MEASUREMENT OF THE $t\bar{t}$ PRODUCTION CROSS SECTION IN THE LEPTON + JETS CHANNEL USING JET PROBABILITY AT CDF

Phys. Rev. D. 74, 072006

Enrique Palencia, Instituto de Física de Cantabria
for the CDF Collaboration

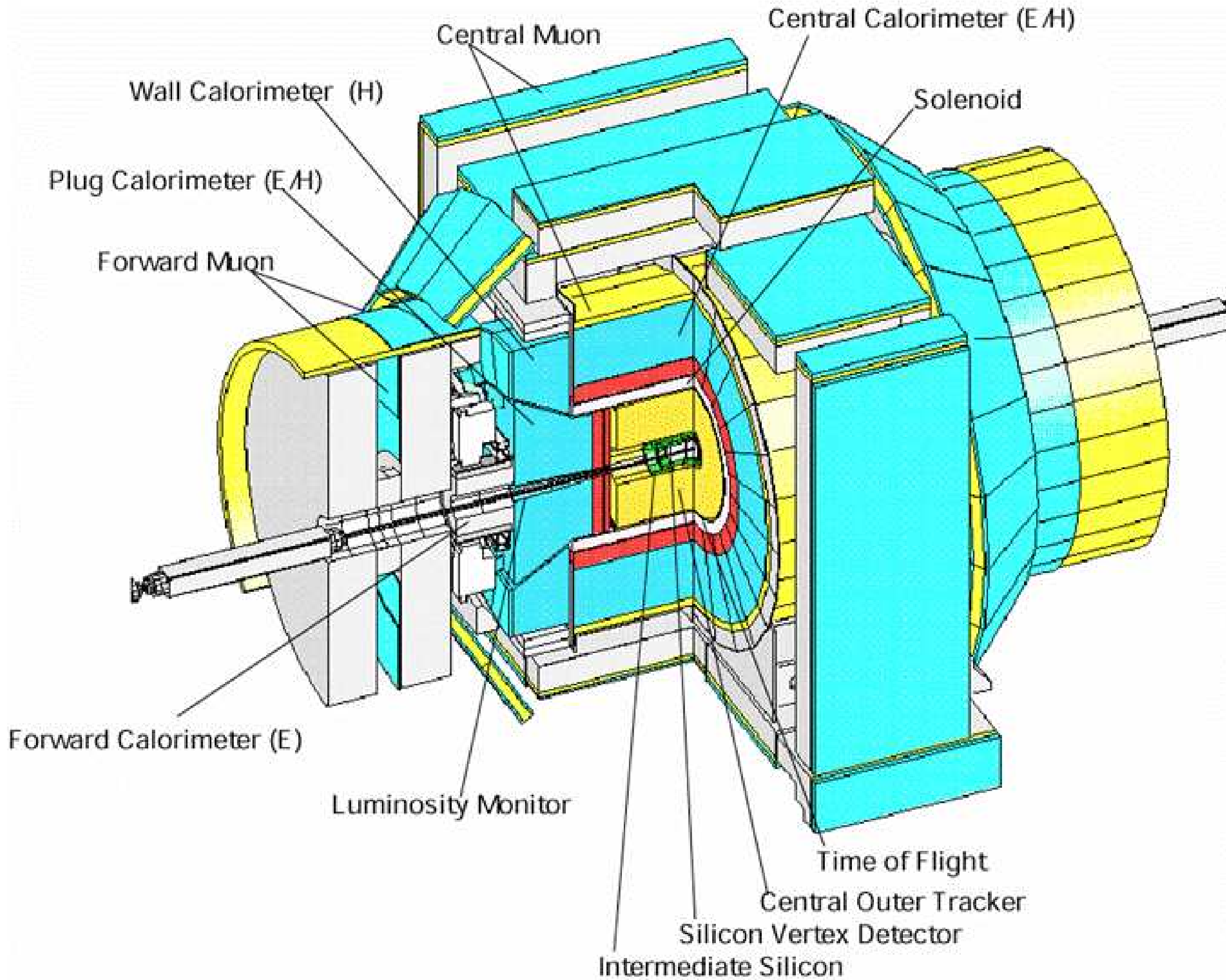
DPF+JPS Meeting, October 29 - November 3, 2006

Outline

- Jet Probability Tagging Algorithm
  - Description of the algorithm
  - Efficiency
  - Mistag Rate
- $t\bar{t}$ Cross Section Measurement
  - Data sample and event selection
  - Acceptance and background estimate
  - Results
- Summary

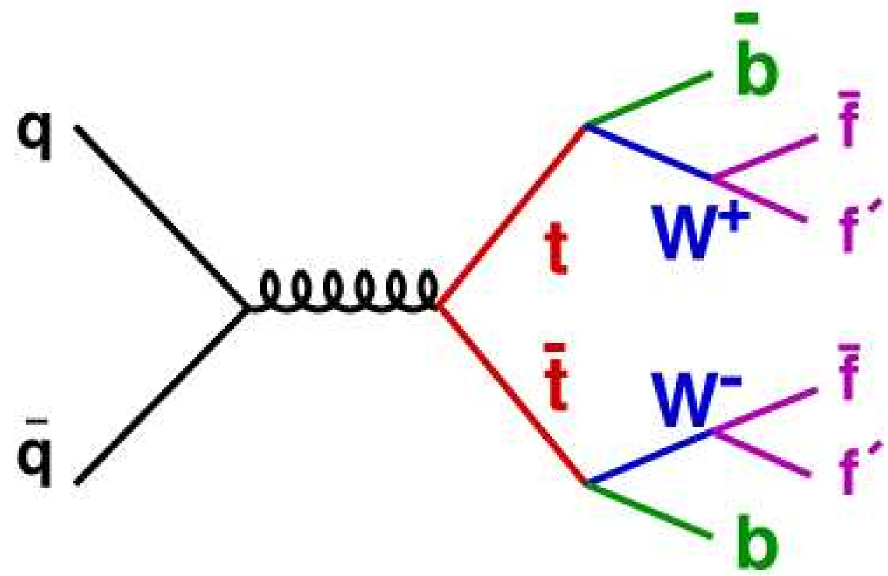
The CDF Detector

- General-purpose particle detector. Cylindrical simmetry
- 3 subsystems: tracking (inside a $1.4\ T$ solenoidal magnetic field), calorimetry and muons systems
- For top physics, the full detector is needed



Jet Probability... Why?

- ... heavy flavor (HF) tagging?
  - ◇ Top signal has 2 b's
  - ◇ $\sim$5% of the main backgrounds has HF
  - $\Rightarrow$ S/B greatly increased



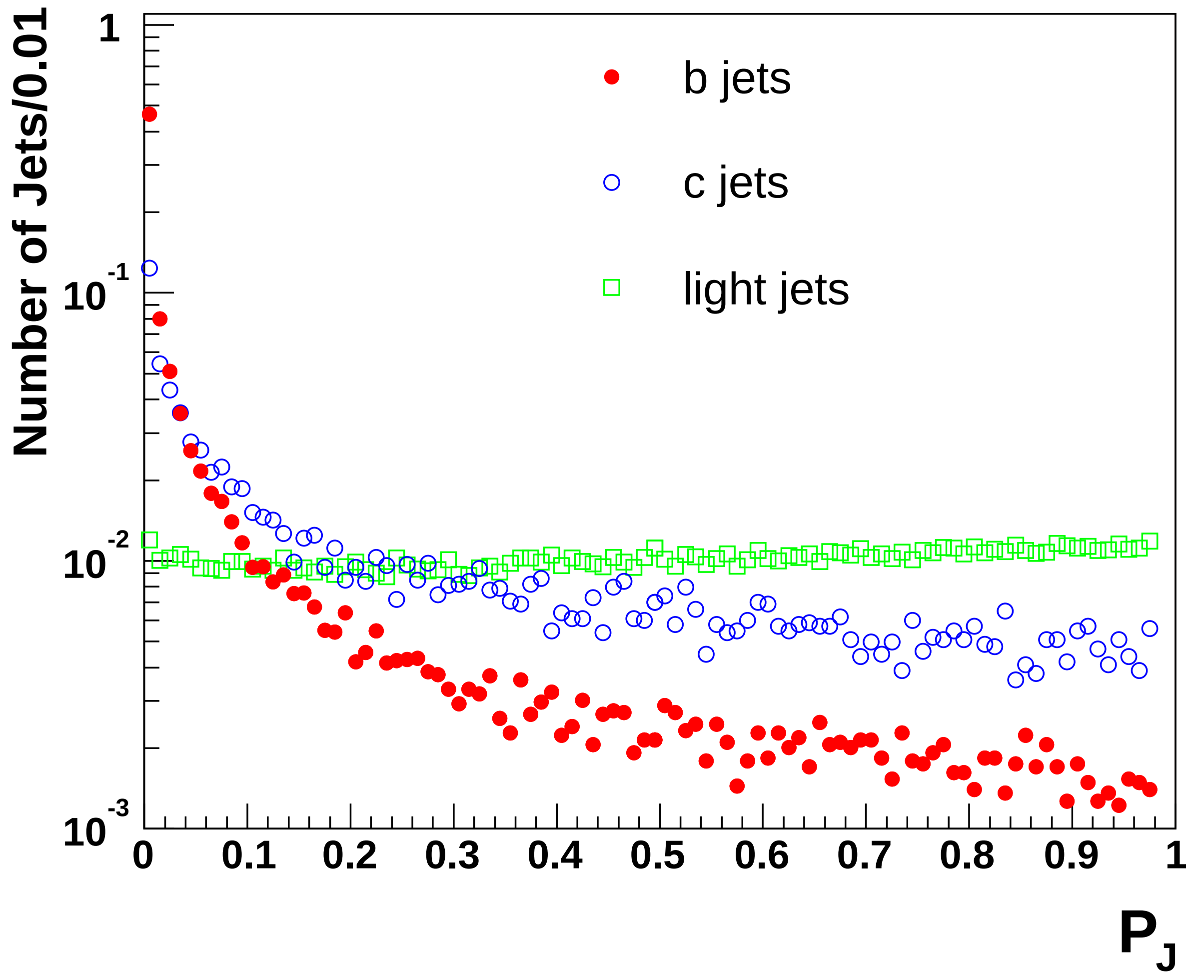
- ... Jet Probability?
  - ◇ Provides a continuous variable $\Rightarrow$ more flexible way to understand the composition of the tagged sample
  - ◇ Can be tuned/optimized differently for other kind of analyses
  - ◇ Potentially, this method can be used to statistically separate b and c heavy flavor contributions

Jet Probability Algorithm (I)

- HF hadrons have long lifetime $\Longrightarrow$ displaced vertices (and tracks) from the primary vertex

- Physically, probability for a jet to come from the primary vertex

- Uniform for light quark or gluon jets. Peaks at 0 for jets containing displaced tracks from HF decays

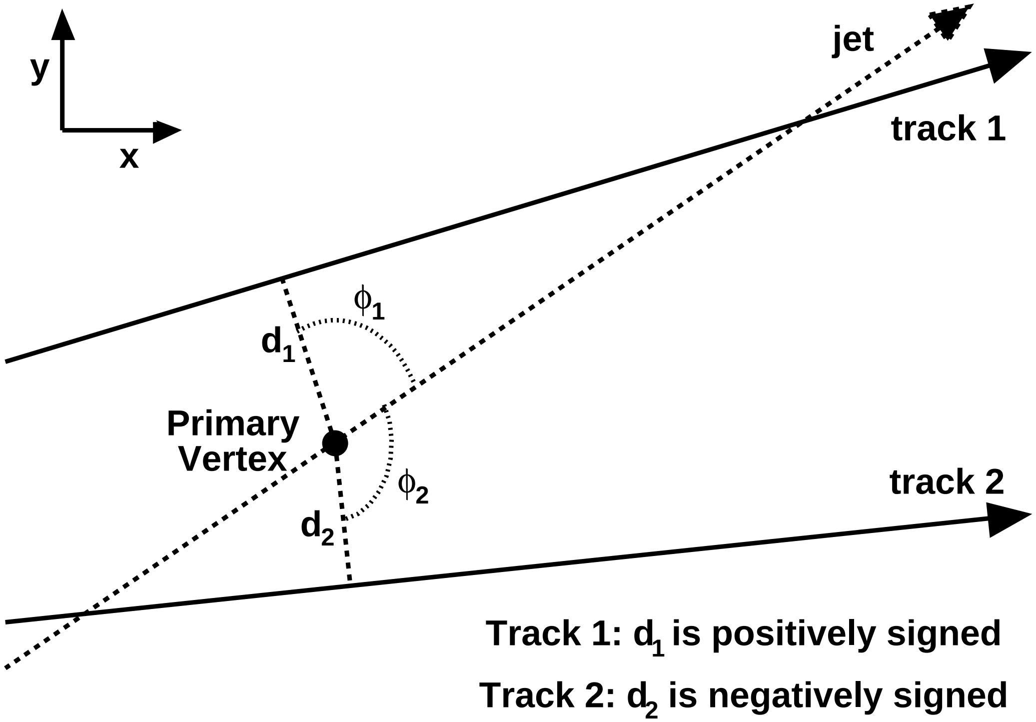


- For the analysis, $P_J < 1\%$ and $P_J < 5\%$

Jet Probability Algorithm (II)

- Signed impact parameter: $d_0 > 0$ if point of closest approach to the primary vertex lies in the same direction as the jet direction ($\cos\phi > 0$)

- + (-) Jet Probability: only tracks with positive (negative) impact parameter
  - ◇ + Jet Probability ⇒ positive tags
  - ◇ - Jet Probability ⇒ mistags



- Track impact parameter significance: $S = D/\sigma_D$

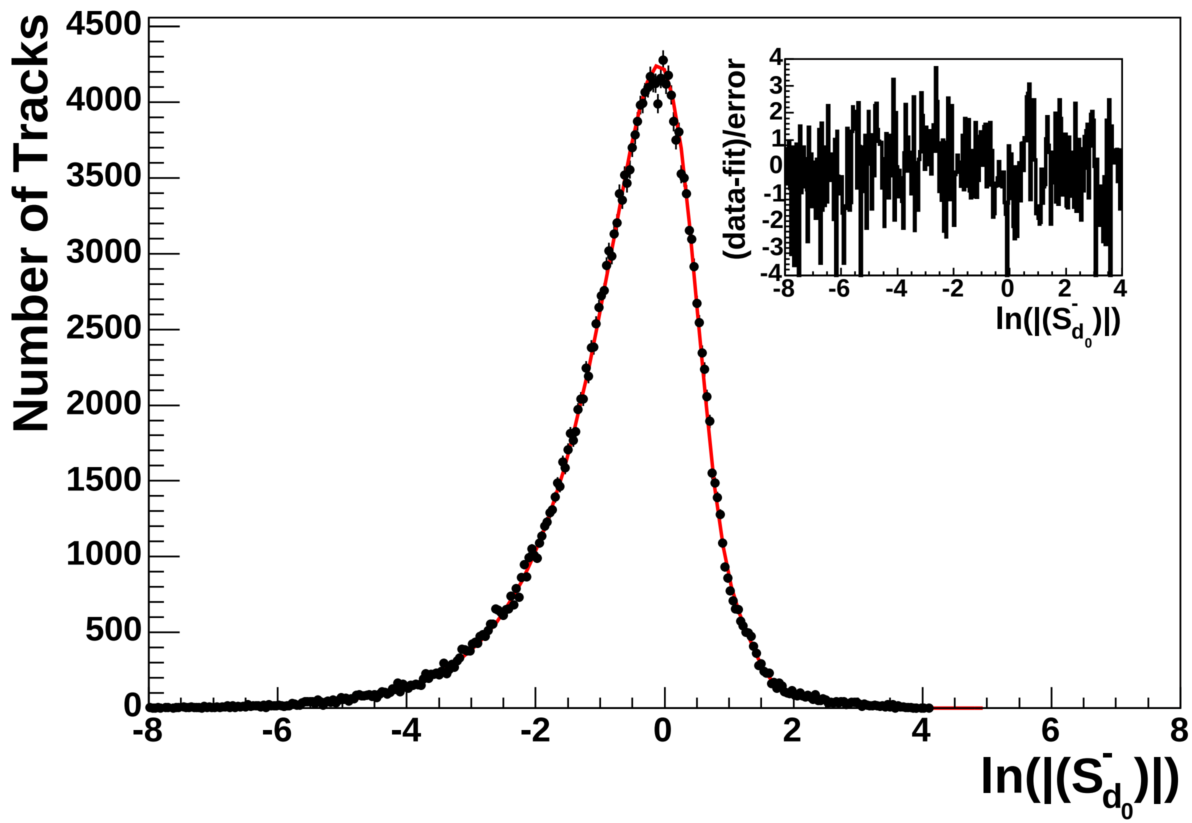
Jet Probability Algorithm (III)

- Fit the distribution of the track impact parameter significance to obtain a resolution function $R(S)$ (different for data and MC)

- Negative side of $R(S)$ used to determine the probability ($P_{tr}(S_0)$) that the impact parameter significance (S_0) of a given track is due to the detector resolution

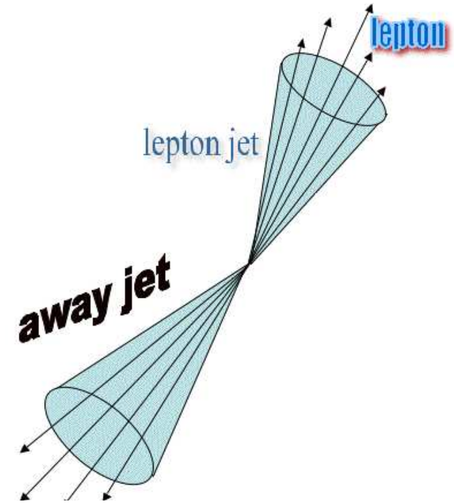
- Probability that a jet is consistent with a zero lifetime hypothesis:

$$\prod_{l=1}^{N_{tr}} P_{tr} \times \sum_{k=0}^{N_{tr}-1} \frac{(-ln \prod_{l=1}^{N_{tr}} P_{tr})^k}{k!}$$



Jet Probability Efficiency

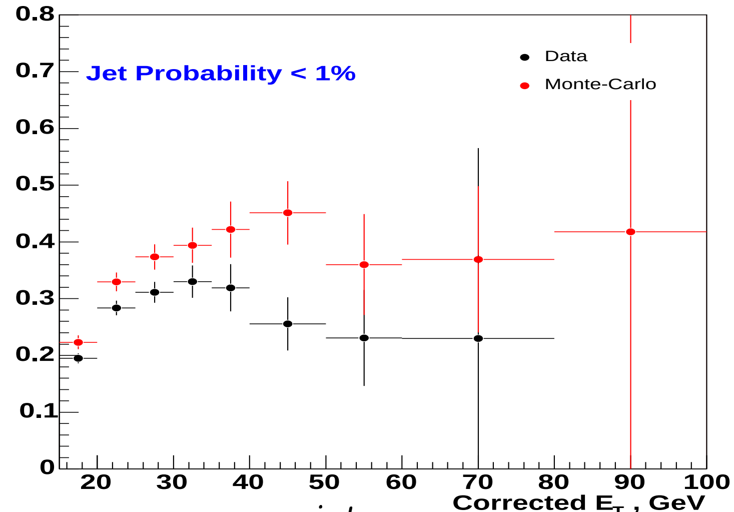
- Measured using an 8 GeV inclusive electron data sample (it is enriched with HF due to the semiloptenic B decays)



- Double tag method: as heavy flavor quarks are mostly produced in pairs, heavy flavor content in one jet is enhanced requiring that the “other” jet (away jet) is tagged

- Efficiencies to tag a HF jet with $E_T > 15$ GeV and 318 pb^{-1}

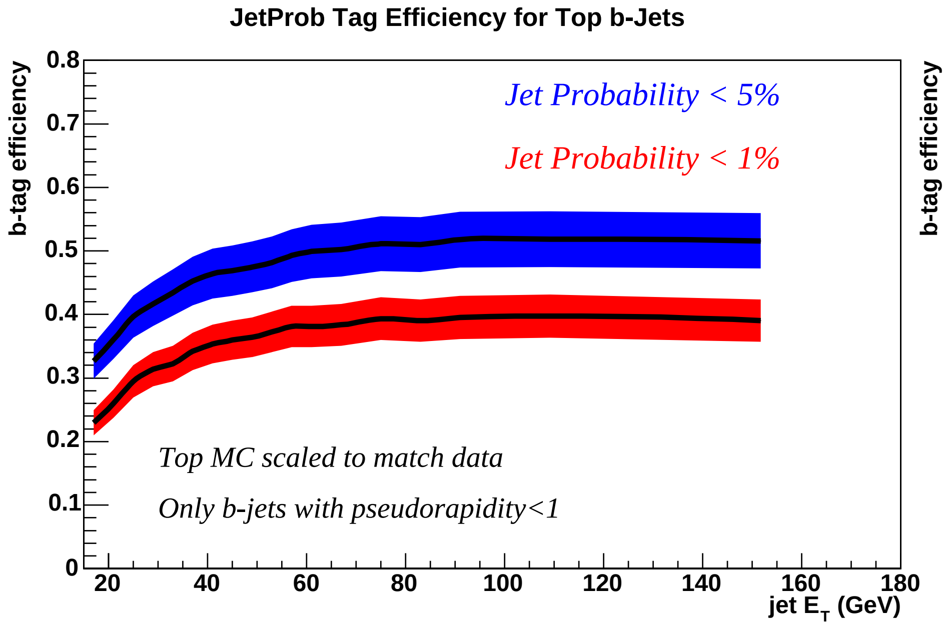
	$P_J < 1\%$	$P_J < 5\%$
ϵ^{data}	0.258 ± 0.018	0.334 ± 0.026
ϵ^{MC}	0.316 ± 0.021	0.392 ± 0.026
Scale Factor (SF)	0.817 ± 0.070	0.852 ± 0.072

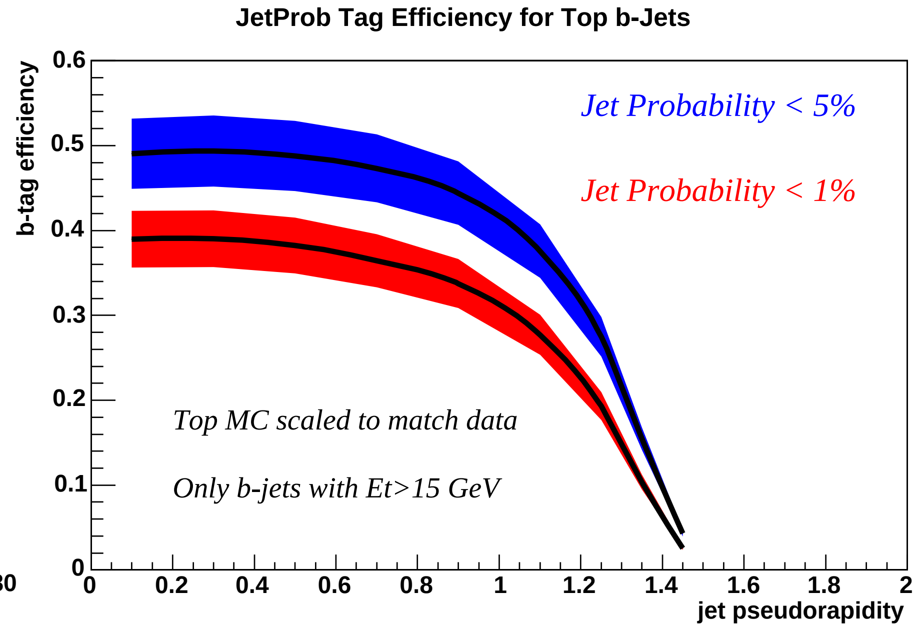


Efficiency vs E_T^{jet} (P_J <1%) in inclusive electron sample

Jet Probability Efficiency in $t\bar{t}$ Events

- b-tagging efficiency (tag rate $\times$ SF) per jet in a top Monte Carlo sample. Bands represent the systematic error due to the scale factor.

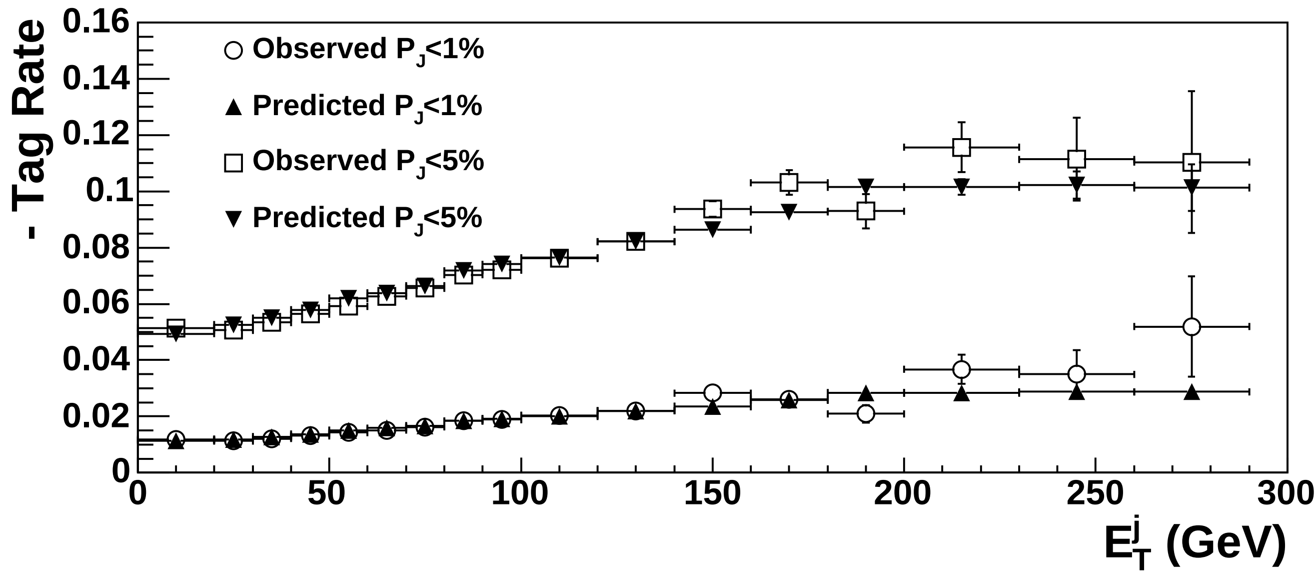




b-tagging efficiency (%)	$P_J < 1\%$	$P_J < 5\%$	Secondary Vertex tagger
per jet	35 ± 3	47 ± 4	40 ± 3
per $t\bar{t}$ event	55 ± 4	69 ± 4	60 ± 3

Jet Probability Mistag Rate

- Mistag rate: probability of tagging a light jet as a heavy flavor one
- Determined using inclusive jet data samples
- Parameterized as a 6 dimensional look-up table (mistag matrix): $E_T, N_{trk}, \sum E_T^j, \eta, Z_{vtx}, \phi$



- Cross check independent samples: observed (multijet trigger) vs prediction (inclusive jet data)
- Results with 318 pb^{-1}

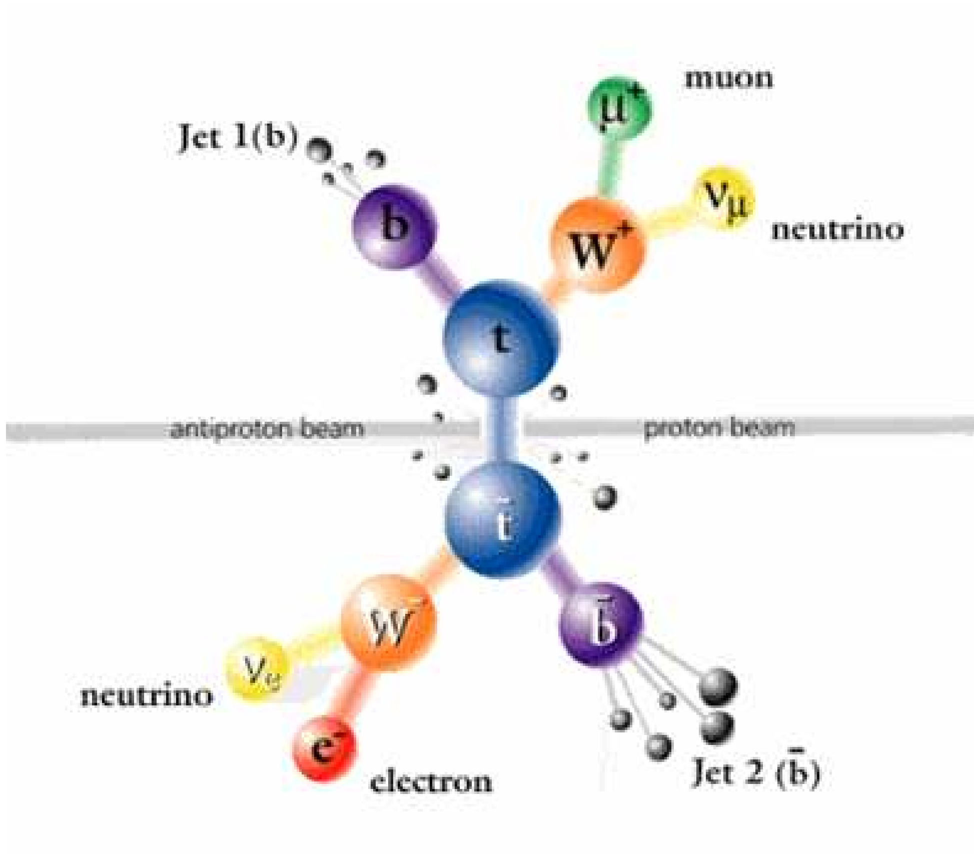
	$P_J < 1\%$	$P_J < 5\%$	Secondary Vertex
Overall - tag rate (%)	1.22 ± 0.08	5.30 ± 0.25	0.48 ± 0.04

$t\bar{t}$ Cross Section Measurement

- Counting experiment: $\sigma_{t\bar{t}} = \frac{N_{obs} - B_{bkg}}{\epsilon_{t\bar{t}} \times \int L dt}$

Why $t\bar{t}$ Production Cross Section?

- Test non-SM top production mechanism
- Look for new physics in the top samples
- Establish the sample for other top properties measurements



- Goal: demonstrate good understanding of backgrounds in control region and observe excess from top in signal region

Data Selection

- Data sample based on Run II data taken untill September 2004

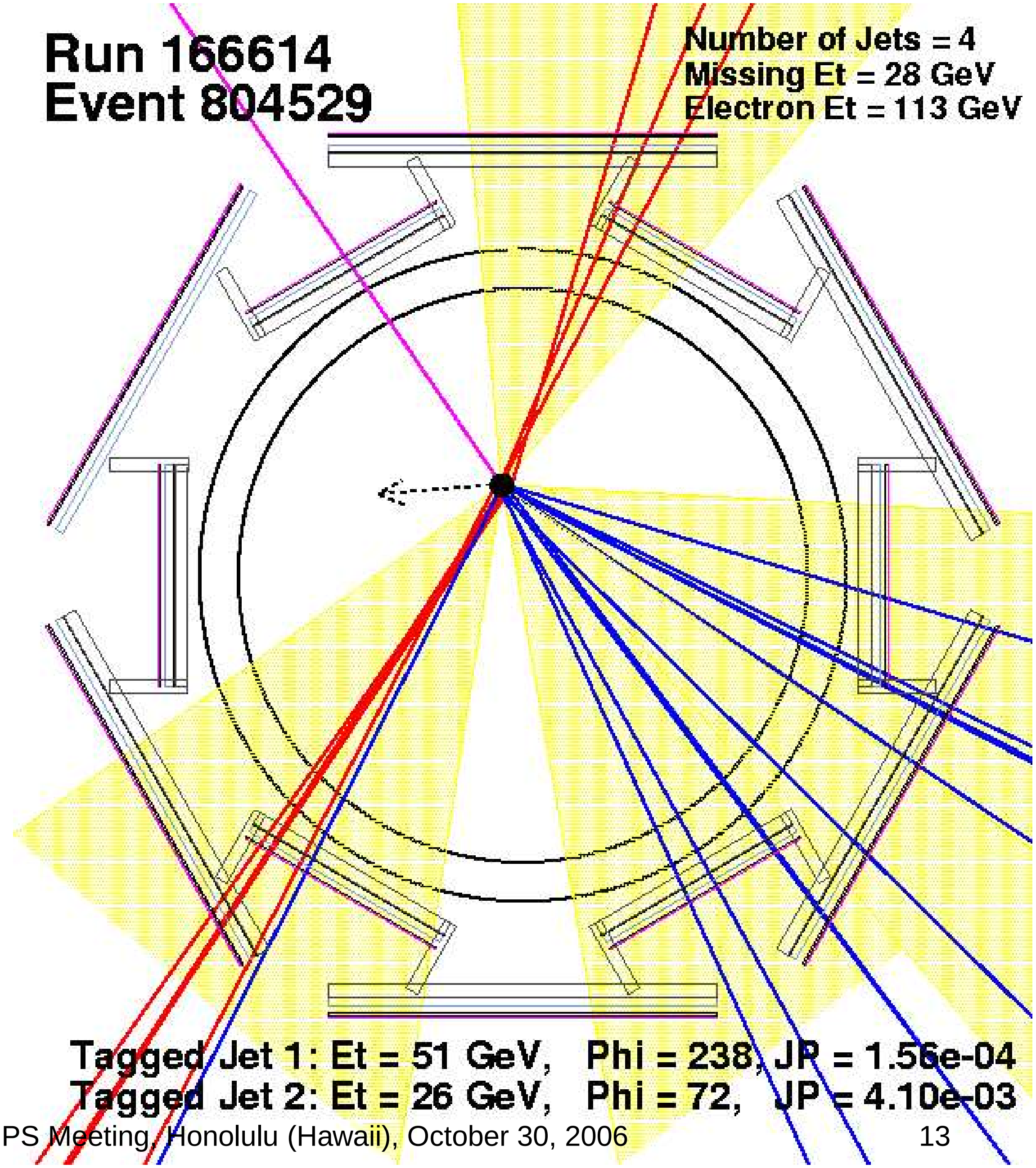
	CEM (Central electrons, $|\eta| < 1$)	CMUP (Central muons, $|\eta| < 0.6$)	CMX (Extension muons, $0.6 < |\eta| < 1$)
Lum (pb^{-1})	318.5 $\pm$ 18.8	318.5 $\pm$ 18.8	305.2 $\pm$ 18.0

- Event selection:
  - ◊ 1 high p_T isolated lepton
  - ◊ high missing transverse energy
  - ◊ $\geq$ 3 energetic jets
  - ◊ vetoes (dilepton, cosmics, conversion, z_{vtx})
  - ◊ $M_T^W > 20\ GeV$ and $H_T > 200\ GeV$
  - ◊ $\geq$ 1 tagged jet (jet with positive P_J <1%)

Jet Multiplicity	1 jet	2 jets	3 jets	$\geq$ 4 jets
Before b-tagging				
# Events	29339	4442	300	166
After b-tagging ($P_J <$ 1%)				
# Events	350	191	52	68

A Top Candidate Event looks like this...

- Jets are represented by yellow hashed cones
- For tagged jets, positive impact parameter tracks are drawn red
- All other (good r-phi) tracks inside jet are drawn blue
- Missing transverse energy direction is the dotted arrow
- Electron track is magenta



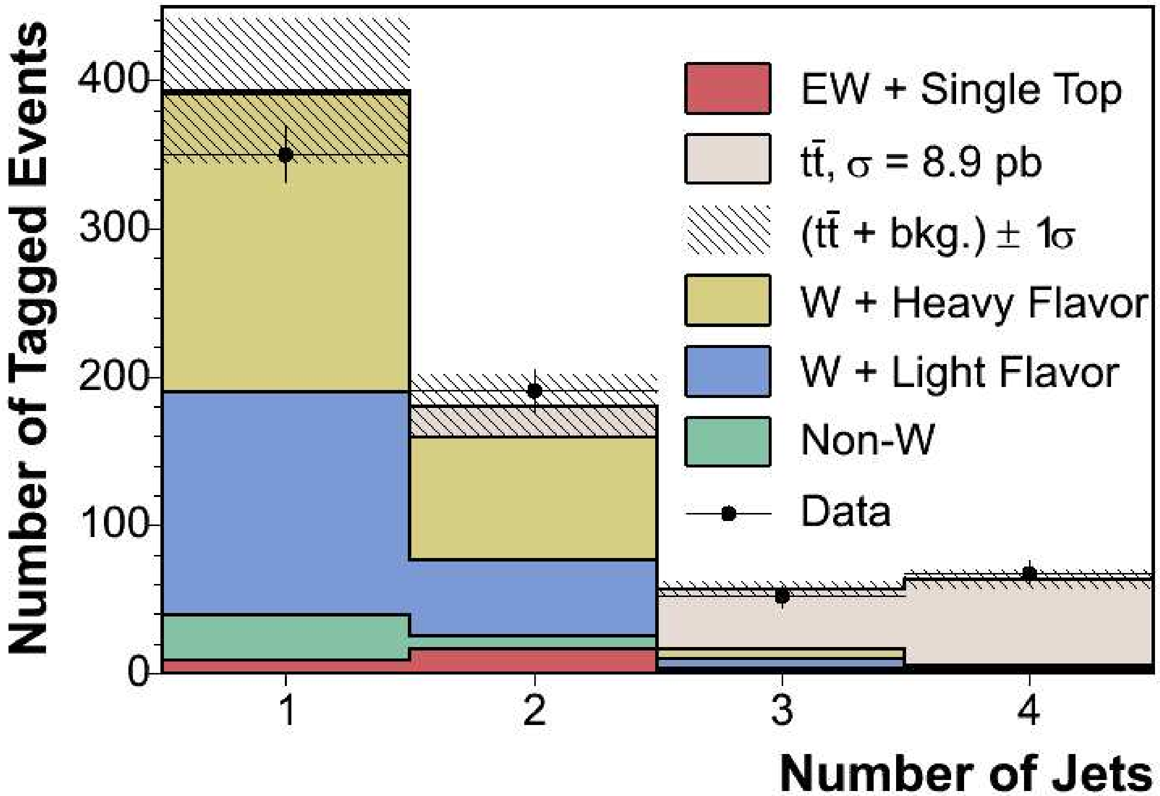
Acceptance and Backgrounds

Quantity	CEM
Single tag, P_J <1% (SF = 0.82 $\pm$ 0.07)	
Acc. No Tag	3.67 $\pm$ 0.02 $\pm$ 0.22
Tag Eff.	54.7 $\pm$ 0.2 $\pm$ 3.6
Acc. with Tag	2.00 $\pm$ 0.01 $\pm$ 0.18
$\epsilon_{t\bar{t}} \int$L dt	6.38 $\pm$ 0.04 $\pm$ 0.68

- Jet Probability tagging efficiencies for $t\bar{t}$ events (PYTHIA Monte Carlo sample with M$_t$ = 178 GeV/c^2)

- Backgrounds estimate:
  - ◊ Mistags (W+light jets): predicted, from data, by the negative tag rate matrix
  - ◊ non-W (QCD production): derived from a control region in data
  - ◊ W+HF: estimated using W+HF MC to
    $\rightarrow$ extract the HF fractions from $\frac{W+HF}{W+Jets}$ MC and the b-tag efficiencies
    $\rightarrow$ normalized to W+jets pretag data
  - ◊ Diboson, Z $\rightarrow \tau\tau$ and single top derived from MC

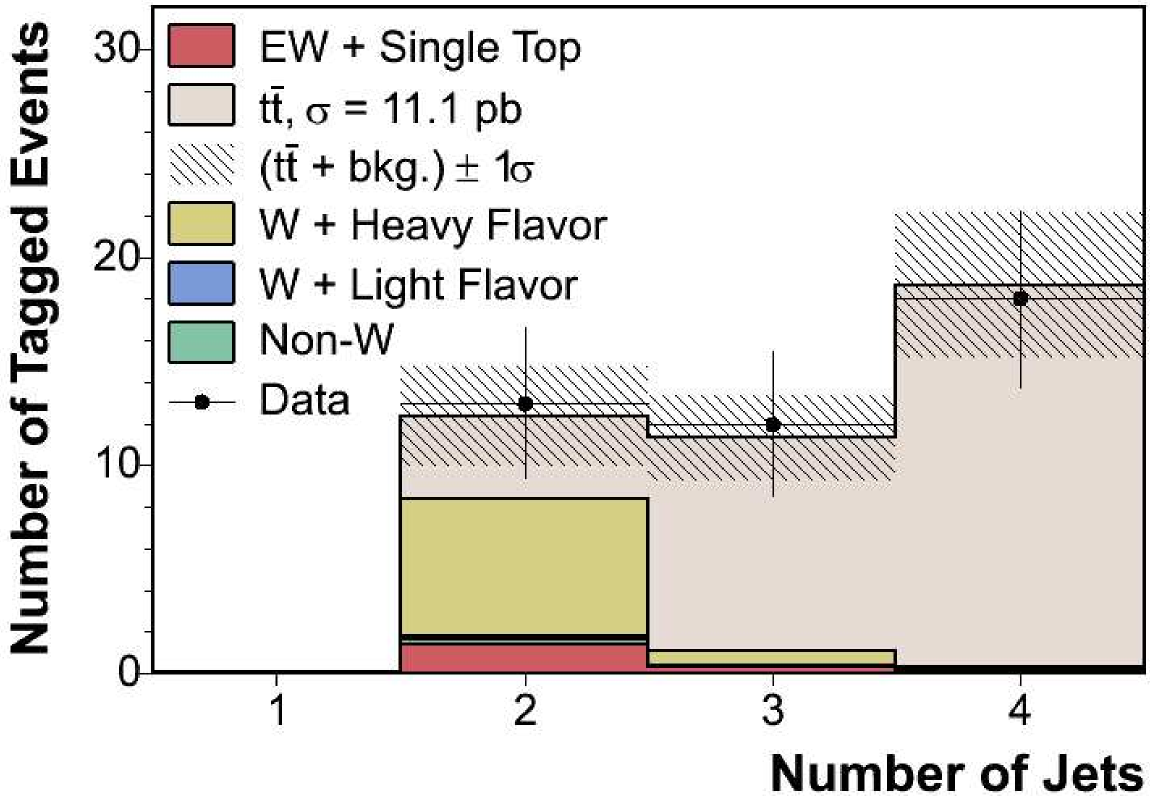
Results for $P_J < 1\%$



$\sigma_{t\bar{t}}$ (pb)	Single Tag
$P_J < 1\%$	8.9 ± 1.0 (stat) $^{+1.1}_{-1.0}$(syst)

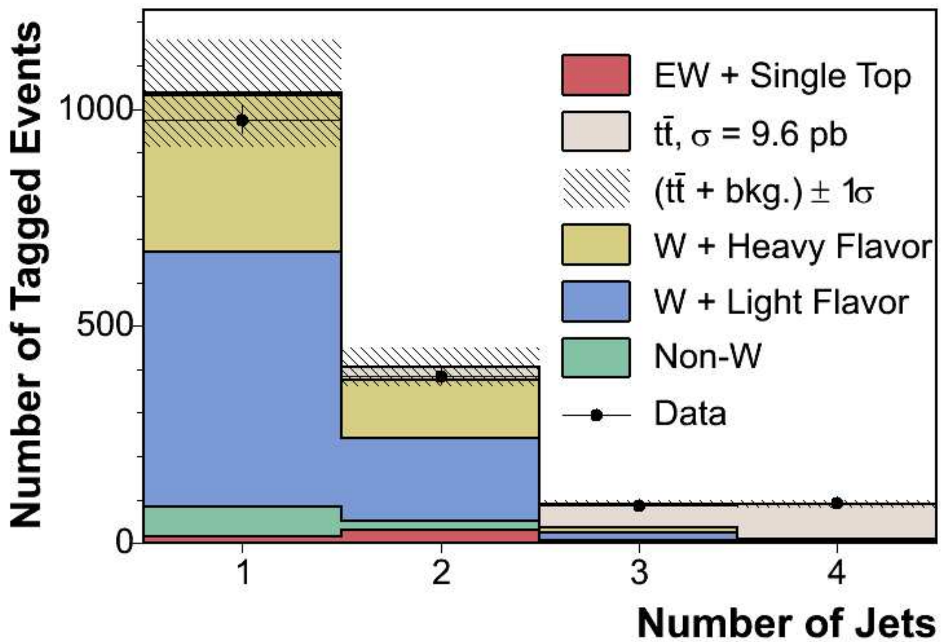
- Largest syst. uncertainty due to the tagging SF ($\epsilon^{data}/\epsilon^{MC}$), $\sim$7%

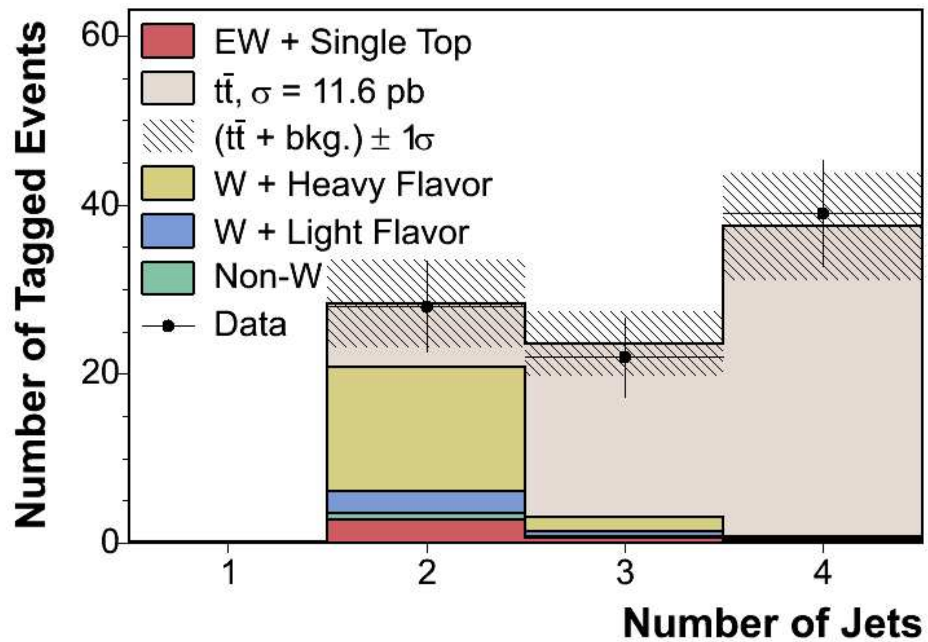
Cross Check (I): Double tag $P_J < 1\%$



$\sigma_{t\bar{t}}$ (pb)	Single Tag	Double Tag
$P_J < 1\%$	8.9 ± 1.0 (stat) $^{+1.1}_{-1.0}$ (syst)	$11.1^{+2.3}_{-1.9}$ (stat) $^{+2.5}_{-1.9}$ (syst)

Cross Check (II): $P_J < 5\%$





$\sigma_{t\bar{t}}$ (pb)	Single Tag	Double Tag
$P_J < 1\%$	8.9 ± 1.0 (stat) $^{+1.1}_{-1.0}$ (syst)	$11.1^{+2.3}_{-1.9}$ (stat) $^{+2.5}_{-1.9}$ (syst)
$P_J < 5\%$	$9.6^{+1.0}_{-0.9}$ (stat) $^{+1.2}_{-1.1}$ (syst)	$11.6^{+1.7}_{-1.5}$ (stat) $^{+2.4}_{-1.8}$ (syst)

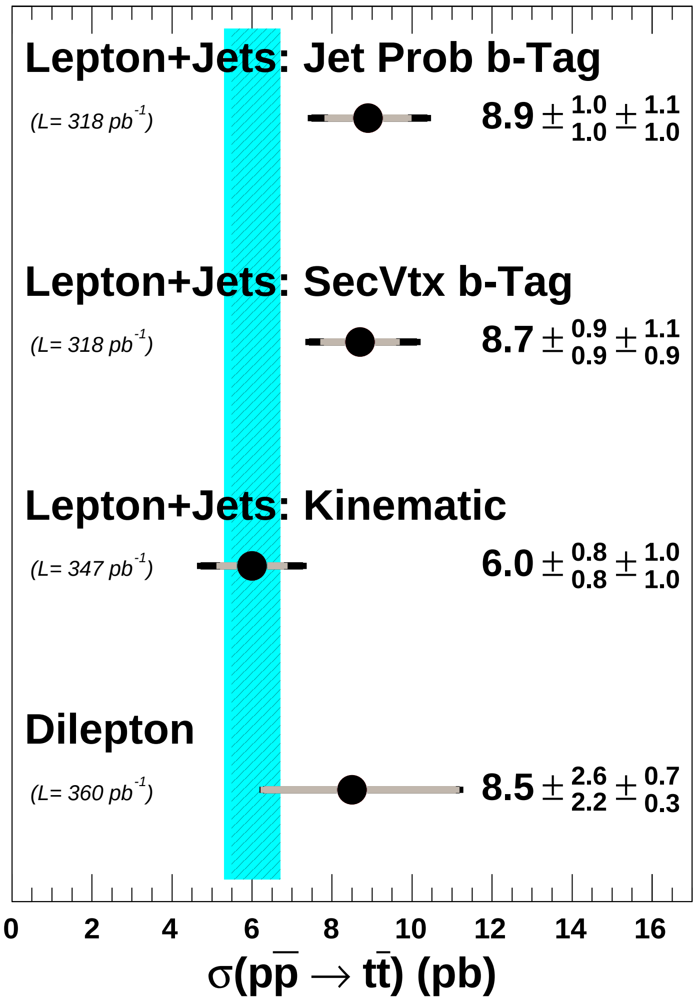
- Prob($\sigma_{meas} > \sigma_{2t} \mid \sigma_{1t}$) = 13.2% (15.6%)

Future Plans

- Repeat the measurement using all data recorded so far
  - $\sim 1.2\ fb$, almost 4 times data used here
  - statistical error will be reduced
- Will try to reduce the systematic uncertainty
  - Measure the scale factor with different methods

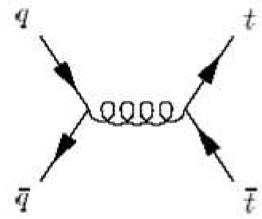
Summary

- We have developed the Jet Probability tagging algorithm for Run II
  - ◇ Based on the track impact parameter information
  - ◇ Continuous variable to discriminate heavy flavor jets
- Characterized the algorithm (efficiency and mistag rate) using data
  - ◇ $54.5 \pm 3.6\%$ efficiency for $t\bar{t}$ events
  - ◇ $1.22 \pm 0.08\%$ mistag rate
- Measured the $t\bar{t}$ production cross section in the Lepton+Jets sample ($M_{Top} = 178\ GeV/c^2$)
  - ◇ $\boldsymbol{\sigma = 8.9 \pm 1.0(stat)^{+1.1}_{-1.0}(syst)}$ pb
- Value consistent with other measurements (and also with the theoretical value)
  - ◇ Total uncertainty of 17%
- Published in Phys. Rev. D. 74, 072006.

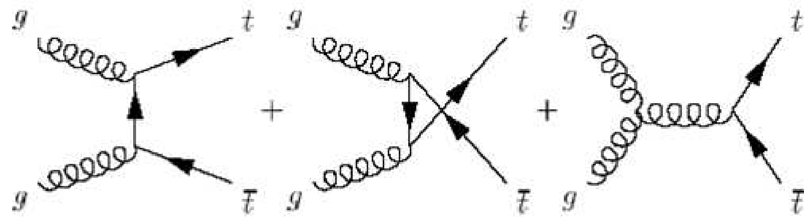


BACK-UP SLIDES

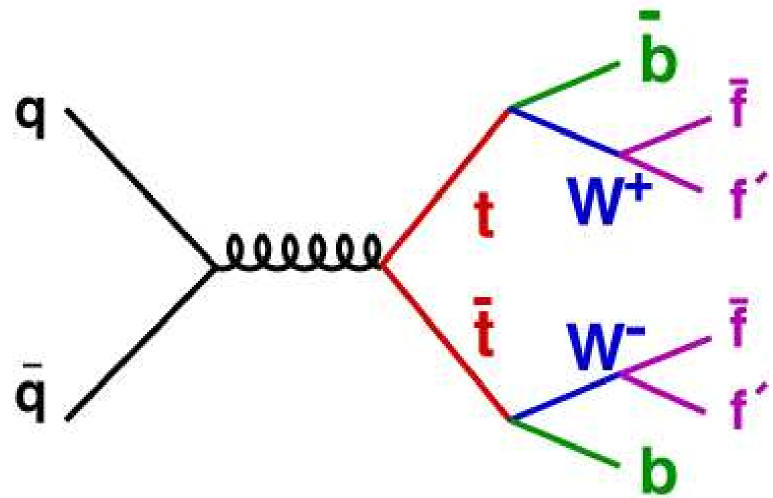
Top Production & Decay Modes



- At Tevatron energies ($\sqrt{s} = 1.96\ TeV$) top quark is mainly produced in pairs via strong interaction
  - ◇ $q\bar{q}$ annihilation (85%) or gluon fusion (15%)
  - ◇ $\sigma(p\bar{p} \rightarrow t\bar{t}\ @\ M_t = 178\ GeV) \approx 6.1\ pb$ $\Rightarrow$ one top event every 10 billion inelastic collisions



- Decays via electroweak interaction $t \rightarrow Wb$
  - ◇ BR($t \rightarrow Wb$)≈1 ⇒ final state given by the $W^{\pm}$ decays
  - ◇ BR($W \rightarrow$ leptons) = 1/3, BR($W \rightarrow$ quarks) = 2/3



lepton ≡ electron or muon

Final State	Dataset	BR	S/B
$l\nu\ l\nu\ bb$	dilepton	∼5%	4/1
$l\nu\ qq\ bb$	lepton+jets	∼30%	2/1
$qq\ qq\ bb$	hadronic	∼44%	1/4

Deduction of the Jet Probability Formula

- If we have a jet with 2 tracks with positive impact parameter which probabilities are P_1 y P_2 and $K \equiv P_1 \cdot P_2$

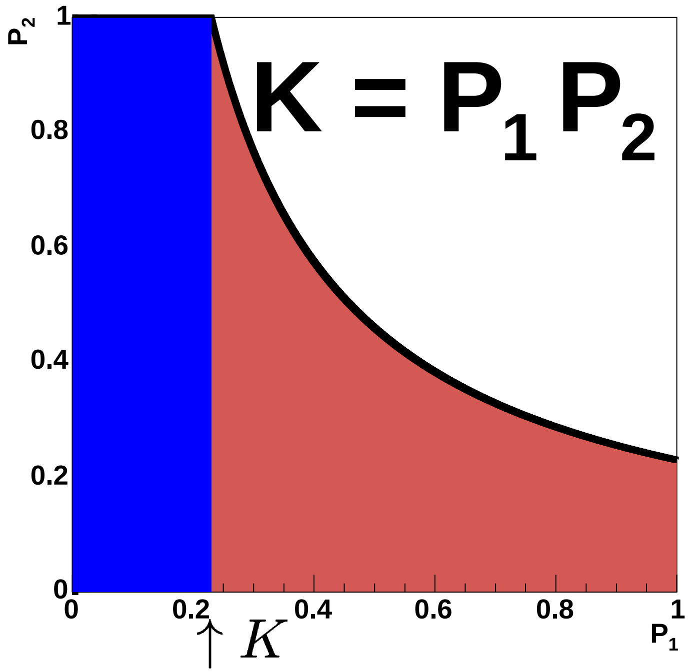
$$0 \leq P_i \leq 1 \text{ i=1,2} \Longrightarrow 0 \leq K \leq 1$$

- The area below and in the left of the curve of constant probability K is the set of combinations, for the 2 tracks, of having a probability less or equal than K. And this area is defined as Jet Probability, P_{jet}

◇ $P_{jet} = A + B, \ A = K \cdot 1$

◇ $B = \int_{x=K}^{x=1} f(x)dx = \int_{x=K}^{x=1} \frac{K}{x}dx = -KLnK$

$\Longrightarrow P_{jet} = K(1 - LnK)$



- En general, se demustra inductivamente que

$$P_{jet} = \prod_{l=1}^{N_{tr}} P_{tr} \times \sum_{r=0}^{N_{tr}-1} \frac{(-ln \prod_{l=1}^{N_{tr}} P_{tr})^r}{r!}$$

Jet Probability Efficiency: Method

- Measured using an 8 GeV inclusive electron data sample and a generic 2→2 Herwig MonteCarlo sample
- Single tag method: $\epsilon = \frac{N_{ej}^{+} - N_{ej}^{-}}{N_{ej}} \cdot \frac{1}{F_B}$
  - ◇ Disadvantage: relays on the correct determination of the heavy flavor fraction in the sample
- Double tag method: sample of events with two jets

$$\epsilon = \frac{(N_{a+}^{e+} - N_{a+}^{e-}) - (N_{a-}^{e+} - N_{a-}^{e-})}{N_{a+} - N_{a-}} \cdot \frac{1}{F_B^a}$$

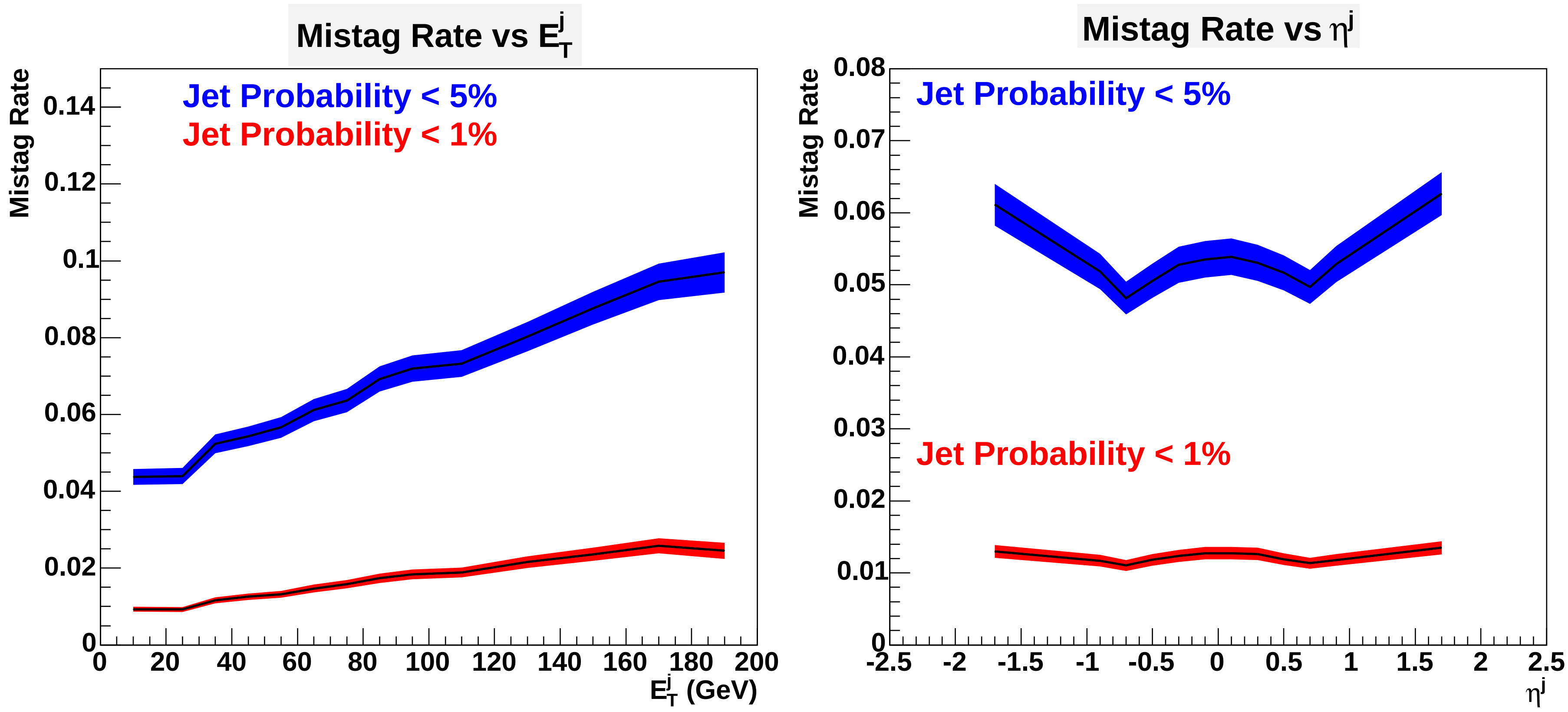
- Calculation of the heavy flavor content in the jet (F_B) has to be corrected for the contribution from charm (determined from MC): $F_B = F_b(1 + \lambda_{c/b})$
  - ◇ F_b from $D^0 \rightarrow K\pi$ decays: $F_b = \frac{N_{D^0}}{N_{ej}} \cdot \frac{1}{\epsilon_{D^0}}$
  - ◇ F_b from cascade muons: select b-hadrons with 2 semiletonic decays (b $\rightarrow$ c $\rightarrow$ X) emitting a pair e-μ with opposite charge:

$$F_b = \frac{1}{\epsilon_\mu} \frac{N_{ej}^{\mu}(OS) - N_{ej}^{\mu}(SS)}{N_{ej}}$$

Tag Rate Matrix Definition

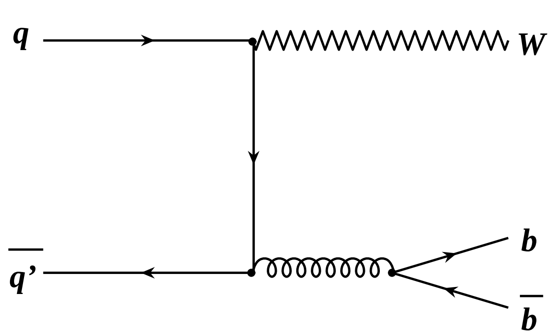
Bin	E_T (GeV)	Trk. Mult.	$\sum E_T^{\text{jets}}$ (GeV)	$\|\eta\|$	$\|Z_{\text{vtx}}\|$ (cm)	ϕ
1	[0,20)	2	[0,80)	[0,1.0)	[0,10)	$[\frac{-\pi}{12}, \frac{\pi}{12})$
2	[20,35)	3	[80,140)	≥ 1.0	[10,20)	$[\frac{\pi}{12}, \frac{3\pi}{12})$
3	[35,50)	4,5	[140,220)		[20,40)	$[\frac{3\pi}{12}, \frac{5\pi}{12})$
4	[50,65)	6,7	≥ 220		[40,50)	$[\frac{5\pi}{12}, \frac{7\pi}{12})$
5	[65,80)	8,9			[50,60)	$[\frac{7\pi}{12}, \frac{9\pi}{12})$
6	[80,100)	10-13			≥ 60	$[\frac{9\pi}{12}, \frac{11\pi}{12})$
7	[100,120)	≥ 14				$[\frac{11\pi}{12}, \frac{13\pi}{12})$
8	[120,150)					$[\frac{13\pi}{12}, \frac{15\pi}{12})$
9	[150,180)					$[\frac{15\pi}{12}, \frac{17\pi}{12})$
10	≥ 180					$[\frac{17\pi}{12}, \frac{19\pi}{12})$
11						$[\frac{19\pi}{12}, \frac{21\pi}{12})$
12						$[\frac{21\pi}{12}, \frac{23\pi}{12})$

Jet Probability Mistag Rate vs E_T and η

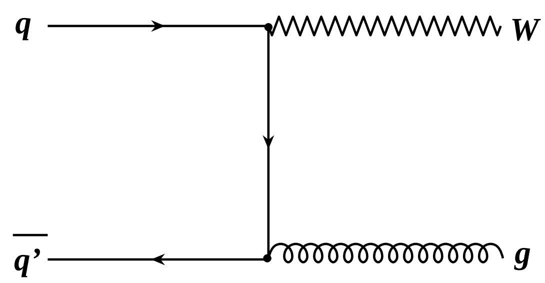


- Bands represent the total uncertainty (statistical and systematic added in quadrature).

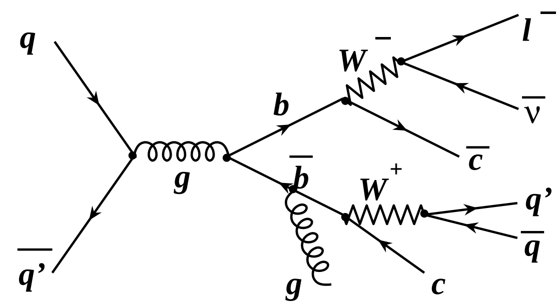
Backgrounds to $t\bar{t}$ lepton plus jets production



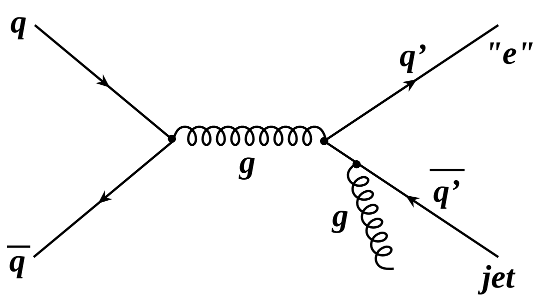
(a)



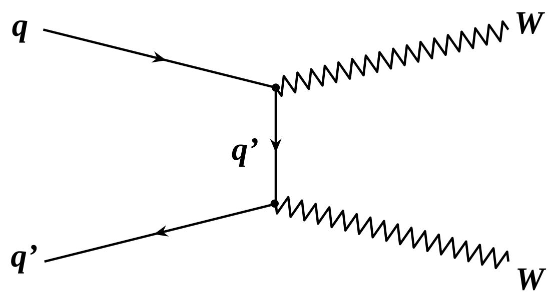
(b)



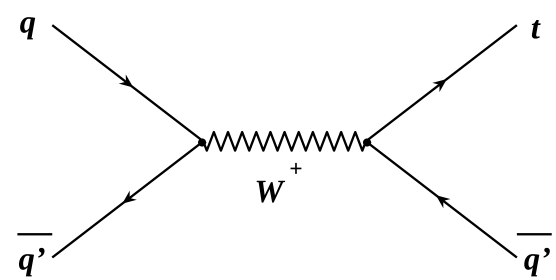
(c)



(d)



(e)



(f)

(a) W + heavy flavor jets, (b) W + light jets, (c) Non-W QCD production from heavy flavor decays, (d) Non-W QCD production from fake leptons, (e) diboson production and (f) single top production, t-channel.

Background Summary, Single tag

Jet Multiplicity	1 jet	2 jets	3 jets	$\geq$ 4 jets
Pretag Data	29339	4442	300	166
		$P_J < 1\%$		
Electroweak	9.3 ± 1.1	16.6 ± 1.8	2.3 ± 0.3	0.71 ± 0.09
$Wb\bar{b}$	83 ± 23	47 ± 13	4.3 ± 1.2	1.1 ± 0.3
$Wc\bar{c}$	31 ± 9	17.3 ± 5.2	1.6 ± 0.5	0.4 ± 0.1
Wc	86 ± 21	19.0 ± 4.9	1.0 ± 0.3	0.21 ± 0.06
Mistag	149 ± 17	51 ± 6	6.1 ± 0.7	2.2 ± 0.3
Non-W	31 ± 16	8.6 ± 4.6	0.9 ± 0.6	0.5 ± 0.5
Total Background	389 ± 49	159 ± 22	16.3 ± 2.0	5.1 ± 0.7
$t\bar{t}$ (8.9 pb)	2.5 ± 0.5	20.6 ± 2.4	40.4 ± 4.5	58.1 ± 6.2
Data	350	191	52	68
		$P_J < 5\%$		
Electroweak	16.3 ± 1.8	28.8 ± 3.0	4.0 ± 0.4	1.4 ± 0.1
$Wb\bar{b}$	111 ± 31	60 ± 17	5.2 ± 1.4	1.1 ± 0.3
$Wc\bar{c}$	68 ± 20	36 ± 11	3.2 ± 1.0	0.76 ± 0.24
Wc	184 ± 45	40 ± 10	2.2 ± 0.6	0.5 ± 0.13
Mistag	585 ± 92	191 ± 30	19.6 ± 3.1	6.1 ± 1.0
Non-W	69 ± 35	21 ± 11	1.3 ± 0.9	0.8 ± 0.7
Total Background	1033 ± 125	377 ± 46	35.5 ± 4.2	10.6 ± 1.4
$t\bar{t}$ (9.6 pb)	3.6 ± 0.6	28.4 ± 3.1	55.1 ± 5.7	78.6 ± 7.8
Data	975	385	87	93

Background Summary, Double tag

Jet Multiplicity	2 jets	3 jets	$\geq$ 4 jets
Pretag Data	4442	300	166
	$P_J <$ 1%		
MC Derived	1.4 $\pm$ 0.3	0.33 $\pm$ 0.06	0.10 $\pm$ 0.02
$Wb\bar{b}$	6.1 $\pm$ 1.9	0.57 $\pm$ 0.19	0.10 $\pm$ 0.03
$Wc\bar{c}$	0.38 $\pm$ 0.17	0.09 $\pm$ 0.04	0.013 $\pm$ 0.008
Wc	0.12 $\pm$ 0.08	0.02 $\pm$ 0.02	0.003 $\pm$ 0.003
Mistag	0.21 $\pm$ 0.05	0.06 $\pm$ 0.01	0.019 $\pm$ 0.004
Non-W	0.19 $\pm$ 0.12	0.03 $\pm$ 0.02	0.05 $\pm$ 0.03
Total Background	8.4 $\pm$ 2.2	1.1 $\pm$ 0.3	0.28 $\pm$ 0.06
$t\bar{t}$ (11.1 pb)	3.9 $\pm$ 0.9	10.2 $\pm$ 2.0	18.4 $\pm$ 3.4
Data	13	12	18
	$P_J <$ 5%		
MC Derived	2.83 $\pm$ 0.51	0.70 $\pm$ 0.12	0.25 $\pm$ 0.05
$Wb\bar{b}$	11.4 $\pm$ 3.6	1.1 $\pm$ 0.3	0.16 $\pm$ 0.05
$Wc\bar{c}$	2.3 $\pm$ 0.9	0.38 $\pm$ 0.15	0.06 $\pm$ 0.03
Wc	0.97 $\pm$ 0.37	0.16 $\pm$ 0.07	0.03 $\pm$ 0.01
Mistag	2.7 $\pm$ 0.8	0.65 $\pm$ 0.20	0.15 $\pm$ 0.05
Non-W	0.63 $\pm$ 0.34	0.09 $\pm$ 0.05	0.14 $\pm$ 0.09
Total Background	20.9 $\pm$ 5.0	3.1 $\pm$ 0.6	0.80 $\pm$ 0.15
$t\bar{t}$ (11.6 pb)	7.5 $\pm$ 1.5	20.5 $\pm$ 3.7	36.6 $\pm$ 6.1
Data	28	22	39