Substructure of Massive Boosted Jets & Boosted Tops at CDF II

Raz Alon, on behalf of the CDF Collaboration

Outline

1. Motivation
2. Pileup Correction
3. Jet Mass
4. Angularity
5. Planar Flow
6. Top Search

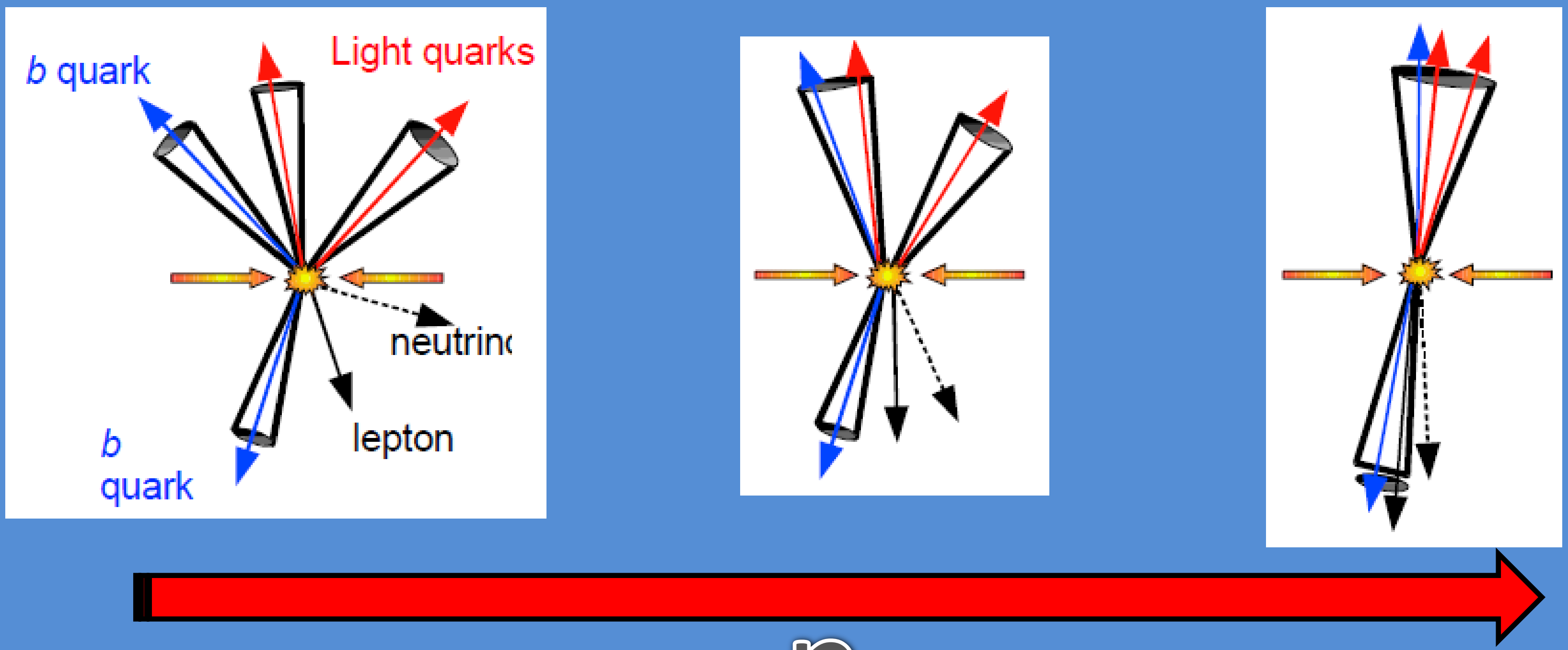
CDF 10199 (arXiv:1106.5952 submitted to PRL)
www-cdf.fnal.gov/physics/new/qcd/BoostedJets
CDF 10234
www-cdf.fnal.gov/physics/new/top/2011/BoostedTops

Motivation

- Jets with high transverse momentum (p_T) and high mass are a playground for testing perturbative QCD (pQCD)

Motivation
Why Substructure?

Resolved **Unresolved**

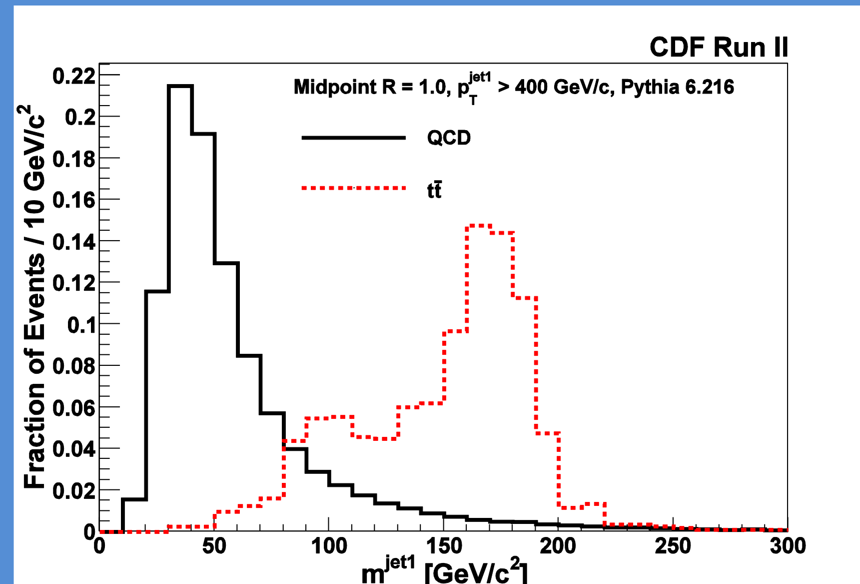


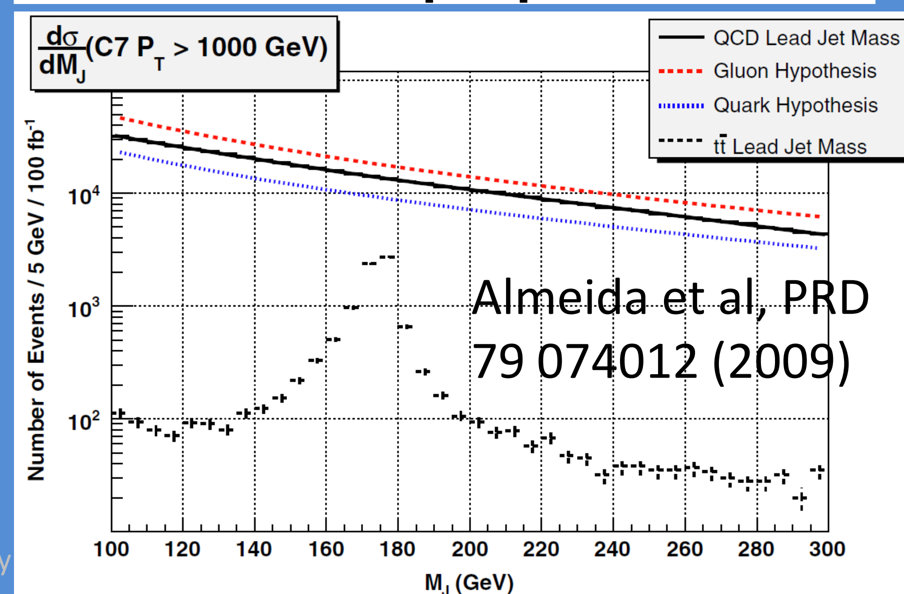
p_T

Almost every jet in this talk has p_T > 400 GeV/c

Motivation

- Jet mass distributions of QCD (light quarks and gluons) and top are distinguishable by shape
- Top events are swamped by QCD events
- QCD jets are a dominant background for searches of
  - Boosted tops
  - Boosted Higgs
  - New physics signatures





Almeida et al, PRD 79 074012 (2009)

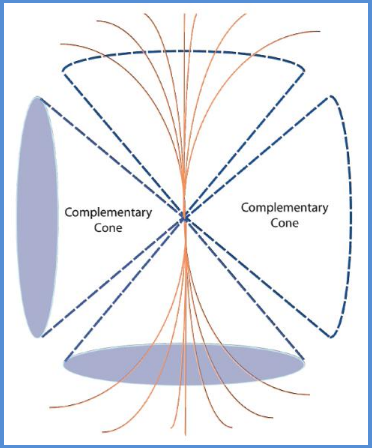
<u>Outline</u>

1. Motivation
2. Pileup Correction
3. Jet Mass
4. Angularity
5. Planar Flow
6. Top Search

Pileup

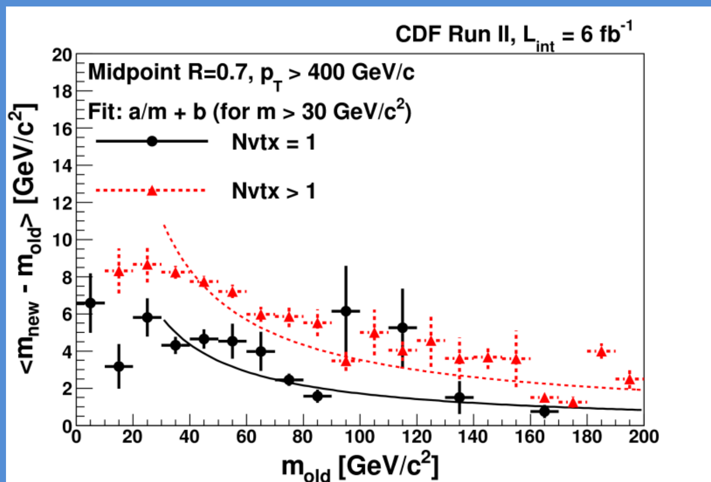
- There are 3 Multiple Interactions (MI) on average in our data in each bunch crossing.
- Can be estimated by the number of reconstructed vertices (Nvtx)
- Introduces extra energy depositions inside the jets
- Critical to jet substructure study
- Creates a problem of comparison to MC

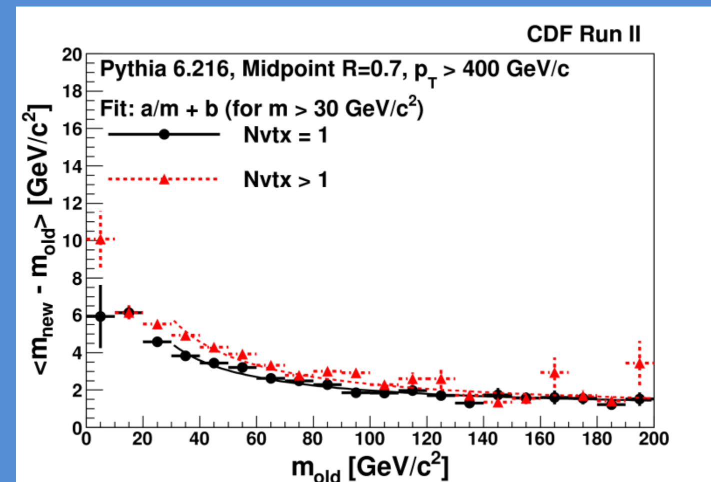
- Complementary Cone
  - 90^o away in ϕ from jet
  - Requiring clear dijet structure
    - $p_T^{jet2} > 100$ GeV/c
    - $\Delta\phi(jet1,jet2) > 166^o$
- Add calorimeter towers from complementary cone to main jet and see how the mass changes
- arXiv:1101.3002



Effect on Jet Mass

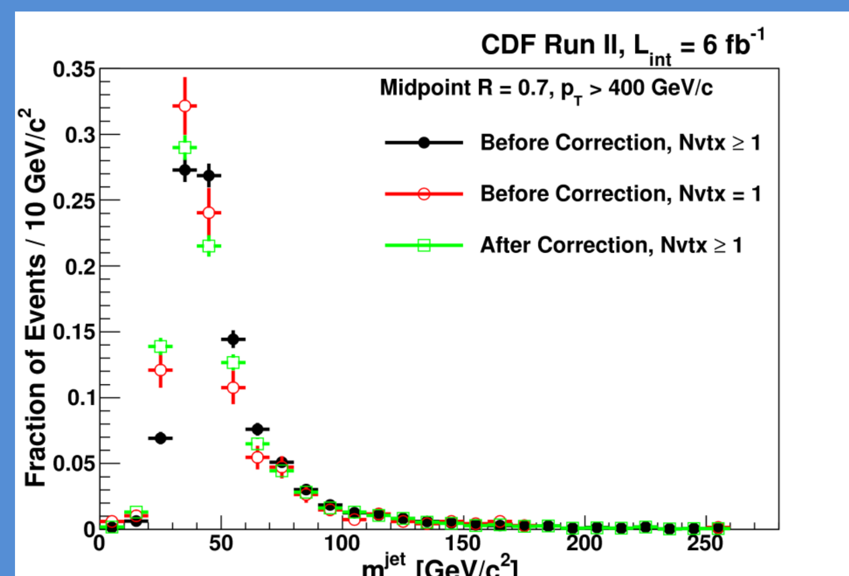
Data





MC

- Small relative effect at high mass
- Brings the low mass peak closer to that of single-vertex events

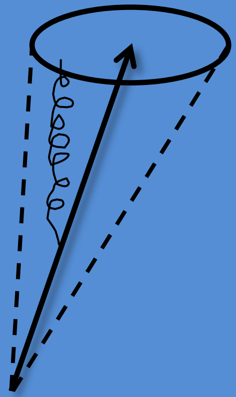


Outline

1. Motivation
2. Pileup Correction
3. Jet Mass
4. Angularity
5. Planar Flow
6. Top Search

Jet Mass

- An analytic calculation of the QCD jet mass distribution



$$\frac{d\sigma(R)}{dp_T dm_J} = \sum_{q,G} J^{q,G}(m_J, p_T, R) \frac{d\hat{\sigma}^{q,G}(R)}{dp_T}$$

$$J(m_J, p_T, R) \simeq \alpha_s(p_T) \frac{4C_{q,G}}{\pi m_J} \log\left(\frac{R\, p_T}{m_J}\right)$$

$$C_{q,G} = 4/3, 3 \text{ for quark and gluon}$$

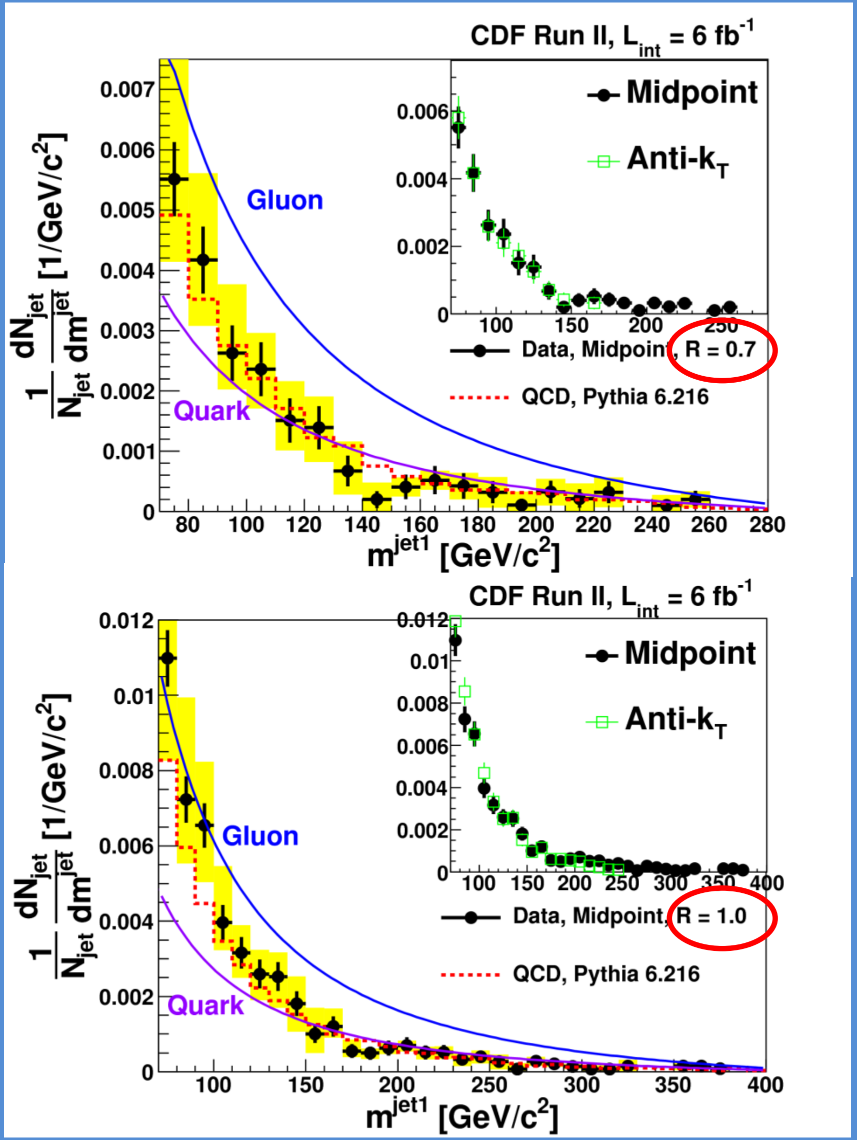
Almeida et al, PRD 79 074012 (2009)

Jet Mass

- Validity: $m^{jet}_{peak} \ll m^{jet} \ll p_T R$
- Leading log approximation. Next to leading log corrections are not fully known but roughly $1/\log(R^2 p_T^2/m^2) \sim 30\%$
- Predicts both
  - Jet mass distribution shape
  - Absolute normalization
- Still, a simple and powerful prediction

Jet Mass

- Good data and MC agreement
- Agreement with quark hypothesis at high mass (consistent with expected ~80% quark jets at high p_T)
- Good agreement between Midpoint (IR_{3+1} G.P. Salam arXiv:0906.1833) and anti-k_T (IR safe)



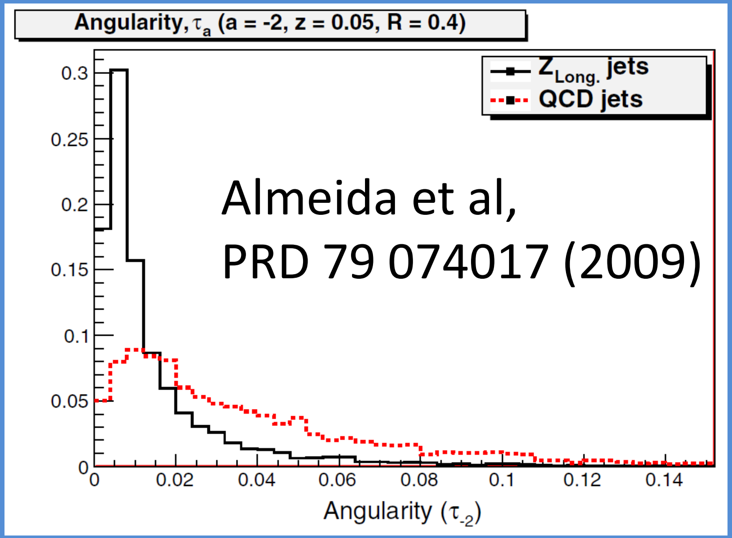
Outline

1. Motivation
2. Pileup Correction
3. Jet Mass
4. Angularity
5. Planar Flow
6. Top Search

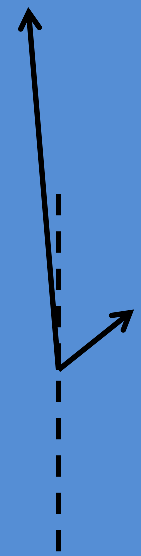
Angularity

$$\tau_a(R, p_T) = \frac{1}{m_J} \sum_{i \in jet} \omega_i \sin^a \theta_i \left[1 - \cos\theta_i\right]^{1-a} \sim \frac{2^{a-1}}{m_J} \sum_{i \in jet} \omega_i \theta_i^{2-a}$$

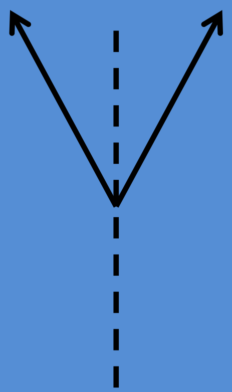
- ω_i component's energy
- θ_i angle w.r.t. the jet axis
- We use a=-2



Kinematic Limits



$$\tau_a^{\max}(z,R) \sim 2^{a-1} R^{-a}$$

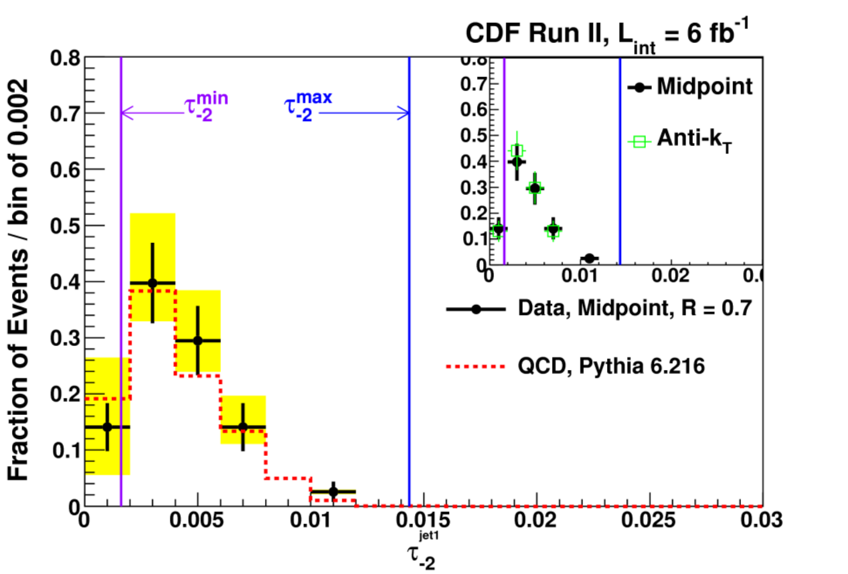


$$\tau_a^{\min}(z) \sim \left(\frac{z}{2}\right)^{1-a}$$

$$\theta_{1,2} = z = m/p_T$$

Angularity

- $90 < m^{jet} < 120$ GeV/c^2
- Good data and MC agreement
- Good agreement with kinematic limits
    - Consistent with two-prong hypothesis
- Good agreement between Midpoint and anti-k_T



Outline

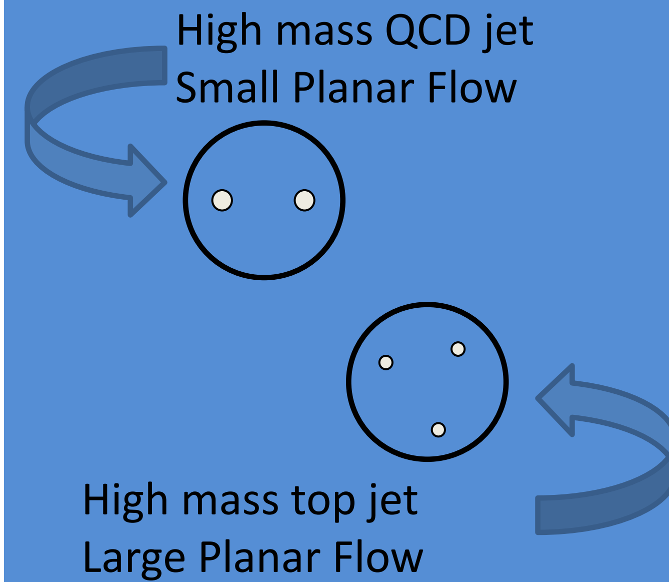
1. Motivation
2. Pileup Correction
3. Jet Mass
4. Angularity
5. Planar Flow
6. Top Search

Planar Flow

$$I_w^{kl} = \frac{1}{m_J} \sum_i w_i \frac{p_{i,k}}{w_i} \frac{p_{i,l}}{w_i}$$

$$Pf = 4\,\frac{\det(\mathrm{I_w})}{\mathrm{tr}(\mathrm{I_w})^2} = \frac{4\lambda_1\lambda_2}{(\lambda_1+\lambda_2)^2}$$

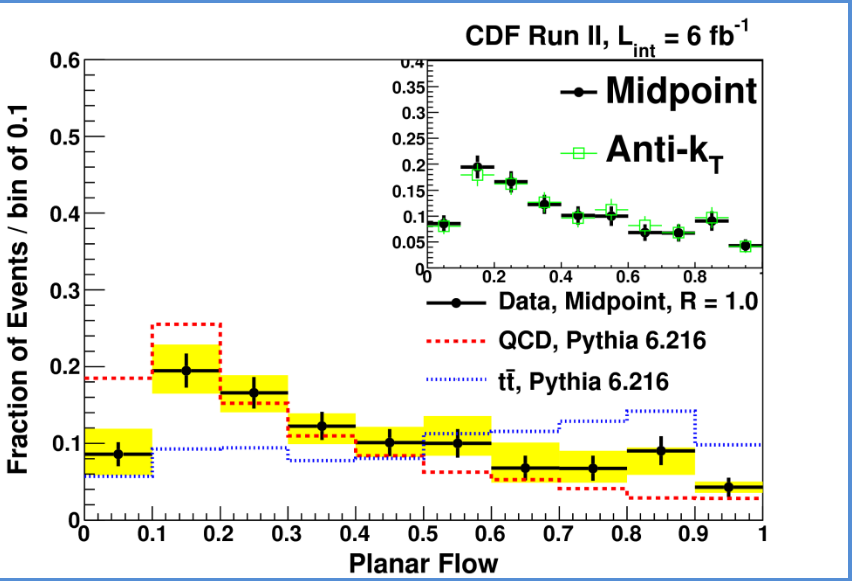
$\lambda_{1,2}$ are the eigenvalues of I_w



- Vanishes for linear shapes and approaches unity for isotropic depositions of energy

Planar Flow

- $130 < m^{jet} < 210\ GeV/c^2$
- Data are QCD like
  - Effective variable not strongly correlated with jet mass
- Good agreement between Midpoint and anti-k_T

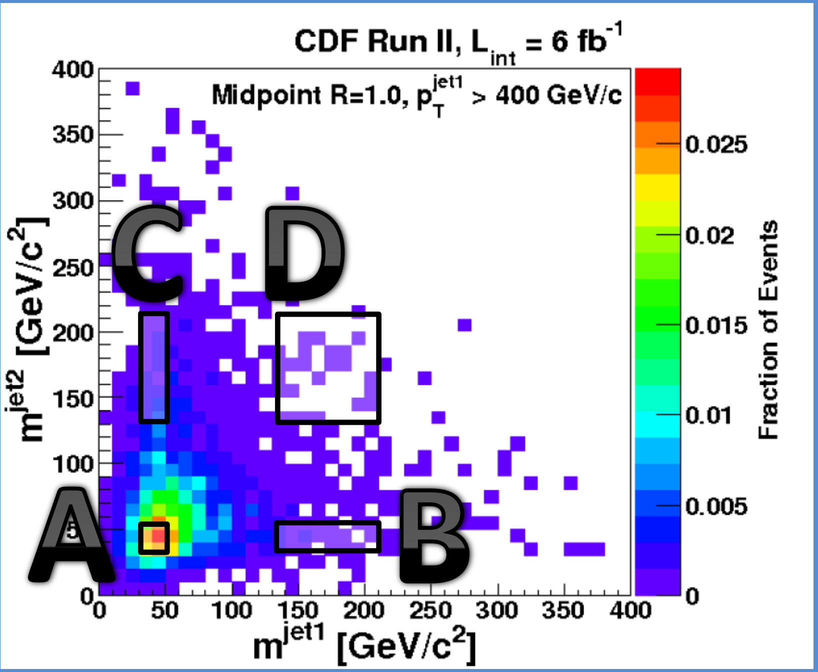


<u>Outline</u>

1. Motivation
2. Pileup Correction
3. Jet Mass
4. Angularity
5. Planar Flow
6. Top Search

Hadronic Channel

- The masses of the two leading jets are independent in QCD
- Predicting the number of events in region D, N_D, from the data
  - $N_D^{pred}=N_B N_C/N_A$
- Recent studies show that $R_{mass}=(N_B N_C)/(N_A N_D) < 1$

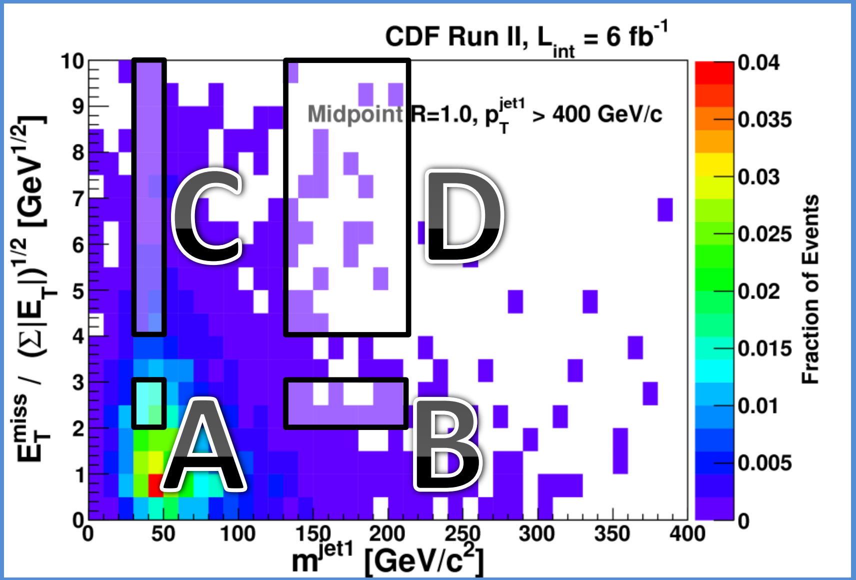


MC tools	Matching	R_{mass}
Sherpa	Yes	0.88 ± 0.03
MadGraph	Yes	0.86 ± 0.04
MadGraph	No	0.76 ± 0.04
Herwig	No	0.86 ± 0.02

Eshel et al, hep-ph 1101.2898
Blum et al, hep-ph 1102.3133

Semileptonic Channel

- Leading massive jet and large missing energy are expected in semileptonic decays
- Using the same technique



Top Search

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			

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)	31	3.03
Predicted QCD in D	14.6±2.76			

<u>Both Channels</u>

- Prediction: 46±8.5(stat)±13.8(syst) with 57 observed events
- $\sigma_{tt}(p_T > 400\ GeV/c) < 38$ fb @ 95% C.L. , expected limit 33 fb @ 95% C.L.

<u>All Hadronic Channel Only</u>

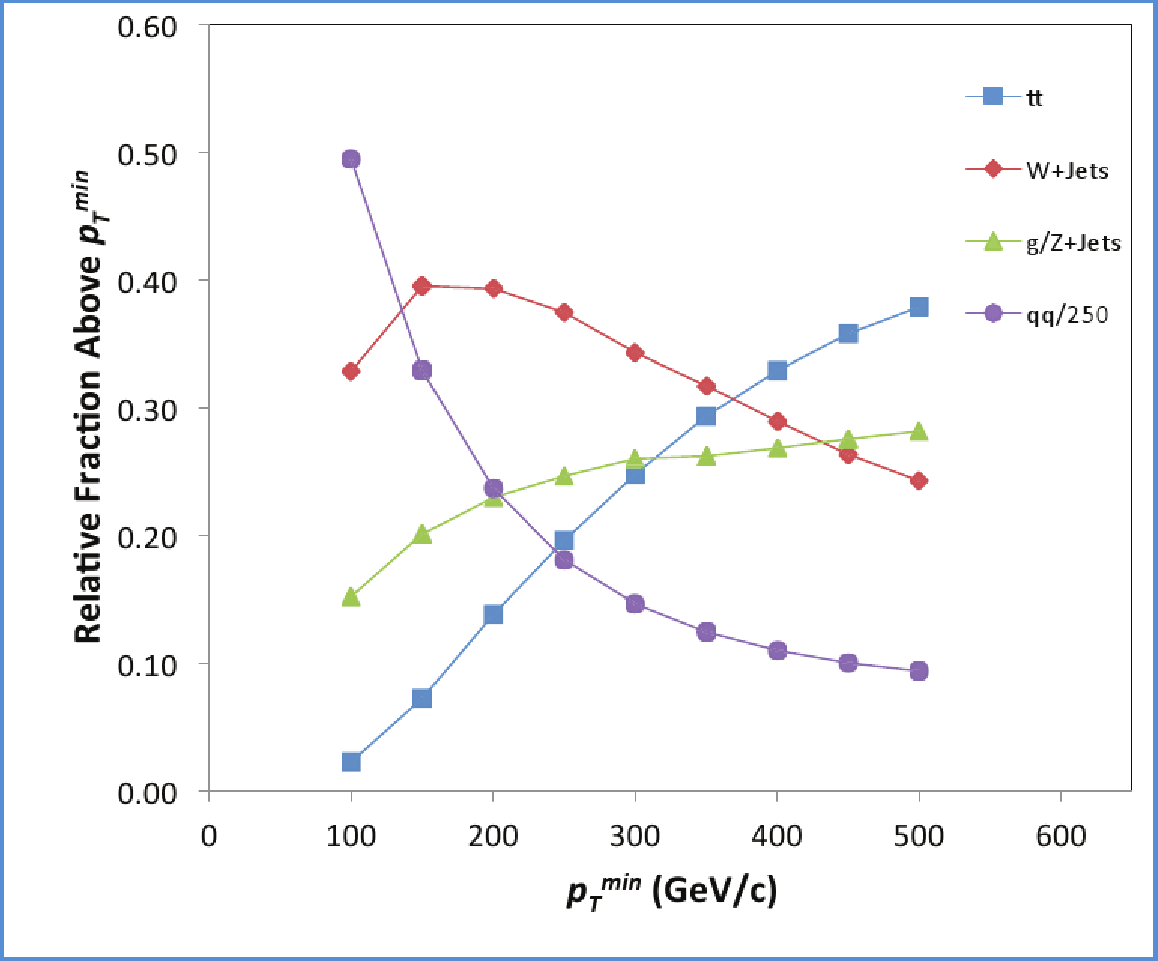
- Prediction: 14.6±2.76(stat)±4.4(syst) with 31 observed events
- Upper limit: 45 fb @ 95% C.L.
- Removing hadronic ttbar branching fraction, upper limit: 20 fb @ 95% C.L.

Summary

- High p_T massive jets were studied in real data for the first time (now confirmed in very recent LHC studies)
- A data-driven pileup correction procedure was developed, tested, and studied analytically
- Good agreement on jet substructure observables between data, MC, and theoretical predictions
- Good agreement between Midpoint and anti-k_T
- A boosted top search yielded an upper limit on the production cross section of high p_T tops
  - Using only jet mass
  - In future searches, other jet substructure observables, *e.g.* planar flow can be used

Backup Slides

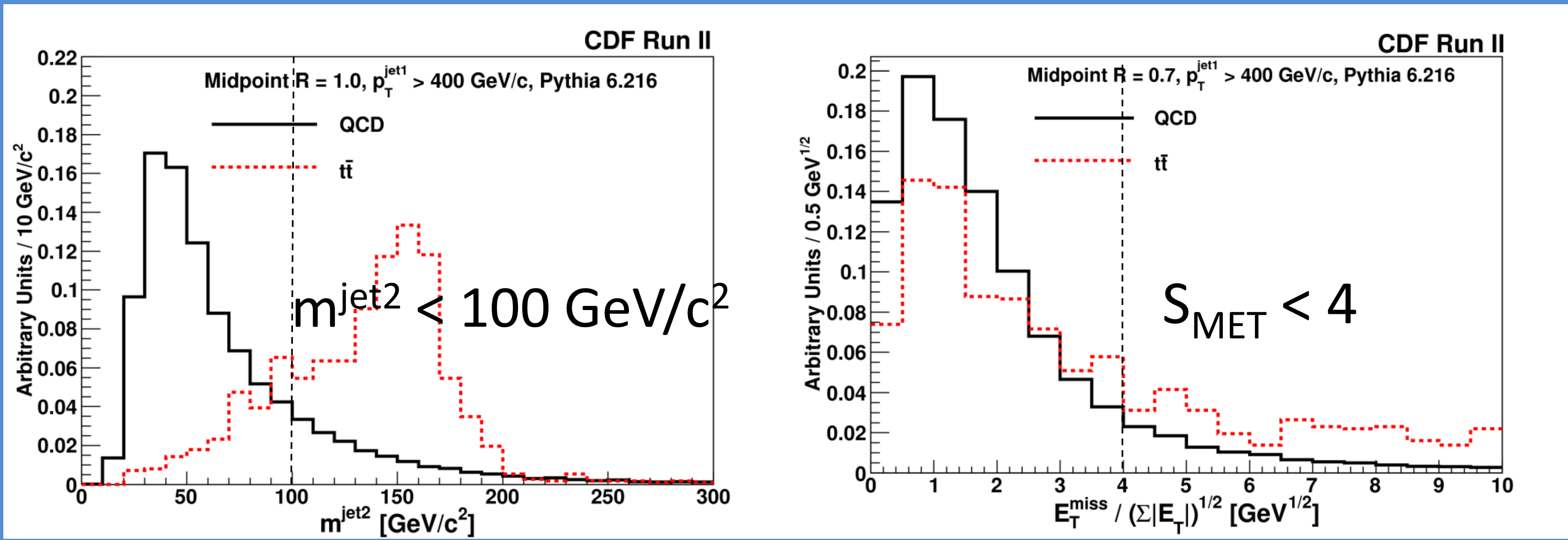
Relative Fraction (Pythia 6.4)



Event Selection

- 6 fb^{-1} of data collected Feb. 2002 – Jan. 2010
- Mainly looking at $p_T^{jet} > 400$ GeV/c
- $0.1 < |\eta| < 0.7$: stay in central parts of CDF
- ≥1 vertex with class 12, $|zvx|<60$cm: quality cuts
- E_T^{miss} Significance = $S_{MET} = E_T^{miss}/\sqrt{\sum E_T} < 10$

Top Rejection/Enhancement Cuts



Efficiency:

~25% for ttbar MC

~80% for QCD MC

Expected ttbar contribution 2.2 fb^{-1}

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

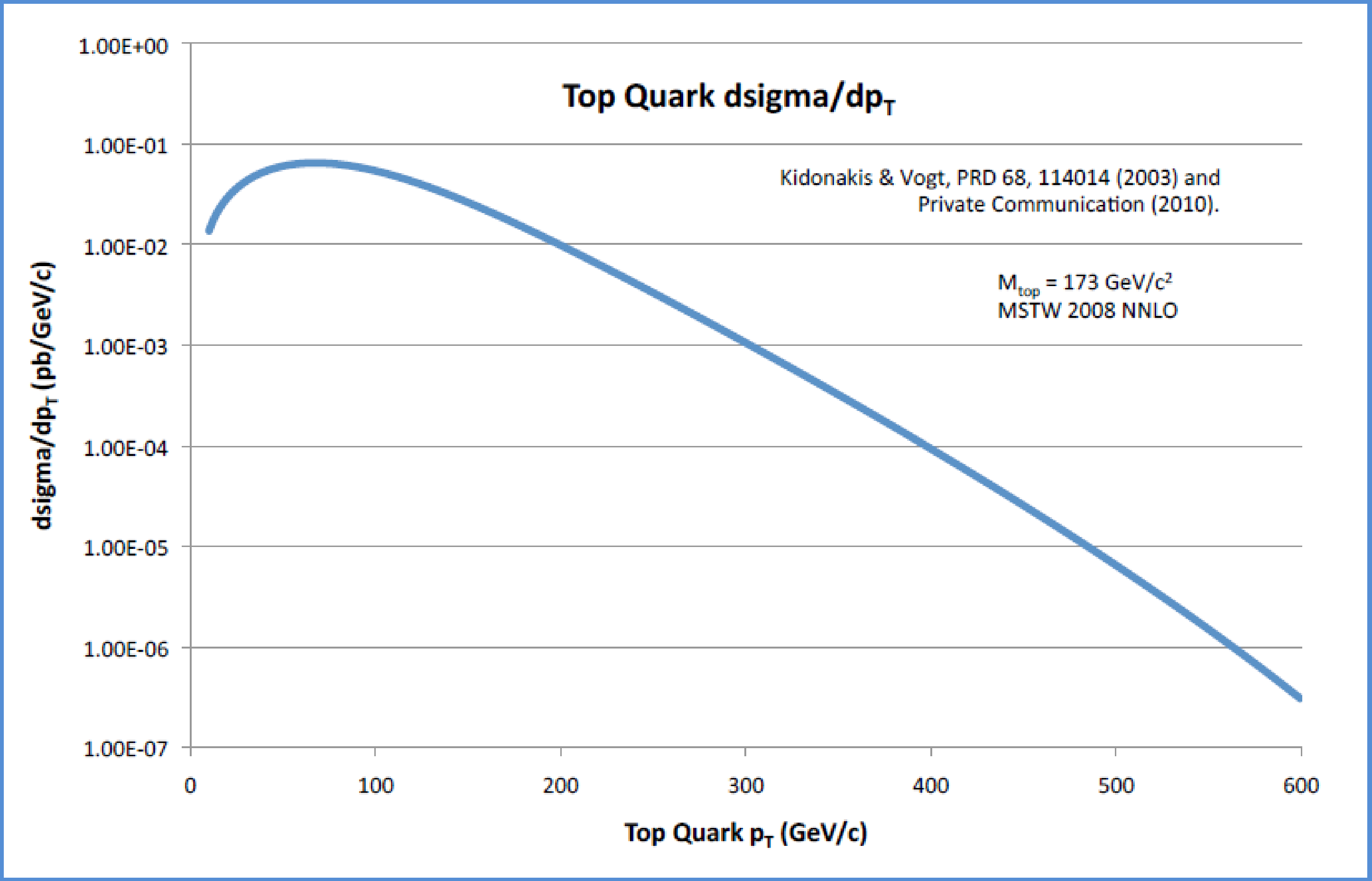
Cut Flow		
	R = 0.4	R = 0.7
All Data, 5.95 fb^{-1}	75,764,270 events	
At least one jet with p_T > 400 GeV/c, 0.1 < $\|\eta\|$ < 0.7, and event quality cuts	2,152	2,699
m^{jet2} < 100 GeV/c^2 and S_{MET} < 4 (with p_T^{jet2} > 100 GeV/c and MI corrections)	1,836	2,108

MI Correction

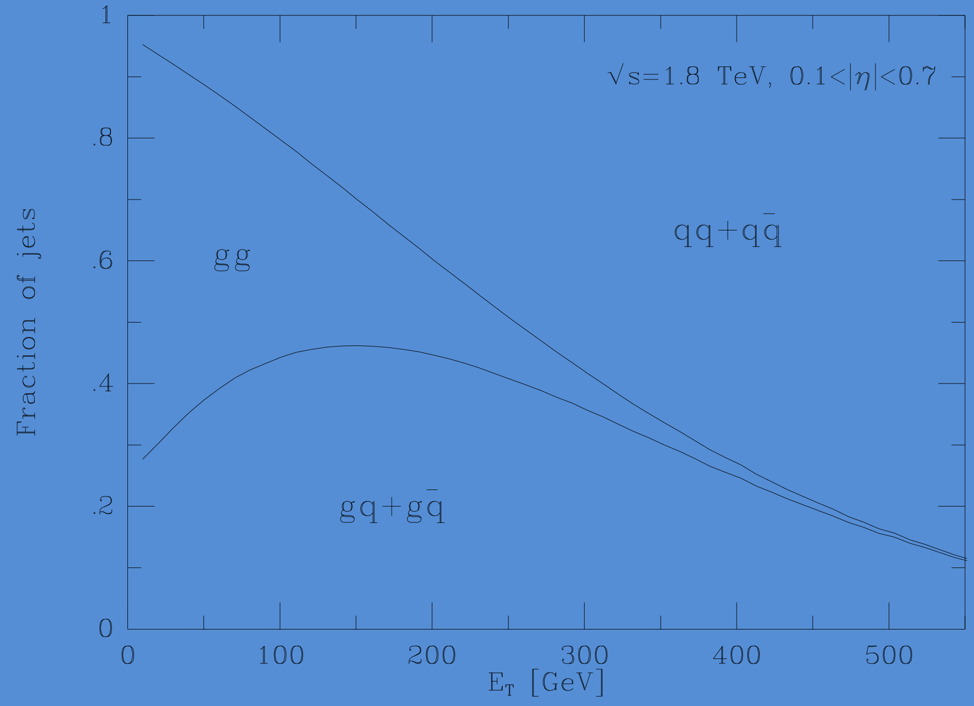
- Scaling of MI correction:
  - Average number of towers in cone ~ R^2
  - Average contribution ~ cosθ ~ R^2
  - Scales like R^4
- Large event by event fluctuations
  - Use the mean shift instead
- Some parts of the underlying event (UE) are coherent with final jet states
  - Nvtx > 1 => UE + MI
    minus
  - Nvtx=1 => UE only
    equals
  - MI Correction

- At a fixed p_T, the mass shift is expected to behave as $\Delta m \sim 1/m$
- A general procedure for deriving the MI contribution for various observables was developed in arXiv:1101.3002v2 by performing a variation of the observable with respect to addition of incoherent energy

Top $d\sigma/dp_T$



Quark and gluon jet fractions at the Tevatron



Ellis, Stirling, Webber, "*QCD and Collider Physics*"

Top Search

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			

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)	31	3.03
Predicted QCD in D	14.6±2.76			

- Both Channels
- Prediction: 46±8.5(stat)±13.8(syst)
- Observed: 57
- $\sigma_{tt}(p_T > 400$ GeV/c$)$ < 38 fb @ 95% C.L.
- Expected number of 4.9 ttbar events in the sample
- Expected limit is 33 fb @ 95% C.L.

- All Hadronic Channel Only
- Predicted: 14.6±2.76(stat)±4.4(syst)
- Observed: 31
- Upper limit: 45 fb @ 95% C.L.
- Removing hadronic ttbar branching fraction
- Upper limit: 20 fb @ 95% C.L.