Substructure of Highly Boosted Massive Jets at CDF II

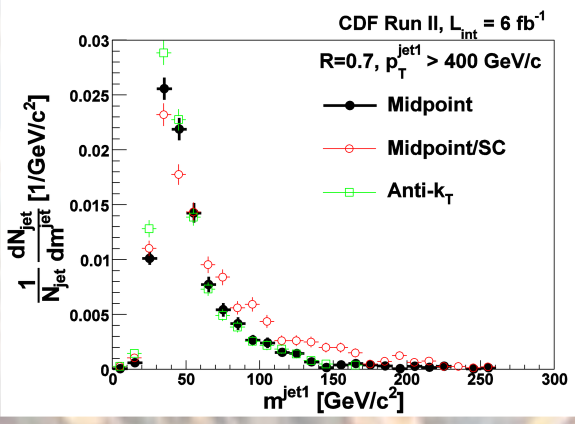
Weizmann Institute of Science

University of Toronto

Ehud Duchovni, Gilad Perez, Pekka K. Sinervo, Raz Alon

1. Introduction

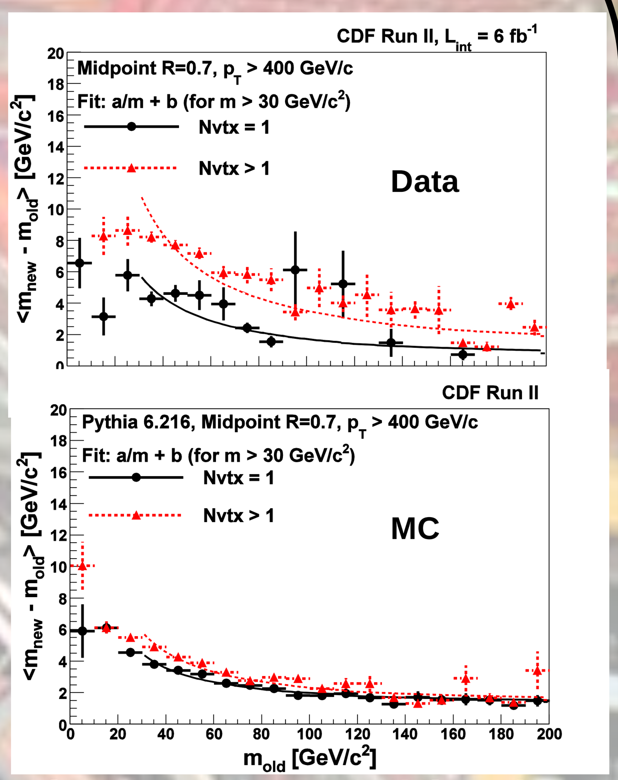
The study of high transverse momentum (p_T) massive jets produced in proton-antiproton interactions provides an important test of perturbative QCD (pQCD). Furthermore, massive boosted jets comprise an important background in searches for various new physics signatures, the Higgs boson, and highly boosted top quark pair production. Hadronic decay products of such objects will typically be detected as a single jet with a substructure that differs statistically from the expected backgrounds arising from pQCD jet production.



Events were selected out of a sample of 6 fb^{-1} and were required to have at least one central jet with p_T>400 GeV/c. Events with a massive recoil jet or large missing energy were rejected, thus casting aside possible hadronic and semileptonic ttbar events. The jet mass spectrum of the leading jets in the remaining 2,108 events is shown in the plot for three different jet algorithms.

2. Pileup Correction

The number of interaction vertices (Nvtx) is a measure of multiple interactions (MI), i.e. additional collisions in the same bunch crossing. These occur in this data sample at an average rate of Nvtx~3 (including the primary interaction). A novel data-driven method for correcting the MI effect on the jet substructure variables was developed. The technique examines the effect of adding into the jet energy deposited in a cone that is 90º away in azimuth from the jet. Furthermore the MI contribution was extracted from the total of MI and underlying event (UE) contributions.



An analytical approximation was derived for the relevant observables by calculating their variation with respect to the addition of incoherent energy (hep-ph 1101.3002v2). The MI contribution for the jet mass was expected to have a $1/m^{jet}$ behavior, as demonstrated in the plots for both data and Monte Carlo (MC). The latter sample contains many more events and thus behaves statistically nicer.

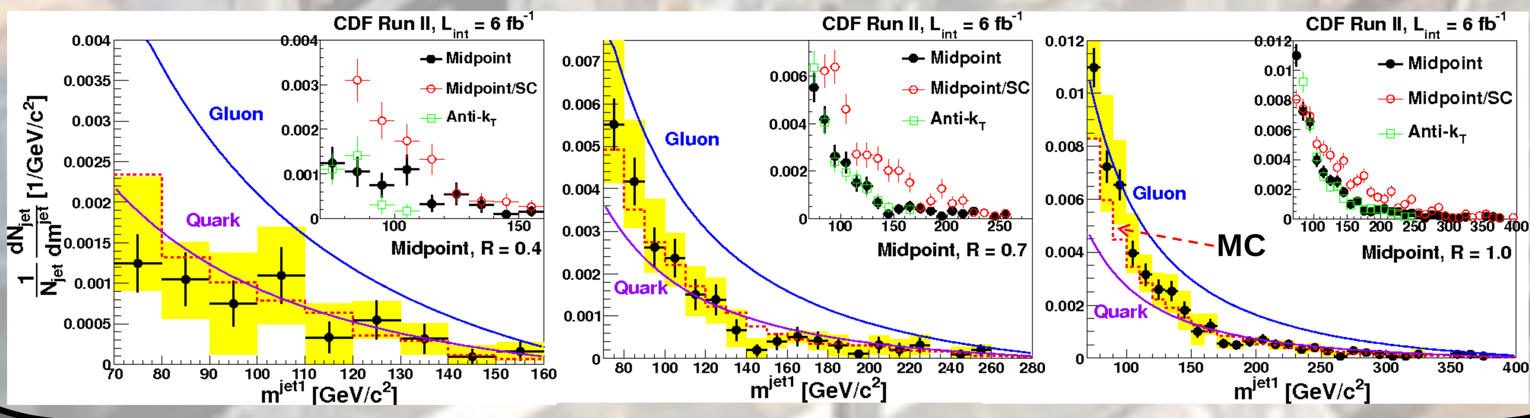
3. Jet Mass

In the leading-log approximation, QCD predicts that a high mass jet acquires that mass, m^{jet}, through a single hard gluon emission. The probability of this process is given by the jet function, for which an approximation is

$$J(m^{jet};p_T;R) \sim \alpha_s(p_T)\ ((4\ C_{q,g}) / (\pi\ m^{jet}))\ \log(R\ p_T / m^{jet})$$

where $\alpha_s(p_T)$ is the strong coupling constant, $C_{q,g}$= 4/3 (3) for quark (gluon) jets, and R is the cone radius used to associate final state particles with the jet. Although uncertainties in this approximation arising from higher-order corrections are~30%, it predicts both the shape of the spectrum and the absolute fraction of high mass jets.

The plots show comparisons of m^{jet} distributions for three different cone sizes with the analytic predictions for the jet function for quark and gluon jets.



4. Angularity

Angularity is defined by:

$$\tau_{-2}(R,p_T) \equiv \frac{1}{m^{jet}} \sum_{i \in jet} E_i \sin^{-2}\theta_i (1-\cos\theta_i)^3$$

where the sum is over the constituents of the jet, E_i are constituent energies, and θ_i are the angles of the constituents relative to the jet axis.

A key prediction of the NLO QCD calculation is that the distribution of angularities of high mass jets has sharp kinematical edges, with minimum and maximum values given by:

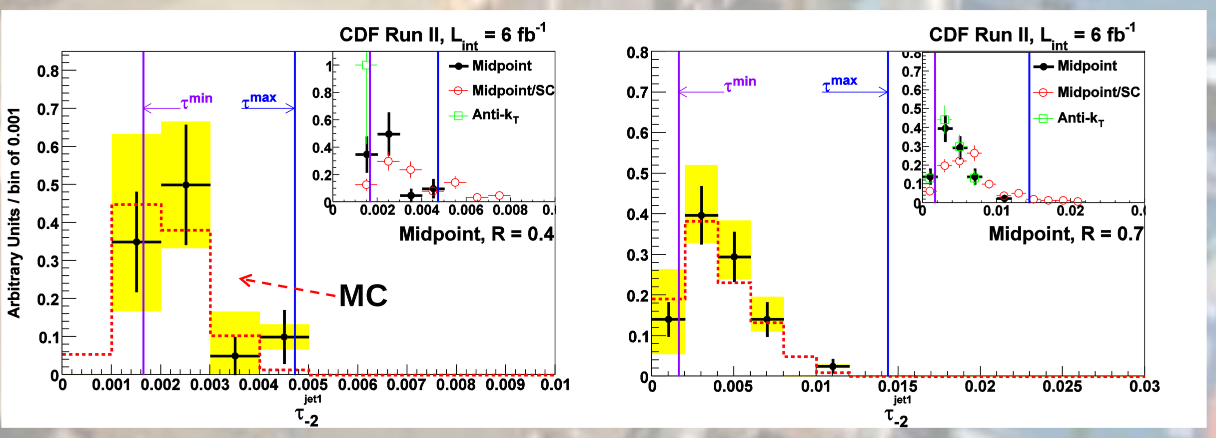
$$\tau_{-2}^{min} \approx \left(\frac{2p_T}{m^{jet}}\right)^{-3}$$

Two equally energetic daughters at the same angle w.r.t. the decaying particle

$$\tau_{-2}^{max} \approx \frac{m^{jet}}{p_T}\frac{R^2}{2^3}$$

One energetic daughter at a small angle and one soft daughter at a large angle w.r.t. the decaying particle

A comparison of the observed angularity distribution with the Pythia prediction is shown for jets with masses in the range 90-120 GeV/c^2.



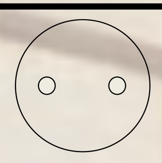
5. Planar Flow

Planar flow is defined by:

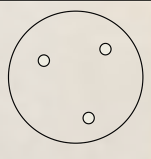
$$P_f \equiv \frac{4\lambda_1\lambda_2}{(\lambda_1+\lambda_2)^2};\ \lambda_{1,2}\ \text{eigenvalues of}\ I^{kl} \equiv \frac{1}{m^{jet}} \sum_{i \in jet} E_i \frac{p_{i,k}}{E_i}\frac{p_{i,l}}{E_i}$$

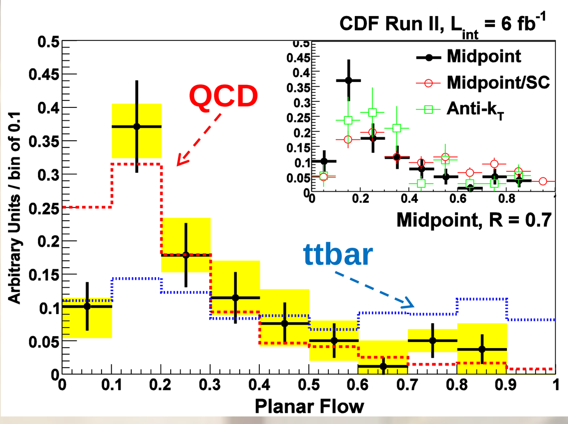
in which $p_{i,k}$ is the k^{th} component of the jet constituent's transverse energy relative to the jet axis. This observable is expected to provide additional separation between QCD jets and boosted objects with multi-prong decay kinematics.

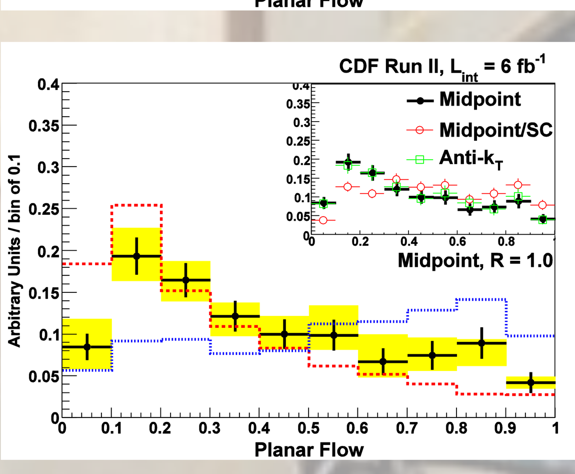
Two-prong massive QCD jet
Low planar flow



Three-prong massive top jet
High planar flow



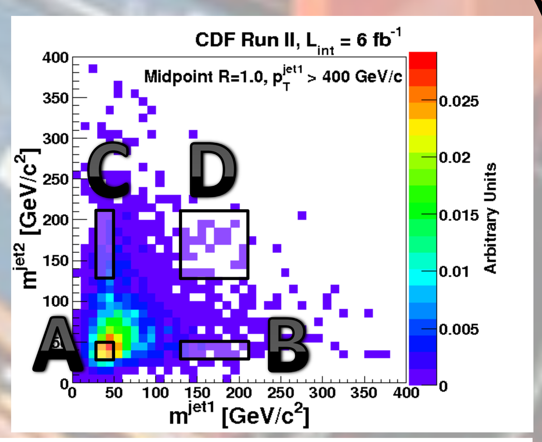




The plots show the planar flow distributions for jets with masses in the range 130-210 GeV/c^2, relevant for jets arising from top quark decays. Comparisons with the Pythia predictions are also shown for both QCD multi-jet and ttbar production.

6. Top Search

ttbar events in which both top quarks decay hadronically are expected to have two massive jets and a low missing energy. Such a correlation should not exist in the dominant background, namely QCD. Consequently the number of events in region D in the top plot, N_D, can be predicted from the data as:



$$N_D^{pred}=N_B N_C/N_A$$

Similarly, in case one of the top quarks decays semileptonically, the event is expected to have only one massive jet and a large amount of missing energy. The tables summarize the counting results together with a MC based ttbar content prediction.

All Hadronic — CDF, L_{int} = 6 fb^{-1}

Region	m^{jet1} (GeV/c^2)	m^{jet2} (GeV/c^2)	Data (events)	ttbar MC (events)
A	(30, 50)	(30, 50)	370	0.00
B	(130, 210)	(30, 50)	47	0.08
C	(30, 50)	(130, 210)	102	0.01
D (signal)	(130, 210)	(130, 210)	32	3.03
Predicted QCD in D	13±2.4			

Semileptonic — CDF, L_{int} = 6 fb^{-1}

Region	m^{jet1} (GeV/c^2)	S_{MET} (GeV$^{1/2}$)	Data (events)	ttbar MC (events)
A	(30, 50)	(2, 3)	256	0.01
B	(130, 210)	(2, 3)	42	1.07
C	(30, 50)	(4, 10)	191	0.03
D (signal)	(130, 210)	(4, 10)	26	1.90
Predicted QCD in D	31.3±8.1			

Using both channels, an upper limit on the cross section of top quark production with p_T > 400 GeV/c is 40 fb at 95% C.L. A recent study (hep-ph 1101.2898v2) shows that $R_{mass}=(N_B N_C)/(N_A N_D)$ < 1. This will increase the predicted number of background events in the signal region.

Conclusion

Distributions of the jet mass, angularity, and planar flow, were measured for the first time for jets with p_T > 400 GeV/c.

Good agreement was found between the Pythia MC prediction, NLO QCD jet function prediction, and the data for the jet mass distribution above 100 GeV/c^2.

The angularity distributions showed that high mass jets coming from light quark and gluon production are consistent with the two-prong final state prediction.

Planar flow was shown to provide separation power between ttbar and high mass QCD jets.

A counting experiment technique was used to search for boosted top jets in an all hadronic channel and a semileptonic channel. This search resulted in an upper bound on the cross section for boosted top quark production.

More information can be found in the following webpages:

www-cdf.fnal.gov/physics/new/qcd/BoostedJets

www-cdf.fnal.gov/physics/new/top/2011/BoostedTops