Search for the Standard Model Higgs Boson Produced in Association with Top Quarks at the Tevatron

J.S. Wilson

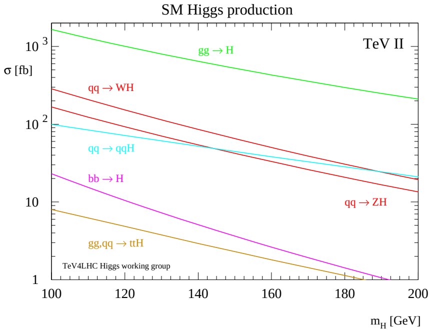
On Behalf of the CDF and DØ Collaborations

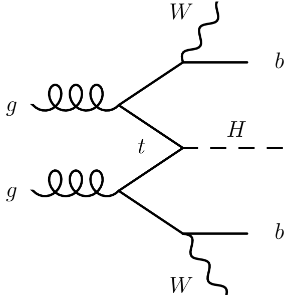
August 30, 2011

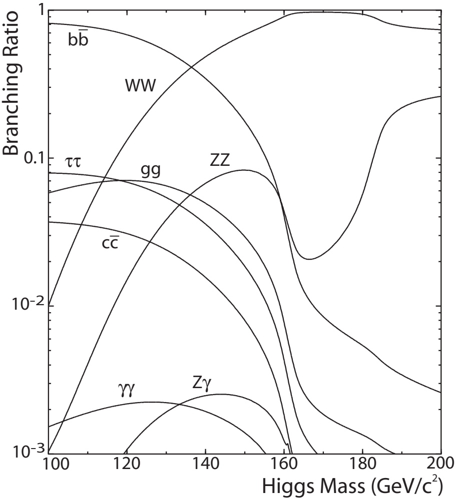
WANTED

FOR SYMMETRY BREAKING

- In the Standard Model, $t\bar{t}H$ cross section is small compared to $gg \to H$, VBF, and VH
- Top quarks are quite distinctive, giving good background discrimination
- If a new particle is discovered, measuring its coupling to top is important in establishing whether or not it is a Standard Model Higgs
- Beyond the Standard Model, many models enhance the $t\bar{t}H$ cross section





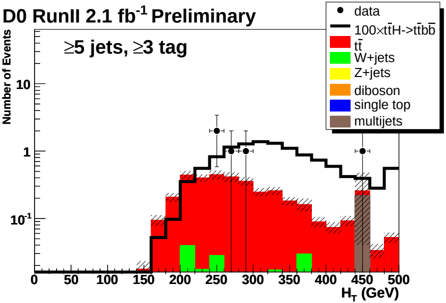


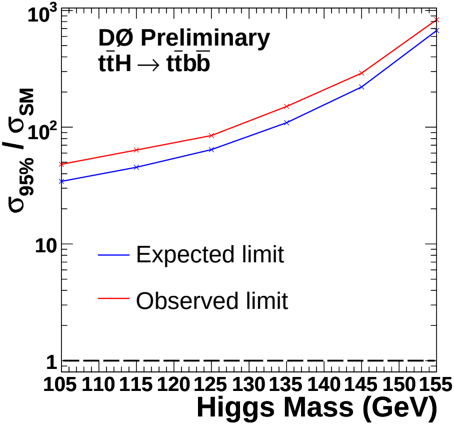
- Tops both decay to Wb, so channels determined by W decays:

$\ell +$ jets	One W decays to e or μ
$\not{E}_T +$ jets	One W decays to τ or to e/μ but lepton is missed
All jets	Both Ws decay hadronically

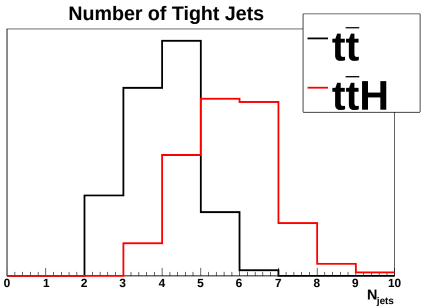
- Although b tagging is employed, we include non-$b\bar{b}$ Higgs decays.
- Additional acceptance from $H \to WW$ and other decays improves sensitivity, especially at higher masses.

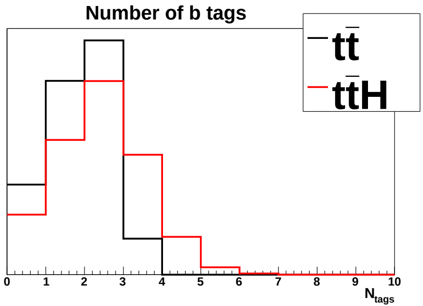
- Before summer 2011, best $t\bar{t}H$ search result was from DØ in 2009.
- Using $2.1\,\text{fb}^{-1}$, a 95 % C.L. upper limit of ~ 45 was set on $\sigma \times \mathcal{BR}/\text{SM}$ at $m_H = 120\,\text{GeV}/c^2$.
- This summer, CDF produced new $t\bar{t}H$ searches in multiple channels.
- I will concentrate on these new results in this talk.



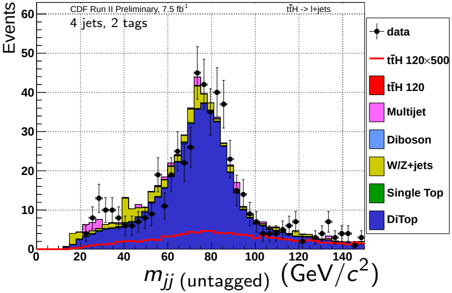


- Top decay final state $\ell\nu bbqq$
- Higgs decay final state bb, $WW/ZZ \to qqqq$, $\tau\tau$, etc.
- Select sample of events with
  - Well-reconstructed e or μ, with $p_T > 20\,\text{GeV}\,c^{-1}$
  - $\not{E}_T > 20\,\text{GeV}$
  - ≥ 4 jets:
    - $E_T > 20\,\text{GeV}$
    - $|\eta| < 2.0$
  - ≥ 2 b tags
- Further divide this into 4 regions with different S/B:
  - exactly 4 jets vs. ≥ 5 jets
  - exactly 2 b tags vs. ≥ 3 b tags

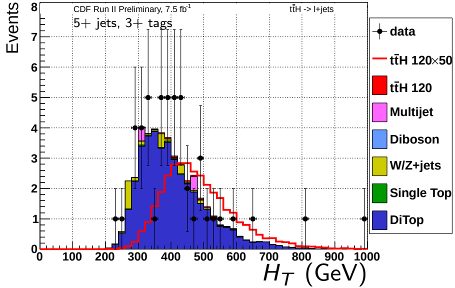


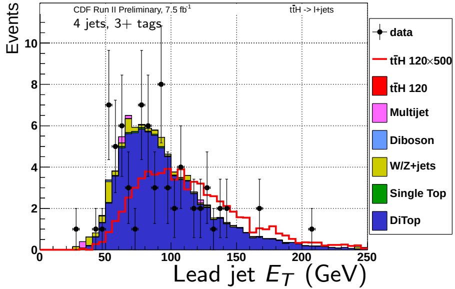


- Signal modeled using Monte Carlo, with assumed Higgs masses 100 – 170 GeV/c^2.
- Most background modeled using Monte Carlo, with $m_{\mathrm{top}} = 172.5\,\mathrm{GeV}/c^2$.
- Multijet production modeled using data driven technique.
- Tagging algorithms applied to MC jets that are matched to generator level b or c.
- Other MC jets tagged using measured light jet tag rate from data.



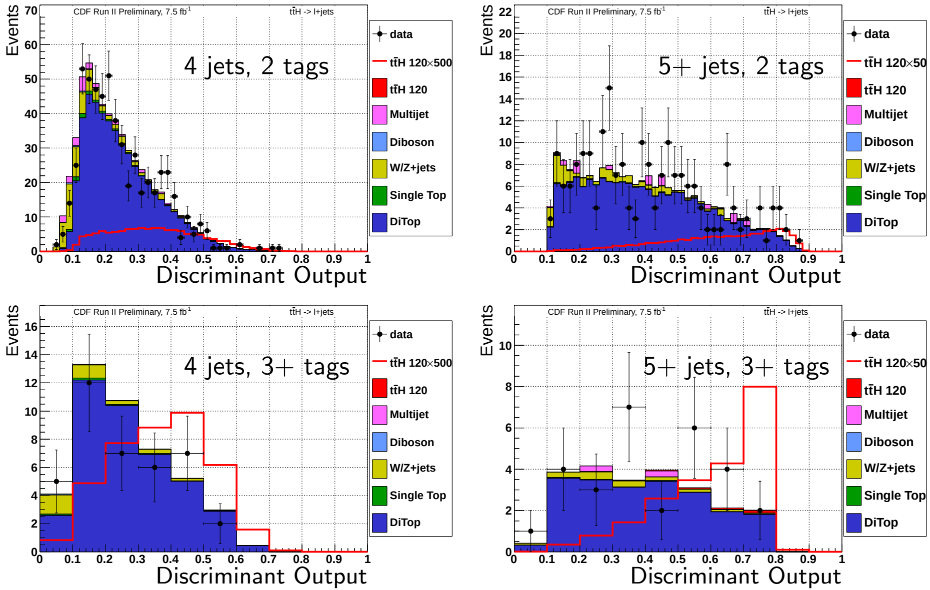
Process	≥ 4 jets, ≥ 2 b tags
$t\bar{t}$	781.42 ± 44.01
W/Z+jets	70.36 ± 11.22
Multijet	25.54 ± 13.08
Single Top	12.28 ± 0.50
Diboson	3.56 ± 0.29
Total Background	893.16 ± 47.27
$t\bar{t}H$	1.83 ± 0.05
Observed	892



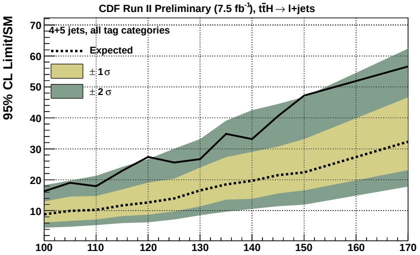


- For each Higgs mass point, train a Bayesian neural network ensemble to discriminate $t\bar{t}H$ from $t\bar{t}$
- 24 input variables, including
  - Dijet mass of leading untagged jets
  - H_T, lead jet E_T, $\not{E}_T$
  - Minimum ΔR between b jets
  - etc.

Outputs $m_H = 115\,\mathrm{GeV}/c^2$



- Since no significant excess observed, set upper bounds on cross section × branching ratio at 95 % confidence level using a standard Bayesian calculation
- Sources of systematic uncertainty:
  - *b* tag efficiency
  - Cross sections
  - Jet energy scale
  - Luminosity measurement
  - Intial- and final-state radiation
  - Mistag rate



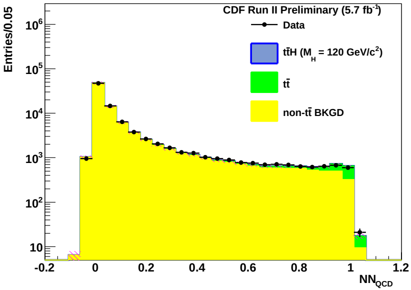
m_H	Obs	-2σ	-1σ	Exp	$+1\sigma$	$+2\sigma$
100	**16.3**	4.5	6.2	**8.9**	13.0	18.3
105	**19.0**	4.8	6.7	**10.0**	14.5	19.7
110	**18.0**	5.4	7.2	**10.3**	14.8	21.3
115	**22.9**	6.0	8.3	**11.7**	16.9	24.1
120	**27.4**	6.3	8.7	**12.7**	19.1	26.7
125	**25.6**	7.2	9.7	**14.0**	20.4	30.1
130	**26.6**	8.5	11.4	**16.6**	24.0	33.1
135	**34.9**	9.7	13.6	**18.5**	27.3	39.1
140	**33.1**	10.6	13.9	**19.7**	29.0	42.5
145	**40.6**	11.5	15.6	**21.5**	30.7	44.5
150	**47.2**	11.9	16.6	**22.4**	33.2	46.7
170	**56.6**	17.8	23.1	**32.3**	46.6	62.4

CDF Run II Preliminary, $7.5\,\mathrm{fb}^{-1}$, $t\bar{t}H \to \ell + \mathrm{jets}$

- Base selection is shared between $\not\!\!E_T +$ jets and all jets channels:
  - high-p_T e/μ veto
  - ≥ 4 jets:
    - $E_T > 15$ GeV
    - $|\eta| < 2.0$
  - leading jet $E_T > 50$ GeV
  - second jet $E_T > 40$ GeV
  - $H_T > 300$ GeV
- $\not\!\!E_T +$ jets channel additional cuts:
  - $\not\!\!E_T/\sqrt{\sum E_T} \geq 2\,\text{GeV}^{1/2}$
  - $5 \leq N_{\text{jets}} \leq 8$
  - $\text{NN}_{\text{QCD}} > 0.8$
- All jets channel additional cuts:
  - $\not\!\!E_T/\sqrt{\sum E_T} < 2\,\text{GeV}^{1/2}$
  - $7 \leq N_{\text{jets}} \leq 10$
  - $\text{NN}_{\text{QCD1}} > 0.9$, $\text{NN}_{\text{QCD2}} > 0.7$

- Base selection is shared between $\not{E}_T +$ jets and all jets channels:
  - high-p_T e/μ veto
  - ≥ 4 jets:
    - $E_T > 15$ GeV
    - $|\eta| < 2.0$
  - leading jet $E_T > 50$ GeV
  - second jet $E_T > 40$ GeV
  - $H_T > 300$ GeV
- $\not{E}_T +$ jets channel additional cuts:
  - $\not{E}_T/\sqrt{\sum E_T} \geq 2\,\text{GeV}^{1/2}$
  - $5 \leq N_{\text{jets}} \leq 8$
  - $\text{NN}_{\text{QCD}} > 0.8$
- All jets channel additional cuts:
  - $\not{E}_T/\sqrt{\sum E_T} < 2\,\text{GeV}^{1/2}$
  - $7 \leq N_{\text{jets}} \leq 10$
  - $\text{NN}_{\text{QCD1}} > 0.9$, $\text{NN}_{\text{QCD2}} > 0.7$

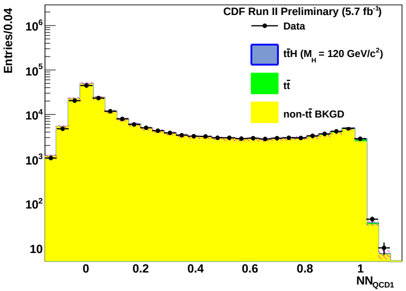
MET+Jets before correction (≥ 1tag)



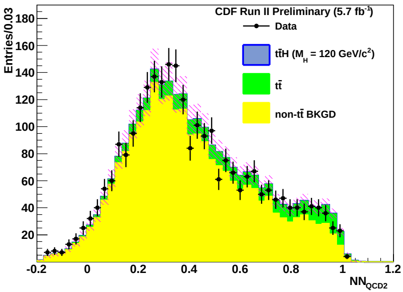
NN_{QCD} trained to discriminate $t\bar{t}H$ from multijet production

- Base selection is shared between $\not\!\!E_T +$ jets and all jets channels:
  - high-p_T e/μ veto
  - ≥ 4 jets:
    - $E_T > 15$ GeV
    - $|\eta| < 2.0$
  - leading jet $E_T > 50$ GeV
  - second jet $E_T > 40$ GeV
  - $H_T > 300$ GeV
- $\not\!\!E_T +$ jets channel additional cuts:
  - $\not\!\!E_T/\sqrt{\sum E_T} \geq 2\,\mathrm{GeV}^{1/2}$
  - $5 \leq N_{\mathrm{jets}} \leq 8$
  - $\mathrm{NN_{QCD}} > 0.8$
- All jets channel additional cuts:
  - $\not\!\!E_T/\sqrt{\sum E_T} < 2\,\mathrm{GeV}^{1/2}$
  - $7 \leq N_{\mathrm{jets}} \leq 10$
  - $\mathrm{NN_{QCD1}} > 0.9$, $\mathrm{NN_{QCD2}} > 0.7$

All jets before correction (≥ 1tag)



All jets pre-signal region (2-tag)

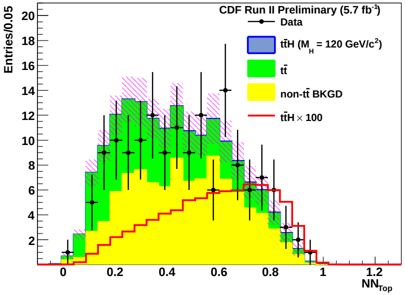


CDF II Preliminary 5.7 fb^{-1}

	2-tag ($\not\!\!E_T$ +jets)	3-tag ($\not\!\!E_T$ +jets)	2-tag (all jets)	3-tag (all jets)
Signal	1.0 ± 0.1	0.8 ± 0.1	0.7 ± 0.1	0.8 ± 0.1
$t\bar{t}$	316.1 ± 43.3	66.6 ± 9.8	120.7 ± 16.5	43.1 ± 6.4
non-$t\bar{t}$	488.9 ± 41.4	98.5 ± 12.6	328.9 ± 35.3	82.7 ± 11.8
Total Expected	806.0 ± 59.8	165.9 ± 16.0	450.3 ± 38.9	126.6 ± 13.4
Observed	756	151	424	133

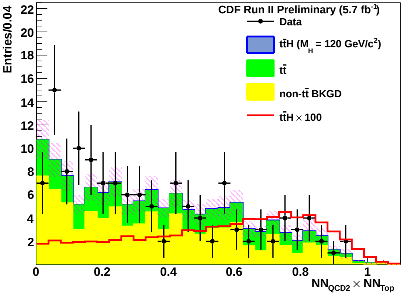
Train another neural network to discriminate $t\bar{t}$ from $t\bar{t}H$ (NN_{Top})

MET+Jets signal region (3-tag)



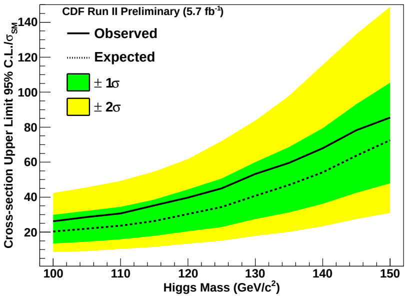
Final discriminant is $NN_{Top} \times NN_{QCD}$ ($NN_{Top} \times NN_{QCD2}$ for all jets)

All jets signal region (3-tag)

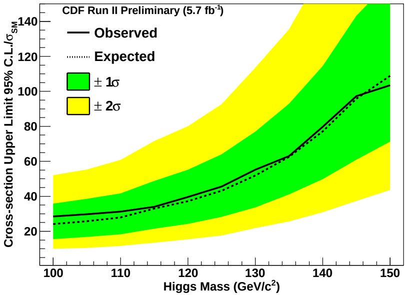


- Again, no significant excess is observed, so set 95 % C.L. upper bounds on production rate.
- Systematic uncertainties from:
    - Cross sections
    - luminosity
    - trigger modeling
    - *b* tag scale factor
    - jet energy scale
    - initial-/final-state radiation
    - PDFs
    - non-$t\bar{t}$ background model
    - *b*-tag categorization

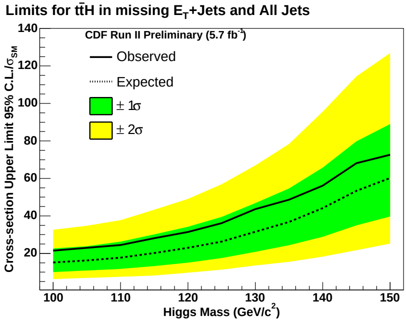
Limits for tt̄H in MET+Jets channel



Limits for tt̄H in All Jets channel



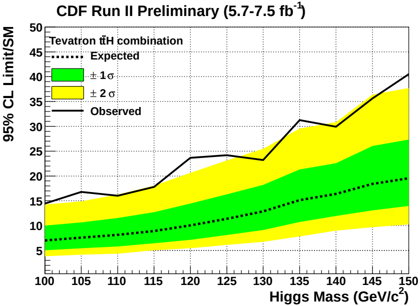
- Again, no significant excess is observed, so set 95 % C.L. upper bounds on production rate.
- Systematic uncertainties from:
  - Cross sections
  - luminosity
  - trigger modeling
  - b tag scale factor
  - jet energy scale
  - initial-/final-state radiation
  - PDFs
  - non-$t\bar{t}$ background model
  - b-tag categorization



CDF II Preliminary 5.7 fb^{-1}

M_H (GeV/c^2)	Expected Limits -2σ	-1σ	**Median**	1σ	2σ	Observed Limits
100	6.3	10.0	**15.2**	22.6	32.6	**21.5**
105	6.9	10.8	**16.3**	23.8	34.8	**23.0**
110	7.6	11.7	**17.8**	26.3	37.7	**24.5**
115	8.2	13.2	**20.2**	30.1	43.3	**28.1**
120	9.7	15.1	**22.9**	34.2	49.1	**31.4**
125	11.3	17.5	**26.2**	39.5	56.9	**36.2**
130	13.5	20.8	**31.5**	46.8	66.9	**43.6**
135	15.4	24.3	**36.8**	54.6	78.4	**48.6**
140	18.3	28.9	**44.2**	65.8	95.6	**56.2**
145	21.7	35.0	**53.4**	79.8	114.5	**68.1**
150	25.2	39.7	**60.3**	89.9	126.9	**72.6**

- We also produce 95 % C.L. upper limits on $t\bar{t}H$ production in combined ℓ + jets, $\not{E}_T$ + jets, and all jets channels.
- This combination makes $t\bar{t}H$ the fourth most sensitive low-mass channel at CDF, behind $WH \rightarrow \ell\nu b\bar{b}$, $VH \rightarrow \not{E}_T b\bar{b}$, and $ZH \rightarrow \ell\ell b\bar{b}$.



m_H	Obs	-2σ	-1σ	Exp	$+1\sigma$	$+2\sigma$
100	**14.4**	3.8	5.0	**7.0**	10.0	14.2
105	**16.8**	4.1	5.5	**7.6**	10.7	14.9
110	**16.0**	4.4	5.8	**8.1**	11.6	16.3
115	**17.8**	5.0	6.4	**8.9**	12.7	18.0
120	**23.7**	5.4	7.1	**10.0**	14.4	20.6
125	**24.2**	6.1	8.1	**11.4**	16.3	23.1
130	**23.2**	6.7	9.1	**12.9**	18.2	25.5
135	**31.2**	7.8	10.7	**15.2**	21.3	29.6
140	**29.9**	9.0	12.0	**16.4**	22.6	30.8
145	**35.6**	9.7	13.1	**18.4**	26.0	36.4
150	**40.5**	10.2	14.0	**19.5**	27.3	37.7

CDF Run II Preliminary, 5.7–7.5 fb^{-1}, $t\bar{t}H$ combination

- We have searched for $t\bar{t}H$ at CDF in three channels: $\ell +$ jets, $\not{E}_T +$ jets, and all jets.
- We set 95 % C.L. upper limits on the $\sigma(t\bar{t}H) \times \mathcal{BR}/\mathrm{SM}$ using the combination of all three channels.
- At $m_H = 120\,\mathrm{GeV}/c^2$, we obtain:
  - Expected limit: 10.4,
  - Observed limit: 22.2.
- These results were included for the first time in the Tevatron Higgs search combination presented at EPS 2011.
- $t\bar{t}H$ is a very sensitive Higgs search channel with potential for observation of BSM physics.
- Observation of $t\bar{t}H$ will be important post discovery at the LHC to verify that the Higgs is the Higgs.

- $t\bar{t}$ and $t\bar{t}H$ modeled using same Monte Carlo as $\ell +$ jets
- Non-$t\bar{t}$ background (mainly multijet) model is data driven
- Need to model multijet rate and shape in 2 and ≥ 3 b tags regions ($b$ness tagger)
- Select 3 (4) jet events in $\not{E}_T +$ jets (all jets) channel
- This sample is uncontaminated by $t\bar{t}$ and $t\bar{t}H$
- Find per-jet b-tag rate, parameterized in three variables
    - jet E_T, N_{tracks} in jet cone
    - $\not{E}_T +$ jets: $\not{E}_T \cos\Delta\phi(\not{E}_T, \text{jet})$
    - all jets: N_{vertices}
- b-tag rate applied to pre-tag data in signal region
- $t\bar{t}$ contribution is iteratively removed from multijet model
- Reweight in N_{jets} using control region $0.05 < \text{NN}_{\text{QCD}} < 0.4$
- Correct for tag rate correlations due to pair production

MET+Jets background region (1-tag)