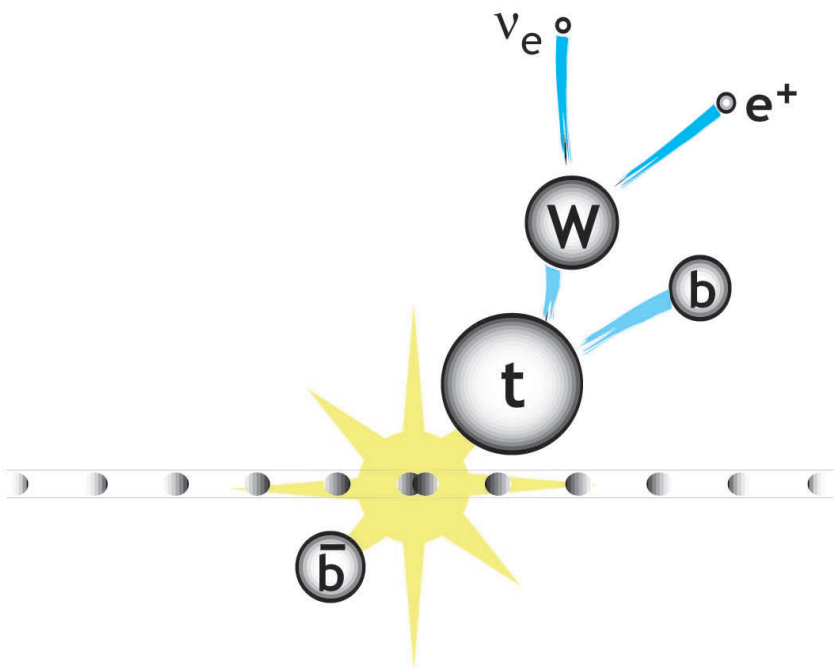
Single Top Perspectives with 1 fb^{-1}

Wolfgang Wagner
for the Single Top Group



Outline:

- Projections for Neural Network Analysis
- Signal MC validation

Disclaimer

Predictions are rather difficult, especially if they concern the future.

Niels Bohr, 1885 – 1962, Nobel Laureate 1922

Scope of this study

- Study prepared for single top workshop on Dec. 2 and 3, 2004. Detailed talk is linked from the single top webpage.
- Based on published analysis, but replacing the cut on $M_{\ell\nu b}$ and the fit to $Q\eta$ by a cut on a neural network discriminant + counting experiment.
- Use 4.9.1/4.11.1 Monte Carlo and data. Well understood data set. All efficiencies and event expectations are available.
- Get approximate idea of potential. $\Rightarrow$ Fine-tuning in 5.3.3.
- Use only 2-jet events $\rightarrow$ comparison with published analysis, several variables not defined for 1-jet bin, MC background samples best describes 2-jet bin.
- Train for "combined search" with t- plus s-channel as signal. Later split into t- and s-channel separately.

Neural Network Analysis
Training Samples

Benefits:

- Learn about correlations.
- Automated pruning of variables with low significance.
- Flexibility of testing and absorbing new ideas.

Training:

- Signal:Background = 35:65
- Realistic background composition.
- caveat non-W: non-isolated lepton sample mis-represents bkg.

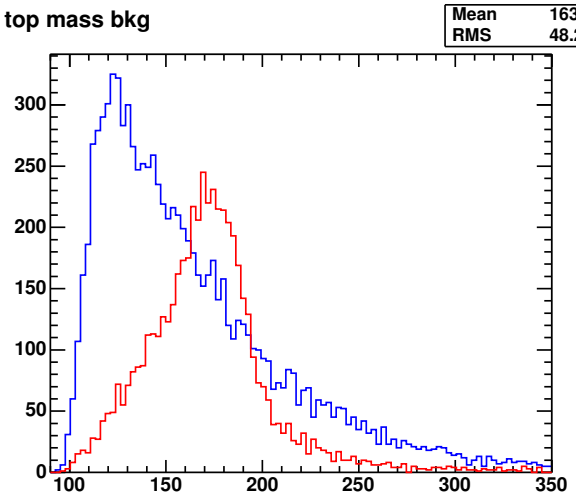
Sample	Pred.	Training
t-channel	3.9%	24.2%
s-channel	3.3%	10.8%
$t\bar{t}$	9.4%	9.1%
WW	1.3%	1.3%
WZ / ZZ	1.4%	1.3%
$Wb\bar{b}$	27.6%	26.9%
$Wc\bar{c}$	9.9%	6.5%
Wc	9.9%	9.6%
mistags	21.5%	1.6%
non-W	12.0%	8.7%

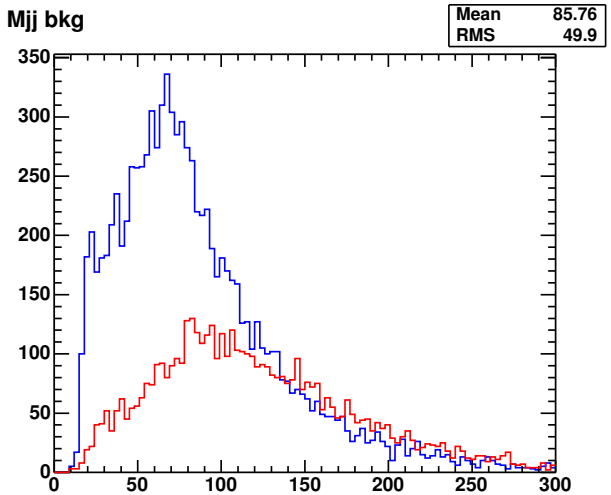
Used Variables

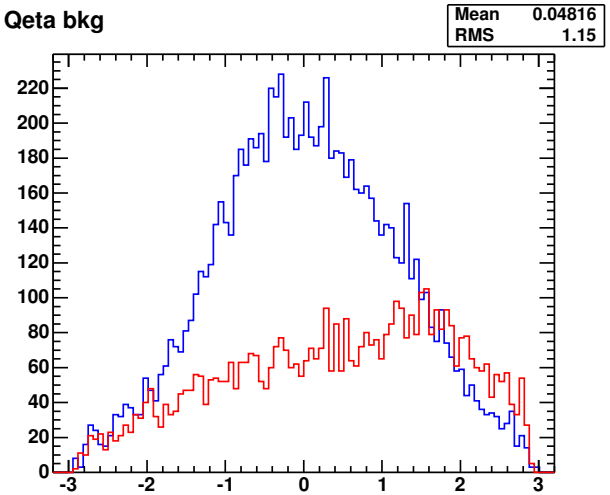
1. $M_{\ell\nu b}$ (top mass)
2. M_{jj} (di-jet mass)
3. $Q \cdot \eta$
4. ANN b-tag (S. Richter)
5. $\cos\theta_W$
6. H_T
7. $\theta_{\ell b}$
8. $\Delta\eta(j_1, j_2)$
9. η_W
10. $\log_{10}(\Delta_{23})$
11. $\sum \eta_j$
12. $\log_{10}$(Sphericity)
13. η_{j1}
14. $\sum E_T(j)$
15. $E_T(j_1)$
16. $\log_{10}(\Delta_{12})$
17. $E_T(j_2)$

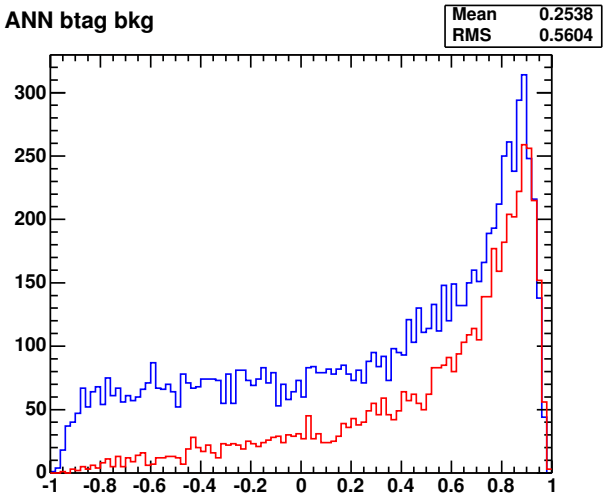
Some Input Distributions

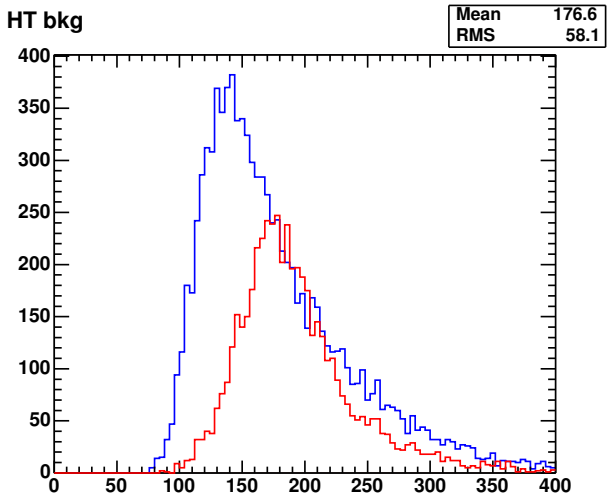
histogam area according to training weights

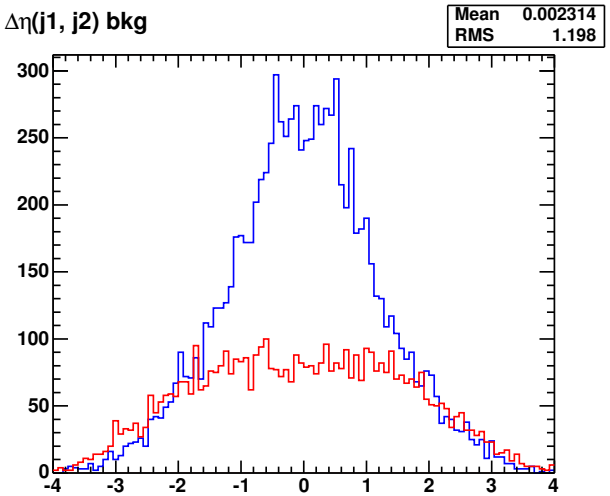










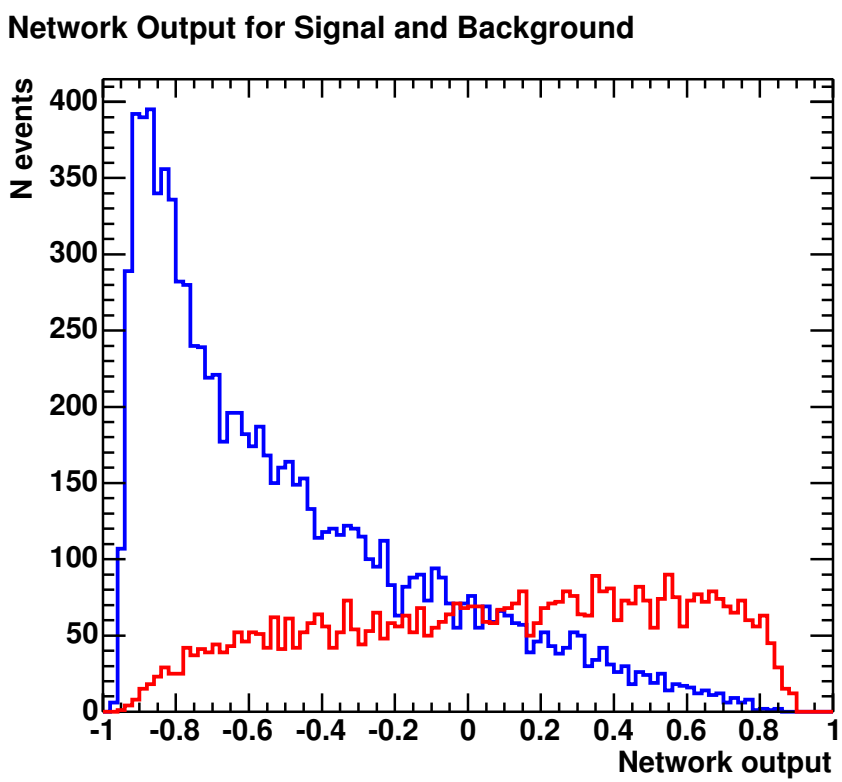


Network Architecture and Settings

- Node architecture: input-middle-output = 18-8-1
  Training is stable against changes in number of nodes in middle layer.
- Size of training sample: 14758 events (sig. + bkg.)
- 50 iterations
- Update weights after 100 events.
- Pre-processing: Decorrelate input variables and rotate all linear dependence with the target (mean) to the first new input variable, 2nd order correlation to 2nd variable etc.
- Prune variables with a significance less than 1 σ.
- Slow learning speed, delayed by factor of 0.5.
- Possible to use binary, discrete or even semi-discrete variables.

Training Performance
Relevance of Individual Variables

Network Output for
Signal (35%) and
Background (65%)



17 variables are used.

Rank	Variable	Relev. σ
1	$M_{\ell\nu b}$	45.0
2	M_{jj}	30.8
3	$Q \cdot \eta$	18.5
4	ANN b-tag	17.6
5	$\cos\theta_W$	16.4
6	H_T	9.8
7	$\theta_{\ell b}$	5.9
8	$\Delta\eta(j_1, j_2)$	4.9
...	....	

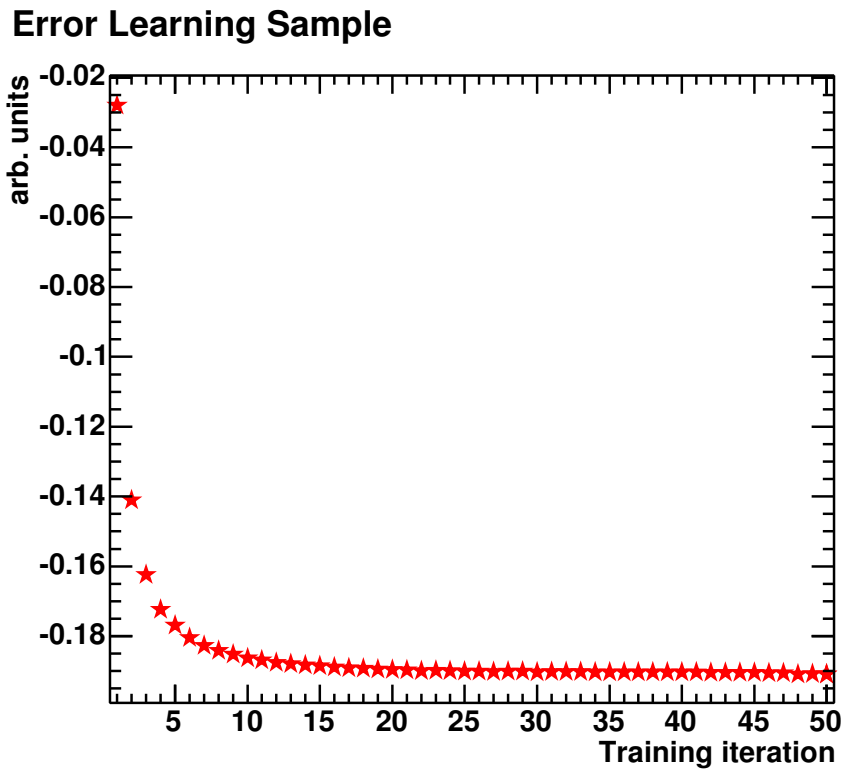
Relevance of Variables (2)

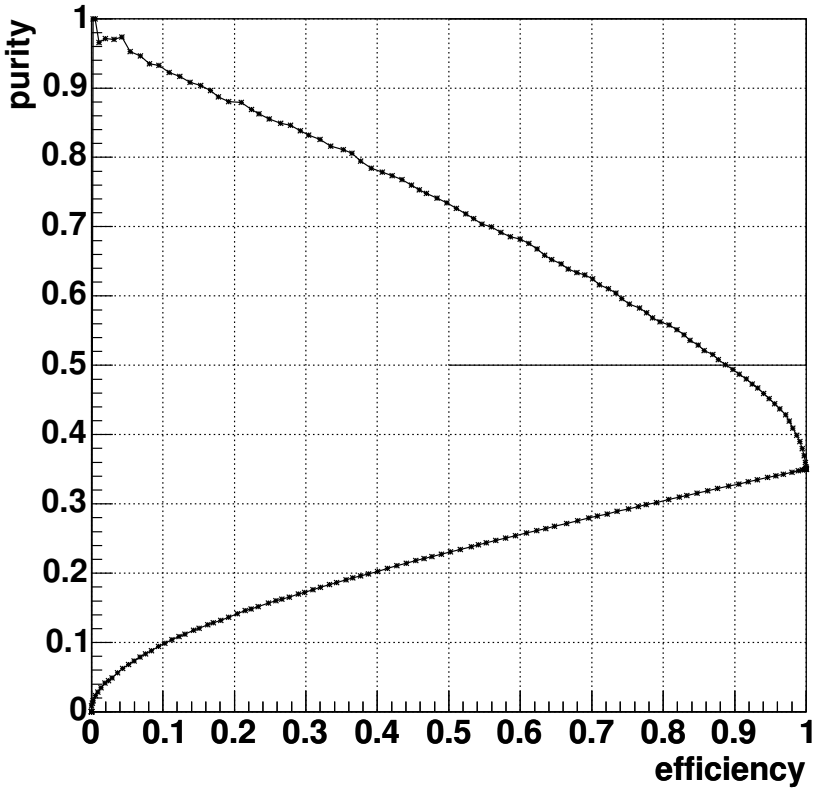
Rank	Variable	Relev. σ
11	$\sum \eta_j$	3.5
12	$\log_{10}$(Sphericity)	1.4
13	η_{j1}	1.2
14	$\sum E_T(j)$	1.2
15	$E_T(j_1)$	1.1
16	$\log_{10}(\Delta_{12})$	0.2
17	$E_T(j_2)$	0.0

Variables below 1 σ relevance are pruned away.

Low relevance does not necessarily correspond to low correlation with the target, but the information is already contained in other variables.

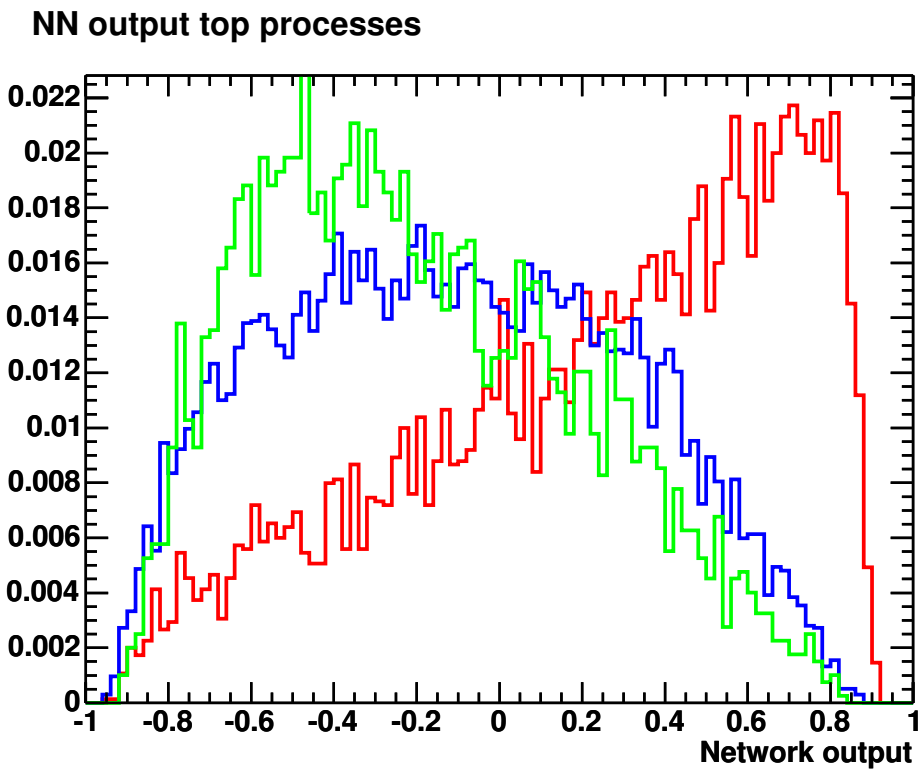
Training Performance

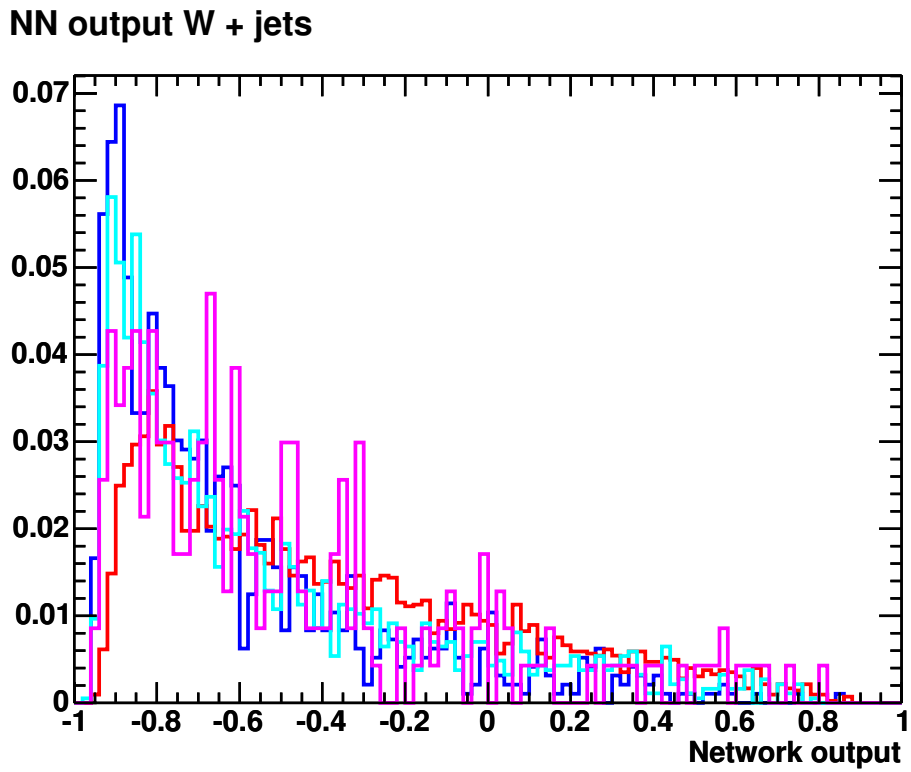




Error stablizes fast and remains stable → no overtraining.

Network Output





t-channel s-channel $t\bar{t}$ W$b\bar{b}$ W$c\bar{c}$ Wc mistags

Significant difference between t- and s-channel

W + non-b jets have same shape within the statistics.

Cut Efficiencies

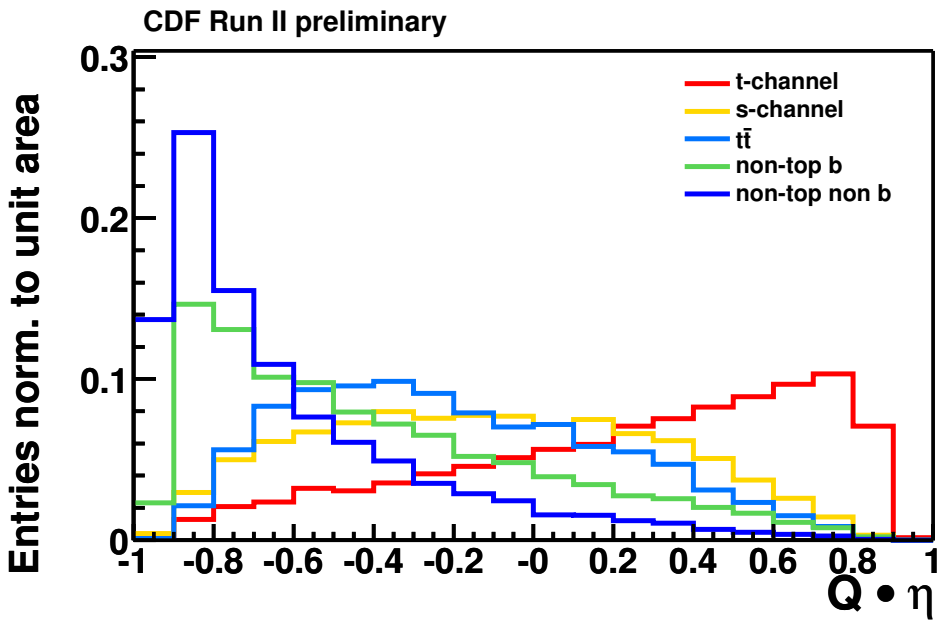
NN cut	-0.6	-0.52	-0.4	-0.3	-0.2	-0.15	-0.1	0.0	0.3	0.5	0.6
t-chan.	0.94	0.92	0.88	0.84	0.8	0.78	0.76	0.71	0.52	0.36	0.27
s-chan.	0.86	0.8	0.72	0.64	0.56	0.52	0.48	0.41	0.19	0.08	0.04
$t\bar{t}$	0.84	0.77	0.65	0.55	0.46	0.42	0.38	0.31	0.13	0.05	0.03
WW	0.33	0.26	0.19	0.13	0.10	0.08	0.06	0.04	0.01	0.0	0.0
WZ	0.56	0.48	0.37	0.29	0.23	0.22	0.18	0.14	0.06	0.02	0.01
ZZ	0.57	0.46	0.38	0.31	0.26	0.2	0.17	0.14	0.051	0.03	0.01
$Wb\bar{b}$	0.6	0.53	0.43	0.36	0.29	0.26	0.24	0.19	0.09	0.04	0.02
$Wc\bar{c}$	0.34	0.29	0.21	0.17	0.14	0.12	0.11	0.08	0.03	0.01	0.0
Wc	0.43	0.36	0.28	0.23	0.19	0.18	0.15	0.12	0.06	0.02	0.01
mistags	0.47	0.41	0.31	0.22	0.2	0.19	0.17	0.12	0.07	0.05	0.03
nonW	0.23	0.18	0.12	0.09	0.06	0.05	0.04	0.03	0.01	0.0	0.0

A peek at Data

Expected event numbers in $162\ \mathrm{pb}^{-1}$

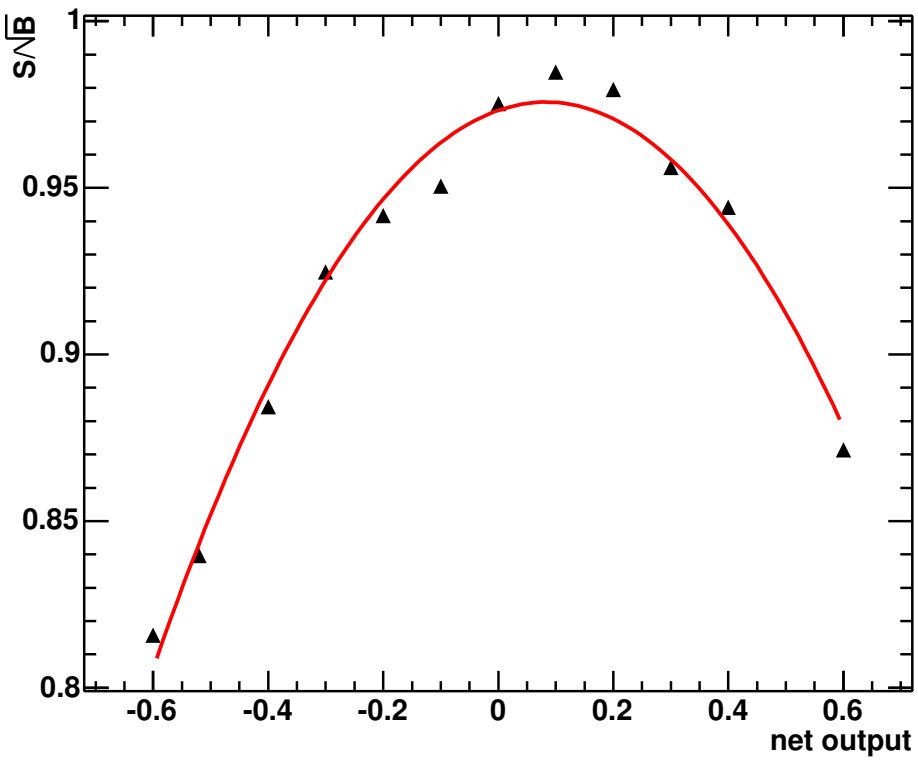
t-channel	3.37
s-channel	2.29
$t\bar{t}$	8.03
di-boson	2.25
W$b\bar{b}$	23.68
W$c\bar{c}$	8.45
Wc	8.45
mistags	18.45
non-W	10.28

Templates



Significance: S/$\sqrt{\mathrm{B}}$ Projections

S / $\sqrt{\mathrm{B}}$ versus cut on netoutput



Assuming SM cross sections

Cut scenarios for 162 pb^{-1}

Maximum: S/$\sqrt{\mathrm{B}} = 0.98$

Simple $M_{\ell\nu b}$ cut: S/$\sqrt{\mathrm{B}} = 0.74$

Improvement by +32% !

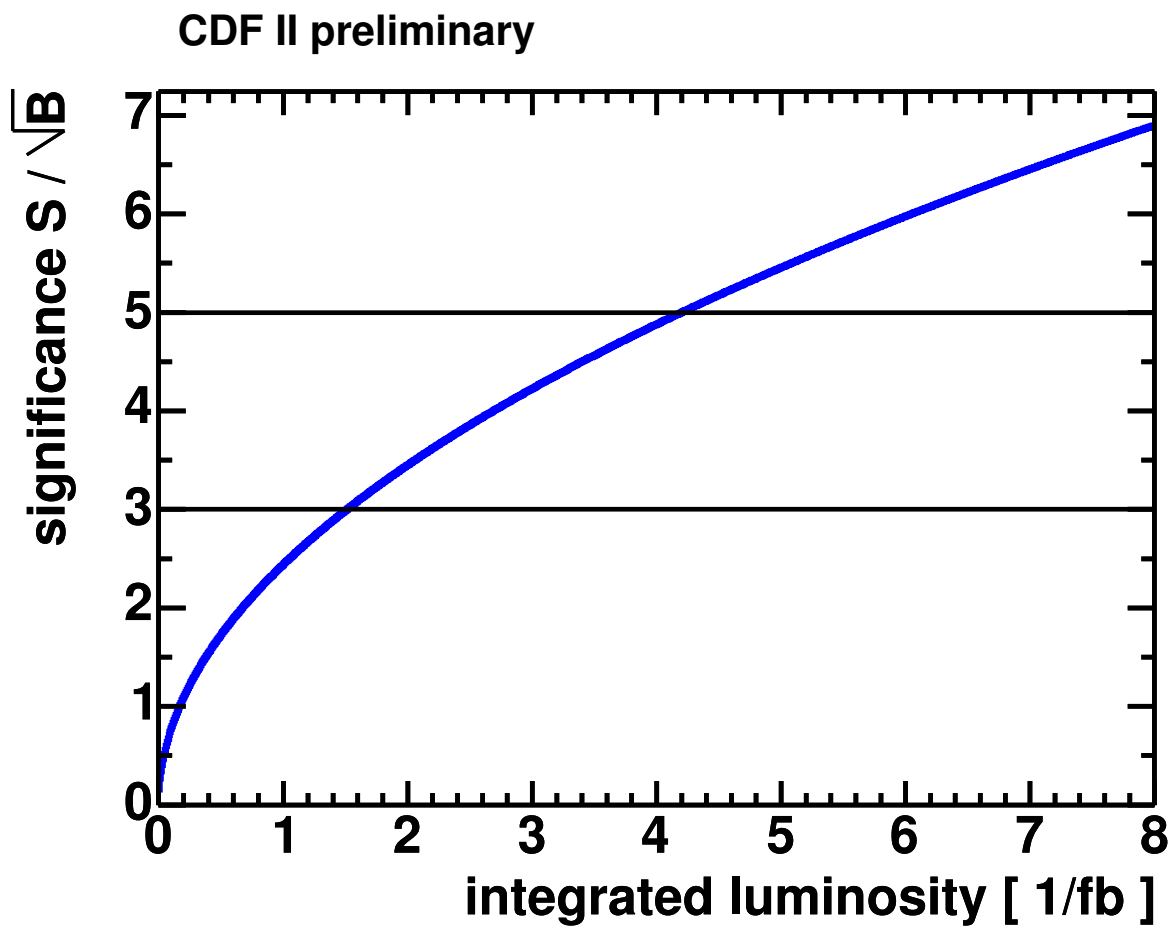
Scale results for 162 pb^{-1} to higher integrated luminosities.

Projections:

With 1 fb^{-1} obtain $\mathbf{S/\sqrt{B} = 2.4}$

Reach S/$\sqrt{\mathrm{B}} = 3.0$ for $\mathbf{1.5\ fb^{-1}}$ at $\mathrm{N_{sig}} = 27.3$ events

Projection Plot



Iterated Disciminant Analysis

Thorsten Walter

Transparent analytical method to combine multiple variables to a test statistic.

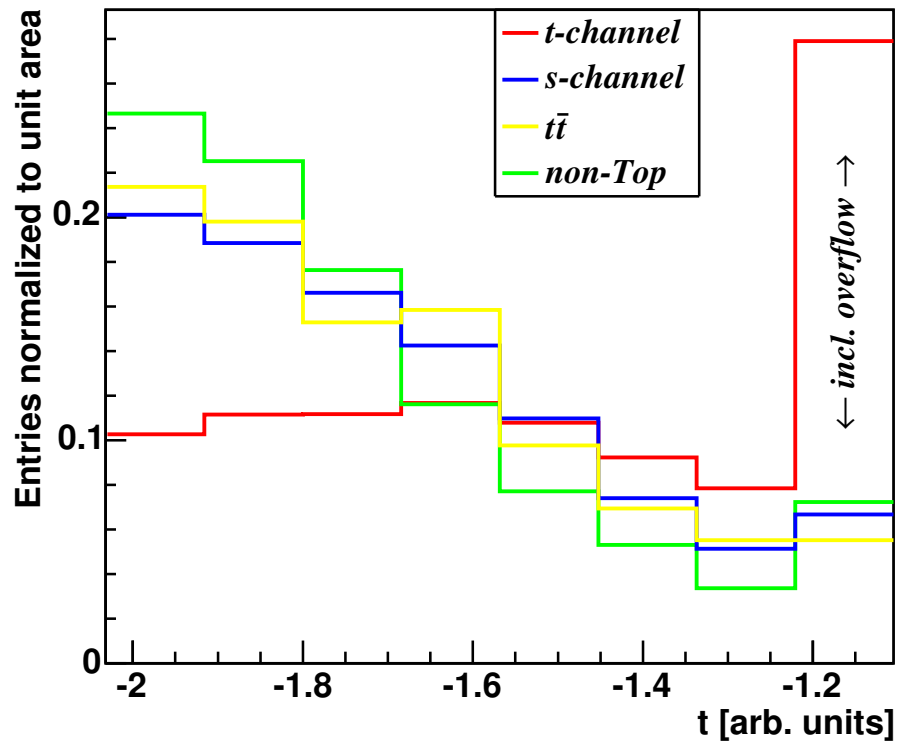
2nd order generalization of Fisher discriminant.

Search optimized for t-channel.

Ph.D. thesis soon available as CDF note.

Study based on 4.11.1 analysis.

Normalized template distributions

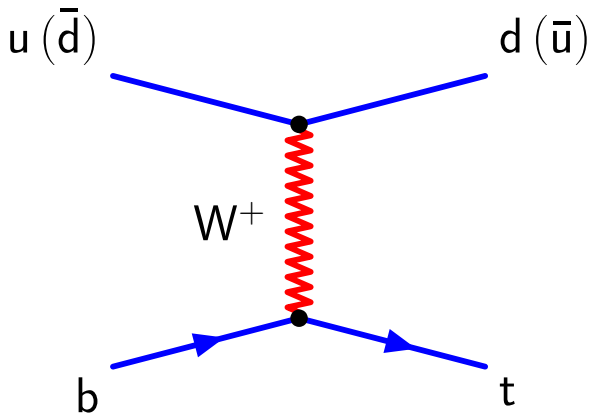


If t- and s-channel are combined:
3 σ evidence with 1.8 fb^{-1}.
With 1 fb^{-1} expect 1.7 σ effect.

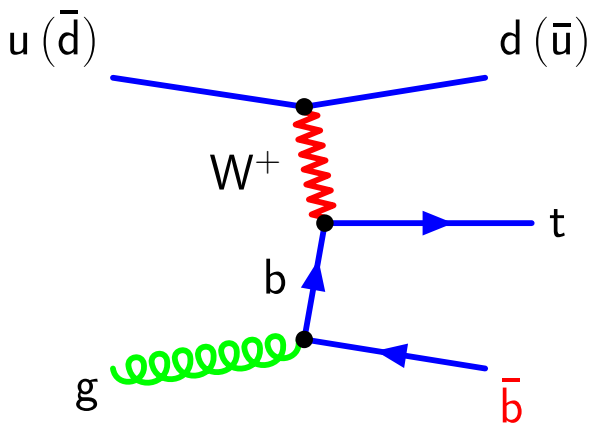
Summary of Projections

- Projections for 3 σ evidence
  t-channel plus s-channel is signal
  - Neural Network: $1.5\ \mathrm{fb}^{-1}$
  - Likelihood Ratio: $1.7\ \mathrm{fb}^{-1}$
  - Iterated Discriminat Analysis: $1.8\ \mathrm{fb}^{-1}$
- Higgs Workshop
  - Present comprehensive summary of projections for other techniques.
  - Update studies to 5.3.3.

Monte Carlo modelling t-channel single top



$2 \to 2$ (LO)



$2 \to 3$ (NLO)

- Leading order process uses b quark structure function.
  $\bar{b}$ quark obtained from backward evolution in parton shower MC (PYTHIA).
  $\Rightarrow p_T(\bar{b})$ too soft, $\eta(\bar{b})$ too far forward.
- Phenomenological solution:
  Generate $2 \to 2$ and $2 \to 3$ events with matrix element MC (MadEvent, TopRex).
  Match samples to improve modelling of $p_T(\bar{b})$ distribution.
  $\Rightarrow$ Matching done for PRD 71, 012005.
- Question: How well do the matched samples agree with NLO differential cross sections?

Validation of matched t-channel sample with ZTOP

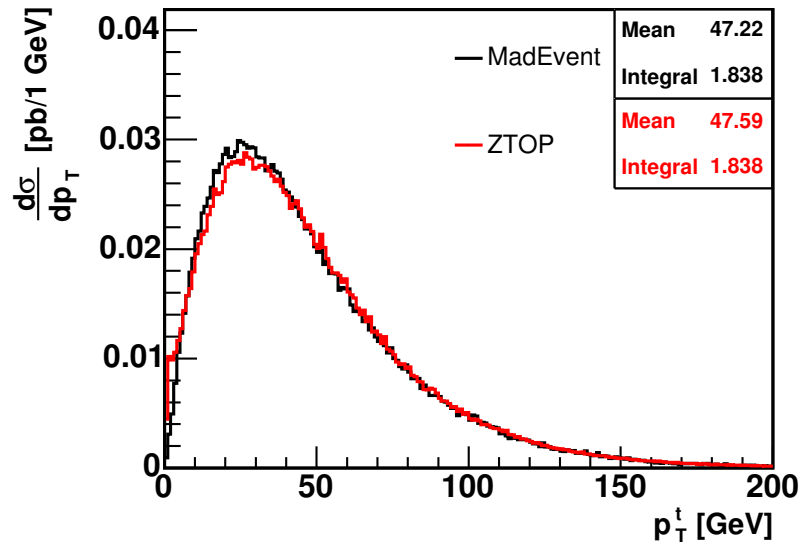
Jan Lück

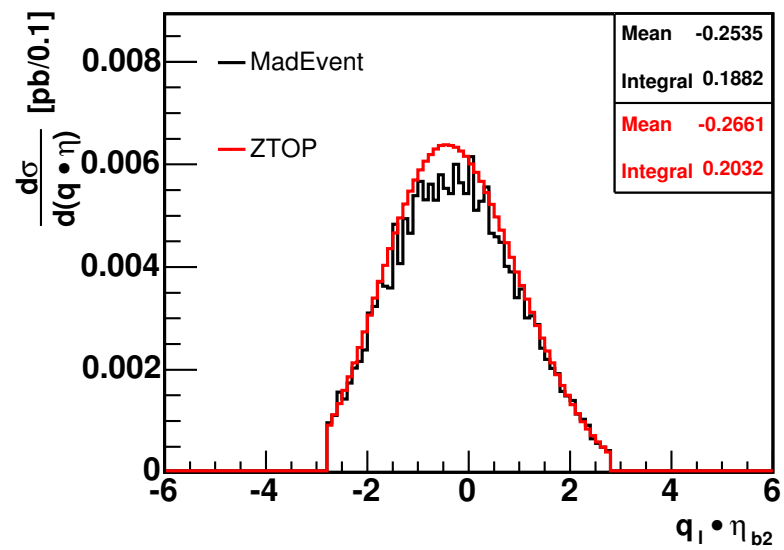
ZTOP: Program by Z. Sullivan
Compare p_T and η of top, light quark jet and b jet.

Big picture:
– Good agreement of rates and shapes.
– Matching solves modelling of 2nd b quark.

Second look:
Improvements are possible, but don't need them now.

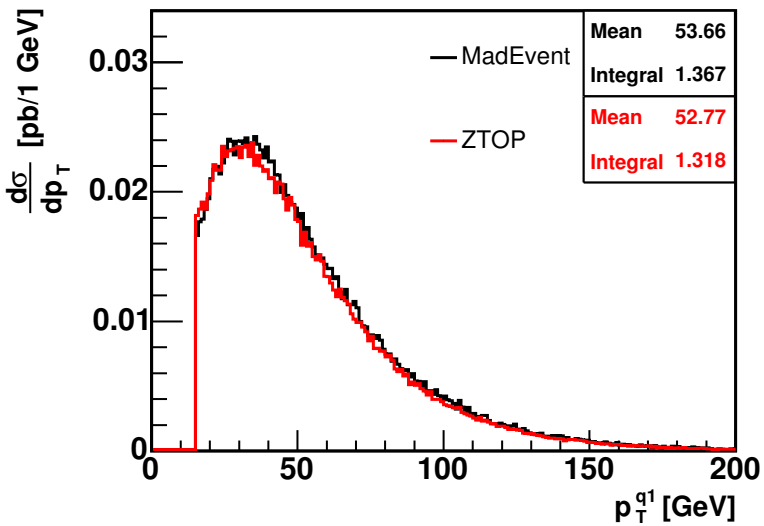
Plan: Use comparison to derive systematic error.
Documentation: CDF 7701



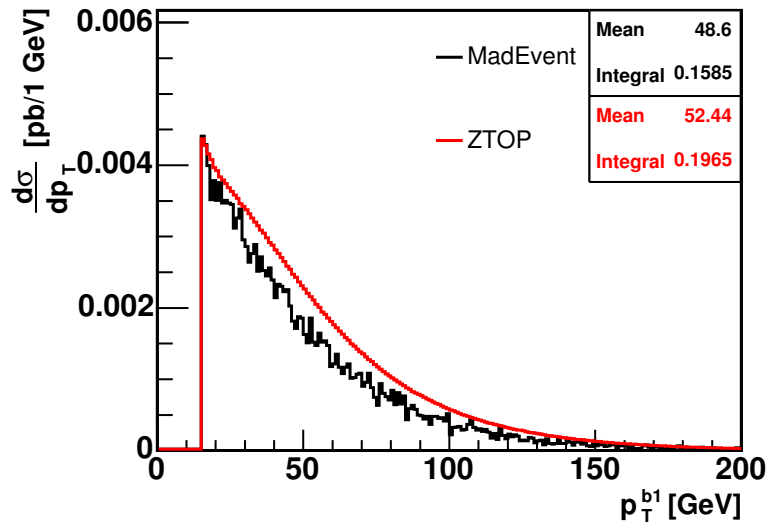


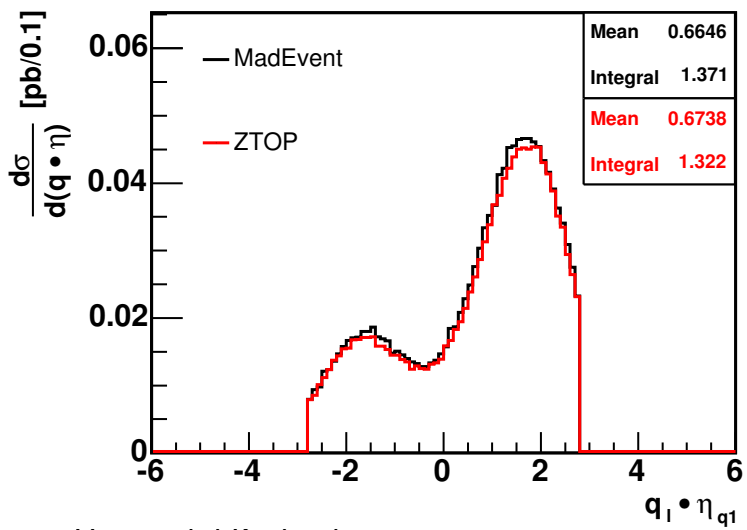
Leading jet distributions

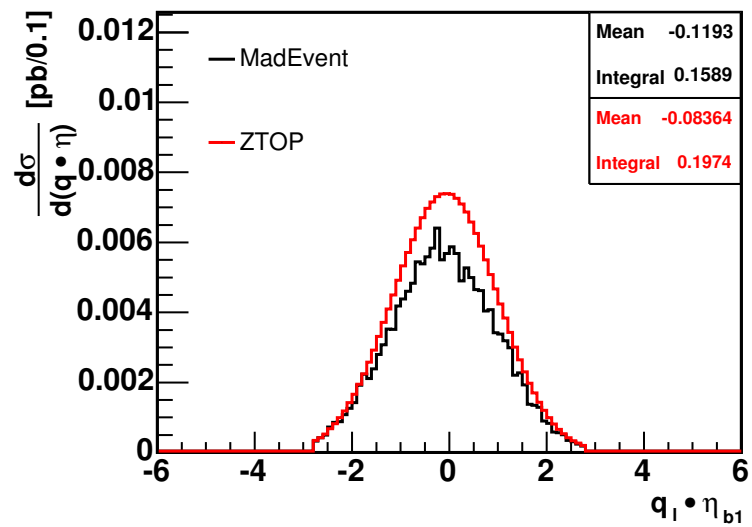
light quark jet



2nd b quark jet



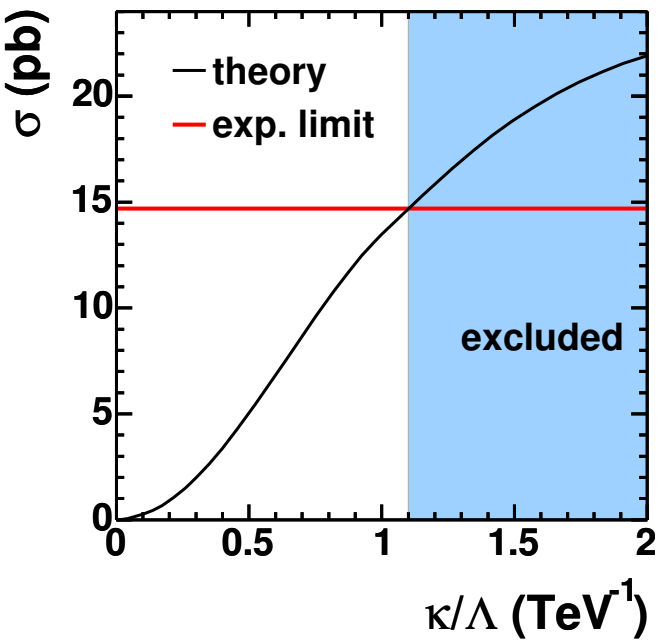




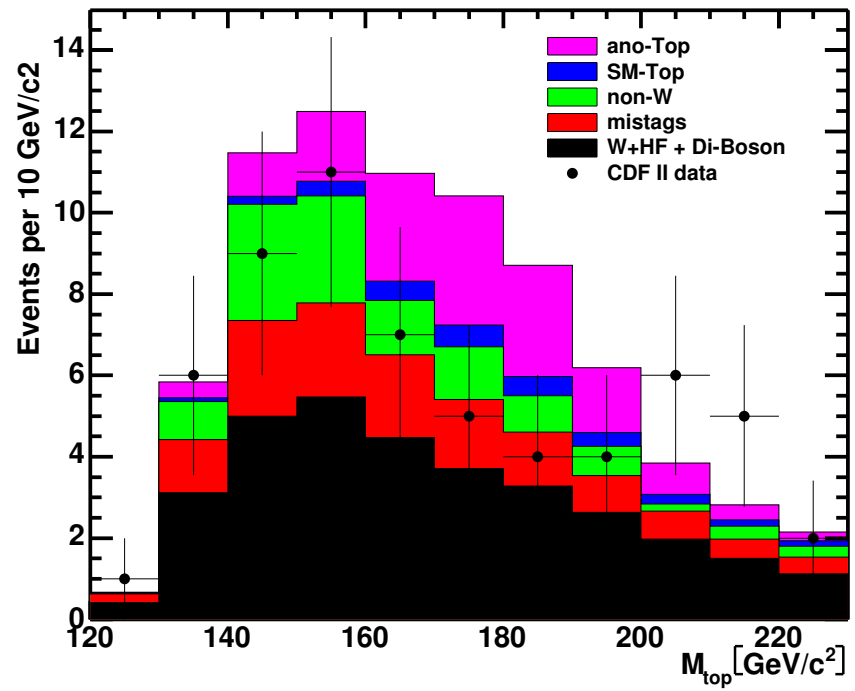
Anomalous Top Production via FCNC

Adonis Papaikonomou

Physics beyond the SM:
Several models predict large FCNC, e.g. ug $\rightarrow$ t.
New idea!
Aim: Give first limits on κ_{tug}.



Look in W+1 jet bin with 162 pb^{-1}.
Exp. Limit: $\sigma < 14.7$ pb



CDF Note: 7686
Now working on TopRex samples.