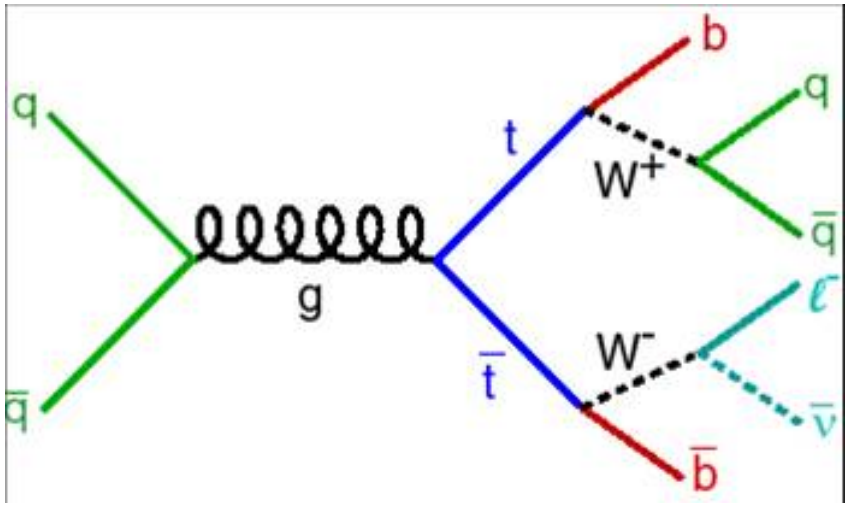
MEASUREMENT OF THE $t\bar{t}$ PRODUCTION CROSS SECTION IN THE LEPTON + JETS CHANNEL USING JET PROBABILITY AT CDF

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

APS Meeting, April 22-25, 2006

Motivation

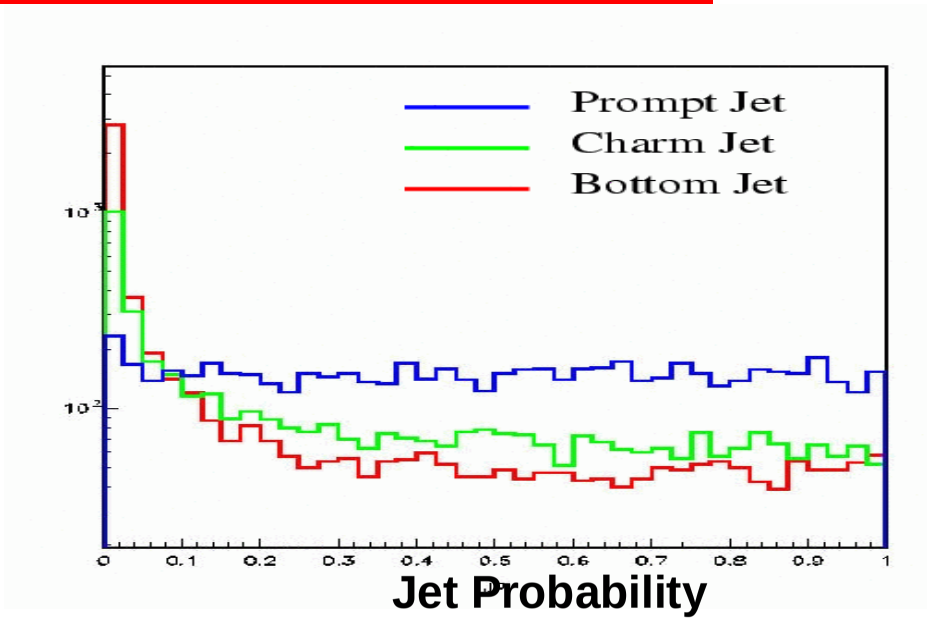
- Why Top?
  - ◇ Heaviest known fundamental particle $\Rightarrow$ special role in EWSB
  - ◇ Might be sensitive to Physics Beyond the Standard Model
- Why Lepton+Jets?
  - ◇ 1 lepton (electron or muon), $\geq$ 3 jets and high $\not\!\!E_T$
  - ◇ S/B = 2/1 (Dilepton 4/1, All hadronic 1/4)
  - ◇ BR = 30% (Dilepton 5%, All hadronic 44%)

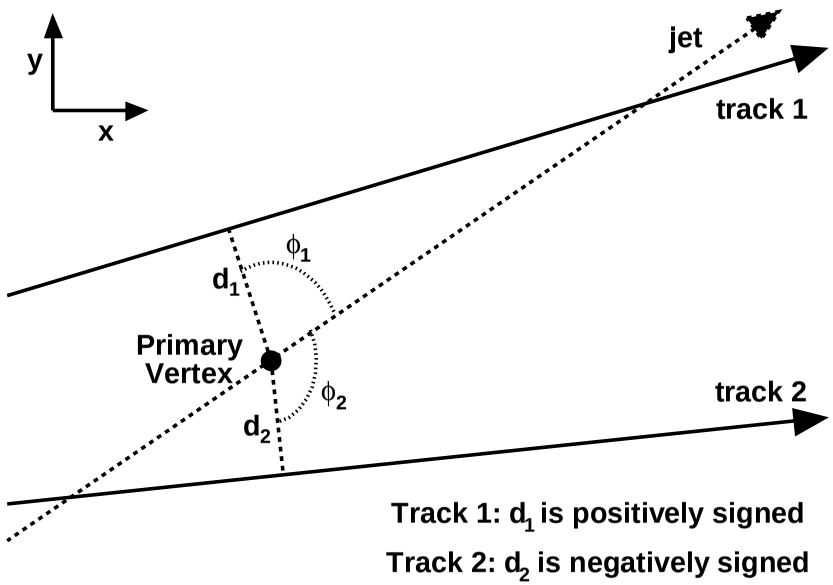


- Why heavy flavor tagging?
  - ◇ Top signal has 2 b's and only $\sim$ 1% of the main backgrounds has HF $\Rightarrow$ S/B greatly increased
- Why Jet Probability?
  - ◇ JP provides (a priori) a more flexible way to understand the composition of the tagged sample by tunning the JP cut
  - ◇ JP can be tuned/optimized differently for other kind of analyses

Jet Probability Algorithm (I)

- Determines whether a jet has been produced at the primary vertex or from the hadronization of a HF quark
- 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
- Signed impact parameter: D>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

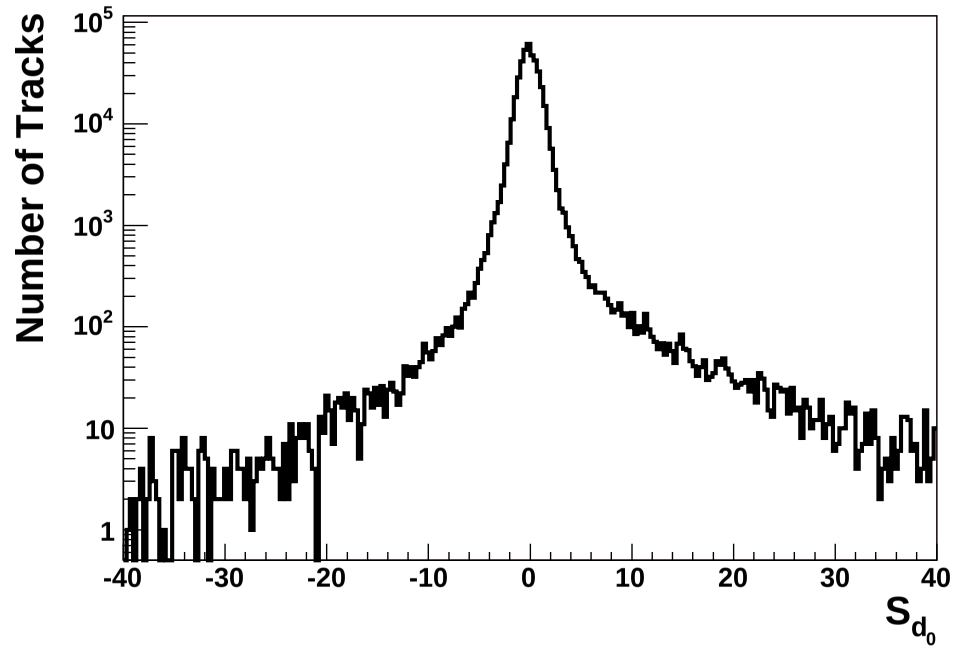




Jet Probability Algorithm (II)

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

$$\boldsymbol{P_{tr}(S_0) = \frac{\int_{-\infty}^{-|S_0|} R(S)dS}{\int_{-\infty}^{0} R(S)dS}}$$



- 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

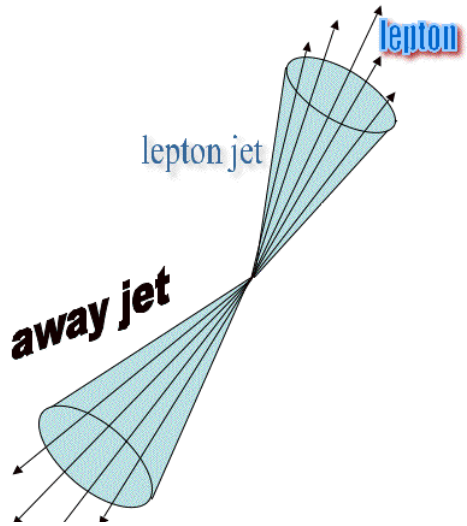
- Measured using an 8 GeV inclusive electron data sample (it is enriched with HF due to the semiloptenic B decays)

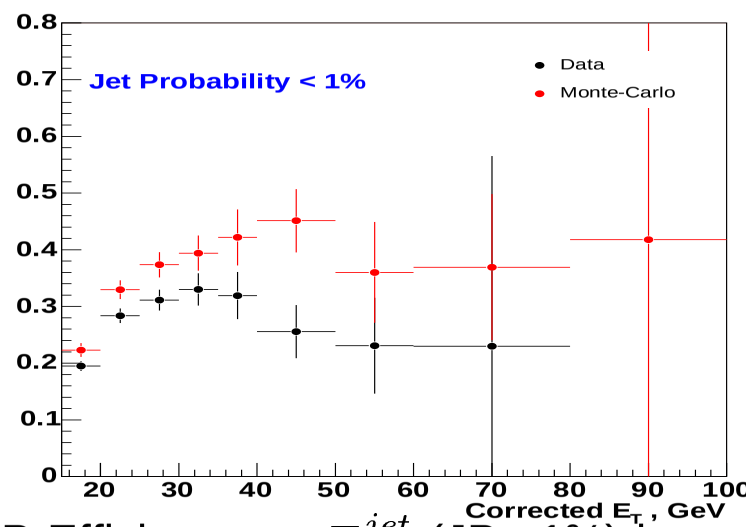
- Single tag method: $\epsilon = \frac{N_{ej}^{+} - N_{ej}^{-}}{N_{ej}} \cdot \frac{1}{F_B}$
  - Disadvantage: relays on the correct determination of the heavy flavor fraction in the sample

- Double tag method: 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}

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

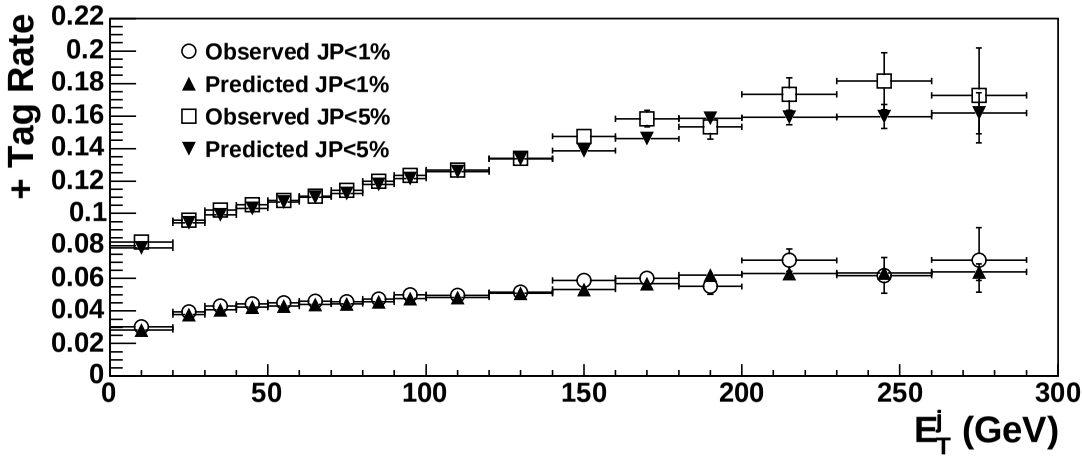


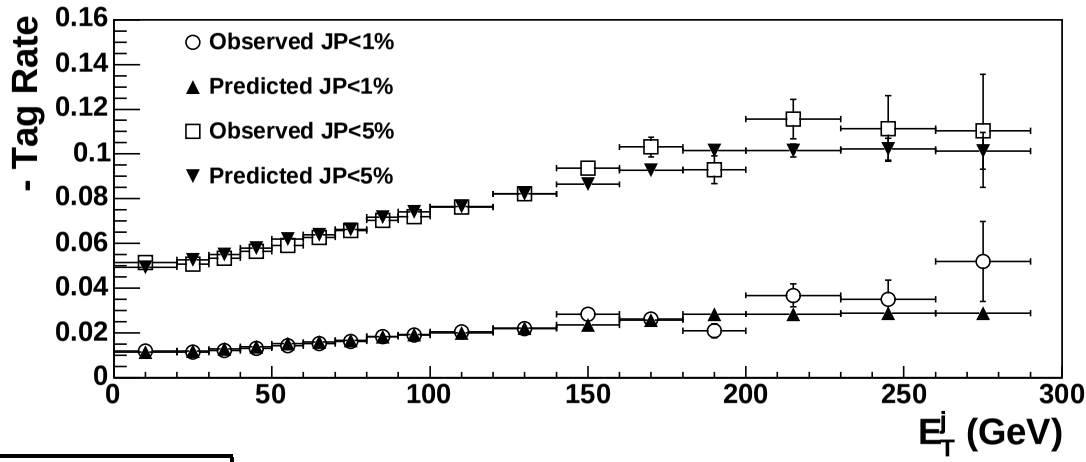


JP Efficiency vs E_T^{jet} (JP$<$1%) in inclusive electron sample

Jet Probability Mistag Rate

- Determined using inclusive jet data samples
- Parameterized as a 6 dimensional matrix: E_T, N_{trk}, $\sum E_T^j$, η, Z_{vtx}, ϕ
- Cross check independent samples: observed (multijet trigger) vs prediction (inclusive jet data)
- Results with 318 pb^{-1}





	JP $<$ 1%	JP $<$ 5%
Overall - tag rate (%)	1.22 $\pm$ 0.08	5.30 $\pm$ 0.25
Overall + tag rate (%)	3.54 $\pm$ 0.18	9.20 $\pm$ 0.26

$t\bar{t}$ Cross Section Measurement

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

Data Selection

- Data sample based on Run II data (with Si) taken untill September 2004

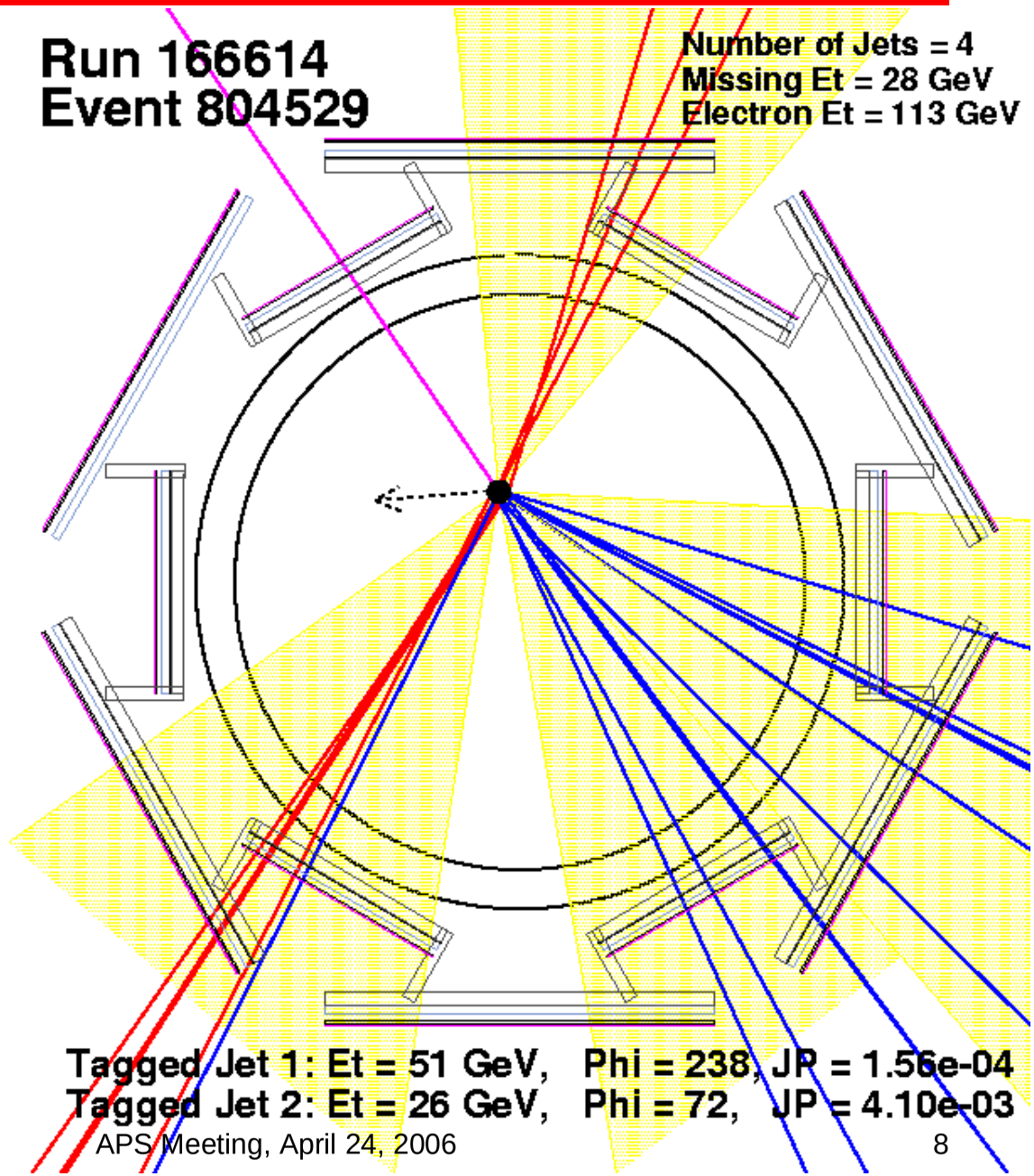
	CEM (Central electrons, $\|\eta\| < 1$)	CMUP (Central muons, $\|\eta\| < 0.6$)	CMX (Extension muons, $0.6 < \|\eta\| < 1$)
Lum (pb^{-1})	318.5 $\pm$ 18.8	318.5 $\pm$ 18.8	305.2 $\pm$ 18.0

- Event selection:
  - ♢ 1 high p_T isolated lepton
  - ♢ high missing transverse energy
  - ♢ $\geq$ 3 energetic jets
  - ♢ $\geq$ 1 (or 2) tagged jet (jet with positive JP$<$1%)

Jet Multiplicity	1 jet	2 jets	3 jets	$\geq$ 4 jets
Before b-tagging				
# Events	29339	4442	300	166
After b-tagging (JP $<$ 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



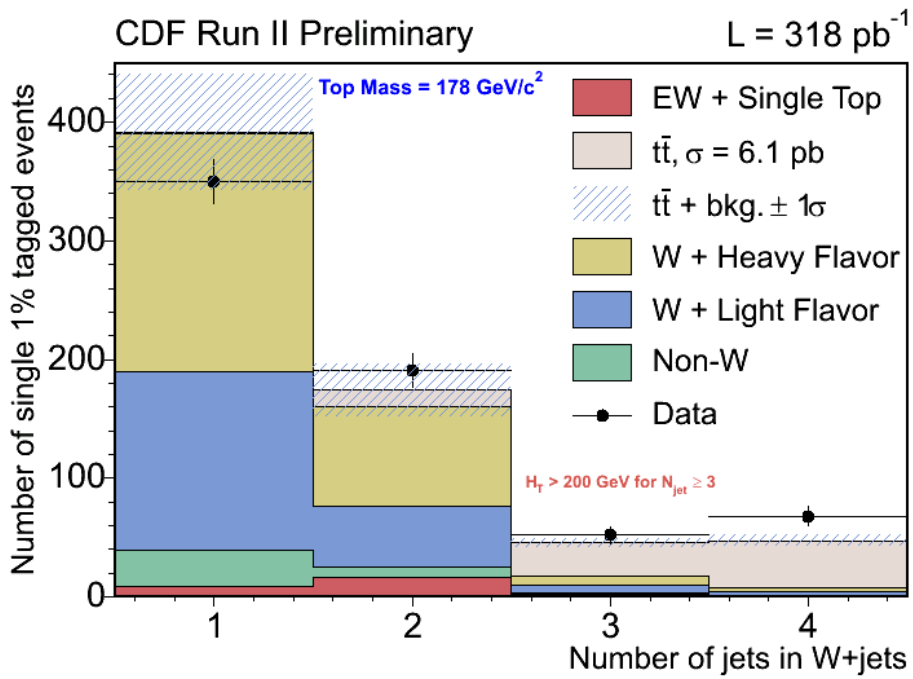
Acceptance and Backgrounds

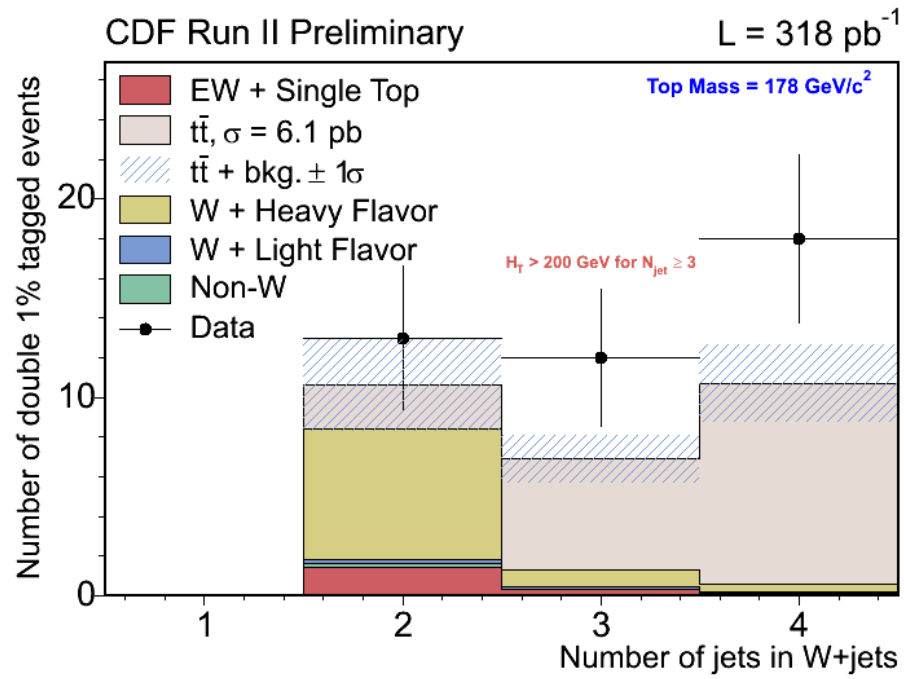
- Jet Probability tagging efficiencies for $t\bar{t}$ events (PYTHIA Monte Carlo sample with M_t = 178 GeV/c^2)

Quantity	CEM
Single tag, JP$<$1% (SF = 0.817 $\pm$ 0.070)	
Acc. No Tag	3.665 $\pm$ 0.017 $\pm$ 0.279
Tag Eff.	54.70 $\pm$ 0.24 $\pm$ 3.59
Acc. with Tag	2.00 $\pm$ 0.01 $\pm$ 0.20
$\epsilon_{t\bar{t}} \int$ L dt	6.38 $\pm$ 0.04 $\pm$ 0.74

- Backgrounds estimate:
  - ♢ Mistags: predicted by the negative tag rate matrix
  - ♢ non-W: derived from the complementary regions of $\not\!\!E_T$ vs lepton isolation
  - ♢ 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 JP $<$ 1%





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

- Largest uncertainty due to the tagging scale factor ($\sim$ 7%)

Conclusions

- We have described the Jet Probability tagging algorithm
  - ♢ Based on the track impact parameter information
  - ♢ Provides a continous 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 (JP $<$ 1%)
- Measured the $t\bar{t}$ production cross section in the Lepton+Jets sample
  - ♢ $\boldsymbol{\sigma = 8.9 \pm 1.0(stat) \pm 1.1(syst)}$ pb ($M_{Top} = 178\ GeV/c^2$)
- Obtained value consistent with other measurements (and also with the theoretical value)
  - ♢ Total uncertainty of 17%

BACK-UP SLIDES

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 $K \equiv P_1 \cdot P_2$

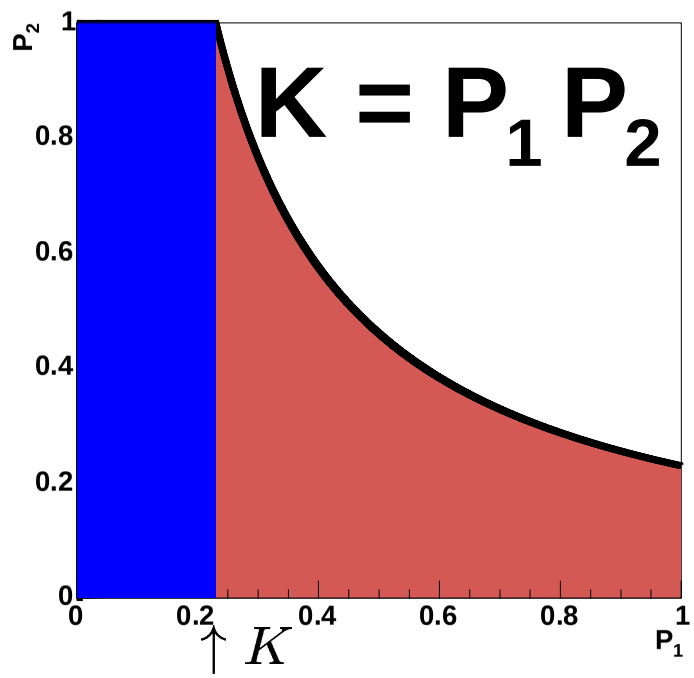
$$0 \leq P_i \leq 1 \text{ i=1,2} \Longrightarrow 0 \leq K \leq 1$$

- 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}

$\diamond$ $P_{jet} = A + B, \; A = K \cdot 1$

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

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



- 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): $F_B = F_b(1 + \lambda_{c/b})$

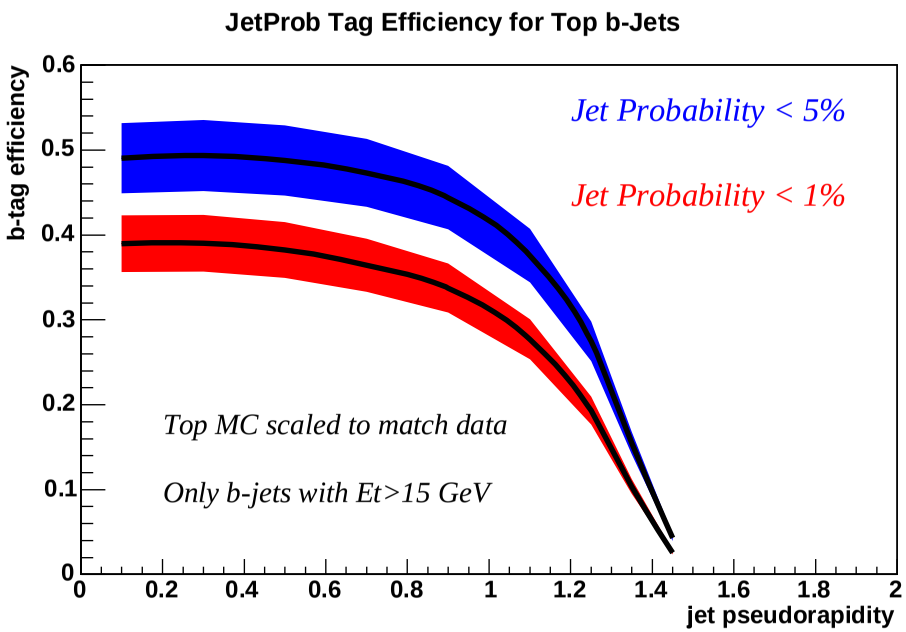
  - F_b from $D^0 \to K\pi$ decays: $F_b = \frac{N_{D^0}}{N_{ej}} \cdot \frac{1}{\epsilon_{D^0}}$

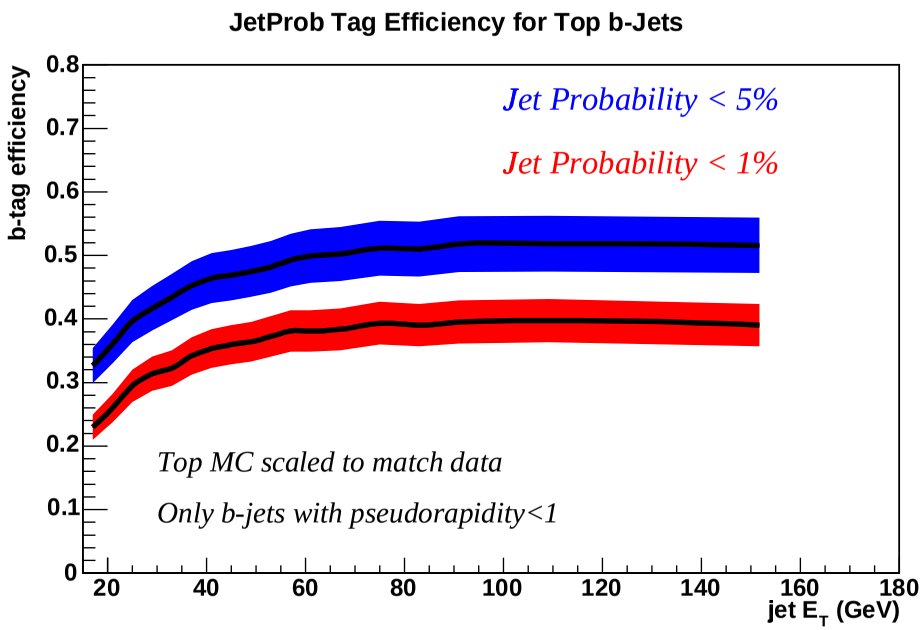
  - F_b from cascade muons: select b-hadrons with 2 semiletonic decays (b → c → X) emitting a pair e-μ with opposite charge:

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

Jet Probability Efficiency

- b-tagging eficiency in a top Monte Carlo sample. Bands represent the systematic error in the scale factor





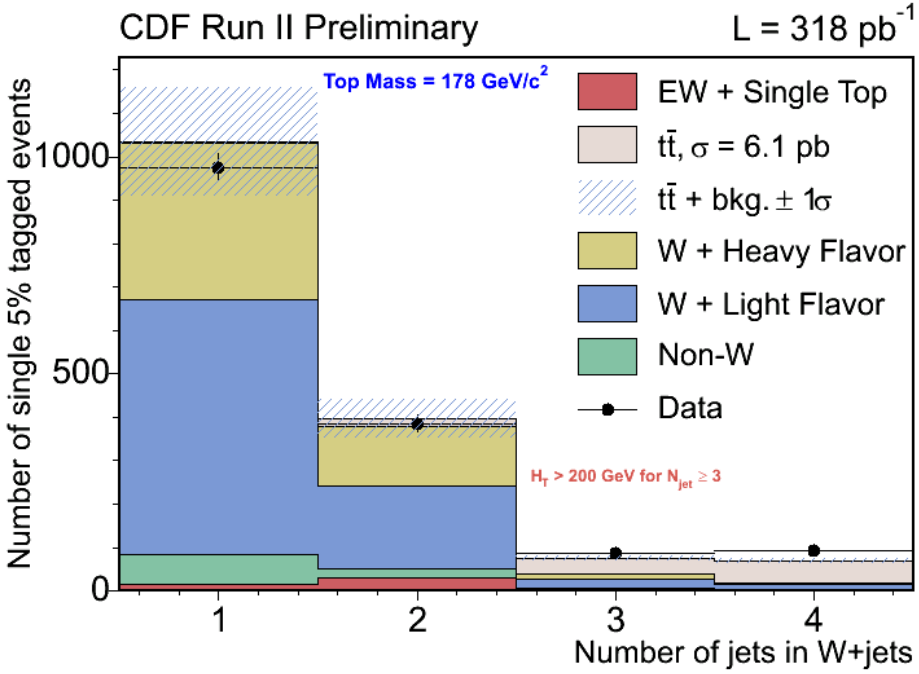
Tag Rate Matrix Definition

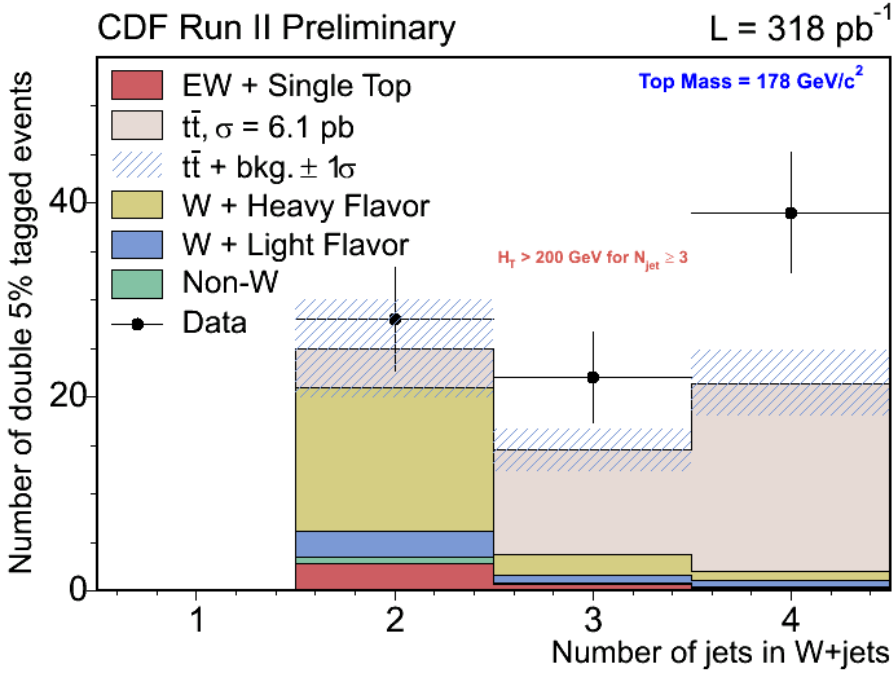
Bin	E_T (GeV)	Trk. Mult.	$\sum E_T^{\text{jets}}$ (GeV)	$\lvert\eta\rvert$	$\lvert Z_{\text{vtx}}\rvert$ (cm)	ϕ
1	[0,20)	2	[0,80)	[0,1.0)	[0,10)	$[\frac{-\pi}{12}, \frac{\pi}{12})$
2	[20,35)	3	[80,140)	≥ 1.0	[10,20)	$[\frac{\pi}{12}, \frac{3\pi}{12})$
3	[35,50)	4,5	[140,220)		[20,40)	$[\frac{3\pi}{12}, \frac{5\pi}{12})$
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						$[\frac{21\pi}{12}, \frac{23\pi}{12})$

Background Summary, JP $<$ 1%

Jet Multiplicity	1 jet	2 jet	3 jet	$\geq$ 4 jets
MC Derived Backgrounds				
WW	2.2 ± 0.3	5.0 ± 0.6	0.7 ± 0.1	0.28 ± 0.06
WZ	1.0 ± 0.1	2.0 ± 0.2	0.23 ± 0.04	0.09 ± 0.02
ZZ	0.027 ± 0.006	0.09 ± 0.01	0.012 ± 0.004	0.007 ± 0.002
Single top $W - g$	4.1 ± 0.5	4.9 ± 0.6	0.7 ± 0.09	0.20 ± 0.03
Single top W^*	1.3 ± 0.2	4.2 ± 0.5	0.60 ± 0.07	0.14 ± 0.02
$Z \to \tau^- \tau^+$	0.74 ± 0.3	0.4 ± 0.2	0.04 ± 0.04	0.0 ± 0.0
Total	9.3 ± 1.1	16.6 ± 1.9	2.3 ± 0.3	0.71 ± 0.09
W + Heavy Flavour				
W$b\bar{b}$	83.0 ± 23.4	47.2 ± 13.2	6.0 ± 1.6	3.3 ± 0.9
W$c\bar{c}$	31.2 ± 9.3	17.5 ± 5.2	2.3 ± 0.7	1.2 ± 0.4
Wc	86.4 ± 21.4	19.2 ± 5.0	1.4 ± 0.4	0.6 ± 0.2
Total	200.6 ± 42.2	83.8 ± 19.8	9.6 ± 2.4	5.2 ± 1.4
Others				
Mistag	149.7 ± 17.9	52.1 ± 6.2	8.6 ± 1.0	6.8 ± 0.8
non W	30.5 ± 15.6	8.6 ± 4.6	0.9 ± 0.6	0.5 ± 0.5
Total Background	390.2 ± 49.1	161.1 ± 22.0	21.5 ± 2.8	13.2 ± 1.7
$t\bar{t}$ (6.1 pb)	1.7 ± 0.4	14.0 ± 1.8	27.5 ± 3.3	39.6 ± 4.6
Data	350	191	52	68

$t\bar{t}$ Cross Section Measurement, JP $<$ 5%





Loose Cut (JP $<$ 5%)	Single Tag	Double Tag
$\sigma_{t\bar{t}}$ (pb)	$9.6^{+1.0}_{-0.9}$ (stat) $\pm$ 1.2 (syst)	$11.6^{+1.7}_{-1.5}$ (stat) $\pm$ 1.9 (syst)