PROCEEDINGS OF SCIENCE

Substructure of Massive Boosted Jets and Boosted Tops at CDF II

Raz Alon[*†]
Weizmann Institute of Science
E-mail: raz.alon@weizmann.ac.il

We present a study of the substructure of jets with transverse momentum greater than 400 GeV/c produced in proton-antiproton collisions at a center-of-mass energy of 1.96 TeV at the Fermilab Tevatron Collider and recorded by the CDF II detector. We measure for the first time the distributions of the jet mass, angularity, and planar flow in a 5.95 fb^{-1} data sample. The observed substructure for high mass jets are found to be consistent with predictions from perturbative quantum chromodynamics. A search for boosted top quarks based on jet mass is presented as well.

The 2011 Europhysics Conference on High Energy Physics-HEP 2011,
July 21-27, 2011
Grenoble, Rhône-Alpes France

*Speaker.
†CDF Collaboration

PoS(EPS-HEP2011)281

1. Introduction

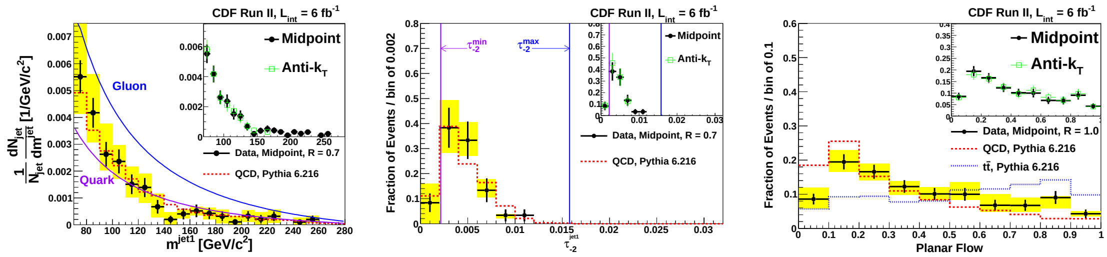
The study of substructure of high transverse momentum (p_T) QCD jets (see *e.g.* [1, 2, 3] for recent reviews) provides tests for perturbative QCD (pQCD) predictions and sheds light on the QCD showering mechanism. Moreover, these jets constitute dominant backgrounds in searches for boosted top quarks, higgs bosons [4], as well as searches for new physics signatures that give rise to massive jets [5, 6, 7, 8]. Available experimental studies at the Tevatron were limited to jets with $p_T < 400\ GeV/c$ [9, 10]. Results on jets with higher p_T produced at the LHC have been published recently [11]. The CDF study presented here [12, 13] was the first to look at substructure of massive jets with $p_T > 400\ GeV/c$. Both Midpoint and *anti-*k_t jet algorithms were used, with typical jet sizes of $R = 0.4,\ 0.7$, and 1.0. A pileup correction technique was used to remove non-coherent energy deposits, critical in substructure studies. Three observables were measured. The jet mass was compared to the theoretical next-to-leading order (NLO) prediction of the jet function [14]. Angularity and planar flow [15] served as further support to the hypothesis of the two-prongness of QCD jets. Finally, a boosted top search [16] based mostly on the jet mass is presented in this proceedings.

2. Pileup Correction

Energy deposits from additional collisions in the same bunch crossing may hinder substructure studies. This problem becomes more severe at the large luminosities present at the LHC. A data driven technique was used to measure the shift in several observables caused by additional incoherent energy in the jet and correct for it [12, 17]. A complementary cone, one with the same size parameter R, is set at 90^o in azimuth away from the leading jet. The energy deposits inside this cone are rotated back and added to the jet. Then the shift in any observable can be measured as a function of the observable value. Underlying event contribution was separated from that of the multiple interactions by exploiting single and multiple vertex events. This technique was implemented on the jet mass, reducing the systematic shift in the jet mass distribution due to pileup. Finally, an approximate analytic calculation to these shifts in the observables is presented in [17], and the relation with the concept of jet area is discussed.

3. Jet Mass

The jet function [14] predicts both the shape of the jet mass distribution, as well as its absolute normalization. In the study presented here, an NLO approximation was used, valid for jet masses above the low mass peak and below $R \cdot p_T$. Higher-order corrections were estimated to be $\sim 30\%$. A comparison between data, Monte Carlo (MC), and the theoretical jet function is shown in Fig. 1(a). The analytical prediction for quark jets approximately described the shape of the distribution and the fraction of jets. A pQCD calculation shows that indeed $\sim 80\%$ of these jets are expected to arise from quarks [18]. The PYTHIA distribution was in reasonable agreement with the data. The data also showed good agreement between the Midpoint and the *anti-*k_t algorithms. This is to be juxtaposed with the fact that the jet mass is an NLO effect since additional gluon emission is needed, while in terms of IR safety, Midpoint is shown to be IR_{3+1} [2] whereas *anti-*k_t is IR-safe.



(a) Mass (b) Angularity (c) Planar Flow

Figure 1: The jet mass (a), angularity (b), and planar flow (c) distributions for central Midpoint jets with $p_T > 400\ GeV/c$. The uncertainties shown are statistical (black lines) and systematic (yellow bars). The theory predictions for the jet function for quarks and gluons are shown as solid curves (purple and blue), as well as the kinematical limits of angularity. Also shown is the PYTHIA MC prediction (red dashed line). The insets compare Midpoint (full black circles) and *anti-k_t* (open green squares) jets [12, 13]

4. Angularity

Angularity was shown to qualitatively distinguish between QCD jets and other two-body decays [15]. As shown in Fig. 1(b) the angularity distribution agrees with the two-prong description of high p_T massive QCD jets, an assumption that the jet function is also based upon. This agreement is manifested by adhering to the two kinematical limits, namely τ_{-2}^{min} which is obtained from symmetric two-body decays, and τ_{-2}^{max} which is obtained when one decay daughter is hard and almost collinear with the mother particle, whereas the second decay daughter is soft and emitted at a large angle. A good agreement was observed between the data and the PYTHIA sample. Furthermore, as in the jet mass case, the angularity distributions obtained when applying Midpoint and *anti-k_t* were in good agreement with each other [12].

5. Planar Flow

Planar flow [14, 15] describes the way energy is deposited on the plane perpendicular to the jet axis. Two-prong boosted massive QCD jets are expected to leave two centers of energy deposits. The soft contribution inside the jet shifts the most probable value of planer flow to a value somewhat higher than the ideal value of zero. On the other hand, three-prong jets coming *e.g.* from hadronic decays of boosted tops, are expected to have a rather uniform planar flow distribution. The mass window of 130 to 210 GeV/c^2 is relevant for searches of boosted tops. The planar flow distribution of data for jets in this mass window exhibited the expected QCD like behavior and peaked at a low planar flow value as shown in Fig. 1(c). The $t\bar{t}$ MC distribution was flatter as expected. Thus planar flow can be used to separate massive boosted QCD jets from top jets. Once again the two jet algorithms showed good agreement.

6. Search for Boosted Tops

The knowledge gained from this study on jet substructure and jet mass in particular can be applied to a boosted top search, using side bands to estimate the dominant QCD background. In principle, the masses of the two leading jets in a typical QCD dijet event are independent of each other. On the other hand, in all hadronic $t\bar{t}$ decays, one expects that both jets are massive. An upper cut on the significance of the missing energy was applied in this case to suppress semileptonic events. The signal region, D, consisted of two massive jets in the range of 130 to 210 GeV/c^2. The definitions of the control regions and the number of counted events in each region can be found in [16]. The background was estimated using three other control regions defined by the masses of the two jets. It included a factor that estimated possible correlations between the two masses as defined in [19].

Missing energy in QCD events comes mostly from instrumental effects and is not correlated to the leading jet mass. In semileptonic $t\bar{t}$ events, one expects a leading massive jet along with large missing energy. The results from this semileptonic channel can be found in [16], as well as the background prediction for the signal region obtained in a similar fashion to that of the fully hadronic case.

Using the data-driven background estimation for these two channels, and 4.9 expected $t\bar{t}$ events, an observed upper limit on the production cross section of Standard Model (SM) $t\bar{t}$ for top quarks with $p_T > 400$ GeV/c was calculated to be 38 fb at 95% C.L. [20]. This is approximately an order of magnitude higher than the estimated SM rate, and is limited by the QCD background rates. It is, however, the most stringent limit on boosted top quark production to date. The expected limit of 33 fb at 95% C.L. is lower than the observed limit since an excess of events was observed in the data in the all hadronic channel. This channel is sensitive to pair production of two massive objects near the mass of the top quark. As the main interest in this case is in beyond-SM contributions to this final state, the background estimate the for the expected $t\bar{t}$ contribution of 3 ± 0.8 events was included. Taking out the top quark hadronic branching fraction of 4/9, the upper limit is 20 fb at 95% C.L.

7. Conclusion

The substructure of high p_T massive jets was studied in real data for the first time. A data-driven pileup correction procedure was developed, tested, and studied analytically. Good agreement between data, MC, and theoretical predictions was observed for all studied jet substructure observables. An interesting agreement between Midpoint and *anti-*k_T jet algorithms was shown. A boosted top search yielded an upper limit on the production cross section of SM high p_T tops. These results can serve as a stepping stone for current and futures studies involving high p_T jets and searches for boosted objects at the LHC.

References

[1] S. D. Ellis, J. Huston, K. Hatakeyama, P. Loch, M. Tonnesmann, *Jets in hadron-hadron collisions*, Prog. Part. Nucl. Phys. **60**, 484-551 (2008) [hep-ph/0712.2447].

[2] G. P. Salam, *Towards Jetography*, Eur. Phys. J. **C67**, 637-686 (2010) [hep-ph/0906.1833].

[3] P. Nath, B. D. Nelson, H. Davoudiasl, B. Dutta, D. Feldman, Z. Liu, T. Han, P. Langacker *et al.*, *The Hunt for New Physics at the Large Hadron Collider*, Nucl. Phys. Proc. Suppl. **200-202**, 185-417 (2010) [hep-ph/1001.2693].

[4] J. M. Butterworth, A. R. Davison, M. Rubin, G. P. Salam, *Jet substructure as a new Higgs search channel at the LHC*, Phys. Rev. Lett. **100**, 242001 (2008) [hep-ph/0802.2470].

[5] J. M. Butterworth, B. E. Cox, J. R. Forshaw, *WW scattering at the CERN LHC*, Phys. Rev. **D65**, 096014 (2002) [hep-ph/0201098].

[6] K. Agashe, A. Belyaev, T. Krupovnickas, G. Perez, J. Virzi, *LHC Signals from Warped Extra Dimensions*, Phys. Rev. **D77**, 015003 (2008) [hep-ph/0612015].

[7] B. Lillie, L. Randall, L. T. Wang, *The Bulk RS KK-gluon at the LHC*, JHEP **0709**, 074 (2007) [hep-ph/0701166].

[8] J. M. Butterworth, J. R. Ellis, A. R. Raklev, G. P. Salam, *Discovering baryon-number violating neutralino decays at the LHC*, Phys. Rev. Lett. **103**, 241803 (2009) [hep-ph/0906.0728].

[9] V. M. Abazov *et al.* (D0 Collaboration), *Subjet multiplicity of gluon and quark jets reconstructed with the k_T algorithm in $p\bar{p}$ collisions*, Phys. Rev. **D65**, 052008 (2002) [hep-ex/0108054].

[10] D. E. Acosta *et al.* (CDF Collaboration), *Study of jet shapes in inclusive jet production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, Phys. Rev. **D71**, 112002 (2005) [hep-ex/0505013].

[11] G. Aad *et al.* (Atlas Collaboration), *Study of Jet Shapes in Inclusive Jet Production in pp Collisions at sqrt(s) = 7 TeV using the ATLAS Detector*, Phys. Rev. **D83**, 052003 (2011) [hep-ex/1101.0070].

[12] T. Aaltonen *et al.* (CDF Collaboration), *Study of Substructure of High Transverse Momentum Jets Produced in Proton-Antiproton Collisions at $\sqrt{s} = 1.96$ TeV*, hep-ex/1106.5952.

[13] T. Aaltonen *et al.* (CDF Collaboration), *The Substructure of High Transverse Momentum Jets Observed by CDF II*, CDF/PUB/JET/PUBLIC/10199

[14] L. G. Almeida, S. J. Lee, G. Perez, I. Sung, J. Virzi, *Top Jets at the LHC*, Phys. Rev. **D79**, 074012 (2009) [hep-ph/0810.0934].

[15] L. G. Almeida, S. J. Lee, G. Perez, G. F. Sterman, I. Sung, J. Virzi, *Substructure of high-p_T Jets at the LHC*, Phys. Rev. **D79**, 074017 (2009) [hep-ph/0807.0234].

[16] T. Aaltonen *et al.* (CDF Collaboration), *A Search for Boosted Top Quarks by CDF II*, CDF/ANAL/TOP/PUBLIC/10234

[17] R. Alon, E. Duchovni, G. Perez, A. P. Pranko, P. K. Sinervo, *A Data-driven method of pile-up correction for the substructure of massive jets*, Phys. Rev. **D84**, 114025 (2011) [hep-ph/1101.3002].

[18] R. K. Ellis, W. J. Stirling, and B. R. Webber, *QCD and Collider Physics*, Cambridge University Press, 1994.

[19] Y. Eshel, O. Gedalia, G. Perez, Y. Soreq, *Implications of the Measurement of Ultra-Massive Boosted Jets at CDF*, Phys. Rev. **D84**, 011505 (2011) [hep-ph/1101.2898].

[20] C. Amsler *et al.* (Particle Data Group Collaboration), *Review of Particle Physics*, Phys. Lett. **B667**, 1-1340 (2008). A Bayesian technique was used similar to the one employed in a recent search for single top production reported in: T. Aaltonen *et al.* (CDF Collaboration), *First Observation of Electroweak Single Top Quark Production*, Phys. Rev. Lett. **103**, 092002 (2009) [hep-ex/0903.0885].