Single-Top Search using Neural Networks with CDF

Important aspects for the LHC

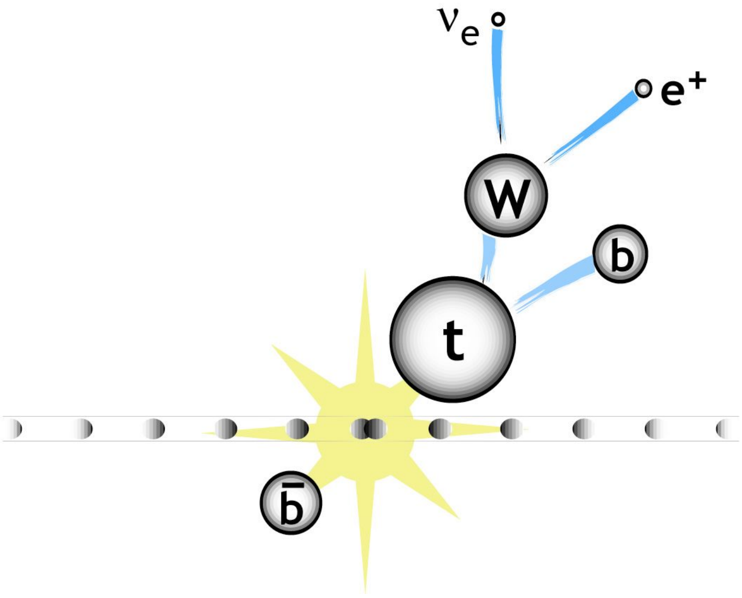
D. Hirschbühl, J. Lück, T. Müller, A. Papaikonomou,
S. Richter, W. Wagner

University of Karlsruhe

LHC D Top-Physik-Workshop

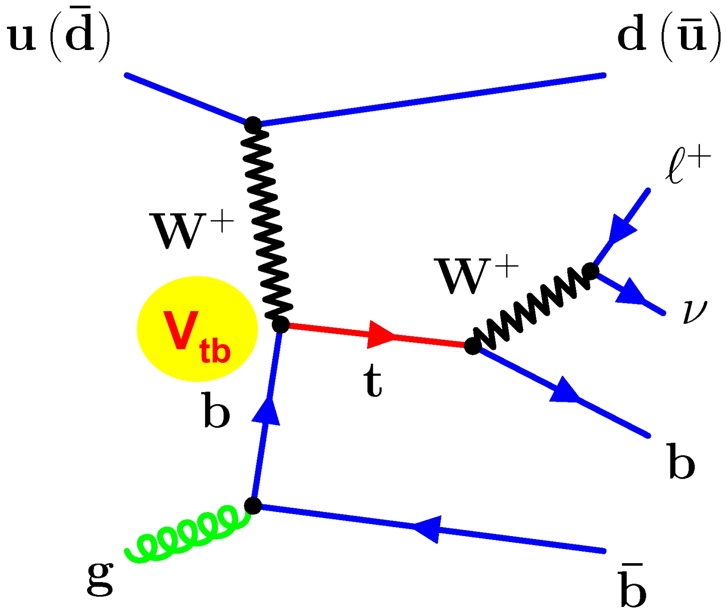
Bad Honnef, January 26, 2007

- Single-Top Monte Carlo Samples
- Data Based Backgrounds
- Neural Network b Tagger
- Measurement Results

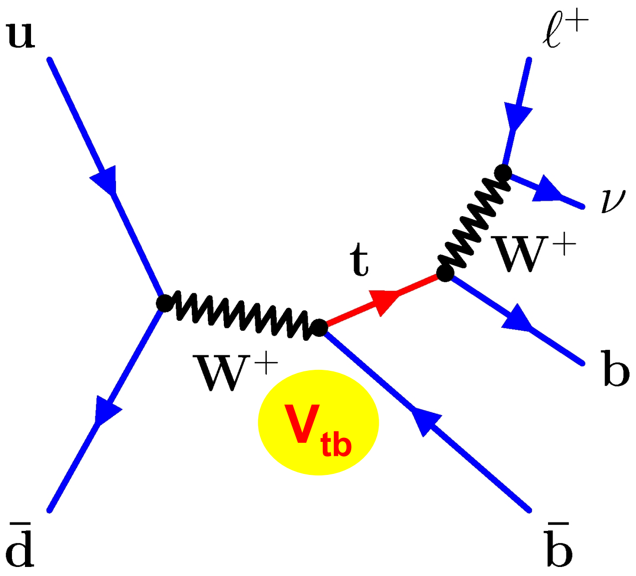


Single-Top Quark Production

top quark production via the weak interaction



t-channel



s-channel

theoretical cross section predictions at $\sqrt{s}$ = 1.96 TeV

1.98 ± 0.25 pb **0.88 ± 0.11 pb**

B.W. Harris et al. Phys. Rev. D 66:054024 (2002)

compatible results: Campbell/Ellis/Tramontano, Phys. Rev. D 70:094012 (2004)

t-channel MC Samples

Problem with leading order process: bq → tq’

Bbar quark from the gluon splitting, the so called 2nd b quark, is not modelled well

→ η distribution is too much forward

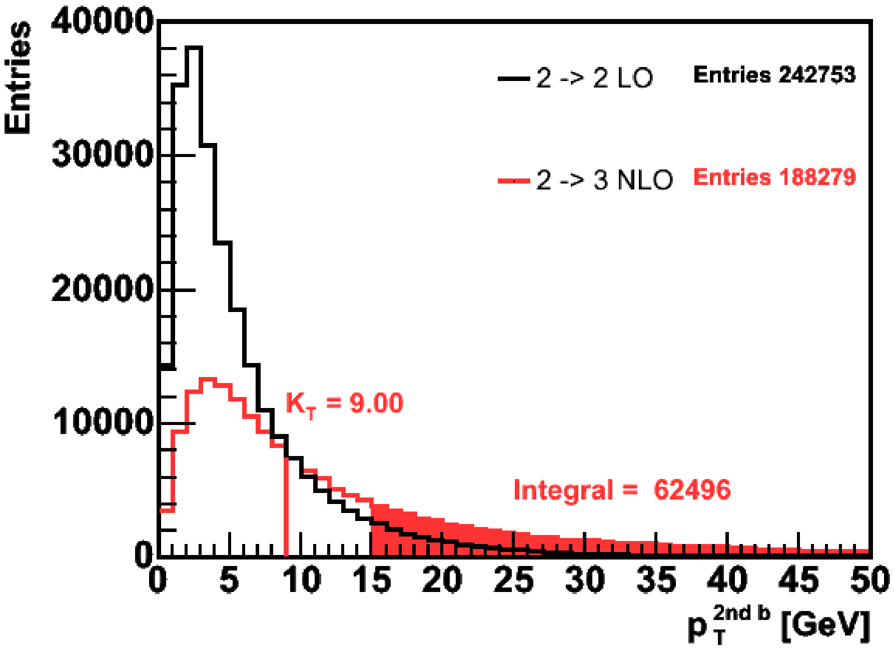
→ pT distribution is too soft

Solution: Use matrix element Monte Carlo. At CDF we use MadEvent.

The t-channel is modeled by two matrix elements:

2 → 2 (LO) : bq → tq’ and 2 → 3 (NLO) : gq → tbq’

In an iterative process a joint t-channel sample is created by matching the cross section for a detectable 2nd b quark ($|\eta| < 2.8$ and $E_T > 15$ GeV) in the MadEvent MC to ZTOP (Z. Sullivan).



Universität Karlsruhe (TH)
Forschungsuniversität • gegründet 1825

Background Estimation

Background estimation and modeling is the most critical part of the analysis. Takes about 80% of the effort. But, often it is not talked about much.

Monte Carlo based background estimate

- top-antitop production
- diboson (WW, WZ, ZZ)
- Z+bBbar

rates are predicted using theoretical cross section and MC acceptances

$$N^{\mathrm{pred}} = \sigma^{\mathrm{theo}} \epsilon_{\mathrm{evt}} \int \mathcal{L}\, dt$$

$$\epsilon_{\mathrm{evt}} = \epsilon_{\mathrm{evt}}^{\mathrm{MC}} \cdot \epsilon_{\mathrm{BR}} \cdot \epsilon_{\mathrm{corr}} \cdot \epsilon_{\mathrm{trigger}}$$

6.6% of the background is MC based

Data based background estimate

W + heavy flavor

fraction: 51.6%
heavy flavor fractions from inclusive jet samples

W + mistagged light quark jets

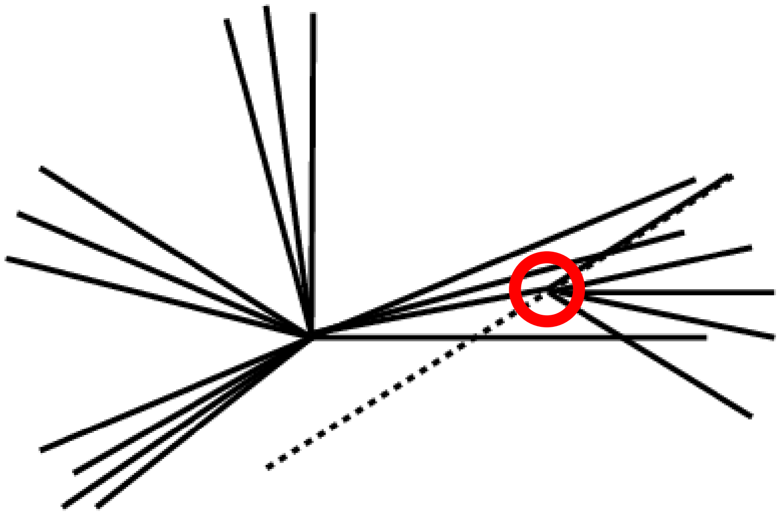
fraction: 23.2%
mistag matrix

QCD multijet (non-W)

fraction: 4.5%
fits to the missing E_T distribution

Mistag Estimate

tagging rates are determined in inclusive jet samples (high statistics control samples)



look at signed 2D decay length

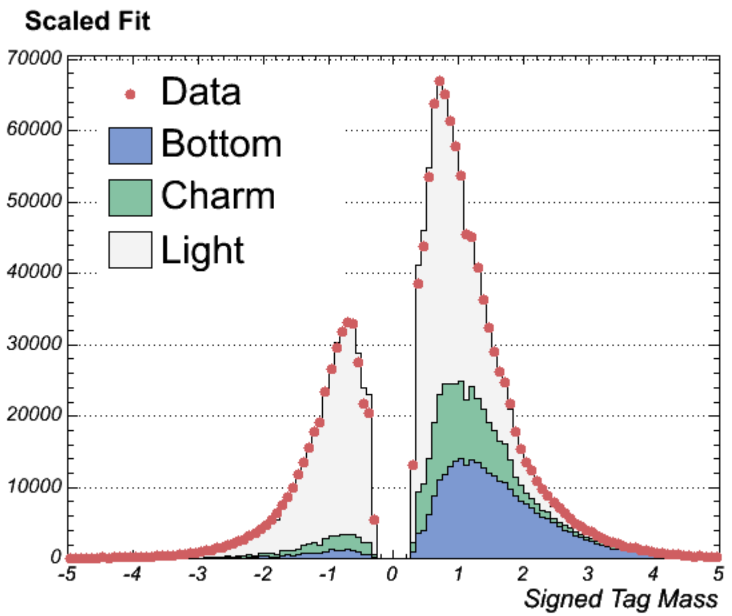
negative tag rates give an estimate on positive mistags due to resolution effects

tag rates are parametrized as a function of 6 variables: jet η, jet E_T, Σ E_T, N_{track} of N_{vertex}, Z_0 of the primary vertex

corrections to positive tag rates:

1.) enhancement of mistags in heavy flavor events

2.) account for long lived particles: K_S, K_L, Λ



non-W Estimation

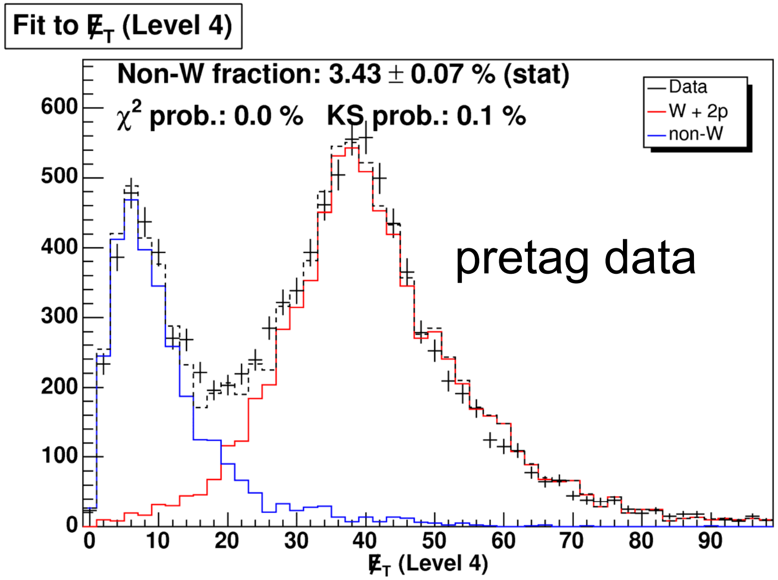
traditional method:
fit of side-band regions of the MET $\otimes$ isolation plane

Problem: estimates are too low
MET and $M_T(W)$ distributions are seriously mismodeled

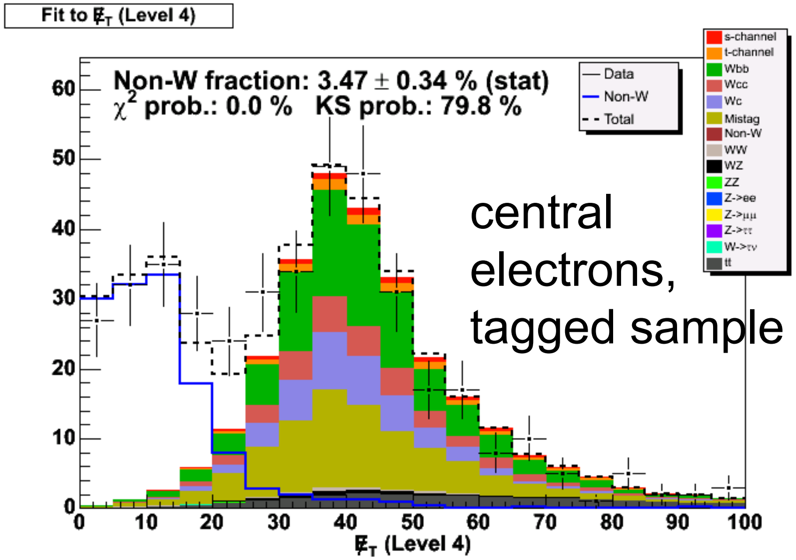
→Tighter cuts: MET > 25 GeV

→ Estimates of the pretag and tagged non-W rates using fits to the MET distribution

→ Still large uncertainties: ±43%



pretag data



central electrons, tagged sample

Result of Background Estimation

CDF II Preliminary 955 pb^{-1}

W + 2 jets	
$Wb\bar{b}$, $Wc\bar{c}$, Wc	303.0 ± 89.6
Mistags	136.1 ± 19.7
QCD multijet	26.2 ± 15.9
$t\bar{t}$	58.4 ± 13.5
Diboson, Z + jets	25.6 ± 6.3
Total Background	549.3 ± 95.2
t-channel	22.4 ± 3.6
s-channel	15.4 ± 2.2
Total Prediction	587.1 ± 96.6
Observation	644

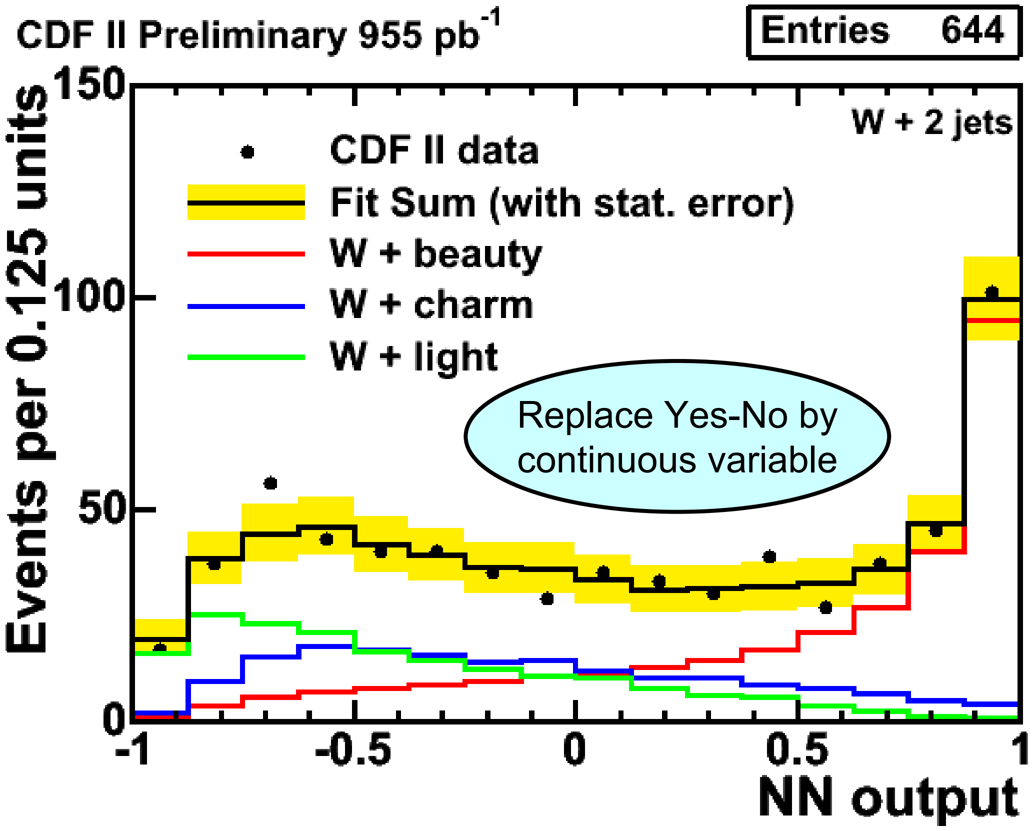
Improved *b* Jet Identification

About 50% of the background in the W+2 jets sample do NOT contain b quarks even though a secondary vertex was required!

Combine jet and track variables to one powerful discriminant using a neural network.
e.g. vertex mass, decay length, track multiplicity, …

New possibility:
In situ measurement of the flavor composition in the W + 2 jets sample

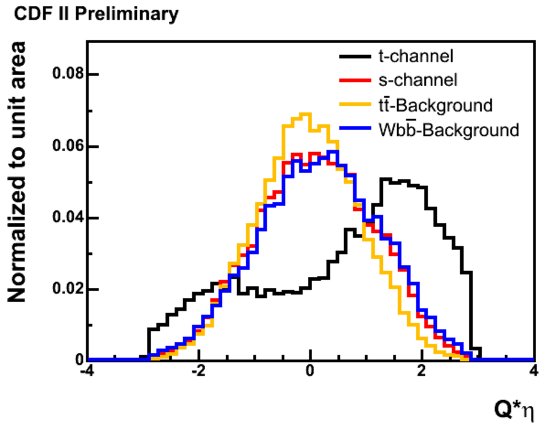
Fit to NN output for W + 2 jets events with one secondary vertex (955 pb^{-1})

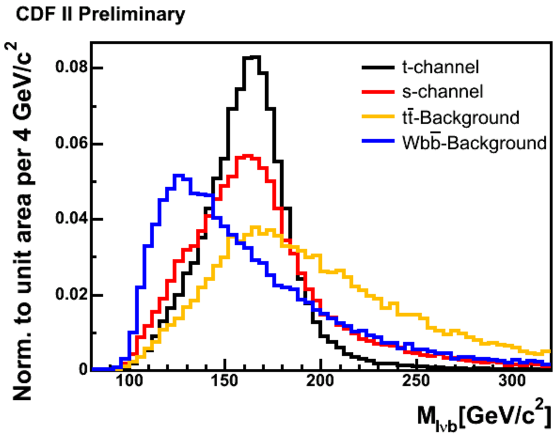


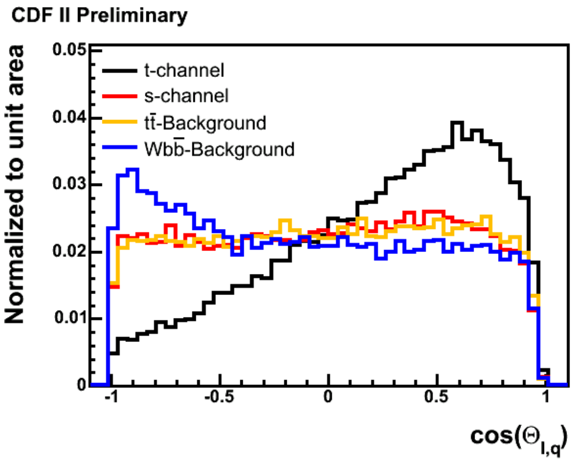
Replace Yes-No by continuous variable

charm-like / **mistags** ……………… **b-like**

Single-Top Neural Network Analysis



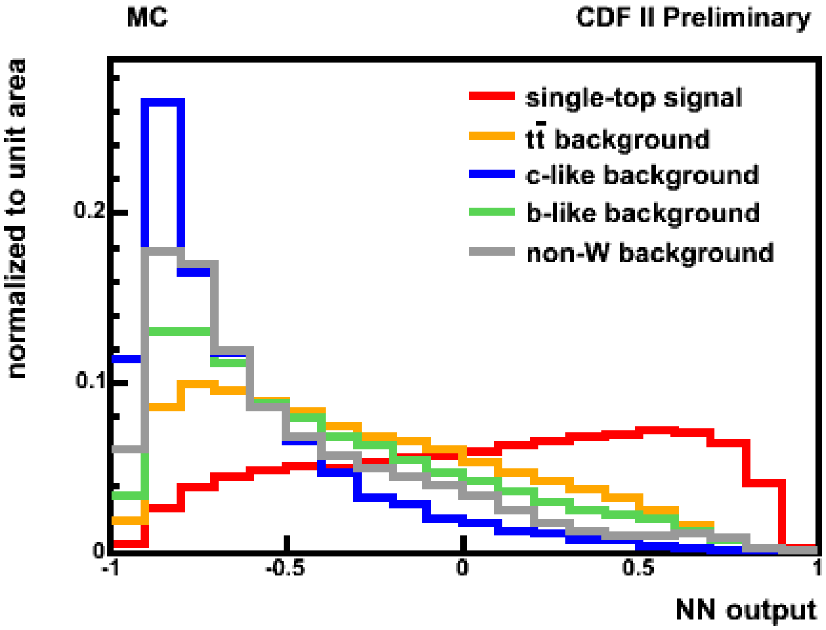




…

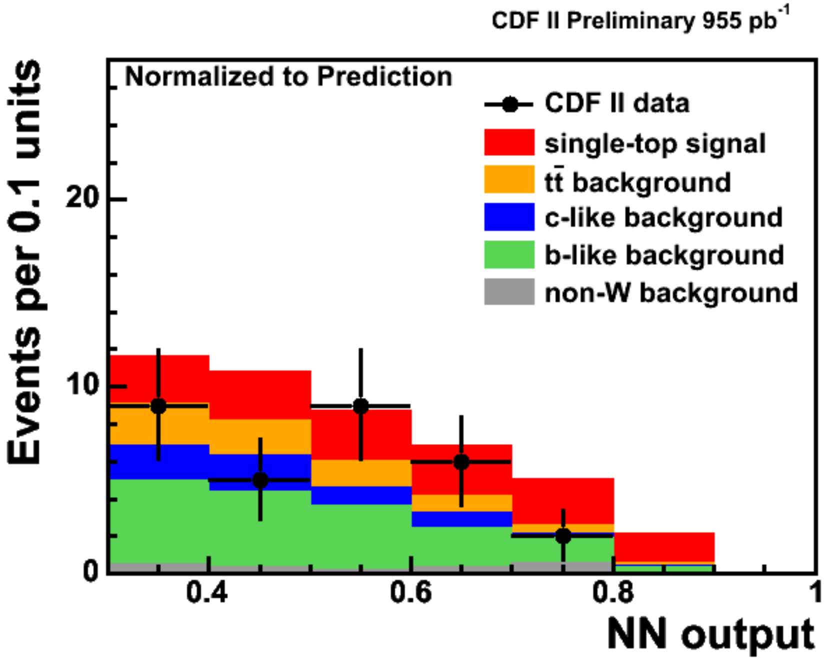
Idea:
combine many variables to one good discriminant

18 variables are used, among them $Q \cdot \eta$, reconstructed top quark mass, top quark polarisation angle, Jet E_T and η, NN b tagger output, W boson η, …

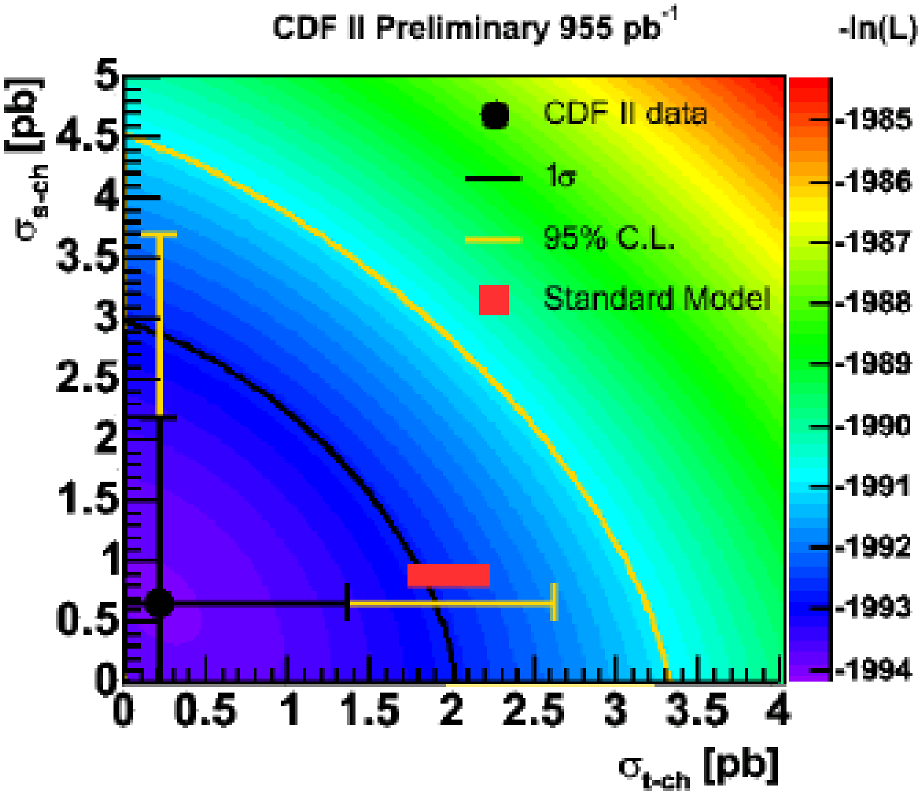


Fit Result

Combined Search



Separate Search



$\sigma_{Fit} = 0.0\ ^{+1.2}_{-0.0}$ (stat. + syst.) pb

$\sigma_{SM} = 2.9 \pm 0.4$ pb (prediction)

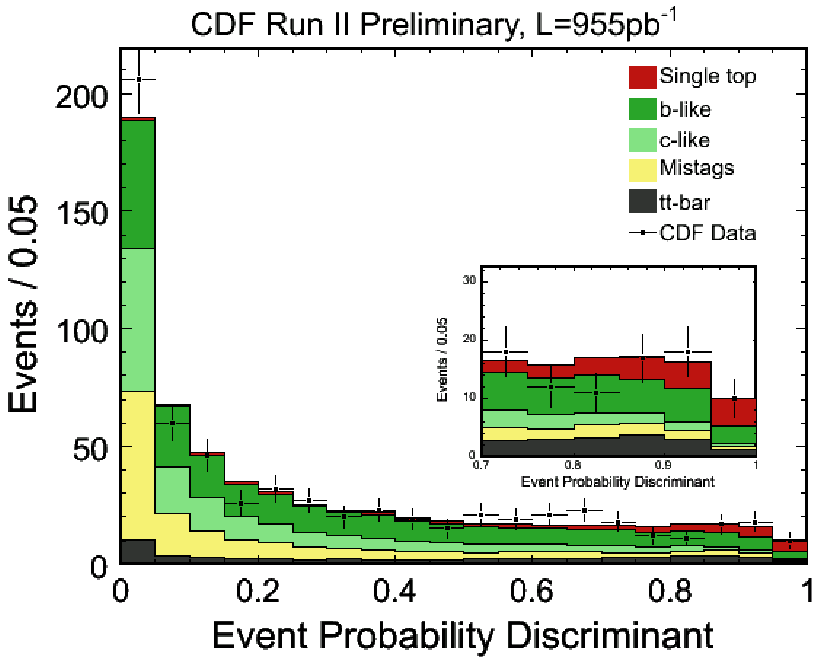
σ (t-Kanal) $= 0.2\ ^{+1.1}_{-0.2}$ pb (SM: 1.98 pb)

σ (s-Kanal) $= 0.7\ ^{+1.5}_{-0.7}$ pb (SM: 0.88 pb)

Universität Karlsruhe (TH)
Forschungsuniversität • gegründet 1825

Two other CDF Analyses

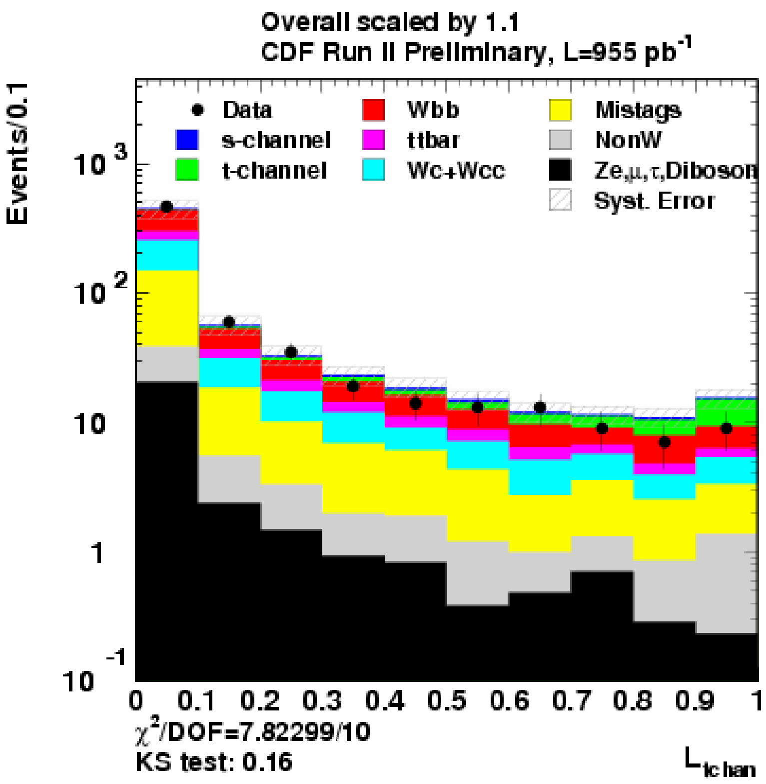
Matrix Element Method



$$\sigma_{Single\ \ Top} = 2.7^{+1.5}_{-1.3}\, pb$$

observed p-value: 1.0% (2.3 σ)

Likelihood Functions



$$\sigma(\mathrm{t-channel}) = 0.2^{+0.9}_{-0.2}\ \mathrm{pb}$$

$$\sigma(\mathrm{s-channel}) = 0.1^{+0.7}_{-0.1}\ \mathrm{pb}$$

CDF Summary

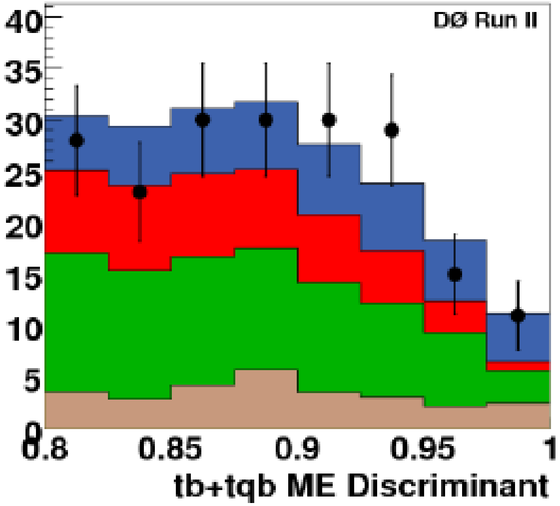
Method	Neural Networks		Matrix Elements	Likelihood Function	
	1D	2D	1D	1D	2D
Expected p-value	0.5%	0.4%	0.6%	2.3%	2.5%
Observed p-value	54.6%	21.9%	1.0%	51.1%	58.5%

consistency of all 6 discriminants based on common pseudo-experiments: 0.6%

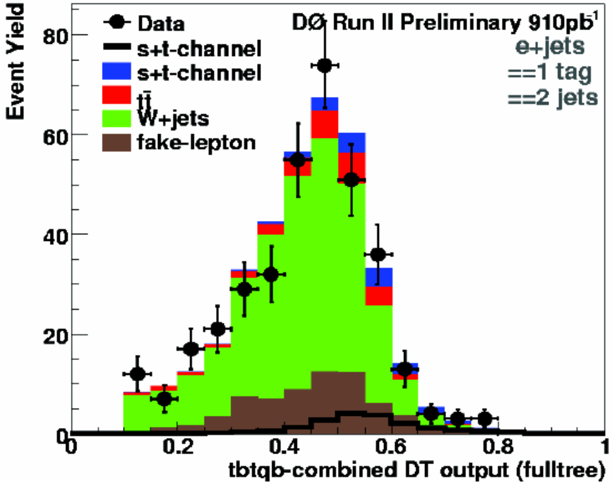
DØ Results

DØ has three analyses: all of them obeserve a signal

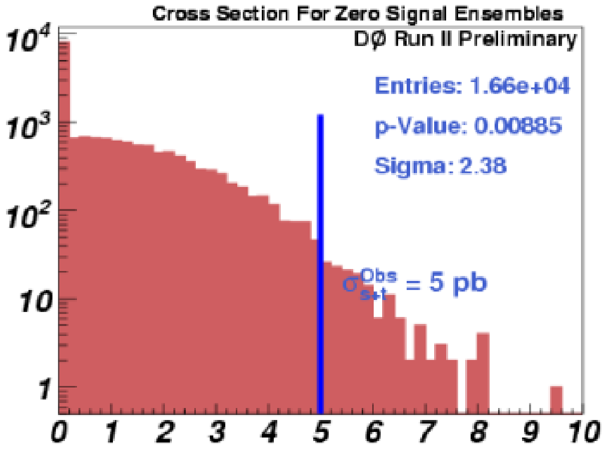
Matrix Element Method



Boosted Decision Trees



Neural Networks



expected p-values

Matrix Element Method	Boosted Decision Trees	Neural Networks
3.7%	1.9%	9.7%

significance of the observed signal

Matrix Element Method	Boosted Decision Trees	Neural Networks
2.9 σ	3.4 σ	2.4 σ

Conclusions

- Modeling of t-channel needs special attention: matching of bq → tq’ and gq → tbq’ Monte Carlo Samples
  Future: use MC@NLO ?
- Estimation and modeling of instrumental backgrounds (mistags and non-W) are one of the major challenges.
  LHC: will allow for different (better) control samples and in-situ measurements
- Neural Network b tagger: usable as continuous variable uses all available information
- Tevatron searches: sensitivity is close to 3 σ
  CDF analyses give inconclusive results.
  More data, that are coming in, will resolve the issues.