Update on

$t\bar{t}$ CROSS-SECTION WITH JET PROBABILITY

Gervasio Gómez, Enrique Palencia, Teresa Rodrigo,
Luca Scodellaro, Rocío Vilar
Instituto de Física de Cantabria IFCA

Lepton + Jets Meeting
July 16, 2004

Outline

- Introduction: motivation and JP tools
- Data sample and event selection
- $t\bar{t}$ Acceptance
- Backgrounds
  - Mistags
  - QCD
  - W + Heavy flavor (method 2)
  - MC Derived Backgrounds (WW, WZ, ZZ and Single Top)
- Cross section estimation
- Summary
- Future plans

Introduction

- Motivation
  - Cross check $t\bar{t}$ x-section measurement with a different tagging algorithm
  - 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 analysis
- Setup of this analysis
  - We use v.4.11 offline (Data sample $\sim$162 pb^{-1}).
    - JP parametrization (provided by the Florida group, CDF note 6315) is only available in this release

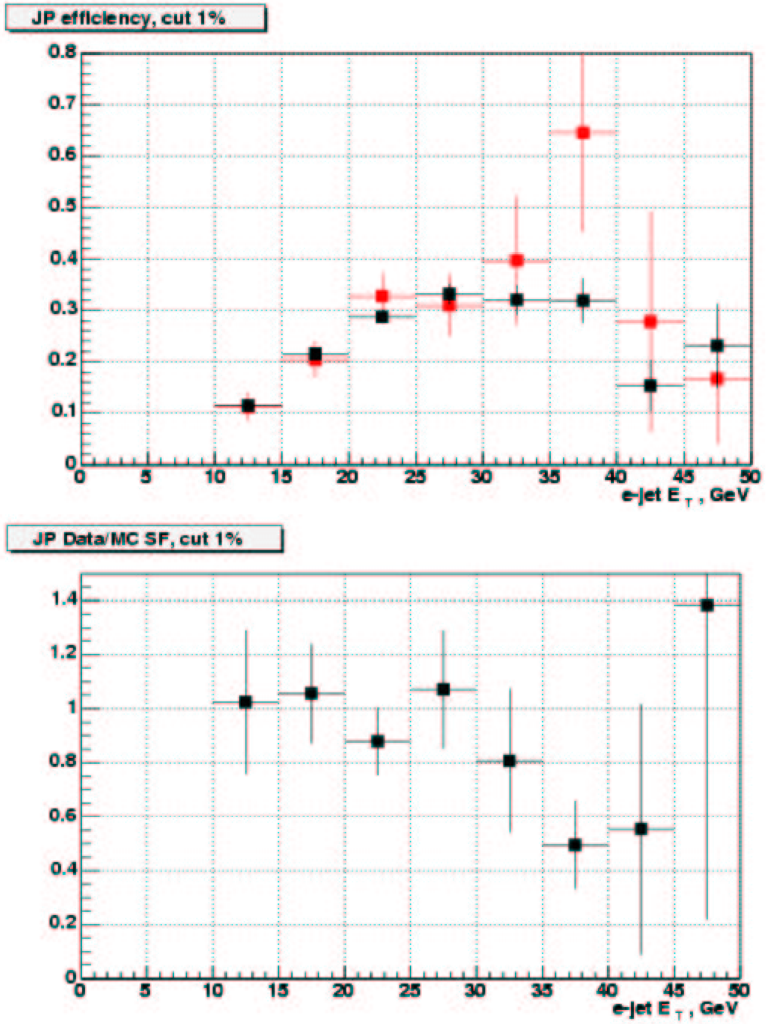
Introduction

⋆ Efficiencies and scale factors are described in CDF note 6931

JP < 0.01

ϵ^{data}=0.197±0.006±0.004±0.01

ϵ^{MC} = 0.25 ± 0.20



Introduction

- The weighted average of the slope of the SF vs E_T^{jet} is compatible with zero

Sample	SF Slope vs E_T^{jet}
Incl. elec.	-0.0173 $\pm$ 0.0076
Jet 50	0.00057 $\pm$ 0.00085
Incl.jets	0.00403 $\pm$ 0.00193
Weighted average	0.00094 $\pm$ 0.00077 GeV

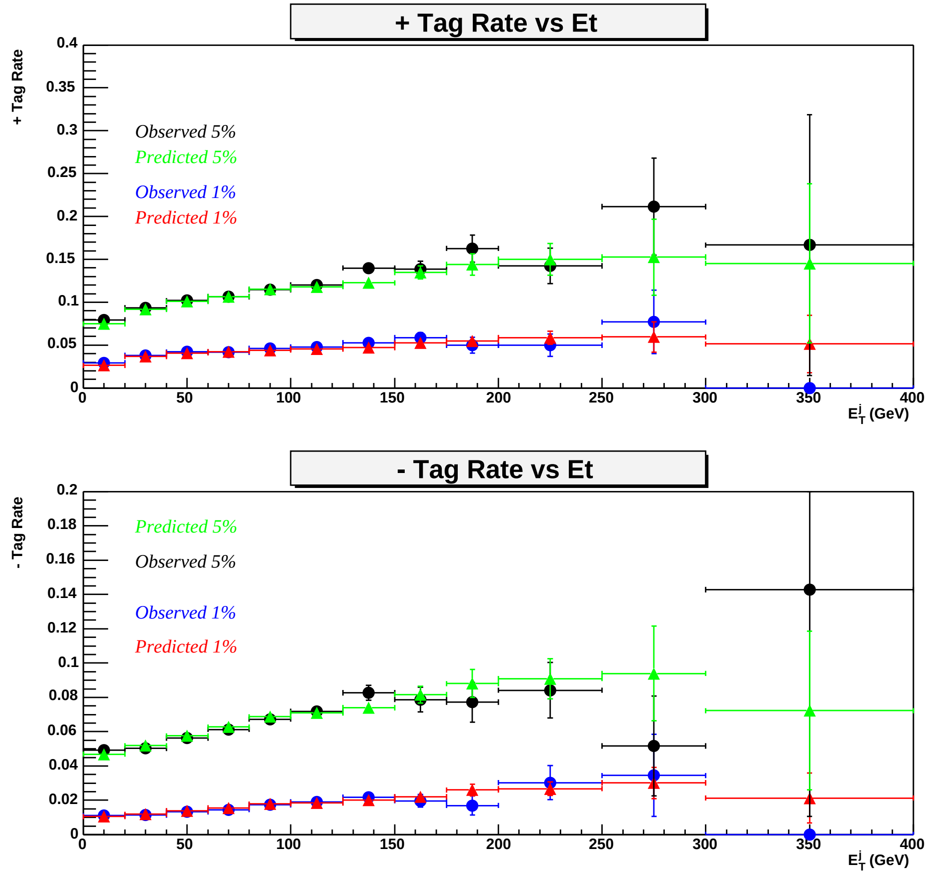
so we consider the SF independent of E_T^{jet} and we use the value

SF = 0.787 $\pm$ 0.105

- Errors include statistical uncertainty, error from F_B measurement and uncertainty due to mistag asymmetry

Introduction

- Mistag Matrix is described in CDF note 6913. The tag rate is parametrized as a 5 $\times$ 1 dimensional matrix of the following variables: (E_T, N_{trk}, $\sum E_T^j$, η, Z_{vtx}) $\times$ ϕ



$\sum E_T$ sample vs prediction (from the inclusive jet data)

Introduction

- From the inclusive jet sample:
  Overall negative tag rate = 1.11% $\pm$ 0.06%
  Overall positive tag rate = 3.22% $\pm$ 0.21%

- This study does not include systematic uncertainties neither from HF contribution to the negative tags nor material interaction distribution to the positive tags.

- Since mistag rate for positively tagged jets is slighty higher than the negative mistag rate, instead of scaling up the negative tag rate we asign an extra 20% error on the mistag which has no basis for JP. It is just taken from SECVTX studies on tagging composition (we plan to do the tagging composition study using gen5)

Introduction

◇ In this first analysis pass, we define as tagged jets those with positive JP$<$0.01 (no attempt has been done to make an optimization cut). We choose this cut to cross-check ourselves with SECVTX since with this cut the performance of JP (efficiency and purity) is similar to SECVTX algorithm

◇ We have remade the official Top ntuples in order to include the JP information for the samples of lepton+jets data, $t\bar{t}$ MC, W+HF and W+jets backgrounds, dibosons and single top ($Z\tau\tau$ not yet ready)

- All the results are (or will be soon) in www-cdf.fnal.gov/internal/physics/top/RunIIWjets/webpages/jetprob/cdfonly.html

Data sample (CDF note 6844)

- The data sample is based on all data taken untill September'03 shutdown (run range 141544 - 167715)

- We use DQM goodrun list version 4.0

- Runs excluded: 164844 (luminosity not well determined), 163463 and 163474 (bad beamline)

- Include runs with good SVX information

- Luminosity (pb^{-1})

CEM	CMUP	CMX
161.6	161.6	149.8

Event selection (CDF note 6844)

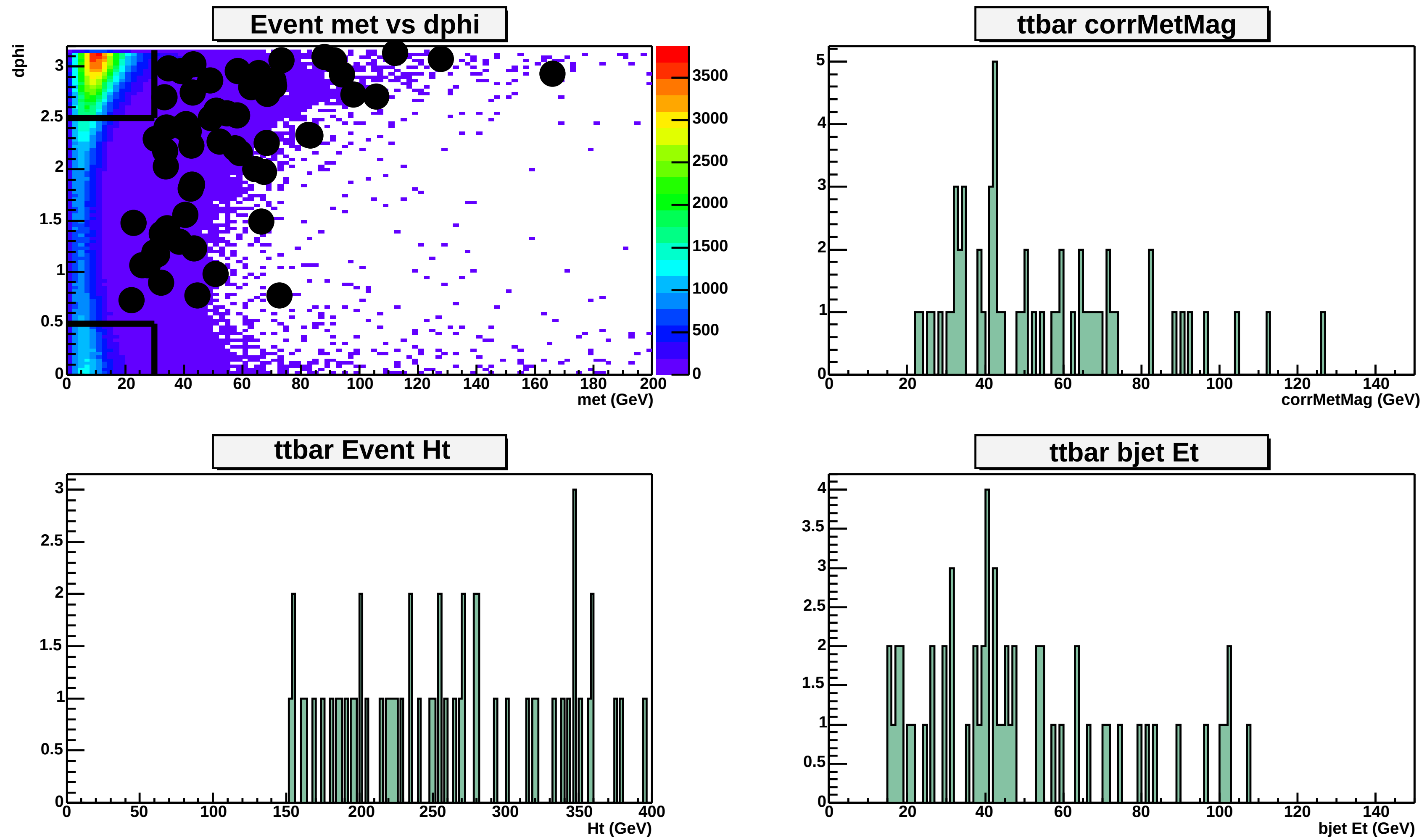
- Our selection is based on the standard Lepton+jets selection
- We apply the QCD veto but not the H$_T$ cut
- Number of selected pretag and tagged events vs jet multiplicity

Jet Multiplicity	1 jet	2 jet	3 jet	$\geq$ 4 jets
Pretag events				
CEM	7819	1202	201	61
CMUP	3758	587	81	27
CMX	1971	293	36	6
Total	13548	2082	318	94
Tagged events				
CEM	78	40	21	17
CMUP	40	30	8	10
CMX	13	11	2	1
Total	131	81	31	28

- Event is tagged if it contains, at least, one jet with positive JP$<$0.01

Event selection (CDF note 6844)

- Some top candidates plots:



Event selection (CDF note 6844)

- Candidates events and % overlap between JP, SECVTX and Loose SECVTX

  - Single tag:

    - JP selects 59 events, SECVTX (*) 47 and Loose SECVTX (**) 91

      - Overlap JP/SECVTX = 35/47 = 74.5%

      - Overlap Loose SECVTX/JP = 47/59 = 79.7%

  (*) For this comparison, we have remade the SECVTX selection including QCD cut
  (**) Loose SECVTX mades a silghtly different kinematical selection (z vertex) and no QCD cut is applied.

Event selection (CDF note 6844)

- Double tag:
  - JP selects 11 events (18.6% of the single tag), SECVTX 8 (17%) and Loose SECVTX 19 (20.6%)
    - Overlap JP/SECVTX = 6/8 = 75%
    - Overlap Loose SECVTX/JP = 10/11 = 91%
    - From the 19 Loose SECVTX double tagged events, 10 are also double tagged by JP and 8 are single tagged by JP

$t\bar{t}$ Acceptance

- From our $t\bar{t}$ ntuples we have reproduced the acceptance calculations of the kinematical selection as in CDF note 6844. Thus, in what follows, we will use the same kinematic efficiencies (and errors)

Quantity	CEM	CMUP	CMX
ϵ_{tt}	0.0409 $\pm$ 0.0033	0.0212 $\pm$ 0.0021	0.0095 $\pm$ 0.0012

- Errors include stat (1%) and syst (8.7%)

$t\bar{t}$ Acceptance

- JP tagging efficiencies for $t\bar{t}$ events (sample of $\sim$685000 PYTHIA Monte Carlo events with top mass 175 GeV/c^2)

Quantity	CEM	CMUP	CMX
JP (QCD cut applied, SF = 0.787 $\pm$ 0.105)			
Acc. No Tag	4.09 $\pm$ 0.04 $\pm$ 0.33	2.12 $\pm$ 0.02 $\pm$ 0.21	0.95 $\pm$ 0.01 $\pm$ 0.12
Tag Eff	56.99 $\pm$ 0.28 $\pm$ 6.66	56.88 $\pm$ 0.36 $\pm$ 6.67	57.84 $\pm$ 0.60 $\pm$ 6.67
Average Tag Eff	57.24 $\pm$ 0.21 $\pm$ 3.85		
Acc. with Tag	2.33 $\pm$ 0.03 $\pm$ 0.33	1.21 $\pm$ 0.01 $\pm$ 0.19	0.55 $\pm$ 0.01 $\pm$ 0.09
Lum (pb^{-1})	161.6 $\pm$ 9.5	161.6 $\pm$ 9.5	149.8 $\pm$ 8.8
$\epsilon_{tt} \int$ L dt	3.77 $\pm$ 0.23 $\pm$ 0.58	1.95 $\pm$ 0.12 $\pm$ 0.32	0.82 $\pm$ 0.05 $\pm$ 0.15
SECVTX (QCD cut not applied, SF = 0.81 $\pm$ 0.07)			
Tag Eff	52.2 $\pm$ 0.386 $\pm$ 3.68	51.9 $\pm$ 0.492 $\pm$ 3.66	53.6 $\pm$ 0.804 $\pm$ 3.78
Average Tag Eff	52.3 $\pm$ 0.284 $\pm$ 3.69		
$\epsilon_{tt} \int$ L dt	3.65 $\pm$ 0.04 $\pm$ 0.47	1.90 $\pm$ 0.02 $\pm$ 0.24	0.80 $\pm$ 0.01 $\pm$ 0.10

Mistags

- We obtain the mistag rate from the mistag matrix prediction applying the mistag estimation for each jet in the pretag data sample and summing the mistag weights for all the jets in the event

Jet Multiplicity	1 jet	2 jets	3 jets	$\geq$ 4 jets
Mistag matrix prediction				
Mistag electrons	24.1 ± 1.8	9.08 ± 0.95	2.87 ± 0.03	1.42 ± 0.03
Mistag muons	17.0 ± 1.6	5.76 ± 1.05	1.57 ± 0.08	0.83 ± 0.01
Total	41.1 ± 2.5	14.84 ± 1.66	4.44 ± 0.09	2.26 ± 0.03

- Errors are only statistical (coming from the matrix)

Mistags

- Corrections:

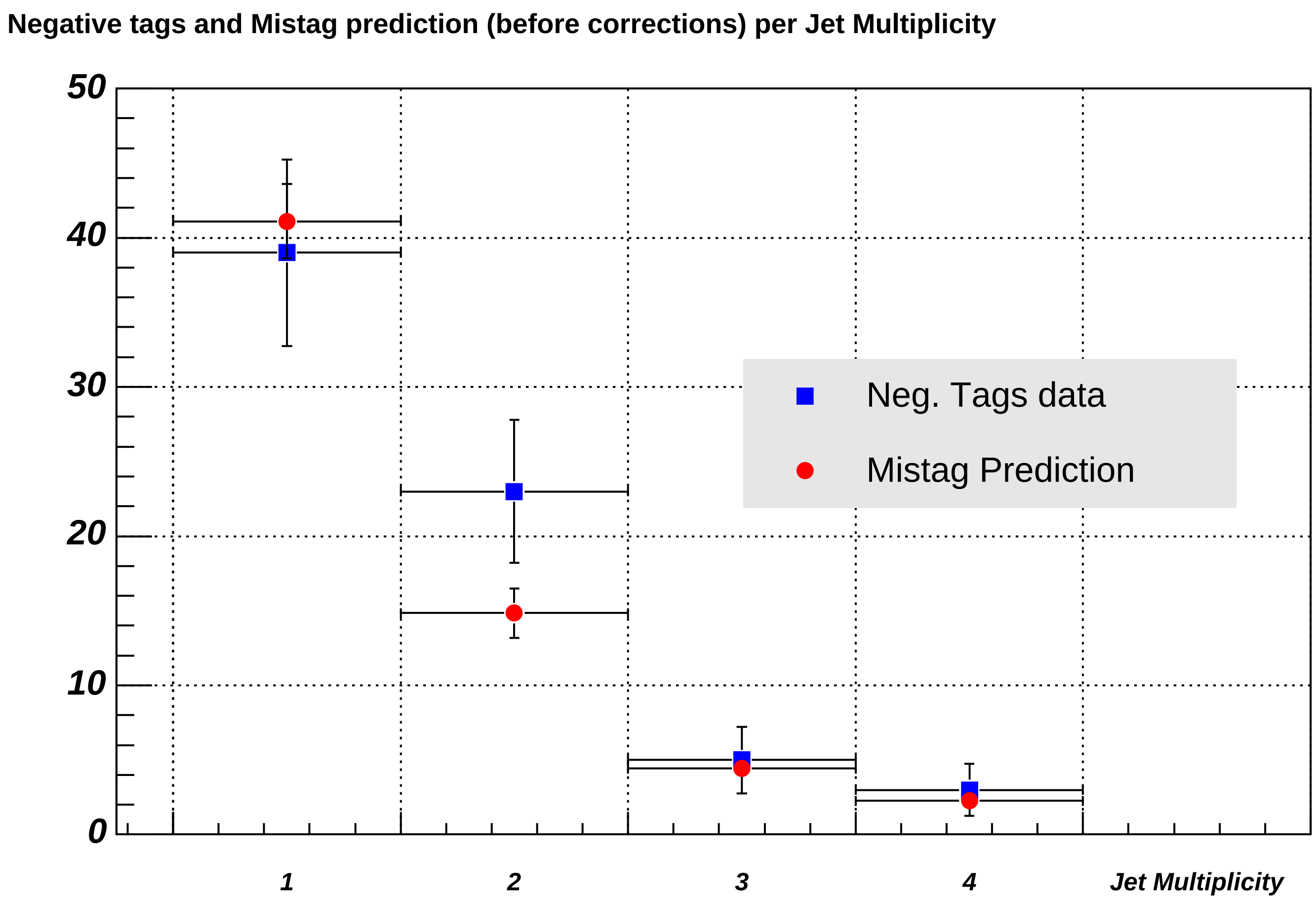
  ◊ mistag rate for positively tagged jets is slighty higher than the negative mistag rate $\Rightarrow$ increase the error by a 20% factor (arbitrary)

  ◊ mistag has to be scaled down by the fraction of pretag events which are due to QCD background

Jet Multiplicity	1 jet	2 jets	3 jets	$\geq$ 4 jets
Prediction $\times$ (1-F_{QCD}) & error$\times$1.2				
Mistag electrons	22.6 ± 4.9	8.20 ± 1.97	2.51 ± 0.66	1.21 ± 0.48
Mistag muons	16.6 ± 3.8	5.59 ± 1.63	1.49 ± 0.51	0.78 ± 0.48
Total	39.2 ± 8.3	13.79 ± 3.38	4.00 ± 0.91	1.99 ± 0.47

Mistags

To compare with the mistag prediction (before corrections), we also estimate the negative tag on the data sample



QCD Background

- Events where the lepton does not come from the decay of a W or Z boson (includes lepton and missing energy fakes, semileptonic B decays...)

- Extracted from the standard "MET vs ISO" method. We define 4 regions in the $\not{E}_T$ vs.isolation plane:

  - Region A: Iso $>$ 0.2 and $\not{E}_T < 15$ GeV

  - Region B: Iso $<$ 0.1 and $\not{E}_T < 15$ GeV

  - Region C: Iso $>$ 0.2 and $\not{E}_T > 20$ GeV

  - Region D: Iso $<$ 0.1 and $\not{E}_T > 20$ GeV

- In the non-isolated region, jets containing the lepton are not counted.

QCD Background

- Pretag event counts in the $\not\!\!E_T$ and isolation sidebands and fraction of non-W background in the pre-tagged sample

Jet Multiplicity	1 jet	2 jets	3 jets	$\geq$ 4 jets
Pretag Electrons				
Region A	23423	3662	536	66
Region B	18517	1782	221	31
Region C	608	239	60	19
Region D	7797	1194	200	60
F_{non-W}	0.062 ± 0.003	0.097 ± 0.008	0.124 ± 0.021	0.149 ± 0.051
Pretag Muons				
Region A	7982	1403	180	31
Region B	3246	304	42	6
Region C	300	112	25	10
Region D	5726	880	115	33
F_{non-W}	0.021 ± 0.001	0.028 ± 0.003	0.051 ± 0.014	0.059 ± 0.034

$$F_{non-W} = (B \times C) / (A \times D)$$

QCD Background

- We estimate this background with 3 different methods (errors are only statistical in all of them):

  - ◊ 1. Extrapolating from region B the tagging efficiency, we predict the expected number of tags from non-W events as

$$N^{tag}_{non-W} = F_{non-W} \times \epsilon_B \times N^{D}_{events}$$

Jet Multiplicity	1 jet	2 jets	3 jets	$\geq$ 4 jets
Electrons				
B event tag rate	0.0148 ± 0.0009	0.030 ± 0.004	0.041 ± 0.013	0.193 ± 0.0710
# events in D	7797	1194	200	60
N^{tag}_{non-W}	7.11 ± 0.52	3.46 ± 0.53	1.01 ± 0.36	1.73 ± 0.84
Muons				
B event tag rate	0.0200 ± 0.0025	0.053 ± 0.013	0.048 ± 0.033	0.167 ± 0.152
# events in D	5726	880	115	33
N^{tag}_{non-W}	2.44 ± 0.34	1.28 ± 0.34	0.28 ± 0.21	0.32 ± 0.34

QCD Background

◊ 2. Calculating the tagged event rates in each sideband and dividing by the number of jets in the bin, the predicted tag rate per jet in the signal region

$$\text{D Pred} = (\text{Tag Rate B} \times \text{Tag Rate C}) / \text{Tag Rate A} \qquad (\text{tag rate / jet})$$

and doing the average over all the bins...

Jet Mult	1 jet	2 jets	3 jets	$\geq$ 4 jets	Average
Electrons					
TR A/j	0.025 ± 0.001	0.025 ± 0.002	0.019 ± 0.003	0.019 ± 0.008	0.0245 ± 0.0009
TR B/j	0.015 ± 0.001	0.015 ± 0.002	0.014 ± 0.004	0.05 ± 0.02	0.0148 ± 0.0008
TR C/j	0.035 ± 0.007	0.027 ± 0.007	0.06 ± 0.02	0.07 ± 0.03	0.034 ± 0.005
D Pred.	0.021 ± 0.005	0.016 ± 0.005	0.04 ± 0.02	0.2 ± 0.1	0.019 ± 0.003
Muons					
R A/j	0.047 ± 0.003	0.043 ± 0.004	0.043 ± 0.008	0.02 ± 0.01	0.045 ± 0.002
R B/j	0.020 ± 0.002	0.026 ± 0.006	0.02 ± 0.01	0.04 ± 0.04	0.021 ± 0.002
R C/j	0.030 ± 0.010	0.06 ± 0.02	0.08 ± 0.03	0.05 ± 0.03	0.043 ± 0.008
D Pred.	0.013 ± 0.005	0.04 ± 0.01	0.03 ± 0.02	0.1 ± 0.2	0.016 ± 0.004

QCD Background

so, finally, the predicted tagged event is

Jet Multiplicity	1 jet	2 jets	3 jets	$\geq$ 4 jets
Electrons				
Pretag	481 ± 20	116 ± 8	25 ± 4	9 ± 3
Tag Rate	0.019 ± 0.003	0.038 ± 0.007	0.06 ± 0.01	0.08 ± 0.01
Predicted Tags	9.2 ± 1.7	4.5 ± 0.8	1.4 ± 0.3	0.7 ± 0.2
Muons				
Pretag	122 ± 7	24 ± 3	6 ± 2	$2 \pm 1.$
Tag Rate	0.016 ± 0.004	0.032 ± 0.008	0.05 ± 0.01	0.06 ± 0.02
Predicted Tags	1.9 ± 0.5	0.8 ± 0.2	0.3 ± 0.1	0.12 ± 0.07

QCD Background

◇ 3. Using tagged events in all regions...

$$N_{non-W}^{tag} = (B/A)_{tagged} \times C_{tagged}$$

Jet Multiplicity	1 jet	2 jets	3 jets	$\geq$ 4 jets
Electrons				
$(B/A)_{tagged}$	0.470 ± 0.034	0.309 ± 0.03		
C_{tagged}	21	13	10	5
N_{non-W}^{tag}	9.87 ± 2.27	4.02 ± 1.19	3.09 ± 1.02	1.55 ± 0.65
Muons				
$(B/A)_{tagged}$	0.173 ± 0.023	0.128 ± 0.029		
C_{tagged}	9	14	6	2
N_{non-W}^{tag}	1.56 ± 0.56	1.79 ± 0.63	0.77 ± 0.36	0.26 ± 0.19
Final result	11.43 ± 2.34	5.81 ± 1.35	3.86 ± 1.08	1.81 ± 1.57

◇ We will use the numbers corresponding to the last method and we will assign a systematic uncertainty of 50% which takes into account the difference between the 3 methods

W + Heavy flavor (method 2)

- By now, we use the heavy flavor fractions in W + jets events from the CDF note 7007 which includes a scale up factor of 1.5 ± 0.4 to correct the gluon splitting discrepancy between Jet Data and MC

Jet Multiplicity	1 jet	2 jet	3 jet	$\geq$ 4 jets
1B	1.0 ± 0.3	1.4 ± 0.4	2.0 ± 0.5	2.2 ± 0.6
2B		1.4 ± 0.4	2.0 ± 0.5	2.6 ± 0.7
1C	1.6 ± 0.4	2.4 ± 0.6	3.4 ± 0.9	3.6 ± 1.0
2C		1.8 ± 0.5	2.7 ± 0.7	3.7 ± 1.0
Wc	4.3 ± 0.9	6.0 ± 1.3	6.3 ± 1.3	6.1 ± 1.3

- Errors are statistical + systematics

W + Heavy flavor (method 2)

- We extract the HF tagging efficiencies applying JP to the Wbb, Wcc and Wc MC sample (includind the SF tagging efficiency)

  - ◊ We assume (by now) the same scale factor for b's and c's

  - ◊ We scale up the error for c-SF by 50%

Jet Multiplicity	1 jet	2 jet	3 jet	$\geq$ 4 jets
1B($\geq$ 1tag)	29.0 ± 3.8	27.5 ± 3.6	28.6 ± 3.7	28.4 ± 3.8
2B($\geq$ 1tag)	0 ± 0	51.8 ± 6.8	48.5 ± 6.4	48.1 ± 6.3
1C($\geq$ 1tag)	8.1 ± 1.6	7.9 ± 1.6	7.8 ± 1.6	7.7 ± 1.6
2C($\geq$ 1tag)	0 ± 0	16.9 ± 3.4	15.5 ± 3.1	15.7 ± 3.2
Wc($\geq$ 1tag)	8.4 ± 1.7	8.4 ± 1.7	8.2 ± 1.6	8.2 ± 1.7

MC derived Backgrounds (WW, WZ, ZZ, Single Top)

- Theoretical Cross Section and sample used for each process

Process	σ_{th}	MC Sample	Events
WW	13.25 ± 0.25	atop4x	597399
WZ	3.96 ± 0.06	atop0y	191011
ZZ	1.58 ± 0.02	atop0z	242500
Single top $W^{\star}$	0.88 ± 0.05	ato0sp	239083
Single top W-g	1.98 ± 0.08	atop1s	262084
Z$\tau\tau$	-	-	-

- We calculate the number of events for each background by

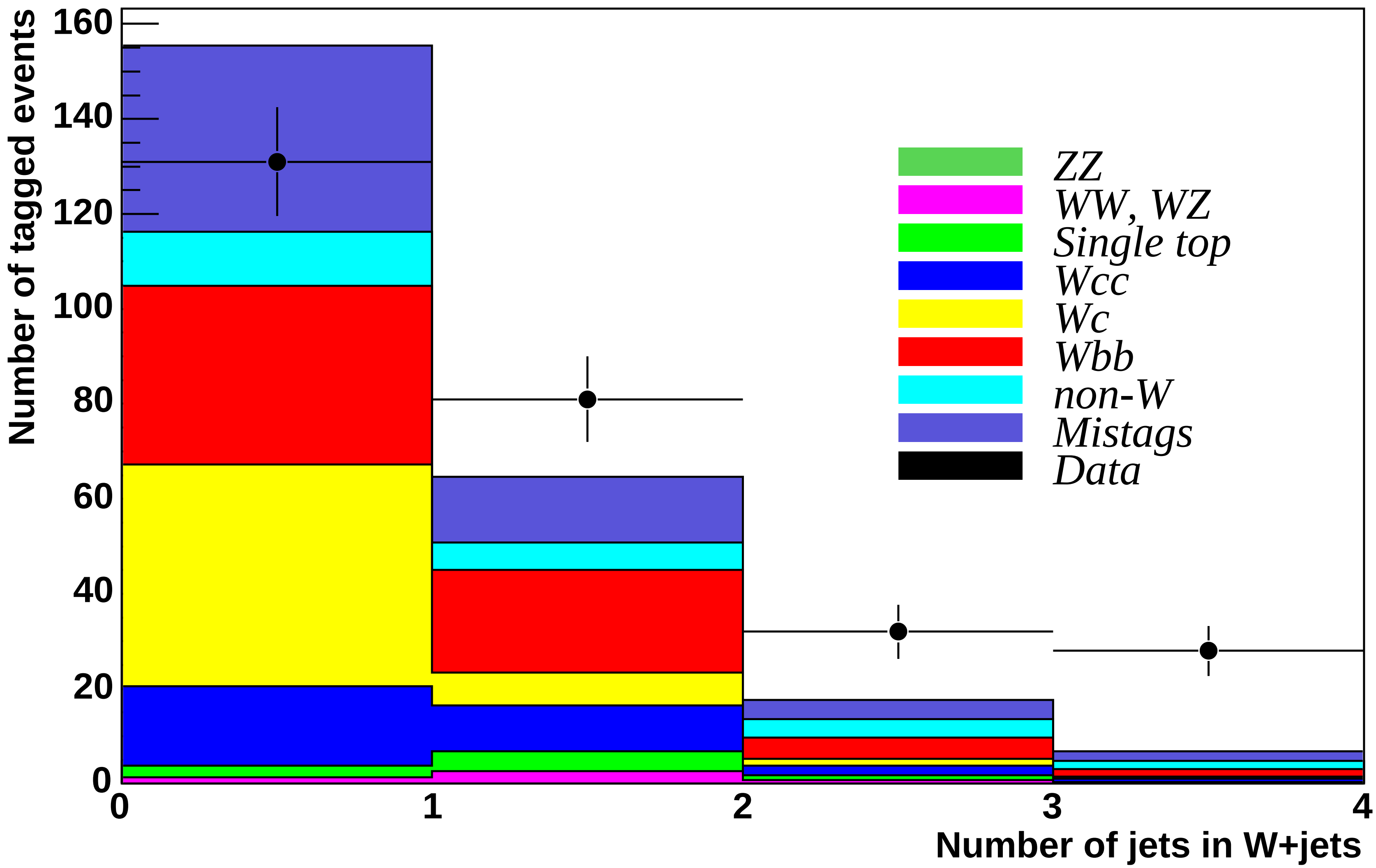
$$N_{events} = \sigma_{th} \times \epsilon_{tot} \times \int L\, dt$$

- The results for these backgrounds are shown in next slide

Summary Backgrounds (tagged sample)

Jet Multiplicity	1 jet	2 jet	3 jet	$\geq$ 4 jets
MC Derived Backgrounds				
WW	0.753 ± 0.127	1.553 ± 0.259	0.437 ± 0.075	0.088 ± 0.017
WZ	0.539 ± 0.095	1.051 ± 0.180	0.319 ± 0.059	0.057 ± 0.015
ZZ	0.036 ± 0.008	0.078 ± 0.015	0.043 ± 0.009	0.009 ± 0.003
Single Top (W^*)	0.538 ± 0.094	1.783 ± 0.312	0.558 ± 0.098	0.131 ± 0.024
Single Top (W-g)	1.907 ± 0.326	2.429 ± 0.414	0.498 ± 0.087	0.075 ± 0.015
Total	3.772 ± 0.752	6.894 ± 1.364	1.855 ± 0.368	0.36 ± 0.073
W + HF				
Wbb	37.53 ± 12.32	21.55 ± 6.81	4.43 ± 1.27	1.56 ± 0.50
W$c\bar{c}$	16.77 ± 5.91	9.58 ± 6.65	1.96 ± 1.45	0.71 ± 0.53
Wc	46.75 ± 13.50	7.01 ± 2.06	1.48 ± 0.43	0.41 ± 0.13
Total	101.1 ± 29.5	38.15 ± 11.88	7.87 ± 2.38	2.68 ± 0.88
Others				
Mistag	39.2 ± 8.3	13.79 ± 3.38	4.00 ± 0.91	1.99 ± 0.47
Non W	11.43 ± 5.74	5.81 ± 3.20	3.86 ± 2.21	1.81 ± 1.81
Total Background	155.50 ± 31.54	64.64 ± 13.41	17.59 ± 3.51	6.84 ± 2.09
DATA	131	81	32	28

Summary Backgrounds (tagged sample)



Cross section estimation

Error	Sys (%)	MC stat (%)	Stat from data (%)
Kinematic efficiency	8.7	1	
Scale Factor (b's/c's)	13/20		
Luminosity	5.9		
QCD fraction	50		17
QCD prediction	50		34
W+HF fraction/tags	30	14	0.8
Mistag	+20.0		9.6
MC derived (σ's)	1.8	2.7	

- We propagate these errors in the expression of the x-section considering correlation for efficiency, scale factor and luminosity. Wbb and Wcc systematics are considered correlated across all the bins. All the other errors are treated as uncorrelated

Cross section estimation

- We calculate the $t\bar{t}$ cross section by

$$\sigma = (N - B)/\ \epsilon \times \int L\,dt$$

and the value we get is

$$\sigma = 5.83 \pm 1.21\ (\text{stat}) \pm 0.25\ (\text{MC stat})\ ^{+1.27}_{-1.28}\ (\text{syst})$$

Summary

- We have made top ntuples (v.4) including JP information

- We have selected 59 top candidates events that pass the JP$<$0.01. We found a 75% overlap with SECVTX candidates and a 79% with Loose SECVTX

- The $t\bar{t}$ acceptance for JP is slighty higher ($\sim 9\%$ than for SECVTX). And we obtain similar fraction of mistags in the total background estimation

- Errors on efficiency and mistag matrix for JP are not optimized for this offline version.

- We obtain a $t\bar{t}$ x-section compatible with previous measurements and with a total error of 30% (22% syst)

Future plans

- Short term goals:
  - ◇ Include Z$\tau\tau$
  - ◇ Cross-check HF contributions
  - ◇ Document the work done
  - ◇ Try to bless this numbers in the coming weeks

- Main goal: Setup and optimize this analysis with gen5 and $\sim$400 pb^{-1} of data and publish the result

- The list of future tasks includes:
  - ◇ Remake (and optimize) JP parametrizacion for gen5 (with the Florida group) We need urgently Jet 50 MC with the right Silicon configuration
  - ◇ Re-calculate efficiencies and mistag matrix and try to improve the errors
  - ◇ Optimize the x-section analysis
  - ◇ The aim is to be ready to start the analysis by the end of this year