Double Tag $t\bar{t}$ Cross Section: Loose Tagger

June 4, 2004

Lepton+Jets Meeting

John Paul Chou, Andrew Foland, Melissa Franklin, João Guimarães da Costa, Ayana Holloway, Lester Miller, **Salvatore Rappoccio**, Daniel Sherman

- High efficiency (loose) SecVtx Tagger in `4.11.2`
  - Scale Factor
  - Mistag Matrix
- Top Cross Section
  - Data samples
  - Backgrounds
    - W+HF
    - Monte Carlo based
    - Mistags
    - Non-W
  - Cross section
  - Kinematic plots
- Discussion and Plans

High Efficiency SecVtx Tagger in 4.11.2

- Attempt to increase efficiency of SecVtx tagger
  - **Double tag analyses require higher efficiency, whereas mistags are not as much of an issue**.
  - Gain by $\sim \epsilon^2$ due to 2 tags.
  - Effective luminosity enhancement of ~45%
- Track selection developed for `4.7.1` tracking too tight for `4.11.2` and `5.3.1`
  - *Loosen* track quality cuts to increase efficiency
  - *Tighten* vertex quality cuts to decrease mistags
  - Use insights from `5.3.1` development to create looser `4.11.2` tagger
  - For loose tagger, specifically:
    - Requiring *two tags*, so mistags are not so important

  $\Rightarrow$ **Only loosening track selection, not tightening vertex cuts** for looser tagger
- Why `4.11.2`
  - Until this week, MC not ready in `5.3.1`
  - Can gain insights into double tag analyses that we can apply in `5.3.1`

High Efficiency SecVtx

- Used simple track selection for `4.11.2` tagger:
  - `defTracks`
  - ≥ 3 SVX $R - \phi$ layers hit
  - $\chi^2_{trk} < 20.0$
  - Reduced $|d_0|$ cut from 0.30 cm to 0.15 cm to reduce material tags
- Included **bug fix** for track error extrapolation
- Included new `ZVertexModule` using beamline as seed.
- Using `Pass1TryHarder` (if a tag isn't found with first seed in Pass 1, keep trying other seeds)
- Increases efficiency in data and MC

BTagging Scale Factor

- Using `b2t120`, `b2t140`, `btop2a`, `btop4a` for 8-GeV data and MC
- Using CDF6912 method for scale factor determination
  - Uses CDF6904 method of using electron dataset with conversions
  - Slight modification : solve directly for the scale factor, derive heavy flavor fraction (F^h) and efficiency (δ)

$\Rightarrow$ No *direct* systematic on the value of F^h (heavy flavor fraction)

- In top sample, SF $= 0.84 \pm 0.067$
- **See Daniel Sherman's talk at High-p_T Btagging meeting today**

Source	Relative Error (%)
Mistag Subtraction	3.0
Conversion-Finding	3.0
c/b Ratio	1.5
Semi-Leptonic Decay	1.4
Charm Scale Factor	0.5
Total	4.7
E_T-Dependence	0.1*($\langle E_T \rangle$ - 25 GeV)

	Loose	**Standard**
Fiducial Efficiencies and Scale Factor		
Per-jet $\epsilon^{MC}_{t\bar{t}}$	46.4%	40.6%
ϵ^{DATA}_{ele}	28.1%	23.6%
ϵ^{MC}_{ele}	33.6%	29.1%
Scale Factor	84 ± 6 %	81 ± 6%
Per-jet $\epsilon^{Data}_{t\bar{t}}$	40 %	33%
Expected Gain in $t\bar{t}$ Significance		
$\mathbf{t}\bar{t}$ $S/\sqrt{S+B}$	3.3	2.8

Expect 40-55% increase in double tag rate in top events

Tag Matrices

- Using same matrix configuration as CDF6533
- Using `gjet08/09` JET20,50,70,100 samples
- Higher tag rates for both positive and negative tags.

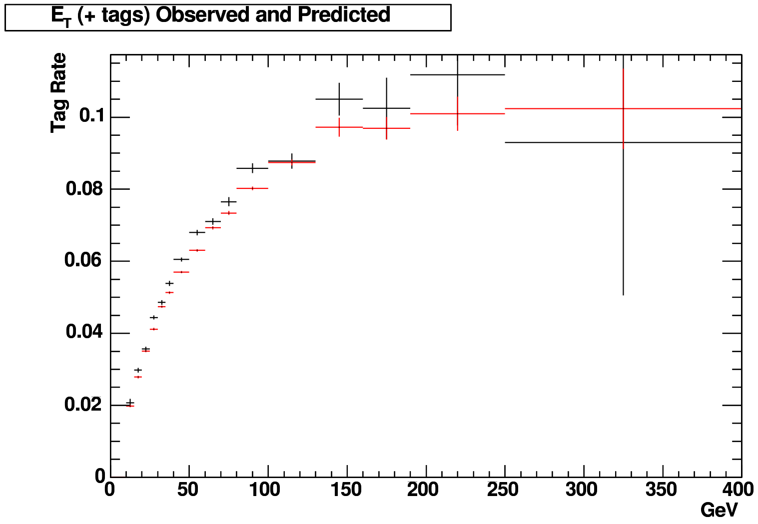
	Loose	**Standard**
Fiducial Rates		
Pos. Tag Rate	$2.1 \pm 0.0\%$	$1.0 \pm 0.0\%$
Neg. Tag Rate	$1.0 \pm 0.0\%$	$0.3 \pm 0.0\%$
$\sum E_T$ Sample Bias	5%	6%

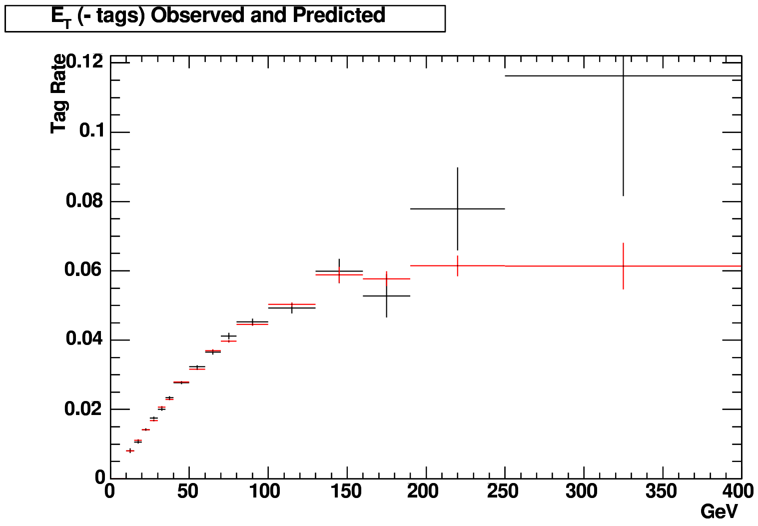
- Dominant systematic for mistag matrix is similar to that of the default tagger

$\Rightarrow$ For now, still assigning same 8% total systematic as standard SecVtx

Tag Matrices

- Examine tag matrix prediction on $\sum E_T$ sample
- **Black is observed in $\sum E_T$ sample, red is predicted from tag matrix derived from** $\mathrm{JET20,50,70,100}$
- Assign 5% systematic bias (see previous page)





Double Tag Cross Section

- Using Standard 4.11.2 Event Selection
  - DQM Good Run List v4.0 with silicon
    - Removed 163463,163474,164844
  - Level 4 jet corrections
  - One tight, isolated trigger lepton
  - $\not{E}_T > 20.0$ GeV (corrected to Level 4)
  - Dilepton, Z boson, conversion, cosmic veto
  - $\Delta\phi_{\not{E}_T-jet}$ QCD removal cut
  - $|\Delta z_{lepton-zvtx}| < 5.0$ cm
  - ≥ 3 jets
  - $\geq 2 +$ SecVtx tags
- CAVEAT!
  - One small difference: ZVertexModule was upgraded in our Vertex packages to input the beamline for the z-vertex
  - Doesn't affect $t\bar{t}$ pretag efficiency
  - However, we will have slightly different pretag samples

	Loose	**Standard**
$\epsilon^{t\bar{t}}_{pretag}\int\mathcal{L}dt$	$11.38 \pm 1.2\ pb^{-1}$	$11.46 \pm 1.2\ pb^{-1}$

- Measure cross section:

$$\sigma_{t\bar{t}} = \frac{N_{obs} - N_{bkg}}{\epsilon_{2\ btag}\epsilon_{pretag}\int\mathcal{L}dt} \tag{1}$$

Data Samples

- Electron/muon signal data: standard REMAKE, reprocessed with loose tagger
- $t\bar{t}$ Monte Carlo: `ttopei`
- Backgrounds

1. Electroweak and Single Top (Monte Carlo backgrounds)
   - WZ (`atop0y`)
   - single top (`ttop0s,ttop1s`)
2. Non-W background (data)
3. W+Heavy Flavor
   - $W + b\bar{b}$ (ALPGEN W+bb+Np)
   - $W + c\bar{c}$ (ALPGEN W+cc+Np)
4. Mistags
   - Real tag + fake tag (ALPGEN W+QQ+Np)
   - Fake tag + fake tag (Data)

Electroweak and Single Top

- Use Monte Carlo to estimate these
- Straightforward calculation using theoretical cross sections for process $p\overline{p} \to X$:

$$N^{X}_{MC} = (\sigma_X \int \mathcal{L} dt)(\epsilon_{pretag})(\epsilon_{2-tag})(SF^2) \quad (2)$$

- Assume MC Derived backgrounds are 100% correlated since btag SF is dominant systematic

N^{exp}_{ev}	2 jet	3 jet	≥ 4 jet
WZ	0.17 $\pm$ 0.04	0.03 $\pm$ 0.01	0.017 $\pm$ 0.010
Single top, s channel	0.47 $\pm$ 0.08	0.14 $\pm$ 0.03	0.038 $\pm$ 0.008
Single top, t channel	0.034 $\pm$ 0.008	0.04 $\pm$ 0.01	0.012 $\pm$ 0.004

Non-W

- Using "standard" $\not\!\!E_T$ vs Isolation method:

$$N_{NON-W} = (N_C^{pretag}) \left(\frac{N_B^{pretag}}{N_A^{pretag}} \right) (\epsilon_{2-tag}) \tag{3}$$

QCD Fractions			
CEM	0.157 $\pm$0.007	0.172$\pm$0.02	0.156$\pm$0.03
CMUP	0.04$\pm$0.003	0.050$\pm$0.008	0.07$\pm$0.03
CMX	0.04$\pm$0.003	0.050$\pm$0.008	0.07$\pm$0.03
QCD 2-Tag Rates			
CEM	0.0009$\pm$0.0005	0.005$\pm$0.003	0.005$\pm$0.003
CMUP	0$\pm$ 0	0$\pm$ 0	0$\pm$ 0
CMX	0$\pm$ 0	0$\pm$ 0	0$\pm$ 0
Total QCD	0.17 $\pm$ 0.10	0.13 $\pm$ 0.09	0.02 $\pm$ 0.02

W+Heavy Flavor

- Using Method 2 estimate:

$$N_{W+HF} = (N_{pretag} - N_{non-W} - N_{MC} - N_{t\bar{t}})(f_{HF})(\epsilon_{2-tag}^{matched})(SF^2) \tag{4}$$

- Using NEW heavy flavor fractions from CDF7007
- Using $SF = 86 \pm 7\%$
- Derive $\epsilon_{2-tag}^{matched}$ from inclusive ("N_{jet}") matching of ALPGEN MC
- Assume $W + b\bar{b}$ and $W + c\bar{c}$ uncorrelated since HFF is dominant systematic

Jet Multiplicity	2 jets	3 jets	$\geq$ 4 jets
N_{pretag}	2107	320	96
N_{non-W}	220	34.2	5.35
N_{MC}	21	4.83	0.842
$N_{t\bar{t}}$	25.8	48.4	55.9
Total Corrected Pretag	1839	233	34
Wbb, HF Frac	0.0143 ± 0.0038	0.0201 ± 0.0054	0.0261 ± 0.007
Wcc, HF Frac	0.018 ± 0.0048	0.0267 ± 0.0072	0.0356 ± 0.0096
Wbb, Pretag	26.3 ± 7.0	4.67 ± 1.3	1.0 ± 0.2
Wcc, Pretag	33.1 ± 8.8	6.21 ± 1.7	1.2 ± 0.3
Wbb, $\epsilon_{2-tag}^{matched} * SF^2$	0.09 ± 0.01	0.0767 ± 0.00933	0.0787 ± 0.00958
Wcc, $\epsilon_{2-tag}^{matched} * SF^2$	0.0046 ± 0.0006	0.0061 ± 0.0007	0.009 ± 0.001
Wbb, N_{Tag}	2.3 ± 0.7	0.4 ± 0.1	0.07 ± 0.02
Wcc, N_{Tag}	0.15 ± 0.06	0.03 ± 0.01	0.012 ± 0.005

Mistags

- Cannot use direct output of mistag matrix because of large heavy flavor content

$\Rightarrow$ Mistag matrix prediction is an overestimate

  - Still a single tag matrix
  - Examining how to correctly use this for double tag analyses

- Stephen Miller's method of estimating this:
  - Try to scale mistag matrix prediction by amount of "fake" mistags (i.e. mistags from resolution)
  - Account for tags from one real HF tag + one fake tag by running the mistag matrix over the W+HF MC
  - Assume these two are correlated and assign large (50%) systematic to this method

$$N^{fake}_{mistag} = \left(\frac{N^{-} + N^{+}_{non-W}}{N^{+}}\right) N^{+-}_{pred} \quad (5)$$

$$N^{W+HF}_{mistag} = (N^{MC}_{pretag})(N^{+-}_{pred:W+HF})(f_{HF})(K) \quad (6)$$

- For simplicity, assume $K = SF^2$, but this is not necessarily the case
- For now, also ignoring Wc (will add later)

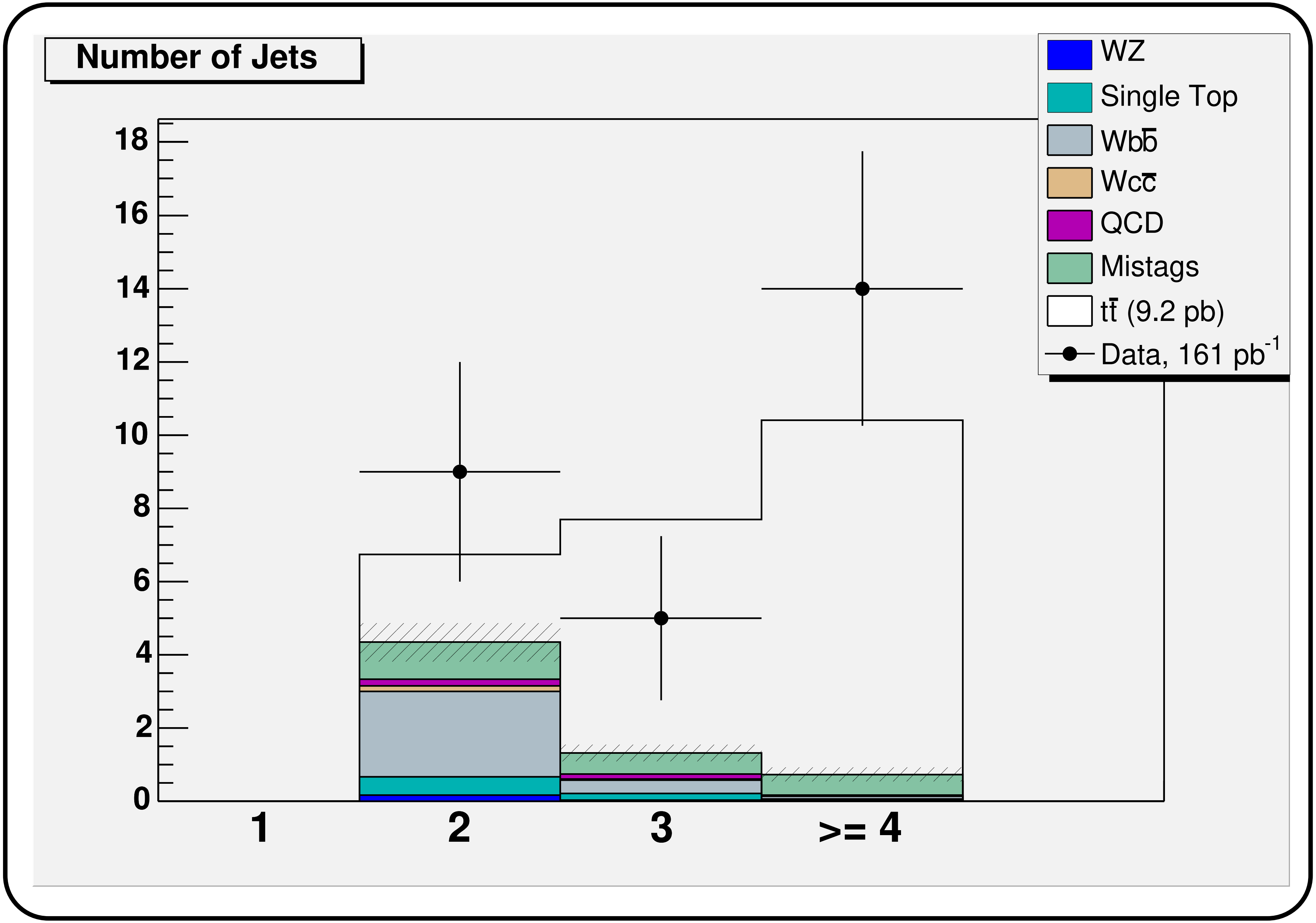
Mistags			
Jet Multiplicity	2 jets	3 jets	$\geq$ 4 jets
$N^{W+HF}_{mistags}$	0.221 $\pm$ 0.07	0.09 $\pm$ 0.03	0.03 $\pm$ 0.01
N^{-}	52	10	8
N^{+}_{non-W}	10 $\pm$ 1	4 $\pm$ 1	0.6 $\pm$ 0.2
N^{+}	127	42	42
N^{+-}_{pred}	1.7 $\pm$ 0.2	1.5 $\pm$ 0.2	2.6 $\pm$ 0.3
N$_{mistag}$	**1.0 $\pm$ 0.5**	**0.6 $\pm$ 0.3**	**0.6 $\pm$ 0.3**

Summary

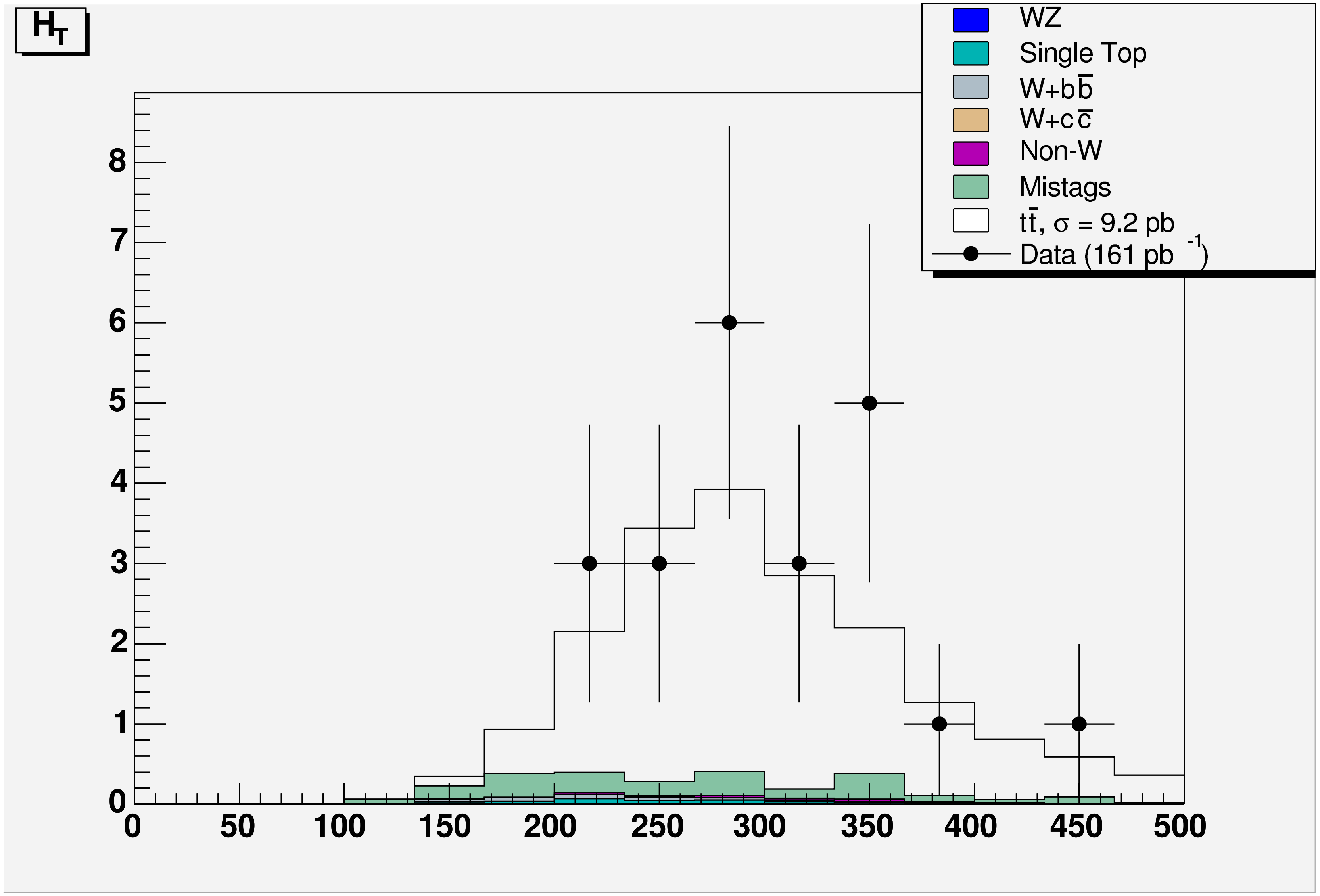
Summary			
	2 jet	3 jet	≥ 4 jet
WZ	0.17 $\pm$ 0.04	0.03 $\pm$ 0.01	0.017 $\pm$ 0.010
Single top, s channel	0.47 $\pm$ 0.08	0.14 $\pm$ 0.03	0.038 $\pm$ 0.008
Single top, t channel	0.034 $\pm$ 0.008	0.04 $\pm$ 0.01	0.011 $\pm$ 0.004
QCD	0.17 $\pm$ 0.10	0.13 $\pm$ 0.09	0.02 $\pm$ 0.02
W$b\bar{b}$	2.3 $\pm$ 0.7	0.36 $\pm$ 0.11	0.07 $\pm$ 0.02
W$c\bar{c}$	0.15 $\pm$ 0.06	0.04 $\pm$ 0.01	0.012 $\pm$ 0.005
Mistags	1.0 $\pm$ 0.5	0.6 $\pm$ 0.3	0.6 $\pm$ 0.3
Total Background	4.4 $\pm$ 0.8	1.3 $\pm$ 0.3	0.7 $\pm$ 0.3
$\epsilon^{t\bar{t}}_{2-tag} * SF^2$	9.1 $\pm$ 1.1 %	13 $\pm$ 1.6%	17.2 $\pm$ 2.1%
$t\bar{t}$ (9.2 pb)	2.4 $\pm$ 0.9	6.4 $\pm$ 2.3	9.7 $\pm$ 3.4
Total Expected	6.8 $\pm$ 1.7	7.7 $\pm$ 2.6	10.4 $\pm$ 3.7
Data	9	5	14
Total $\epsilon^{t\bar{t}}_{2-tag}\epsilon^{t\bar{t}}_{pretag}\int \mathcal{L}dt$	1.85 $\pm$ 0.4 (sys) $\pm$ 0.1 (lum)		
$\sigma_{t\bar{t}}$	**9.2 $\pm$ 2.4 (stat) $\pm$ 1.9 (sys) $\pm$ 0.5 (lum) pb**		

Systematics

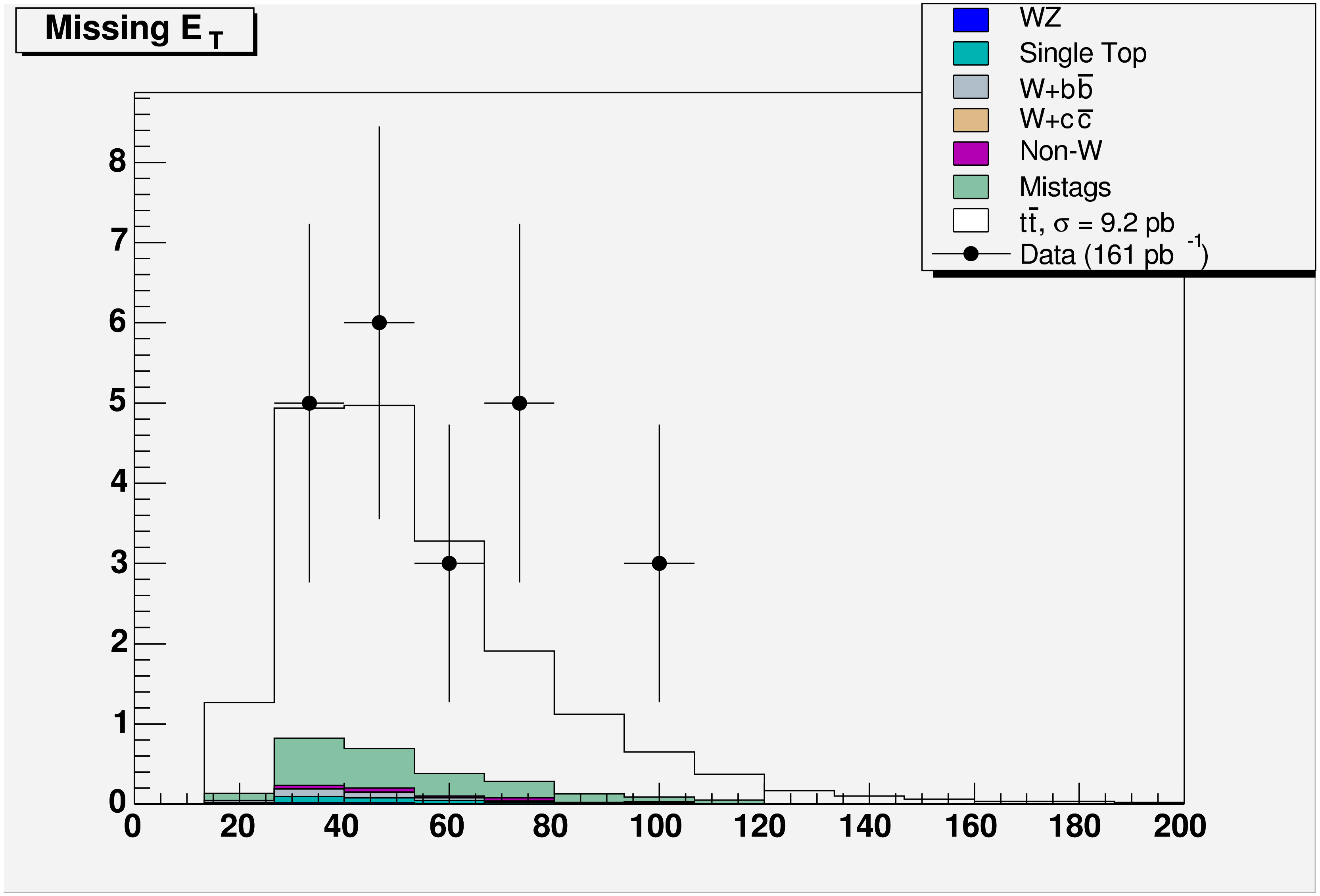
Summary	
Quantity	Relative Uncertainty
ϵ_{2-tag}	16%
ϵ_{pretag}	11%
$\int Ldt$	6%
Background	3%
Total	20%

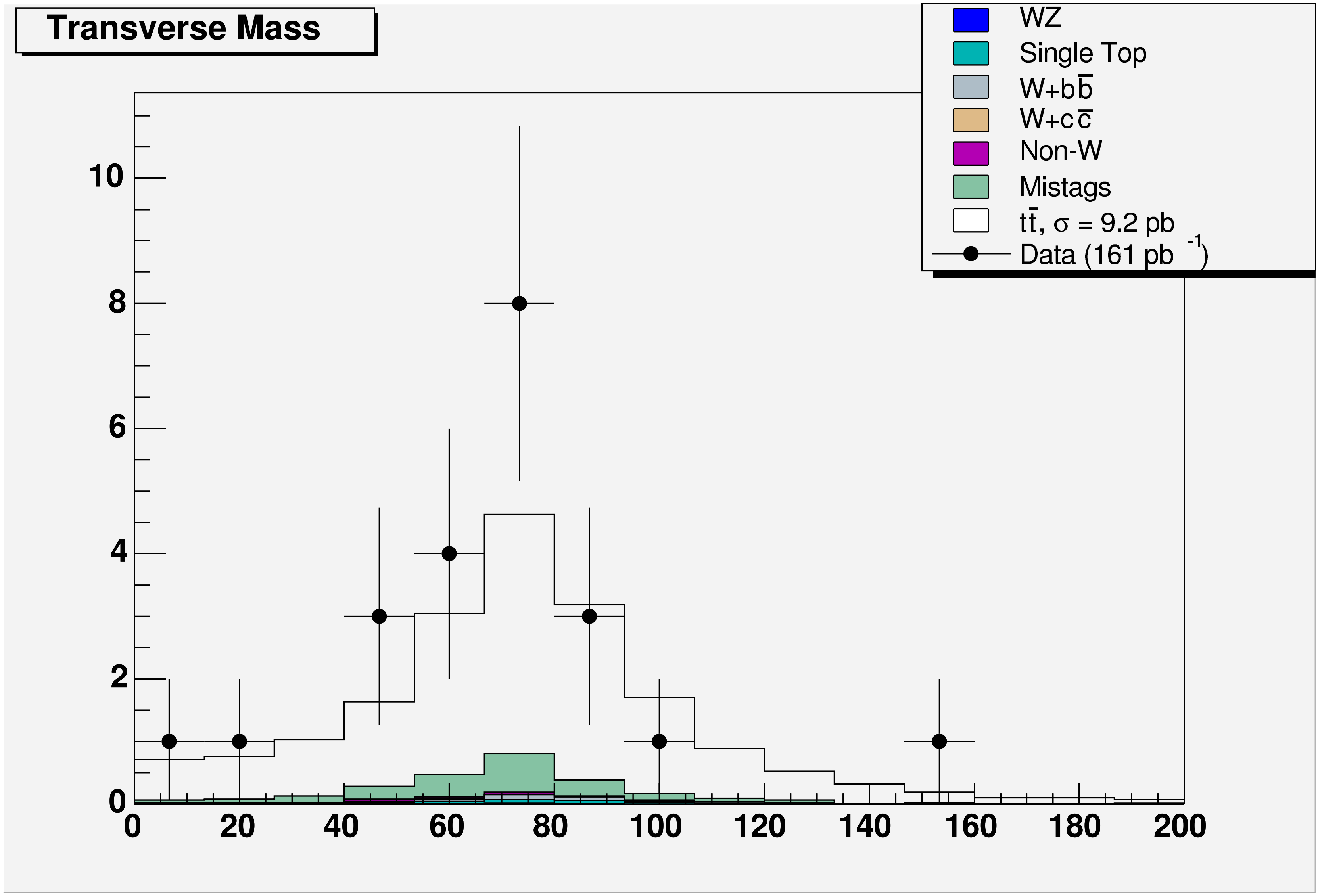


Number of Jets
WZ
Single Top
Wb$\bar{b}$
Wc$\bar{c}$
QCD
Mistags
t$\bar{t}$ (9.2 pb)
Data, 161 pb^{-1}
18
16
14
12
10
8
6
4
2
0
1
2
3
>= 4

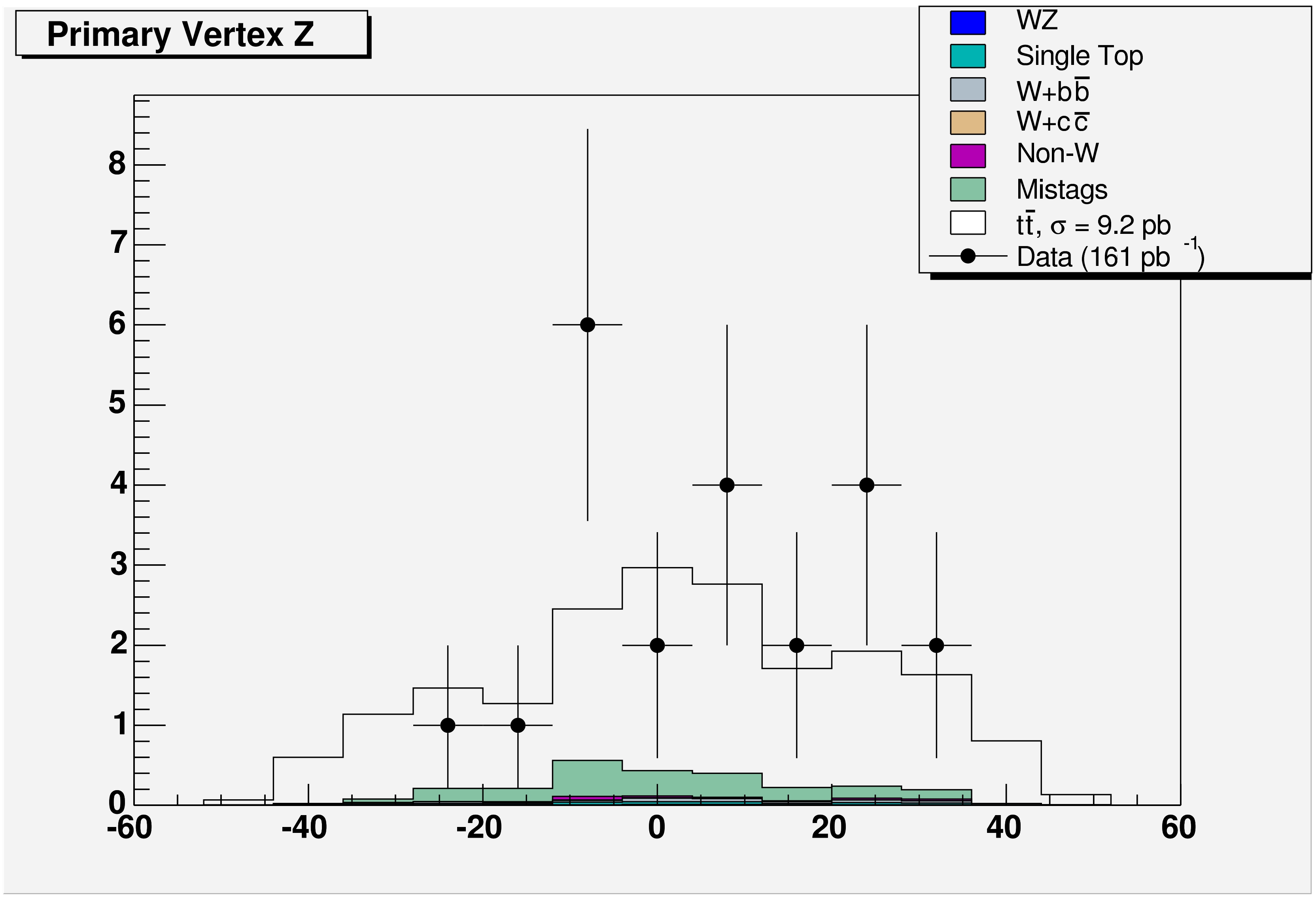


H_T
WZ
Single Top
W+b b̄
W+c c̄
Non-W
Mistags
t t̄, σ = 9.2 pb
Data (161 pb^-1)
0
1
2
3
4
5
6
7
8
0
50
100
150
200
250
300
350
400
450
500


Missing E T
WZ
Single Top
W+b b̄
W+c c̄
Non-W
Mistags
t t̄, σ = 9.2 pb
Data (161 pb -1)
0
1
2
3
4
5
6
7
8
0
20
40
60
80
100
120
140
160
180
200



Transverse Mass
WZ
Single Top
W+b$\bar{b}$
W+c$\bar{c}$
Non-W
Mistags
$t\bar{t}$, σ = 9.2 pb
Data (161 pb^{-1})
0
2
4
6
8
10
0
20
40
60
80
100
120
140
160
180
200



Primary Vertex Z
WZ
Single Top
W+b$\bar{b}$
W+c$\bar{c}$
Non-W
Mistags
t$\bar{t}$, σ = 9.2 pb
Data (161 pb^{-1})
0
1
2
3
4
5
6
7
8
-60
-40
-20
0
20
40
60

Conclusions and Plans

- Loose SecVtx for 4.11.2 ready to begin blessing process
  - Scale factor calculated
  - Have mistag matrices
- Top cross section with double tags benefitting from larger statistics

$$\sigma_{t\bar{t}} = 9.2 \pm 2.4\ (stat) \pm 1.9\ (sys) \pm 0.5\ (lum)\ pb \tag{7}$$

- Need to understand mistags in context of double tag analyses
- Will prebless/bless 4.11.2 tagger along with cross section in June/July
- Tools in place to move to 5.3.1. Waiting on Monte Carlo generation to update analysis and include new data.