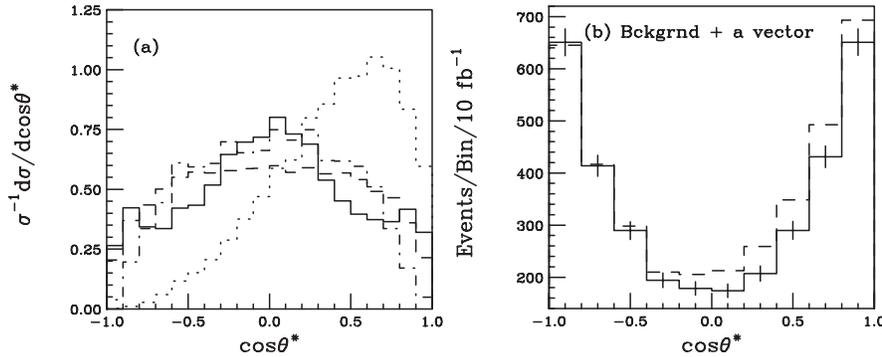
FIG. 4.   Polar angular distributions for the top quark in the c.m. frame; (a) signal only by the (M_W, m_t) scheme for a scalar (dashed), a vector (dots), and a graviton from $q\bar{q}$ (solid) or from gg (dot-dashed); (b) number of events for the SM $t\bar{t}$ background (solid) with 1σ statistical error bars, and the background plus a vector resonance (dashed).

chiral couplings of the particle to the top quark. With the identification of the charged leptons, one may even form kinematical triple products to test the *CP* properties of the couplings [20]. As for distinguishing the other spin states, it is a question of statistical significance. A narrow graviton may be relatively easy to confirm, achievable with a few tens of fb^{-1} luminosity; a broad scalar may require the highest luminosity, more than 300 fb^{-1}.

In summary, we investigated two ways to reconstruct semileptonic $t\bar{t}$ events at high $t\bar{t}$ invariant mass and showed the utility of each in discovering new physics in the form of integer-spin resonances. The angular distributions of the top in the reconstructed c.m. frame reveal the spin of the resonance, and relative contribution from the initial states $q\bar{q}$ or gg. The forward-backward asymmetry and *CP*-odd variables can be constructed to further differentiate models. Because SM top-quark physics is well predicted, high invariant mass top pair production may provide an early indicator for new physics beyond the standard model at the LHC.

*barger@physics.wisc.edu
†than@physics.wisc.edu
‡walker@physics.wisc.edu

[1] S. Dimopoulos and H. Georgi, Nucl. Phys. **B193**, 150 (1981).

[2] S. Weinberg, Phys. Rev. D **13**, 974 (1976); L. Susskind, Phys. Rev. D **20**, 2619 (1979).

[3] C. T. Hill, Phys. Rev. D **24**, 691 (1981); C. T. Hill, Phys. Lett. B **266**, 419 (1991); R. S. Chivukula, B. A. Dobrescu, H. Georgi, and C. T. Hill, Phys. Rev. D **59**, 075003 (1999).

[4] B. A. Dobrescu and C. T. Hill, Phys. Rev. Lett. **81**, 2634 (1998).

[5] For a survey of models that feature strong dynamics for electroweak physics see C. T. Hill and E. H. Simmons, Phys. Rep. **381**, 235 (2003); [**390**, 553(E) (2004)], and the references therein.

[6] D. B. Kaplan, H. Georgi, and S. Dimopoulos, Phys. Lett. B **136**, 187 (1984).

[7] N. Arkani-Hamed, A. G. Cohen, and H. Georgi, Phys. Lett. B **513**, 232 (2001); N. Arkani-Hamed, A. G. Cohen, E. Katz, and A. E. Nelson, J. High Energy Phys. 07 (2002) 034; For a review on Little Higgs models see M. Schmaltz and D. Tucker-Smith, Annu. Rev. Nucl. Part. Sci. **55**, 229 (2005).

[8] N. Arkani-Hamed, S. Dimopoulos, and G. R. Dvali, Phys. Lett. B **429**, 263 (1998).

[9] L. Randall and R. Sundrum, Phys. Rev. Lett. **83**, 3370 (1999).

[10] J. L. Hewett and T. G. Rizzo, Phys. Rep. **183**, 193 (1989); M. Cvetic and P. Langacker, Phys. Rev. D **54**, 3570 (1996); D. Berenstein and S. Pinansky, Phys. Rev. D **75**, 095009 (2007).

[11] K. D. Lane, Phys. Rev. D **52**, 1546 (1995);

[12] B. Abbott *et al.* (D0 Collaboration), Phys. Rev. D **58**, 052001 (1998); F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **80**, 2767 (1998).

[13] V. D. Barger, J. Ohnemus, and R. J. N. Phillips, Int. J. Mod. Phys. A **4**, 617 (1989); M. Arai, N. Okada, K. Smolek, and V. Simak, Phys. Rev. D **70**, 115015 (2004).

[14] Following ATLAS and CMS, we adopt the kinematical cuts as $E_T^j > 20$ GeV, $|\eta_j| < 2.5$; $p_T^\ell > 20$ GeV, $|\eta_\ell| < 2.5$, and $\not{E}_T > 20$ GeV. The lepton-jet and jet-jet isolation is $\Delta R > 0.4$.

[15] ATLAS Technical design report. Vol. 2, Report No. CERN-LHCC-99-15; CMS Technical Design Report V.2. Report No. CERN-LHCC-2006-021.

[16] T. Han, J. D. Lykken, and R.-J. Zhang, Phys. Rev. D **59**, 105006 (1999).

[17] H. Davoudiasl, J. L. Hewett, and T. G. Rizzo, Phys. Rev. D **63**, 075004 (2001).

[18] E. Laenen, J. Smith, and W. L. van Neerven, Nucl. Phys. **B369**, 543 (1992); E. L. Berger and H. Contopanagos, Phys. Rev. D **54**, 3085 (1996); M. Cacciari, S. Frixione, M. L. Mangano, P. Nason, and G. Ridolfi, J. High Energy Phys. 04 (2004) 068.

[19] For the definition and convention of the *d* functions, we follow the PDG, W.-M. Yao *et al.*, J. Phys. G **33**, 1 (2006).

[20] D. Chang and W. Y. Keung, Phys. Lett. B **305**, 261 (1993); G. Valencia and Y. Wang, Phys. Rev. D **73**, 053009 (2006).