Single Top Quark Production at CDF

Jeannine
Wagner-Kuhr

on behalf of the CDF *Collaboration*

Universität
Karlsruhe

KIT
Karlsruhe Institute of Technology

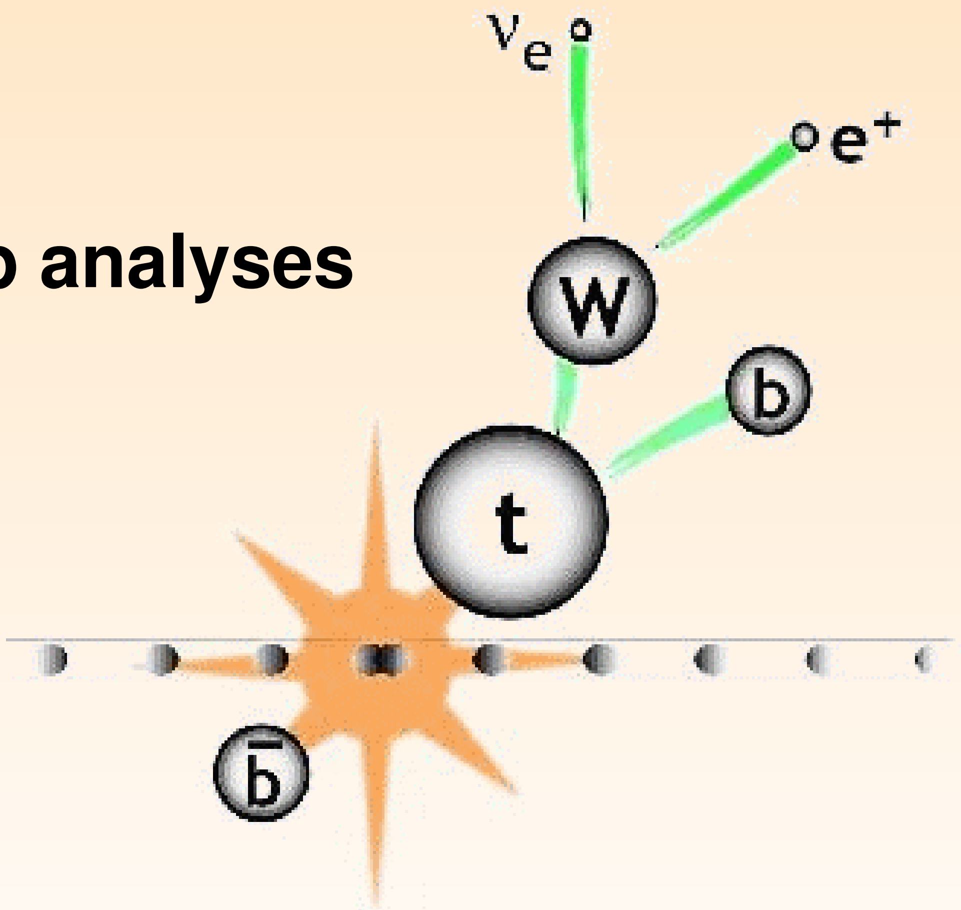
CERN seminar

22.4.2008

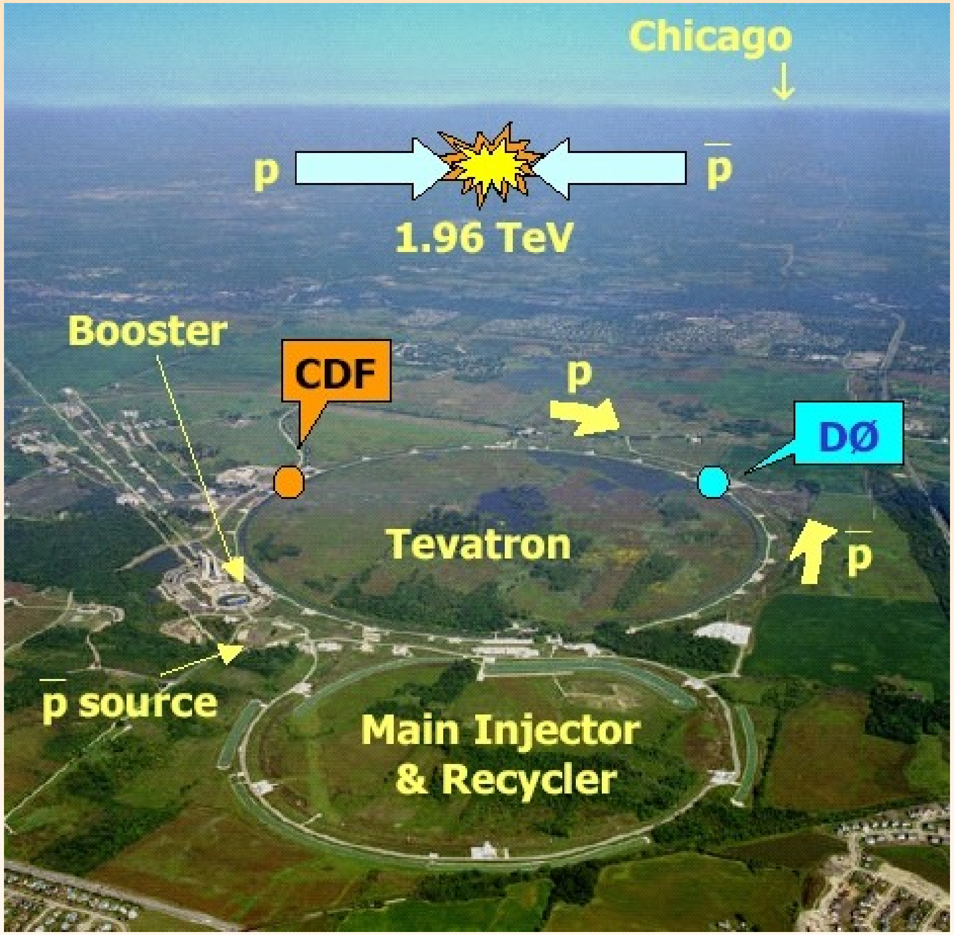
Photo: Reidar Hahn

Overview

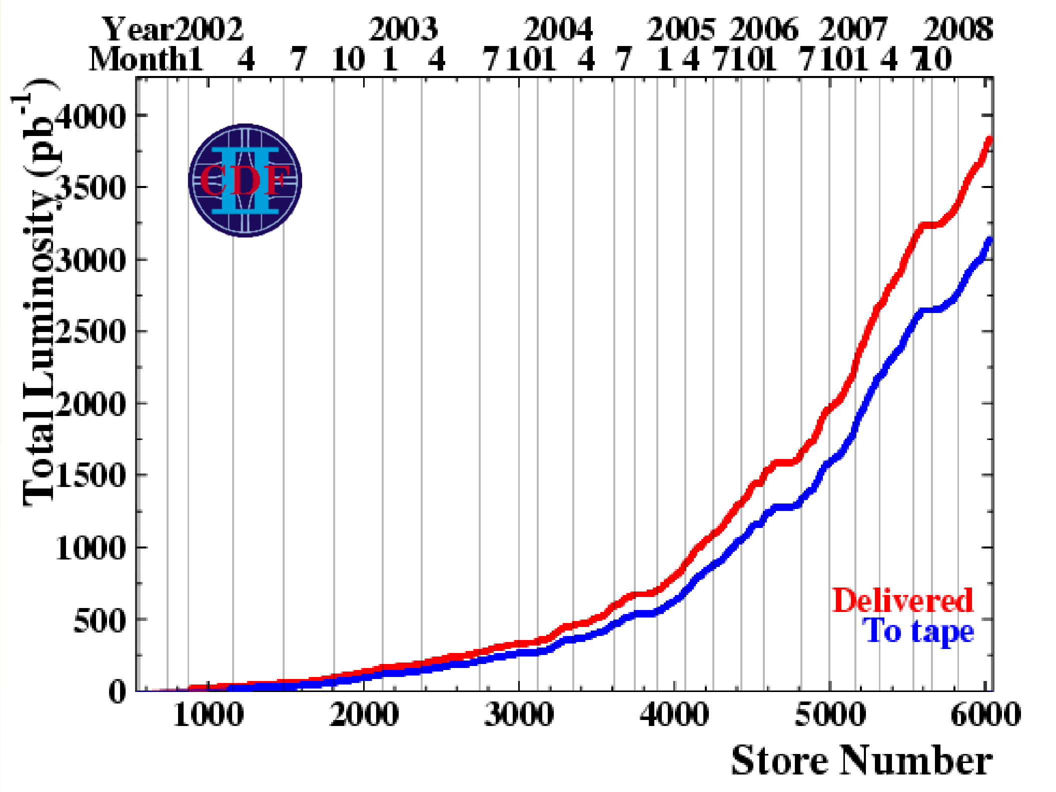
- Tevatron and CDF experiment
- Motivation
- Ingredients for single top analyses
- Multivariate analyses
- Statistical analysis
- Results from CDF and comparison with results from DØ



Tevatron Accelerator



- **Delivered luminosity 3.8 fb^{-1}**
  (3.1 fb^{-1} on tape)
- **Record luminosity: ~$3.2 \cdot 10^{32} cm^{-2} s^{-1}$**

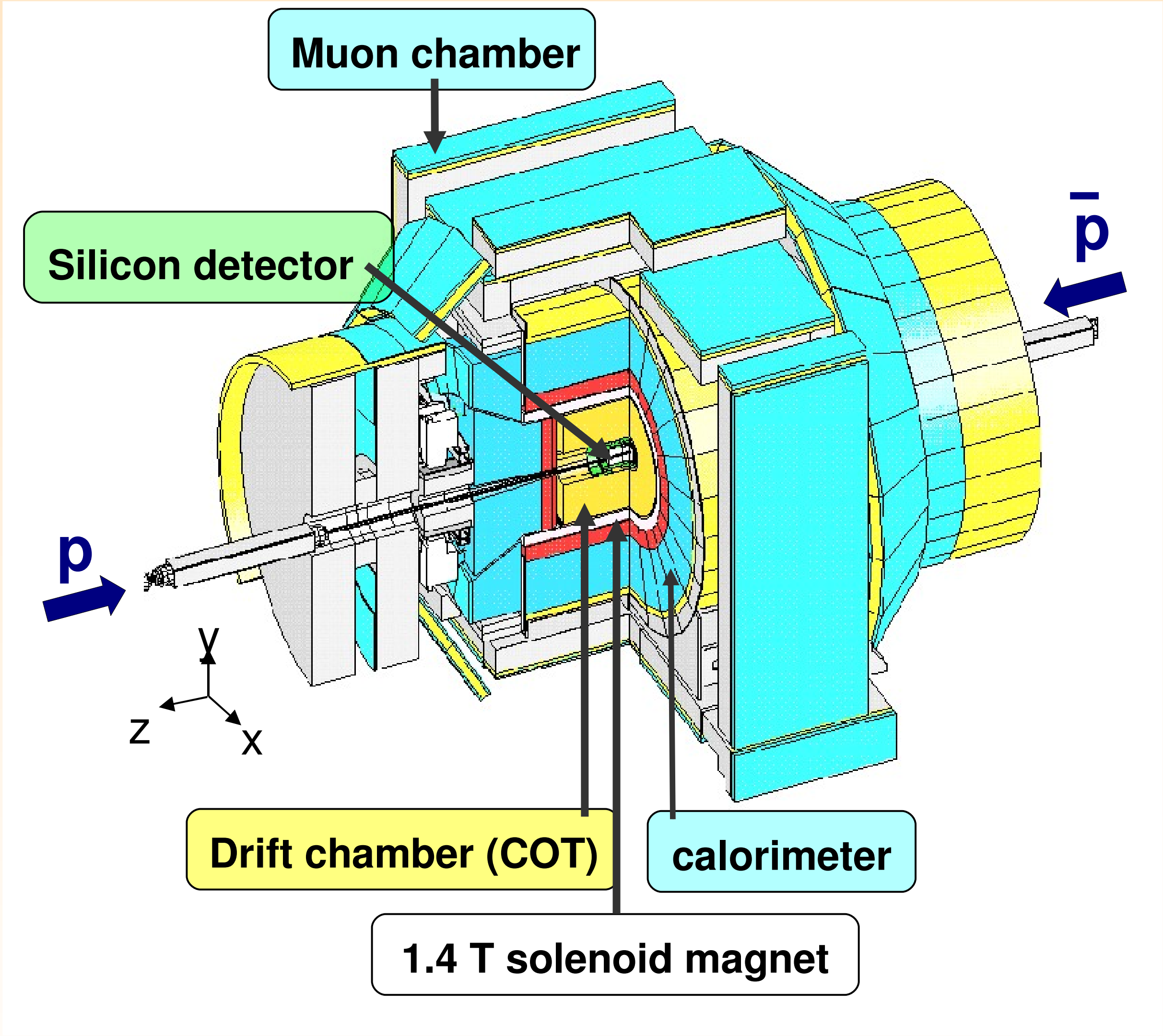


Single top quark analyses:

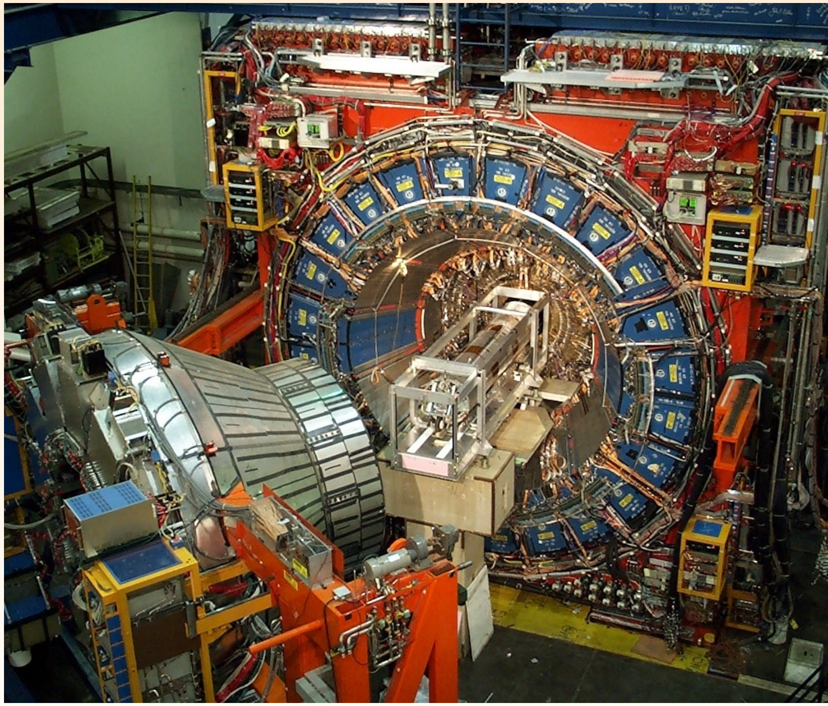
L_{int} = (1.9-2.2) fb^{-1}

CDF Detector

Multi purpose detector



- Very good momentum resolution
- Precise vertex reconstruction



Installation of the silicon detector

Why are Top Quarks Interesting?

- The top quark is very heavy - Why?

  $M_t >> M_b > M_c >> M_s$, $m_t \approx 173$ GeV/c^2

- M_t of same order as EW symmetry breaking scale ➡ *special top dynamics?*

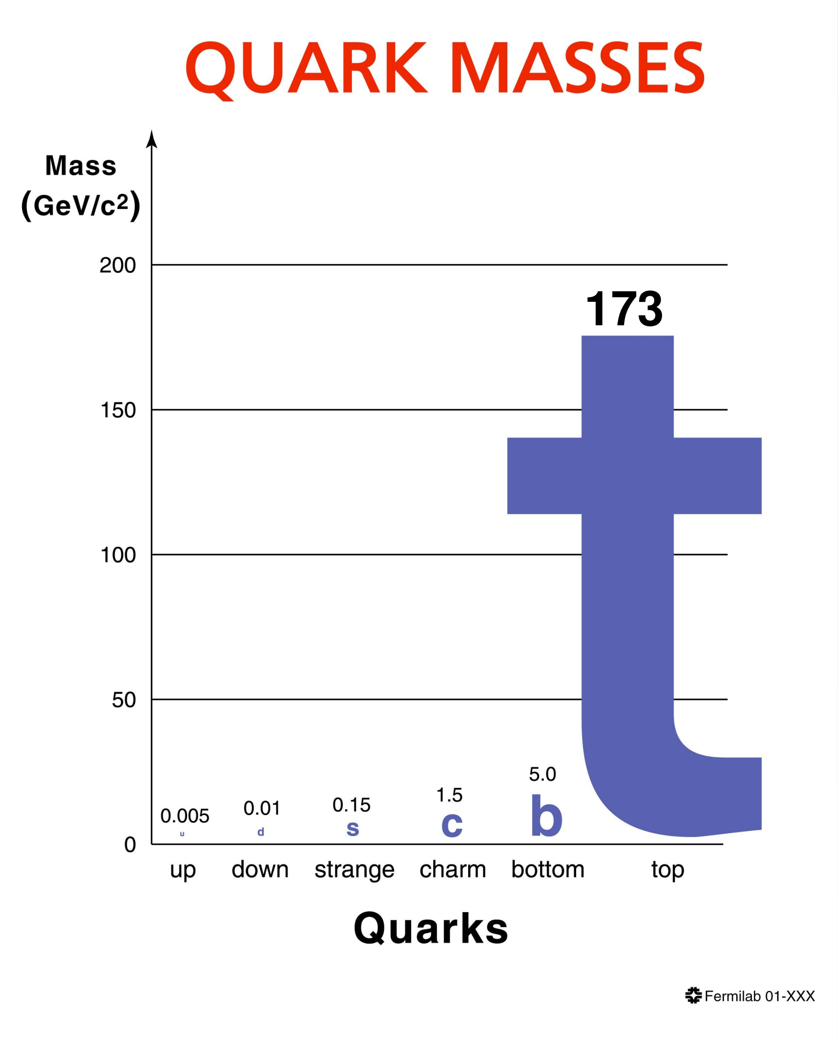
- Top mass allows together with W mass for predictions on Higgs mass

- Top quark decays before it hadronizes

  No top-hadrons, no spectroscopy

  But: quasi free particle

  (spin measurement possible)

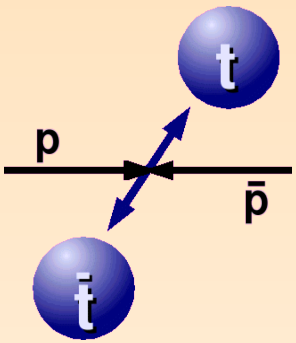


5 orders of magnitude between quark masses

Knowledge about Top-Quark

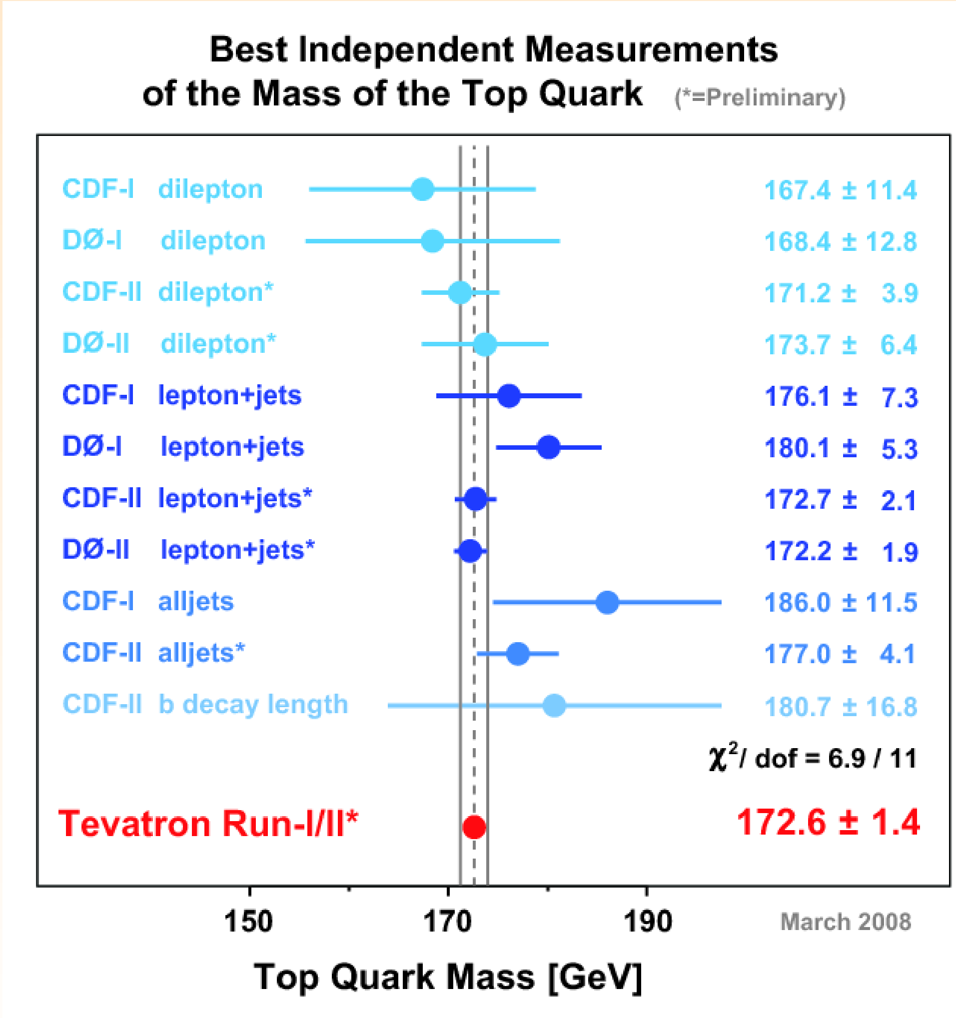
1995: Discovery of top pair production

Strong interaction (QCD)

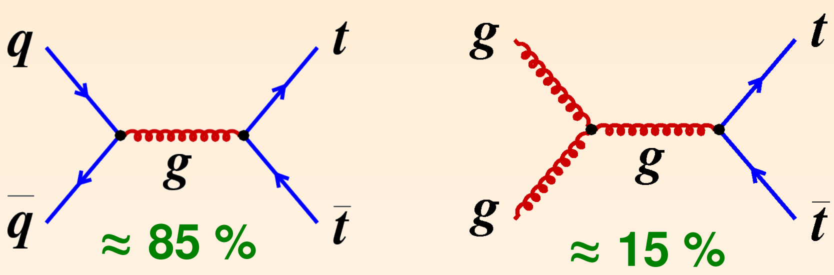


Rel. top mass uncertainty $\lesssim$ 0.9%

(Tevatron combination)



Tevatron:



SM NLO prediction: σ = 6.7 ± 0.8 pb

M. Cacciari et al., JHEP 0404, 068 (2004)

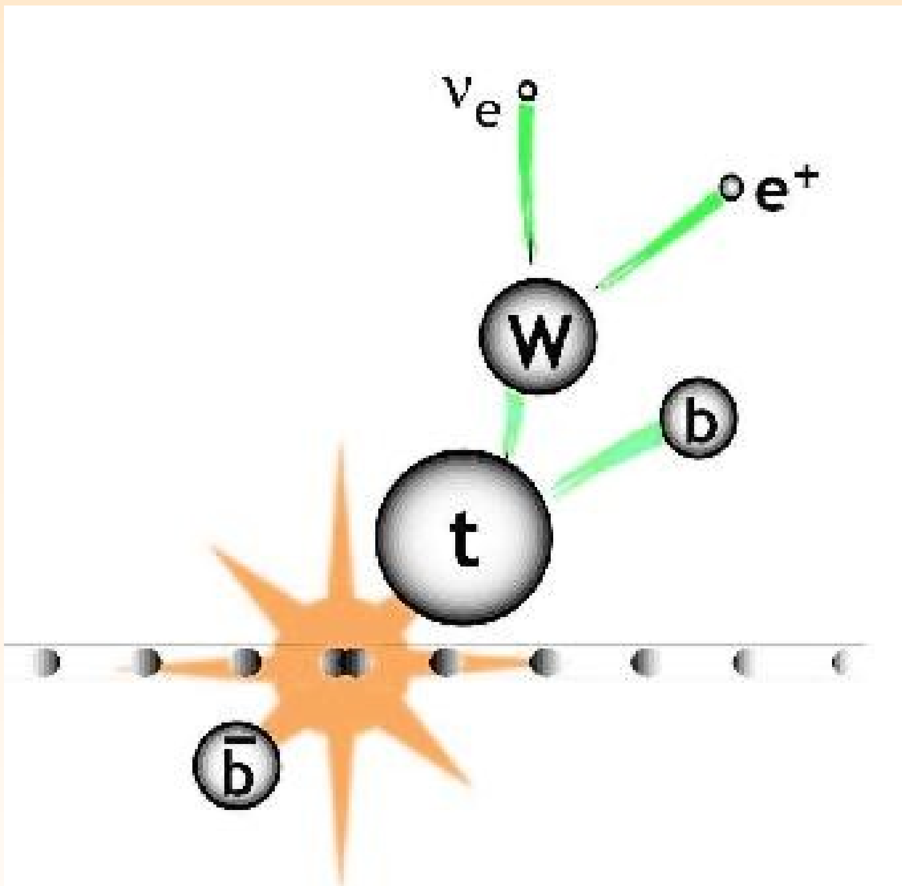
- **Well established prod. mechanism**
- **Measurement of top mass and other properties**

So far, no hint for a non-SM top quark

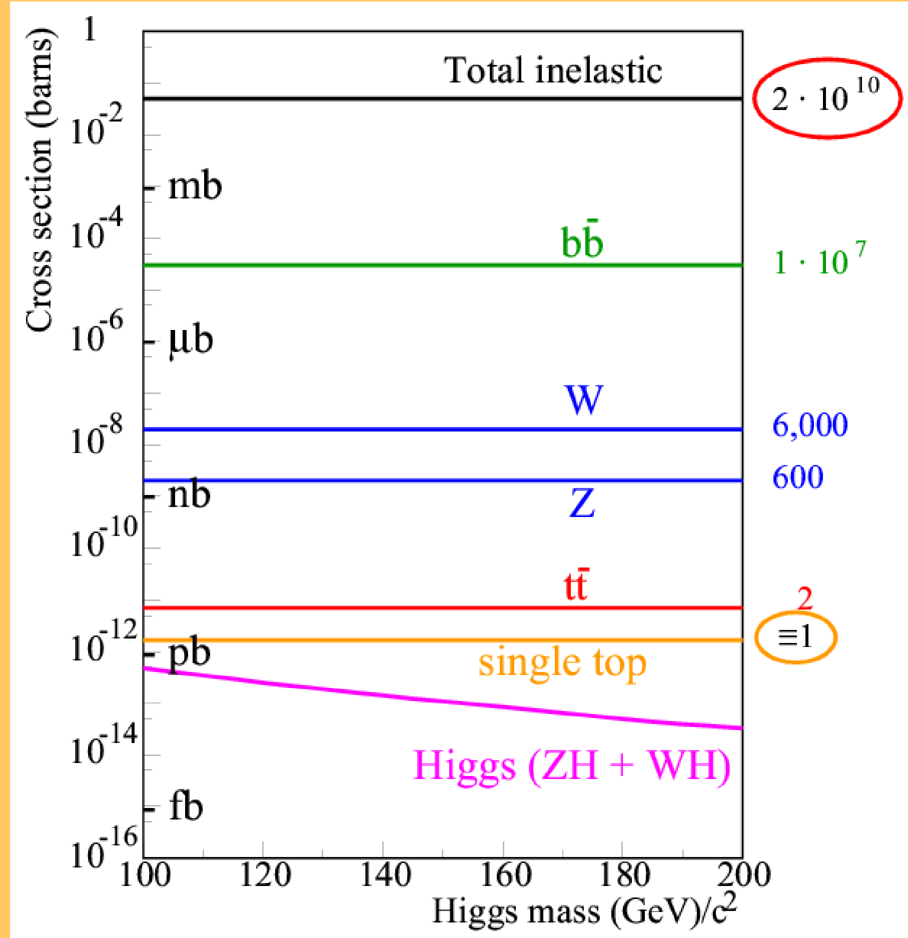
Top Quark – EW Production

Production of single top quarks:

Electroweak interaction



Evidence since 2007 (DØ)

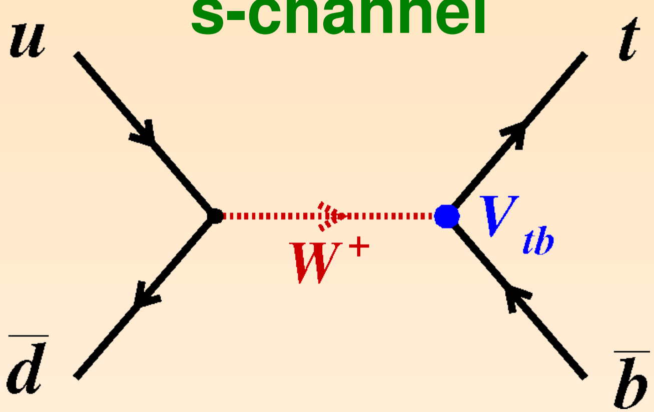


This talk:

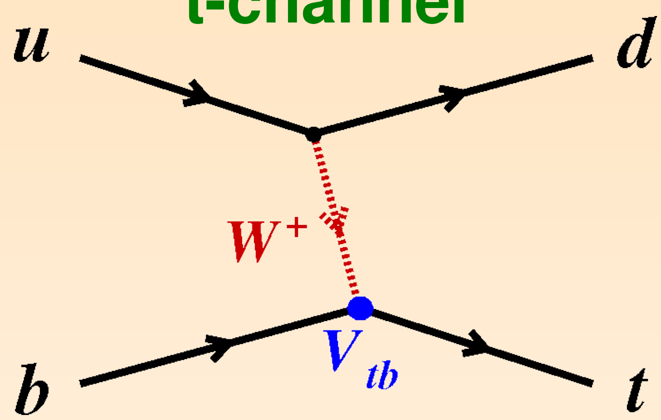
Update of CDF analyses

Single Top Quark Production

s-channel



t-channel



LHC:

$\sigma_s = (11 \pm 1)$ pb

$\sigma_t = (247 \pm 10)$ pb

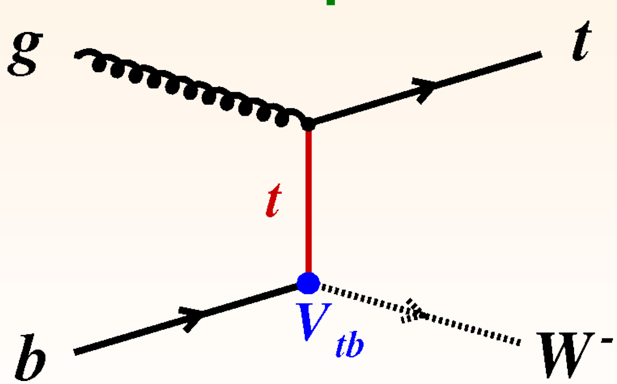
$\sigma_a = (56 \pm 6)$ pb

SM NLO predictions (Tevatron):

$\sigma_s = (0.88 \pm 0.11)$ pb $\quad\quad \sigma_t = (1.98 \pm 0.25)$ pb

B.W. Harris et al., Phys. Rev. D 66, 054024 (2002), Z. Sullivan, Phys. Rev. D 70, 114012 (2004), N. Kidonakis, Phys. Rev. D74, 114012 (2006)

associated production



negligible

(σ_a ~ 0.3 pb)

$\sigma_{s+t} = 2.9 \pm 0.4$ pb

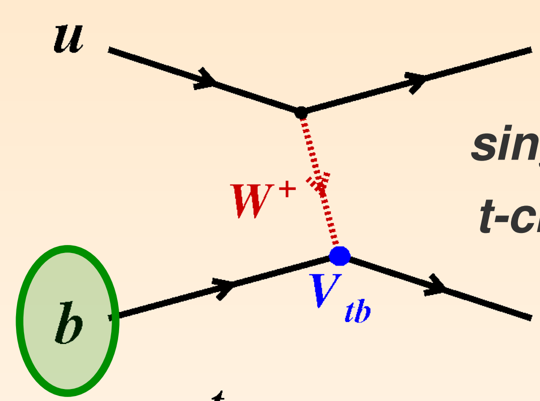
$\sigma_{s+t} \approx 0.4\, \sigma_{top\ pair}$

Single Top Quarks – Why Interesting?

Production of single top quarks:

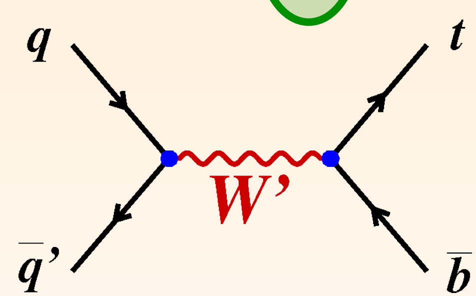
- **Test of SM prediction**

- $\sigma_{\text{single top}} \sim |V_{tb}|^2$ → **measurement of V_{tb}**



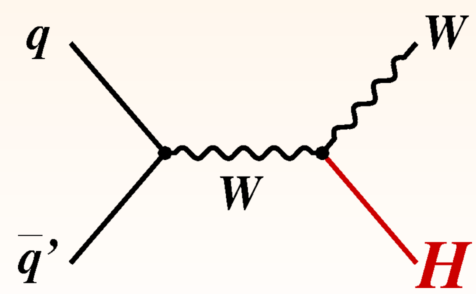
single top
t-channel

- **Test of b-quark structure function**
  (DGLAP evolution)



- **Search for non-SM contributions**
  (W' or H^+)

- **Technical motivation - WH, $H \rightarrow b\bar{b}$**
  (Same final state, $\sigma_{WH} \approx 0.1\ \sigma_{\text{single top}}$)



CKM Matrix Element V_{tb}

Cabbibo-Kobayashi-Maskawa matrix: Rotation of down type quark mass eigenstates into weak eigenstates

directly measured

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

ratio now well measured from B_s mixing

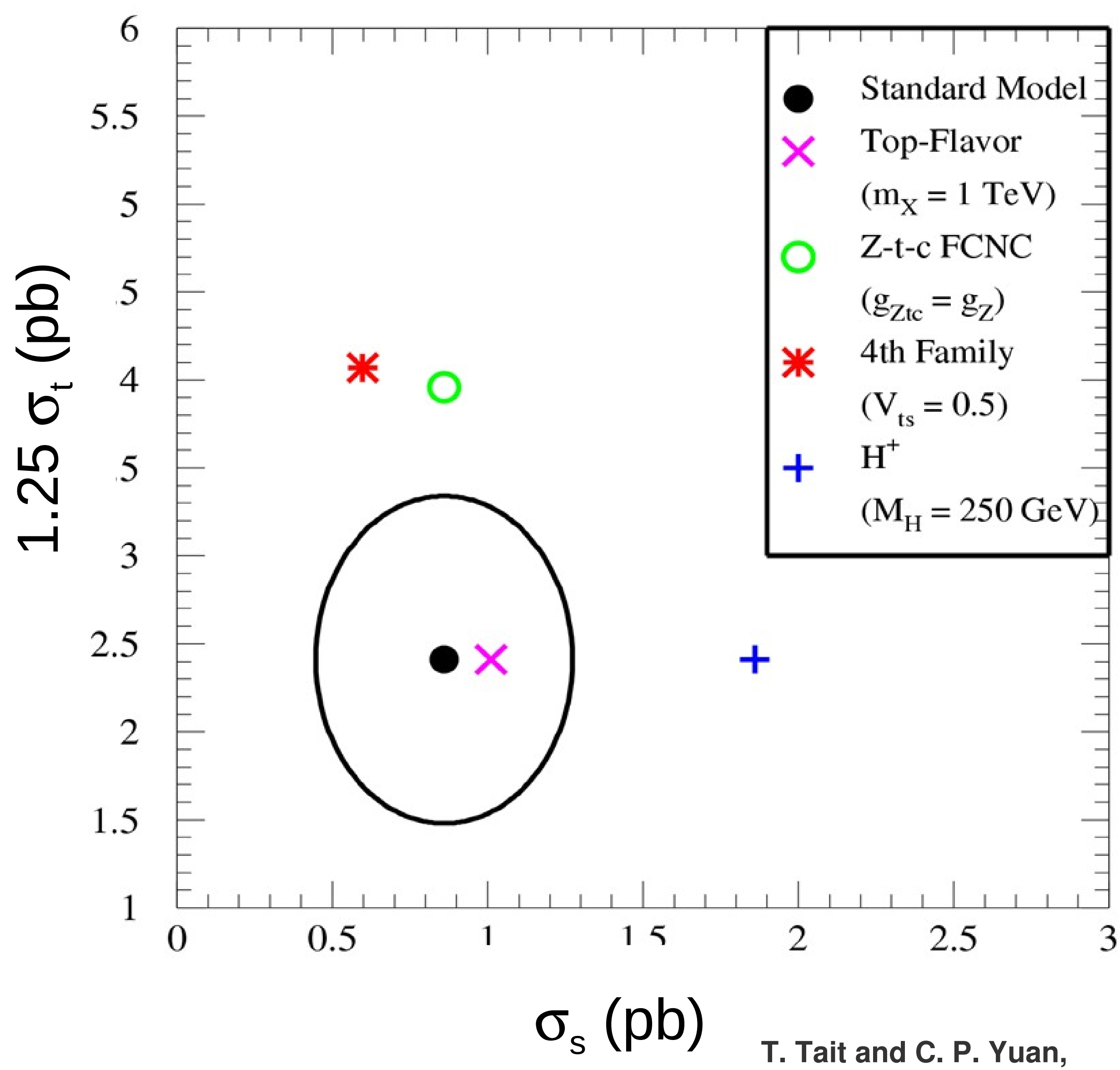
only indirectly known

Direct measurement only via single top-quark production

Are unitarity relations valid?

e.g.: $V_{ub}^2 + V_{cb}^2 + V_{tb}^2 \stackrel{?}{=} 1$ ⟺ *Hints for the existence of a 4th generation ?*

Single Top Quarks - Predictions

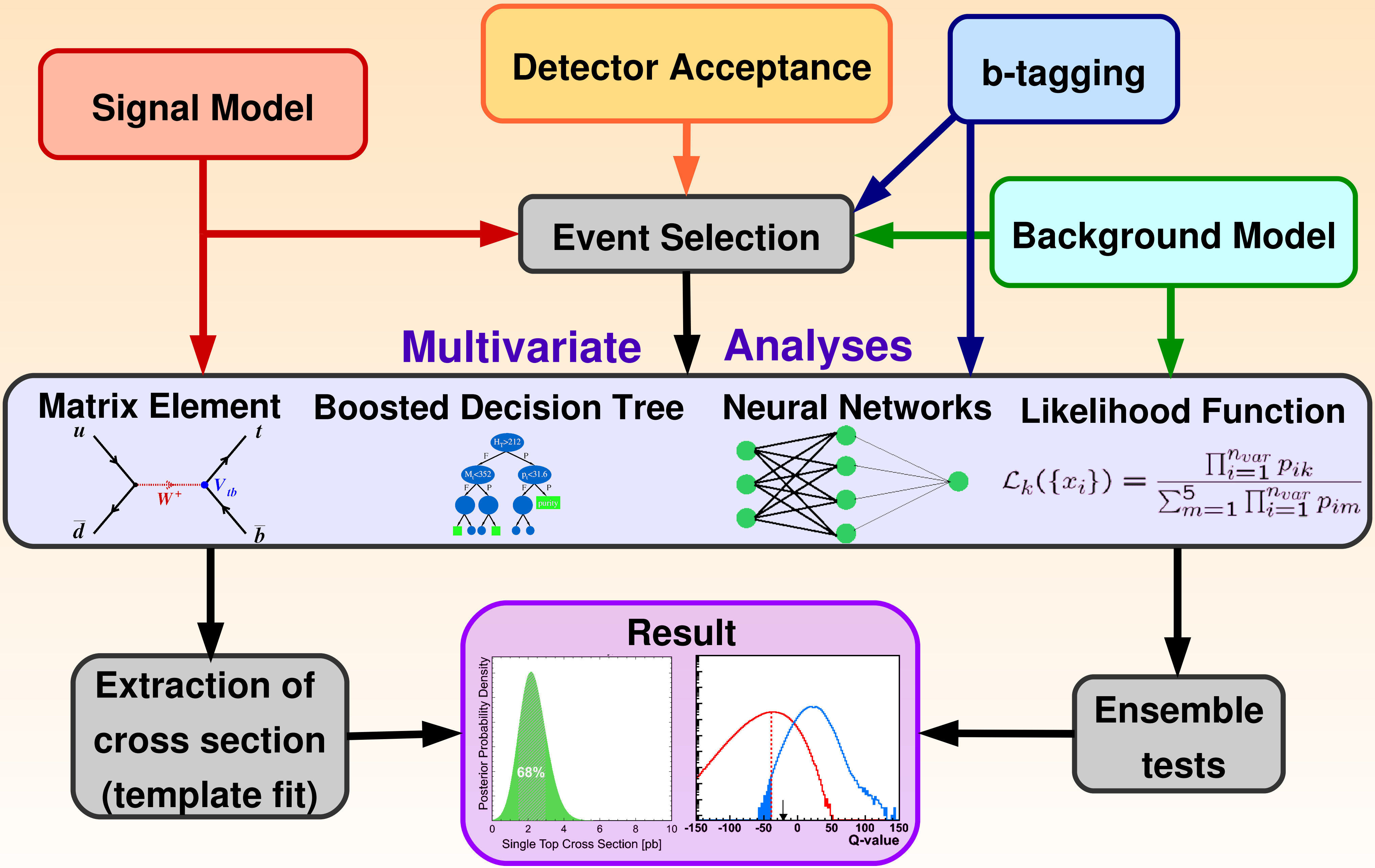


T. Tait and C. P. Yuan, Phys.Rev.D63:014018 (2001)

Single top quark production:

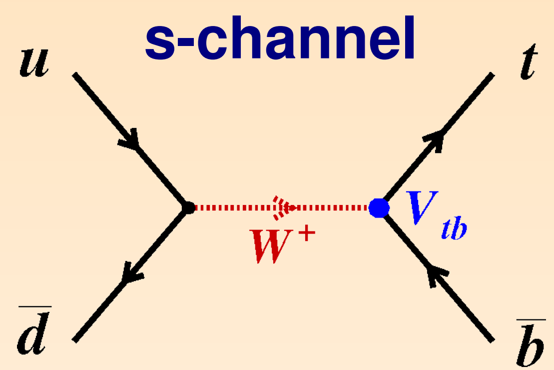
s- and t-channel rates are differently sensitive to new physics

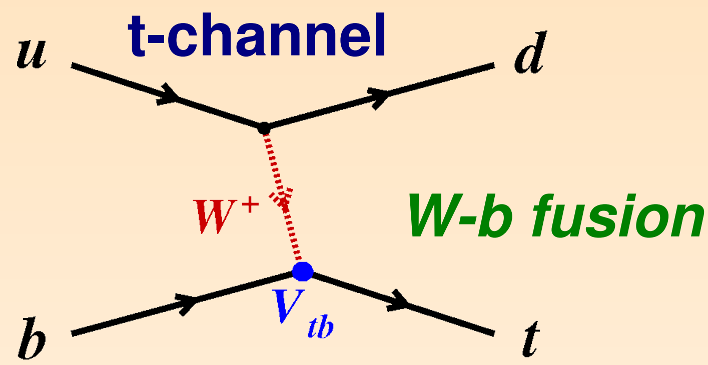
Analysis Roadmap



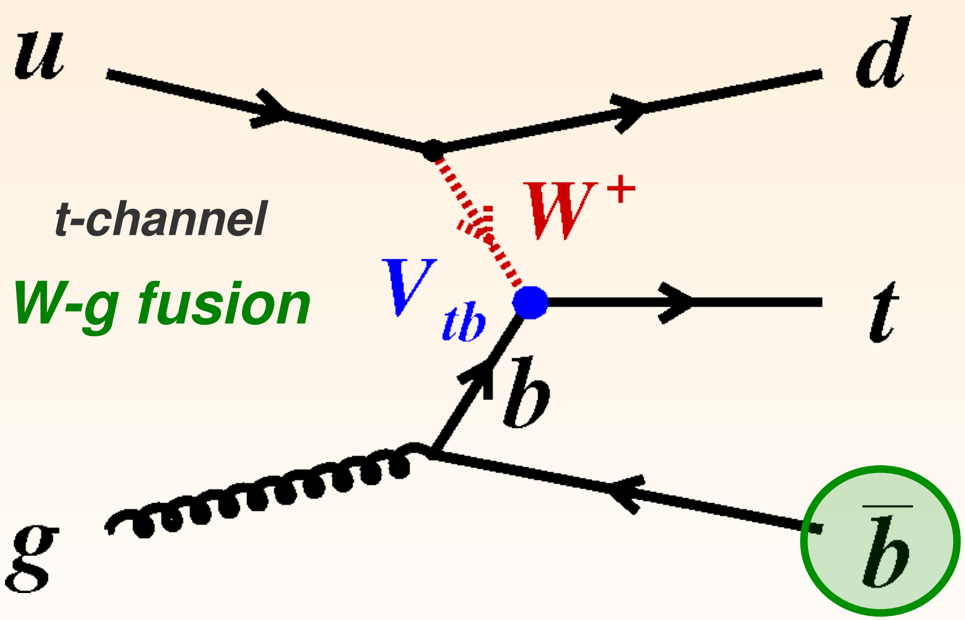
Signal Model

LO diagrams





Most important NLO diagram



Important features:

Polarization of top quarks, W-g fusion process

Modeling:

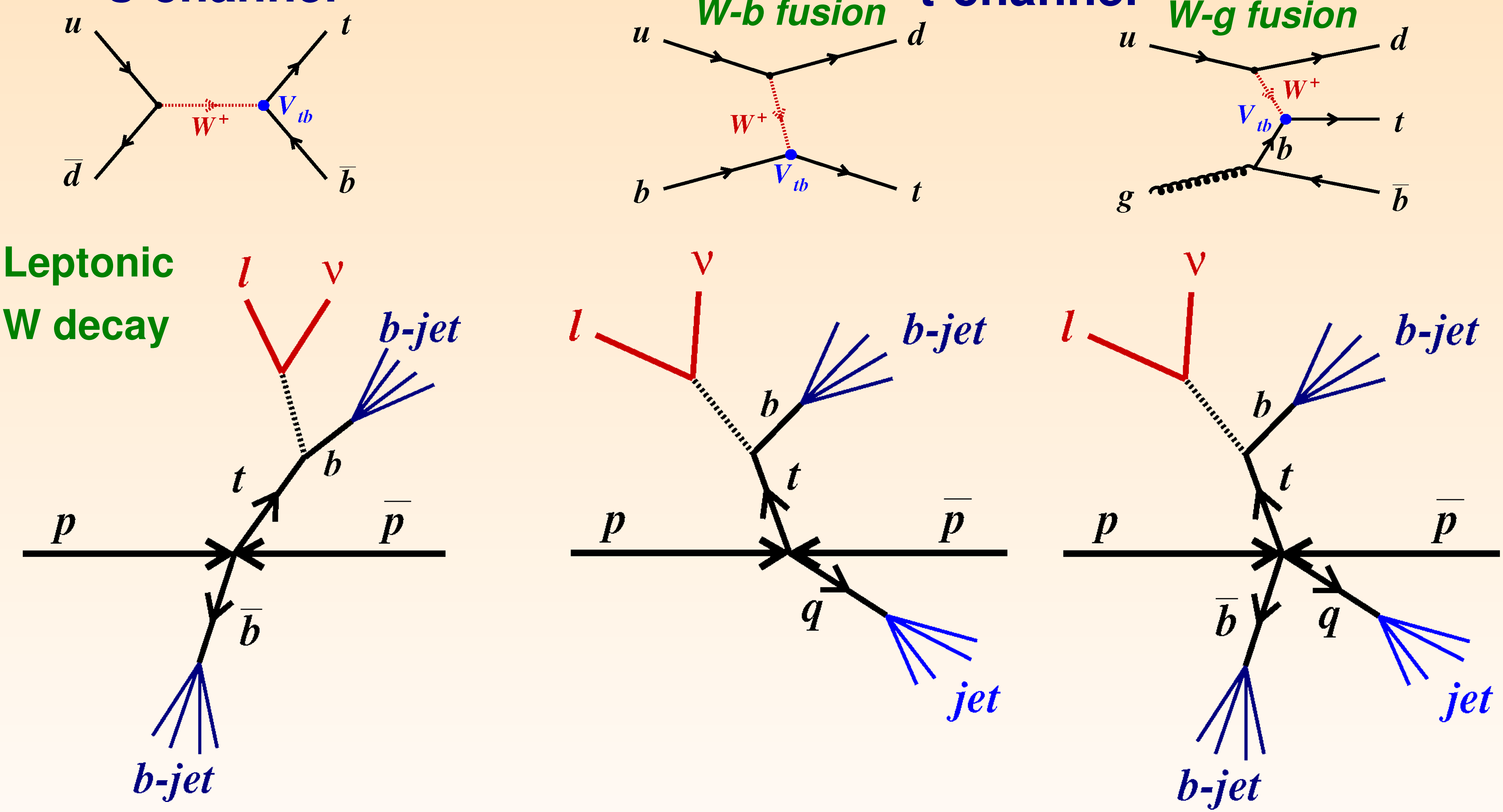
- **MADEVENT + PYTHIA for showering**
- **W-b and W-g fusion processes generated separately and matched in b p_T to match ZTOP NLO calc.**

ZTOP: PRD66,054024 (2002); MADEVENT: JHEP 0709:028 (2007)

Event Topology

s-channel | **t-channel**

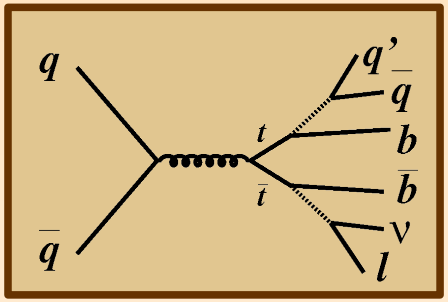
Leptonic W decay

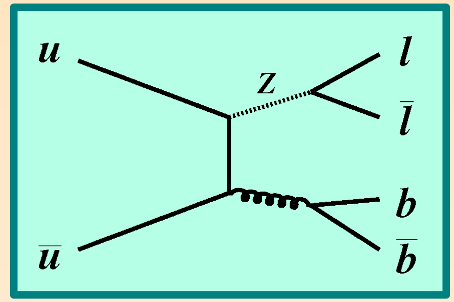


Signature: Exactly one isolated high p_T e or μ, missing E_T, ≥ 2 jets, ≥ 1b-jet

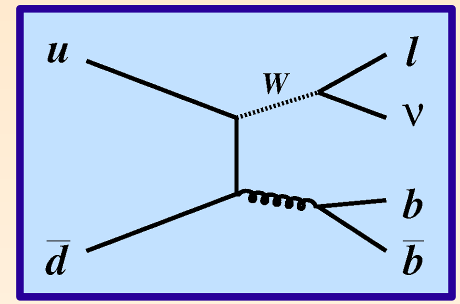
Background Processes

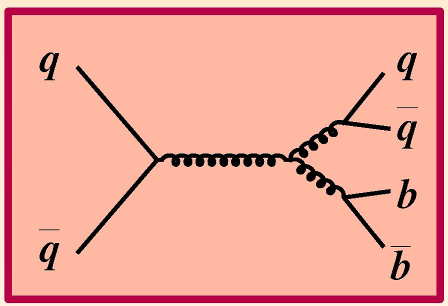
N_{jets}: 2 or 3

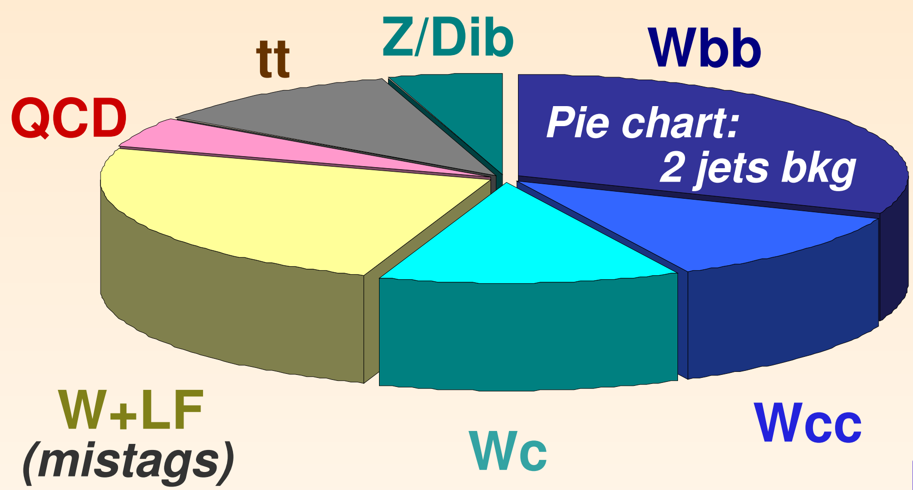




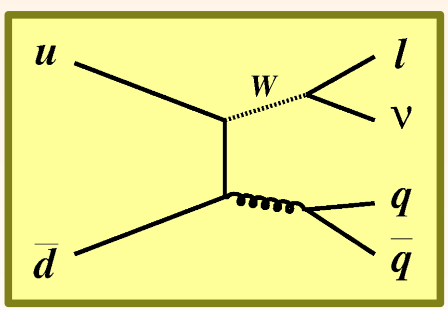
Apply a Z-veto

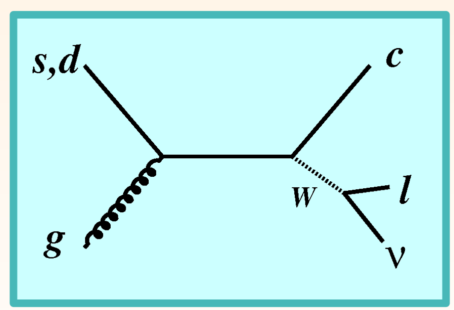


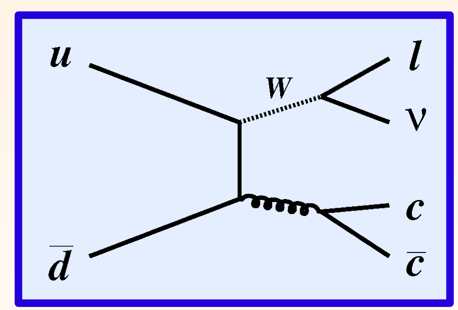




Apply a QCD-veto







Require at least one jet with a secondary vertex

Event Selection

- **1 Lepton (e or μ)**

  High p_T lepton triggers,

  MET+jets trigger (ME, BDT analyses)

  $p_T > 20$ GeV/c^2, $|\eta| < 2.0$

- **Missing E_T (MET)**

  MET > 25 GeV

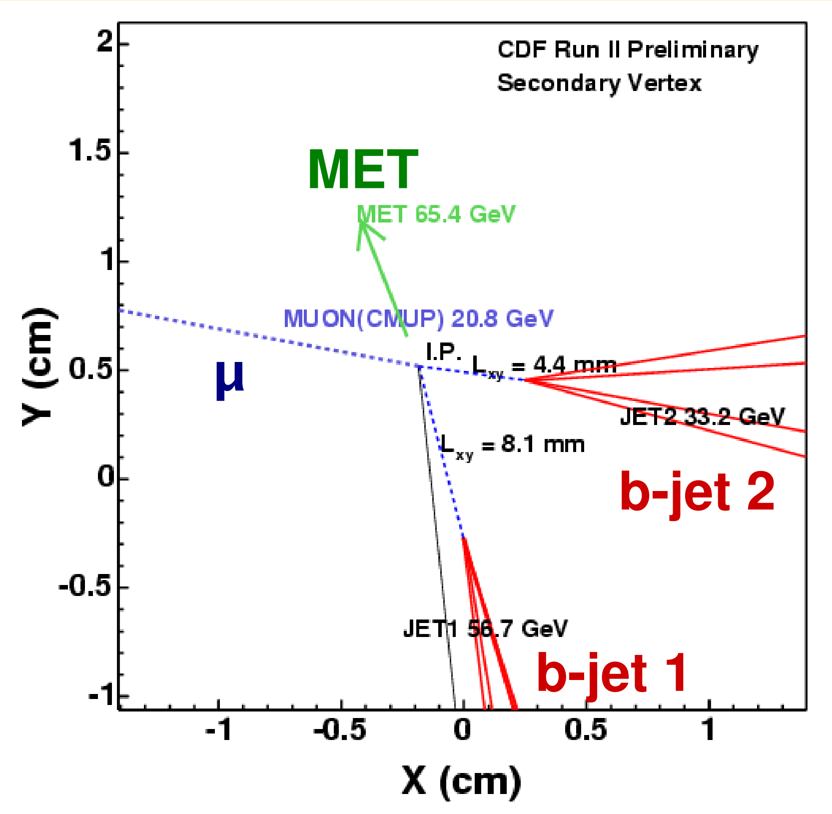
- **2 or 3 jets (hadron level)**

  $E_T > 20$ GeV, $|\eta| < 2.8$

- **At least one b-tagged jet**

  secondary vertex tag

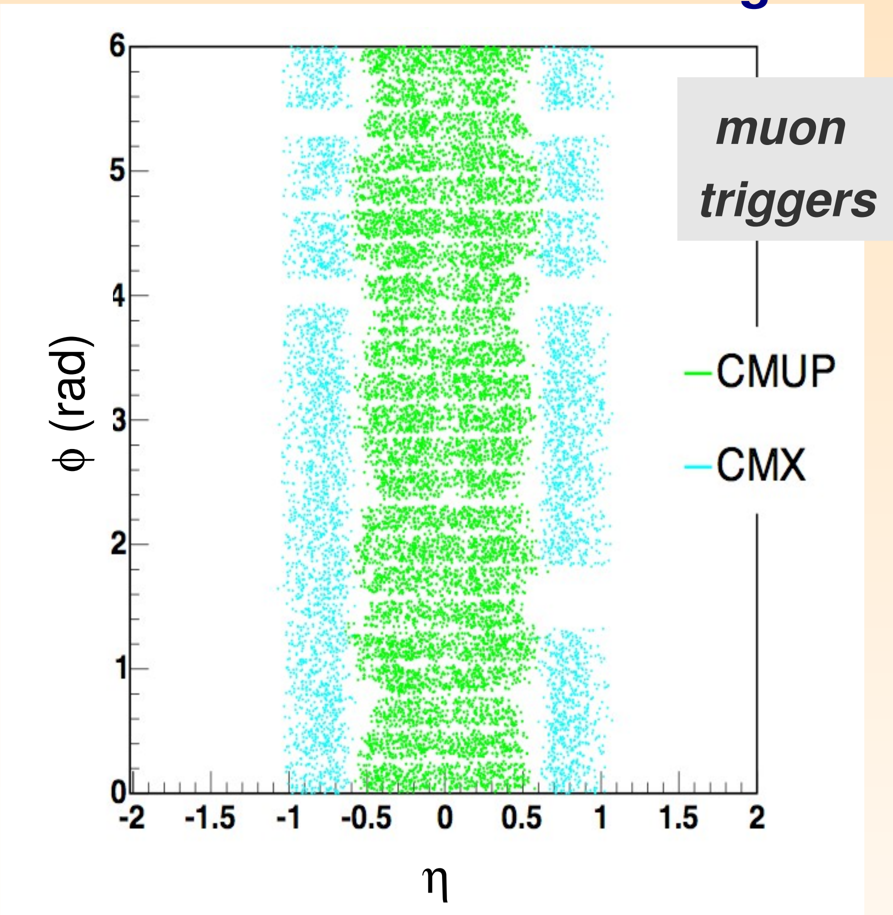
- **Z-veto and QCD-veto**



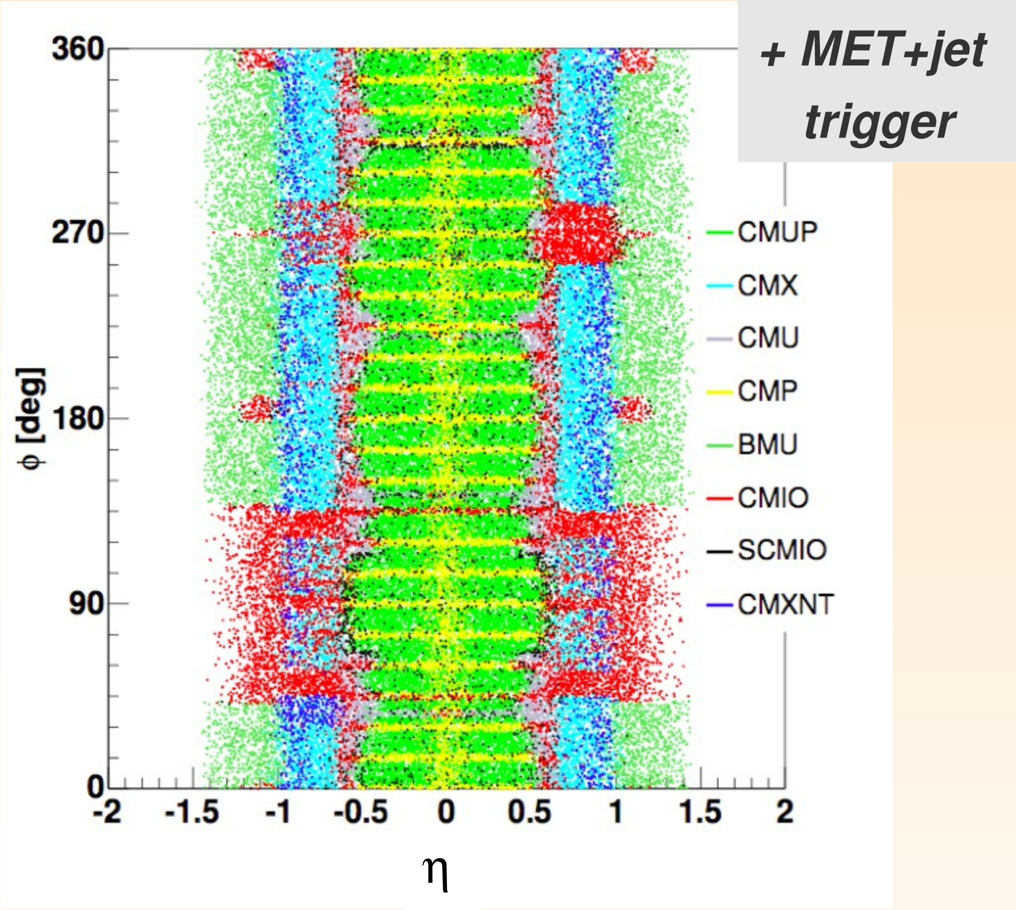
	W+2jets	W+3jets
s-channel	42 ± 6	14 ± 2
t-channel	62 ± 9	18 ± 3
Total pred.	103 ± 15	22 ± 5

Acceptance Gain for Muons

Standard CDF muon coverage



Extended Muon coverage

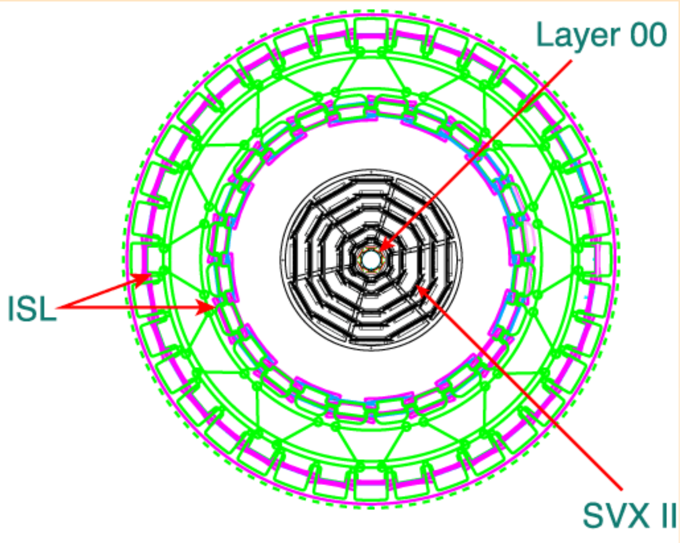


Gain from missing-E_T+jets triggered events:

- ~30% gain in muon acceptance
- ~12% gain in total signal acceptance (with electrons)

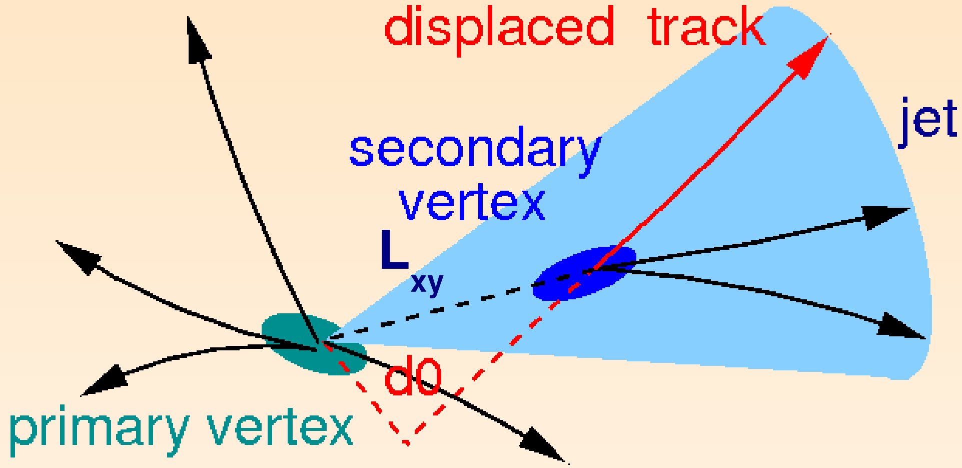
b-Tagging at CDF

Silicon detectors



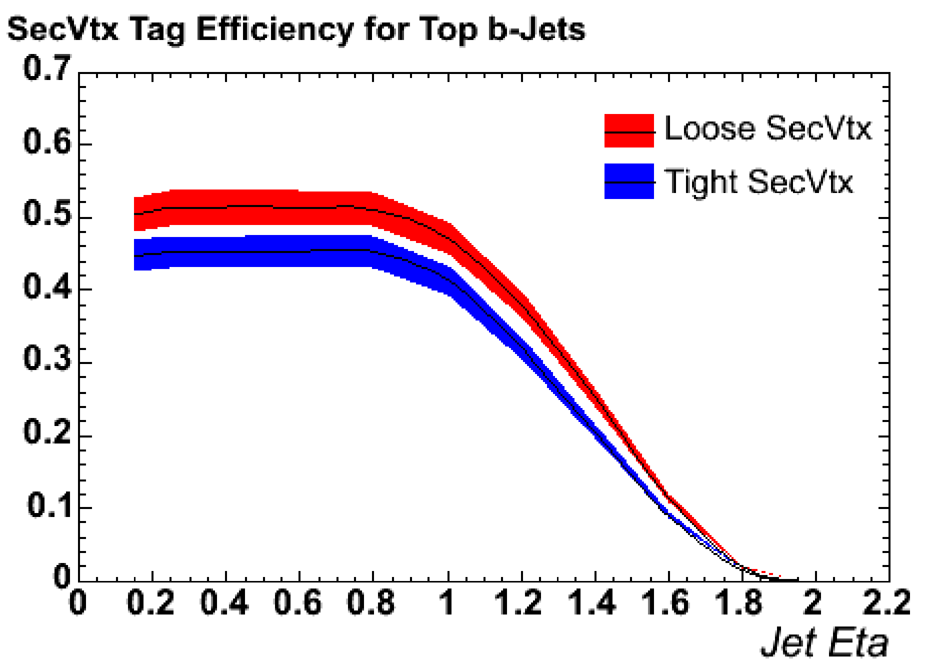
Impact parameter d_0 resolution for high-p_T tracks ~18µm

Decay length L_{xy} :



$\tau_B \approx 1.5$ ps $\rightarrow c\tau \approx 450$ µm

Typical L_{xy} in CDF: O(mm)



Tight b-tag: $L_{xy} / \sigma(L_{xy}) > 7.5$

$\epsilon(b) \approx 45\%$

$\epsilon(c) \approx 9\%$

$\epsilon(q) \approx 1\%$

B-Tagging Tool - Flavor Separator

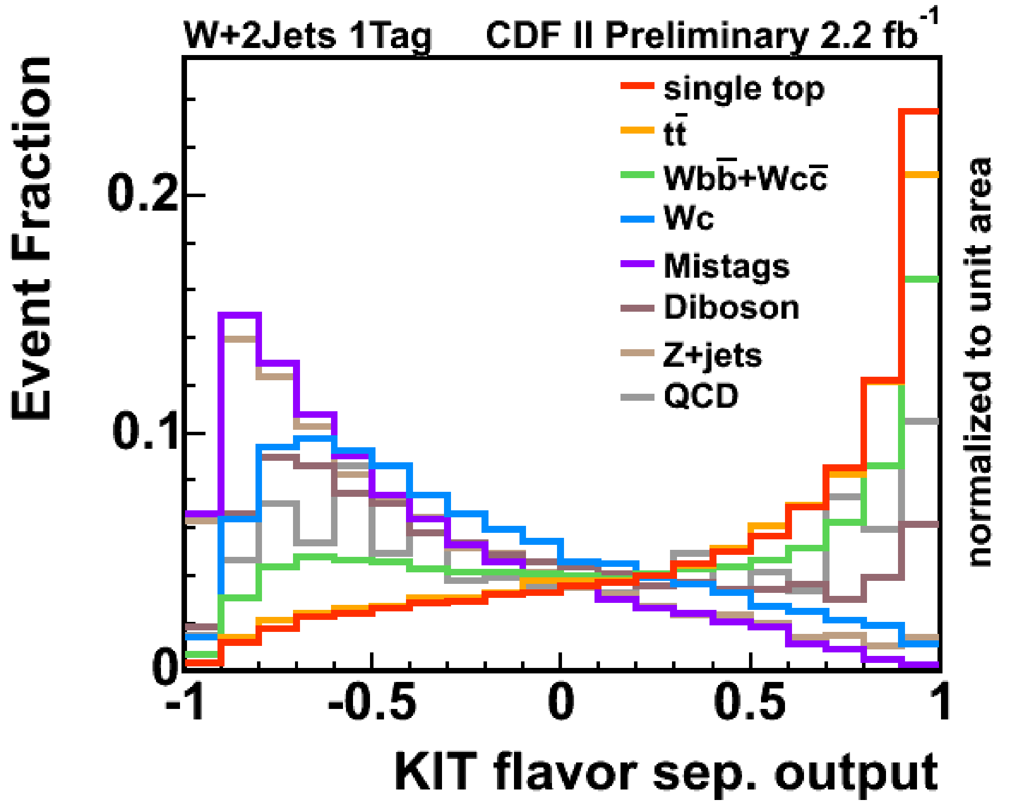
- **Identified sec. vertex tags have a significant charm and mistag contamination**

- **Combination of jet and track variables in NN:**

  Vertex mass, decay length, track multiplicity, ..

  ⇒ powerful discriminant

- **Gain in sensitivity for single top analyses: ≈ 15-20 %**



Mistags / W+charm .... beauty

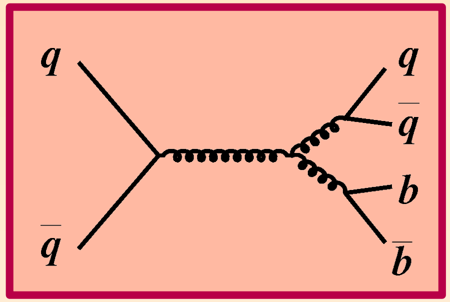
Very powerful at separating out residual mistags and charm from b-tagged sample

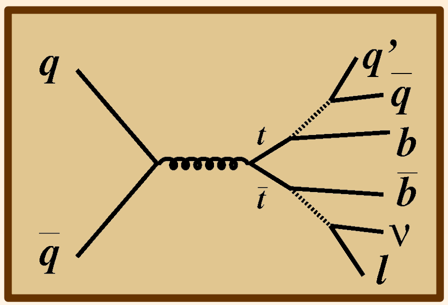
Background Estimation - Strategy

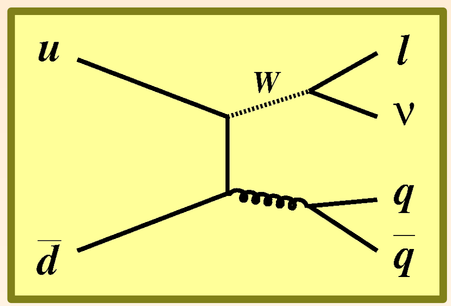
Diboson-, Z+jets- and top pair production:

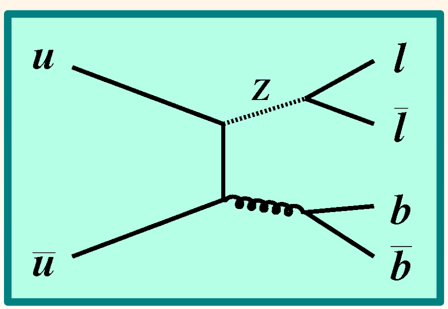
Determined with MC

$$N_{pred} = \sigma_{theo} \cdot \epsilon_{evt} \cdot \int \mathcal{L}\, dt$$









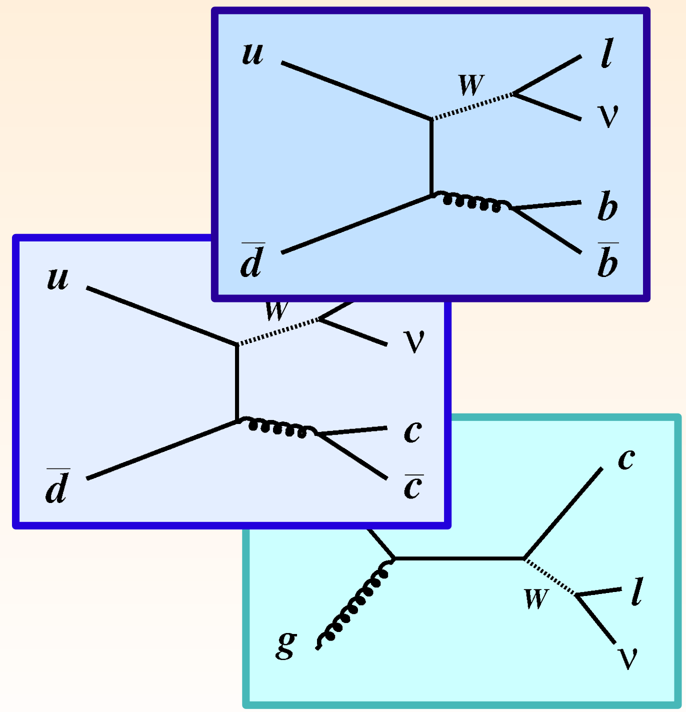
Mistags (W+LF jets) and QCD:

Determined from data

W+HF jets:

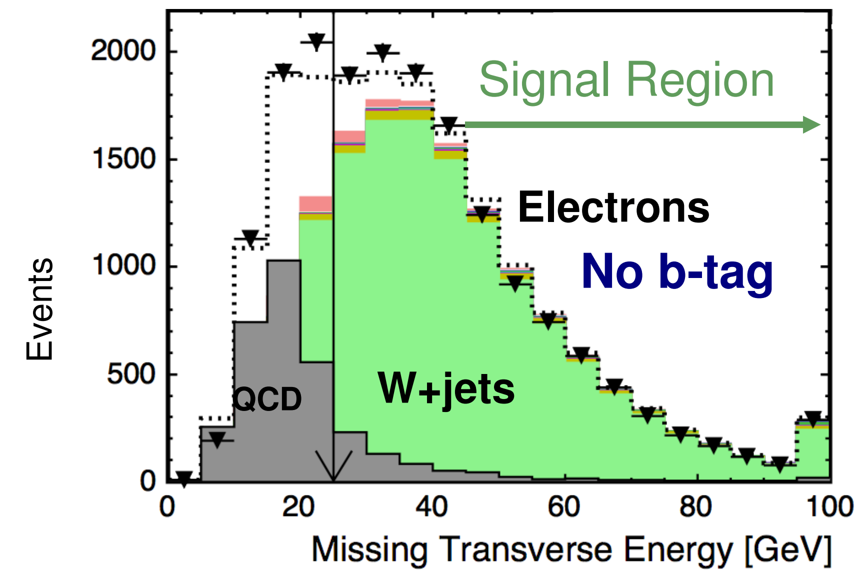
W+jets normalization from data, HF-fraction from ALPGEN MC

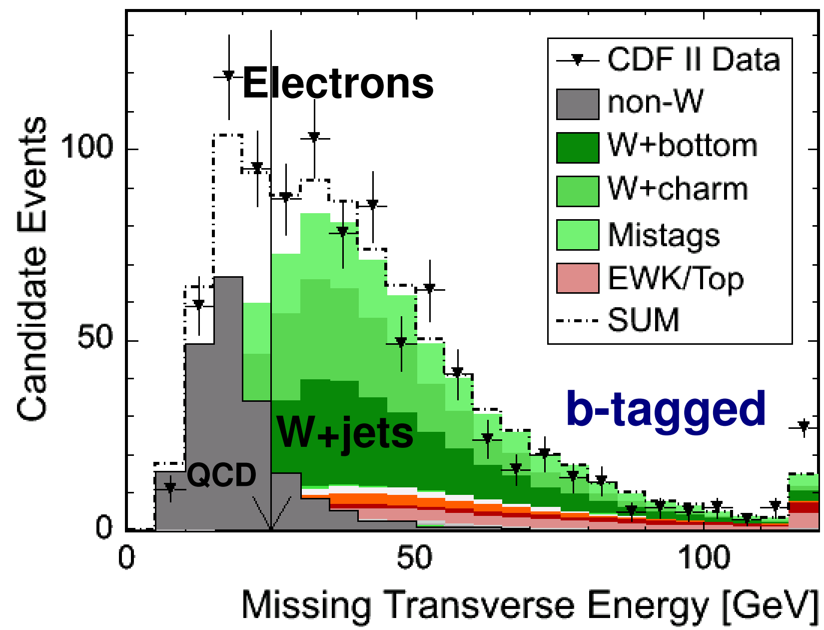
(calibrated in W+1jet events with b-tag)



QCD Background

- **MC models of inclusive jet+MET production not precise enough**
- **Apply QCD-veto to suppress QCD bkg**
  cuts on MET, $M_{T,W}$, Δφ(jet, MET), Δφ(l,jet)
- **Use data samples to model kinematic:**
  - ***"Anti-Electrons": Electron candidates which pass all but two of the selection requirements***
  - *"Jet-Electrons": Multijet events where one jet fakes the electron*
- **Perform fits to MET distribution**
- **Uncertainty on QCD rate: ± 40%**
  Shape uncertainty on HF content



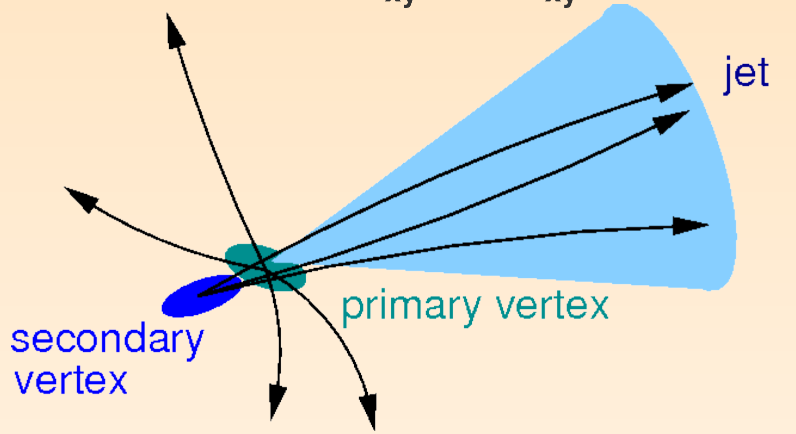


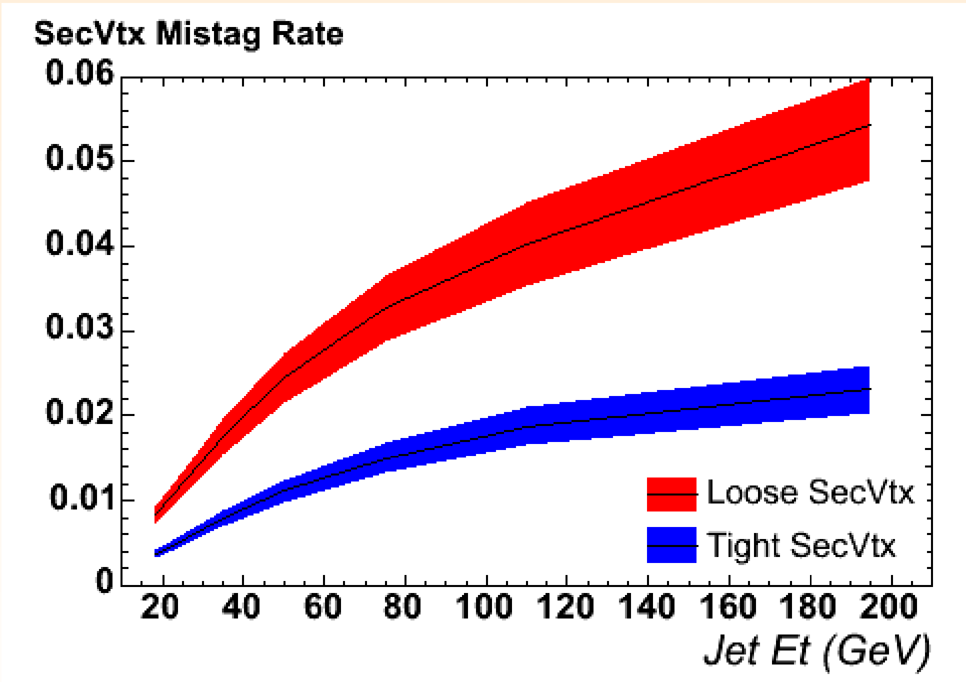
Background Estimation - W+LF jets

- **ALPGEN W+LF used to predict kinematic distributions**
- **Determine tagging rate for W+LF in inclusive multijet events using events with negative L_{xy}:**
  - **Parametrize rate of negatively tagged events as a function of**

    $n_{trks,jet}$, $E_{T,jet}$, $|\eta_{jet}|$, $n_{primary\text{-}vertices}$, $\Sigma\ E_T(event)$, $z_{primary}$
  - **Apply "asymmetry correction" to obtain positive tagging rate**
- **Apply tagging rates to W+LF MC**

Negative tag: L_{xy} / $\sigma(L_{xy}) < -7.5$

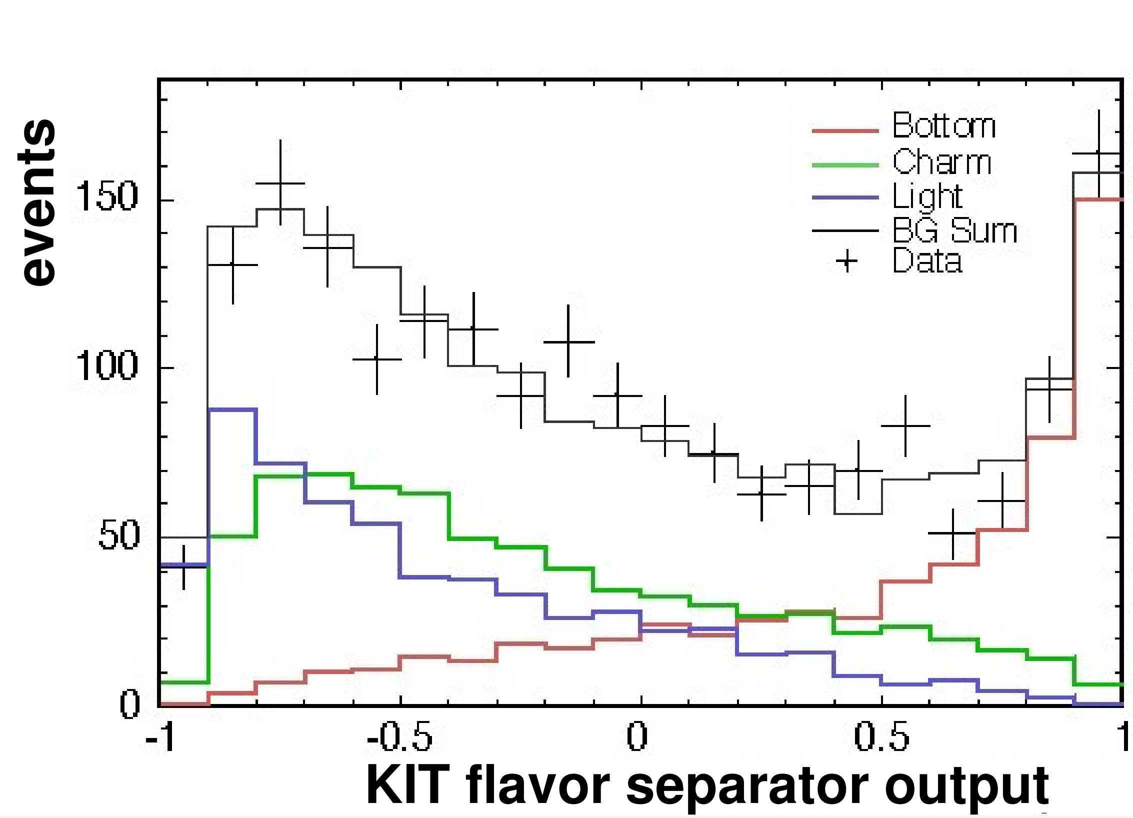




W+jets: Heavy Flavor Calibration

- **Use ALPGEN predictions of kinematic distribution but mistrust ALPGEN rates**
- **Use W+1-jet b-tagged data control sample to estimate the HF fractions**
- **Three-parameter fit to bottom/charm/light templates of KIT flavor separator distribution**

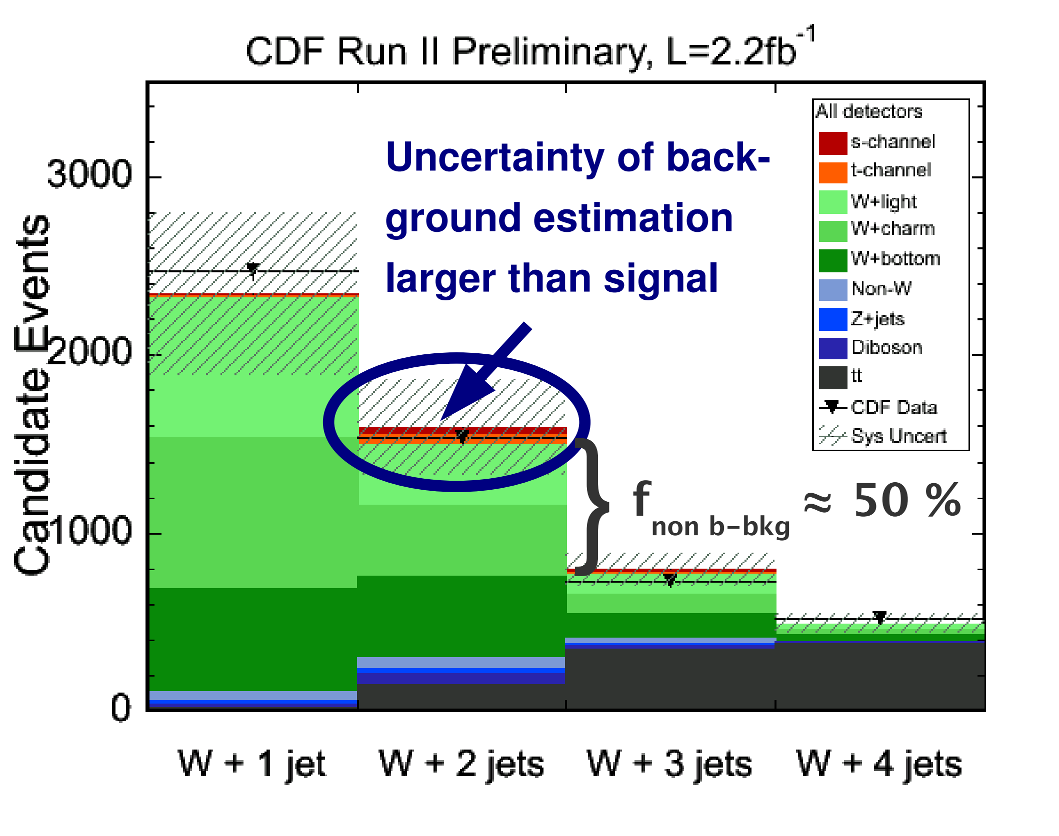
Cross check: displaced vertex mass



Apply correction factor for c and b:

$$K_{HF} = 1.4 \pm 0.4$$

Candidate Events



	W+2jets	W+3jets
Total pred. bkg	1492 ± 269	755 ± 91
pred. single top	103 ± 15	22 ± 5
Total pred.	1595 ± 269	777 ± 91
Observation	1535	712

S:B ratio:

- **W+2jets: 1 : 14.5**
- **W+3jets: 1 : 34**

(dominant bkg: top pair prod., 45 %)

- Counting experiment im-possible
  → Multivariate analyses
- W+2 jets: $f_{\text{non b-bkg}} \approx 50$ %
  → Use KIT flavor separator in multivariate analyses

Search Strategy

Full selected data set

split in subsets of different purity

- **W + 2 jets**
  - 1 tag
  - ≥ 2 tags
- **W + 3 jets**
  - 1 tag
  - ≥ 2 tags

Multivariate methods

Matrix elements | Boosted decision tree | Neural networks | Likelihood function

Combined search

- t- + s-channel = one single-top signal
- σ ratio is fixed to SM value
- important for „discovery“

Separate search

- σ ratio is not fixed to SM value
- Sensitive to new physics processes

Old LF (1.5/fb), NN (1.0/fb) results;

New s-channel meas. (optimized sel., LF)

Matrix Element Analysis

Idea: Compute event probability for signal and background hypotheses

$$\mathrm{P}_i(x) = \frac{1}{\sigma_i} \int d^n\sigma_i(y)\, dq_1\, dq_2\, f(q_1)\, f(q_2)\, W(x|y)$$

x: lepton and jet 4-vectors

LO differential cross section

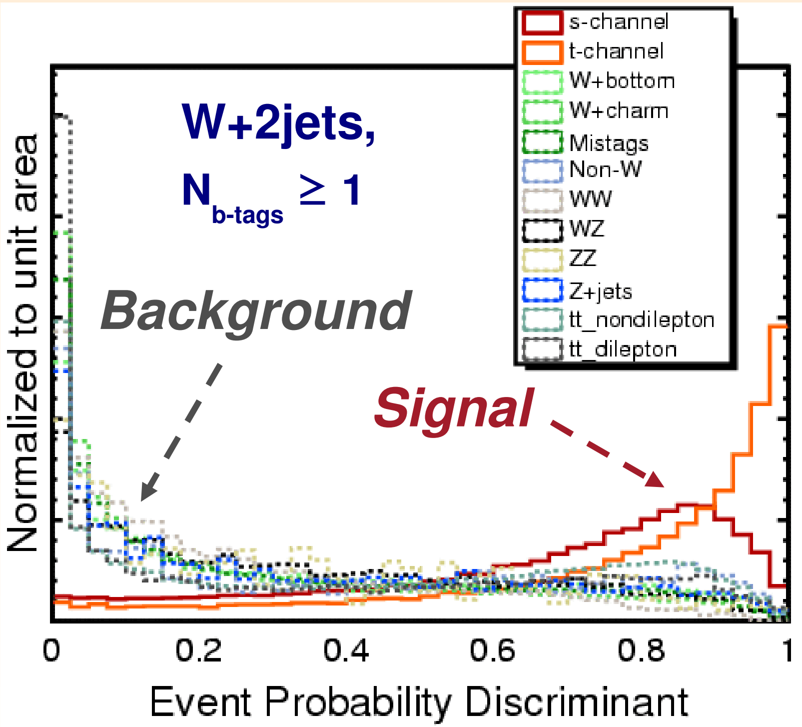
PDF's

Transfer function: Parton (y) → Rec. Object (x)

- Use full kinematic information of event, integrate over unknown or poorly measured quantities
- Calculate probability densities P for s- and t-channel and main backgrounds (*Wbb, Wcc, Wc+jets, top pair)*
- Discriminant:

KIT flavor separator output

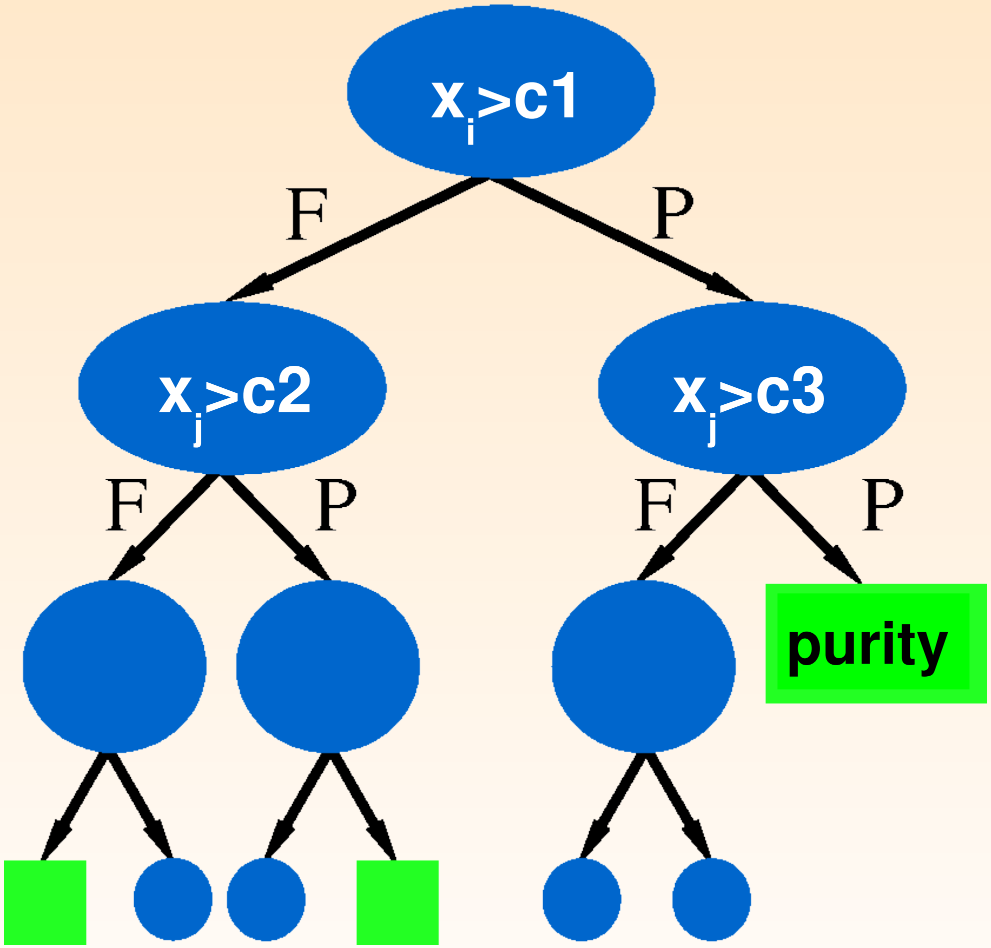
$$EPD = \frac{b \cdot P_{signal}(x)}{b \cdot P_{signal}(x) + b \cdot P_{b-bkg}(x) + (1-b) \cdot P_{non\ b-bkg}(x)}$$



Boosted Decision Tree

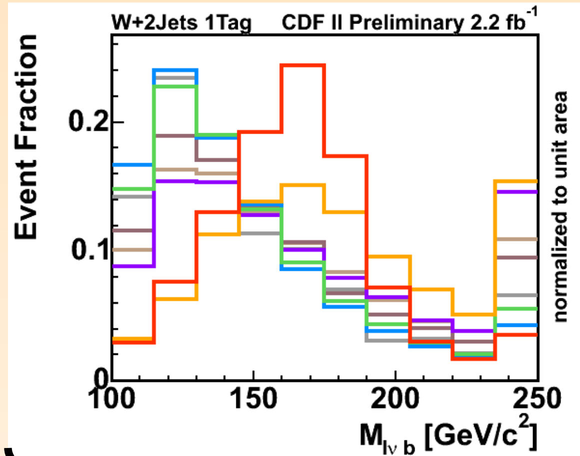
Idea: Effective extension of a cut-based analysis

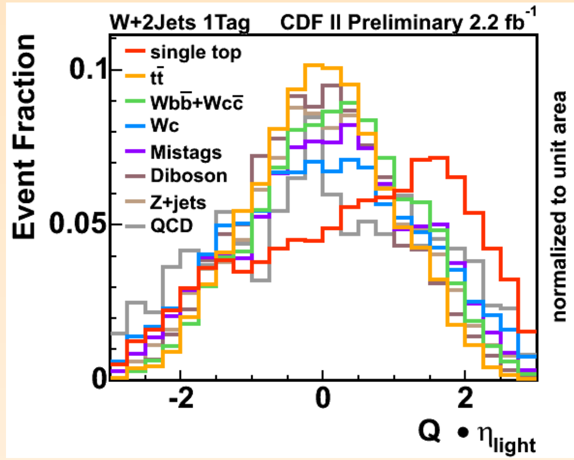
- Use many input variables (20)
- Non-discriminating variables are automatically ignored, but don't degrade the performance
- Optimize series of binary cuts with training sample
- Calculate for each leaf purity $p = s/(s+b)$
- Sort events by output purity
- Create series of “boosted“ trees by reweighting based on value of misclassification

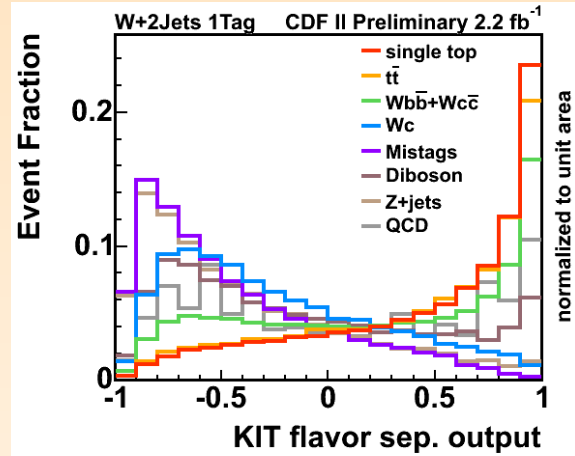


Neural Network Analysis

Idea: Combine many variables into one more powerful discriminant





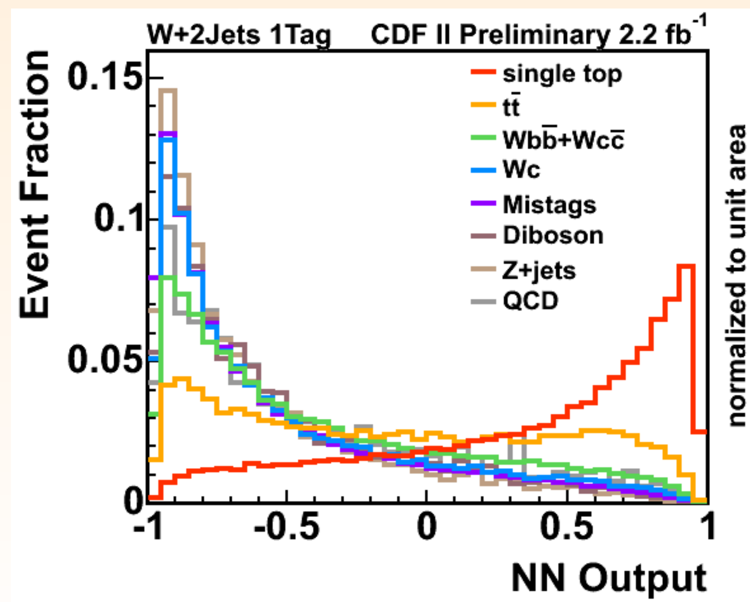


...

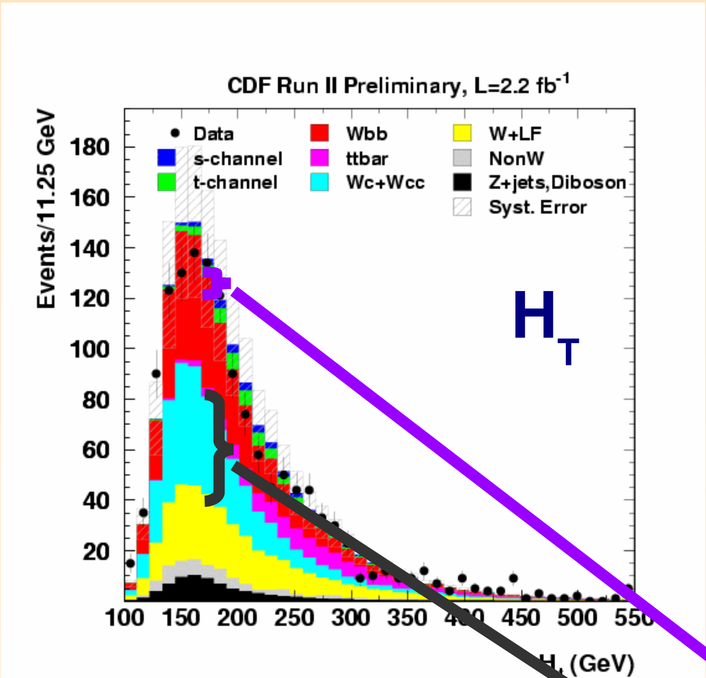
- **Network: NeuroBayes**

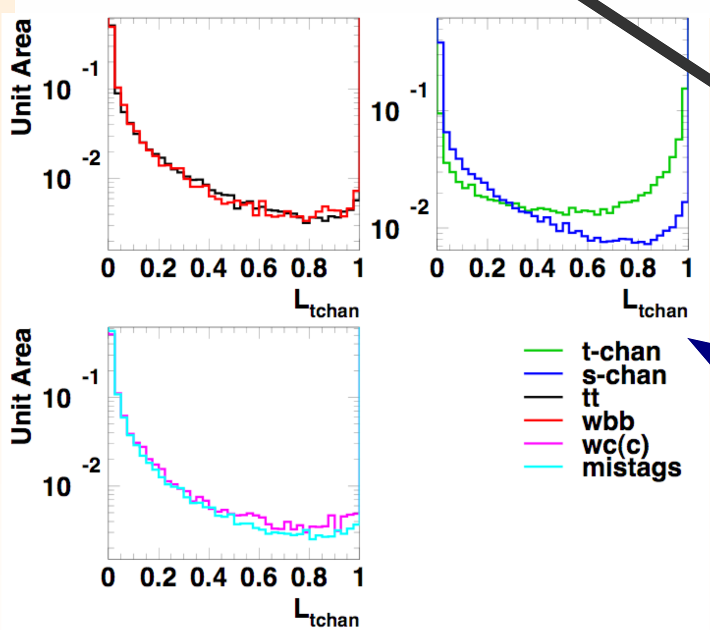
  (3-layered feed-forward net combined with automated preprocessing of input variables)
- **Exploits correlation between variables**
- **Realistic mixture of backgrounds**
- **Use 11-18 variables:**

  $M_{l\nu b}$, $Q \cdot \eta$, KIT flavor separator, ...



Likelihood Function





Idea: Combine many variables into one more powerful discriminant

- **Binned likelihood function** *(LEP technique)*
- **Correlations between variables not taken advantage of**
- **Use 7-10 variabes**

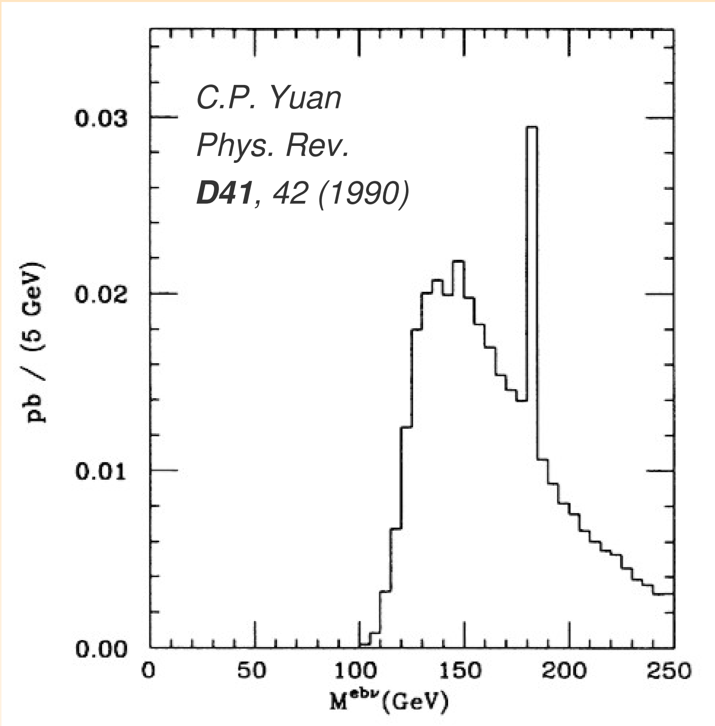
$$N_i^{sig} \quad N_{ij}^{bkg} \rightarrow p_i^{sig} = \frac{N_i^{sig}}{N_i^{sig} + \sum_{j=1}^{n_{bkg}} N_{ij}^{bkg}}$$

Norm. values

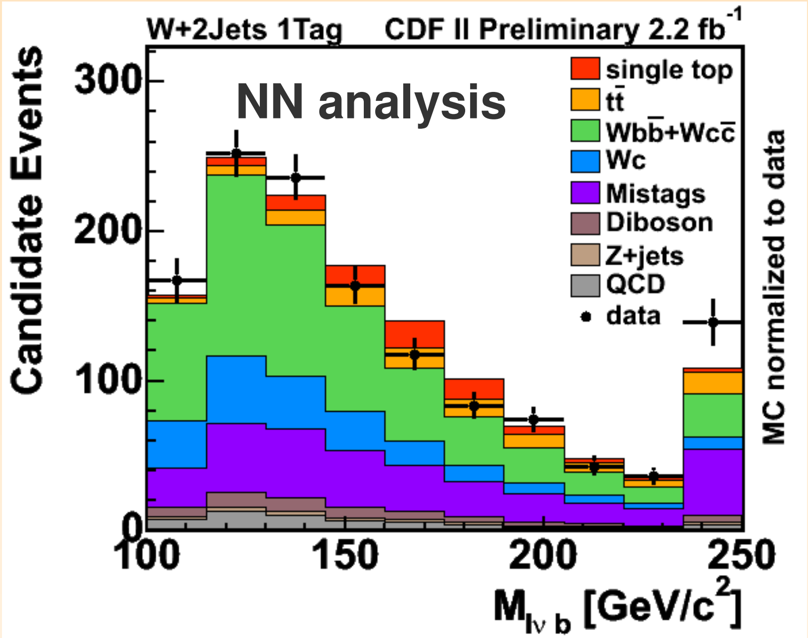
Bkg: Wbb,top pair, Wcc/Wc, mistags

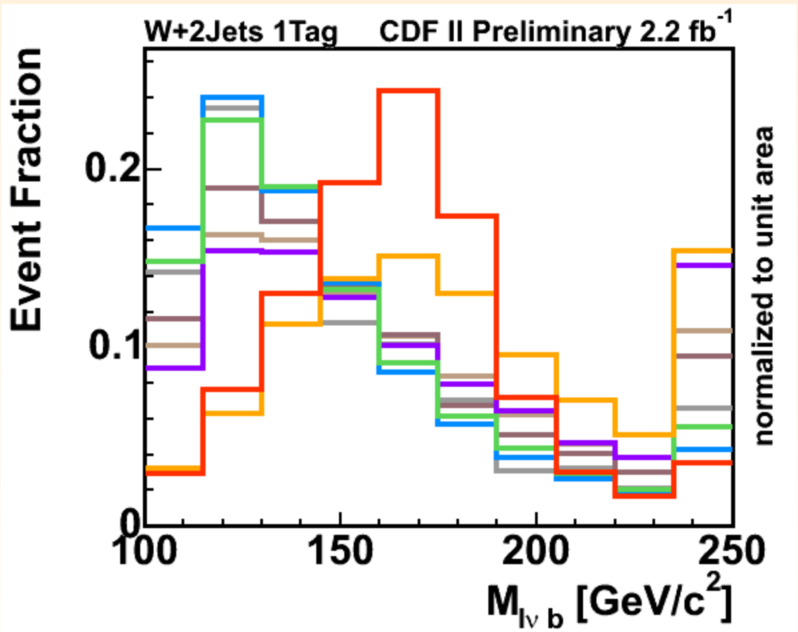
$$L(x) = \frac{\prod_{i=1}^{n_{var}} p_i^{sig}(x_i)}{\prod_{i=1}^{n_{var}} p_i^{sig}(x_i) + \sum_j^{n_{bkg}} \prod_{i=1}^{n_{var}} p_i^{bkg}(x_i)}$$

Separation Variable – $M_{l\nu b}$



No experimental resolution effects taken into account

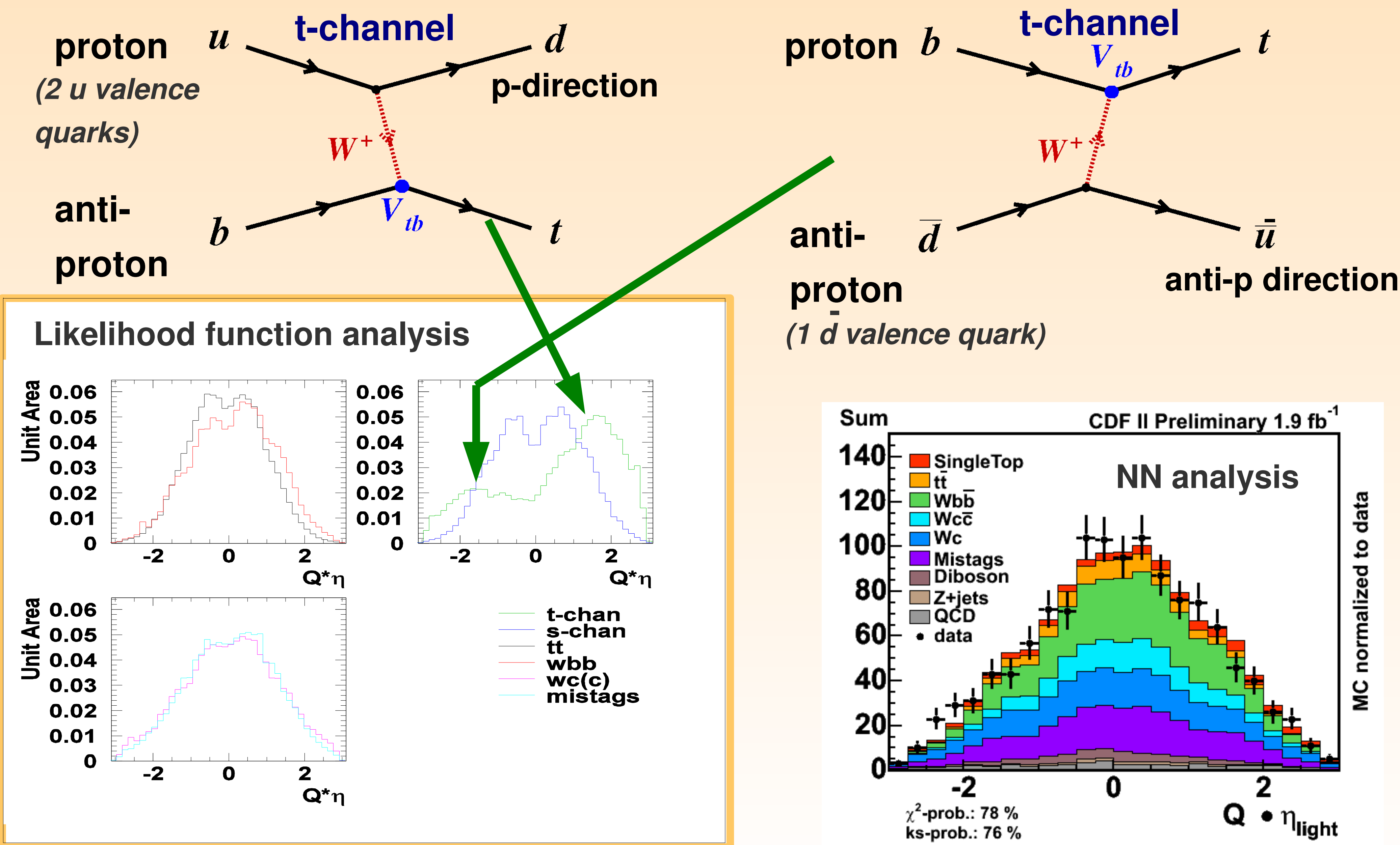




In reality signal peak degraded by:

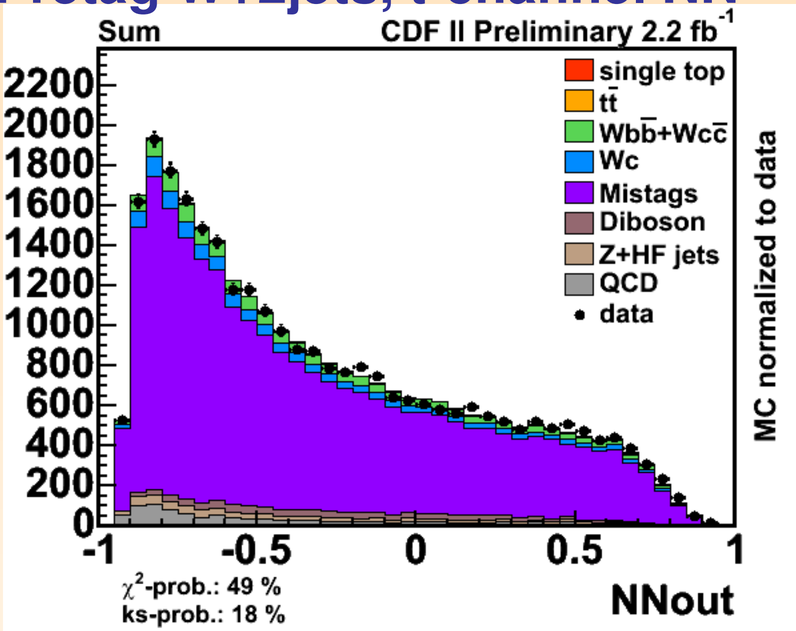
- **MET and jet resolution**
- **wrong jet assignment**
- **wrong $p_{z,\nu}$ choice**

Separation Variable - Q · η

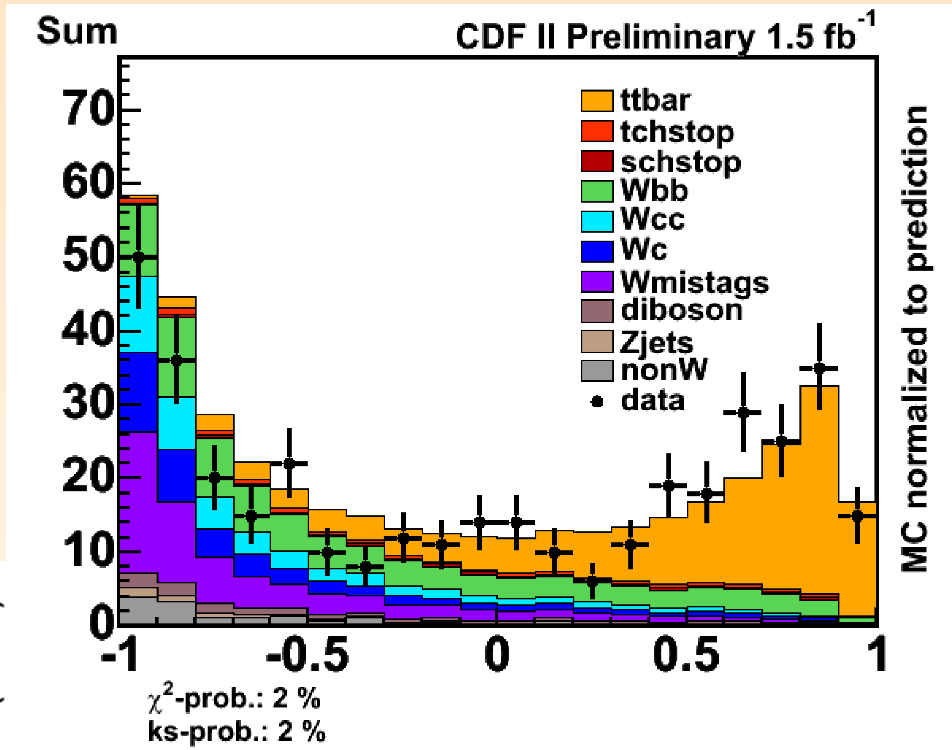


Validation of MVA - Neural Networks

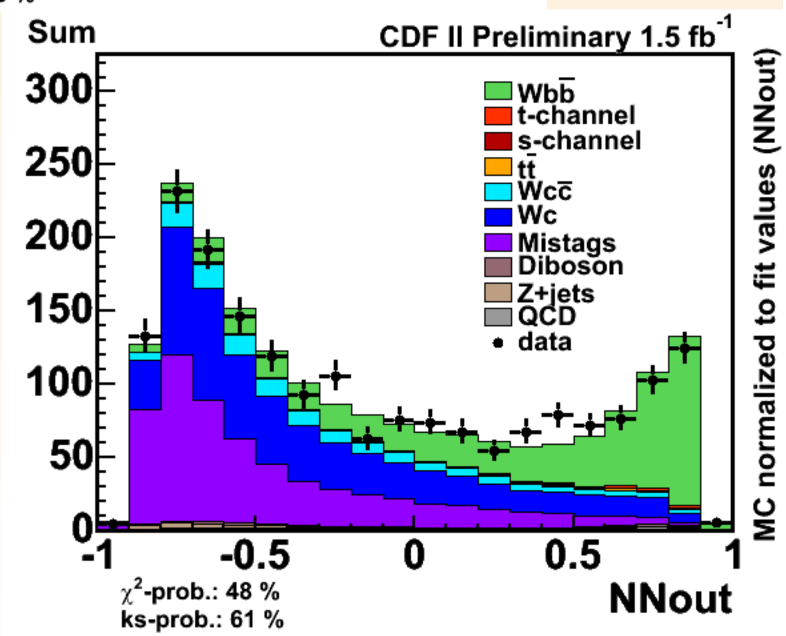
Pretag W+2jets, t-channel NN



W + 3 Jets, b-tagged, top pair NN



W + 1 Jet, b-tagged, Wbb NN



$\sigma_{\text{top pair}} \approx$ 7.5 ± 0.8 (stat.) pb

fit value in good agreement with CDF-average

Statistical Analysis

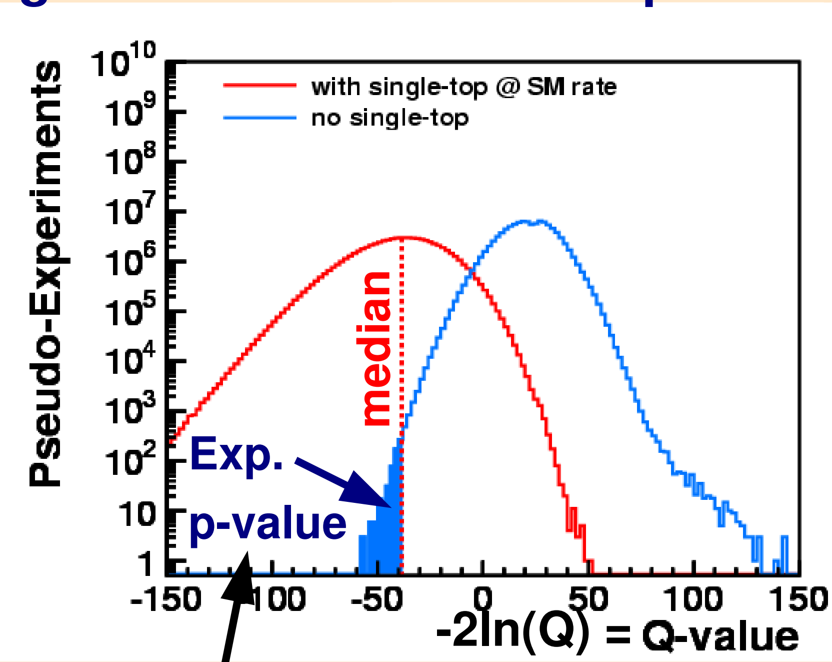
Cross section:

- **Bayesian treatment** *(flat prior in σ_{s+t})*
- **Binned likelihood fit including all rate and shape syst. uncertainties**

Significance:

- **Modified frequentist approach**
- **Perform pseudo experiments (PE) with and without SM single top**
  - *Fluctuate all syst. rate and shape uncertainties in PE*
  - *Binned likelihood fit for each PE (fit W+LF and W+HF norm. parameter)*
- $-2\ln(Q)$, $Q = L_{reduced}(\text{SM }\sigma_{s+t})/L_{reduced}(\sigma_{s+t}=0)$

Significance: Definition of p-value



p-value defined via likelihood ratio is the most powerful criteria to distinguish 2 hypotheses, *Neyman-Pearson Lemma*

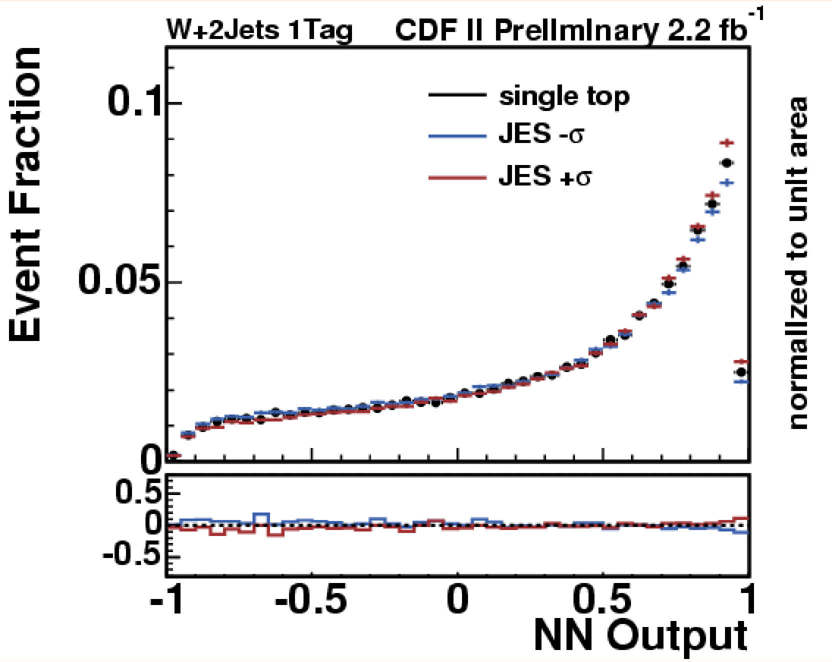
Systematic Uncertainties

Syst. Uncertainty	Rate	Shape
Jet Energy Scale	0...16%	✓
Initial state radiation	0...11%	✓
Final state radiation	0...15%	✓
Parton Distribution Function	2...3%	✓
MC Generator	1...5%	
Event Detection Efficiency	0...9%	
Luminosity	6%	
NN Flavor-Separator		✓
Mistag model		✓
Q2 scale in ALPGEN MC		✓
Input variable mismodeling		✓
Wbb+Wcc normalization	30%	
Wc normalization	30%	
Mistag normalization	17...29%	
Top-pair normalization & mtop	23%	✓

Different uncertainties for each subset

Shape uncertainty - example

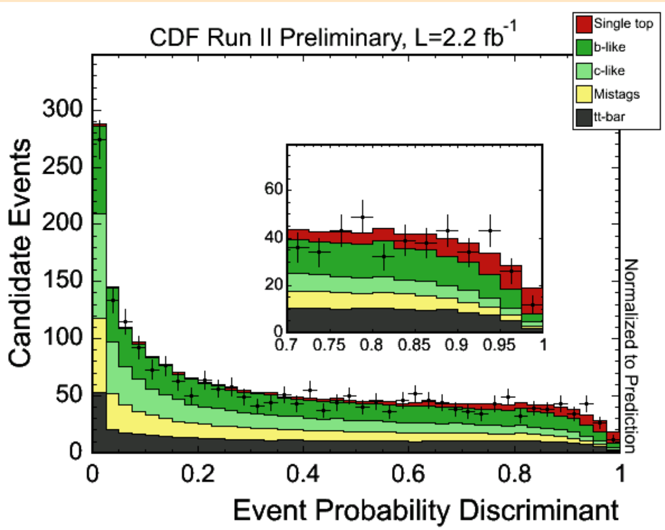
Jet energy scale, single top



Results of CDF

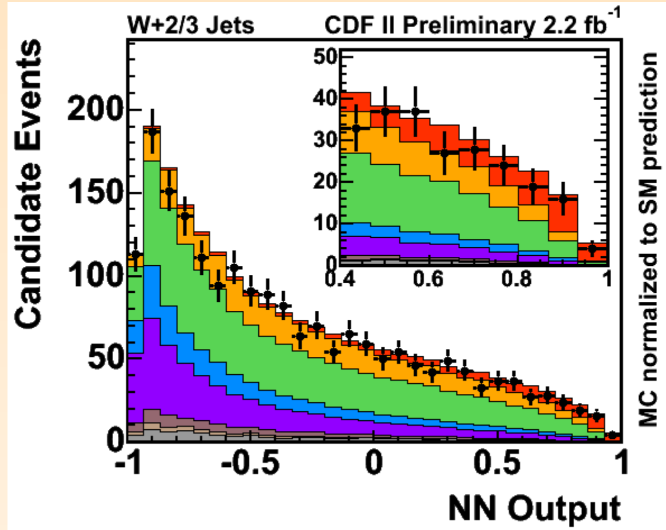
Measured Cross Sections

Matrix element method



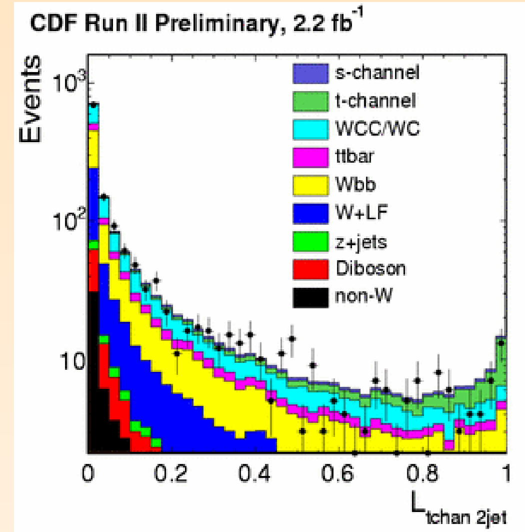
$$\sigma_{s+t} = 2.2^{+0.8}_{-0.7} \text{ pb}$$

Neural networks

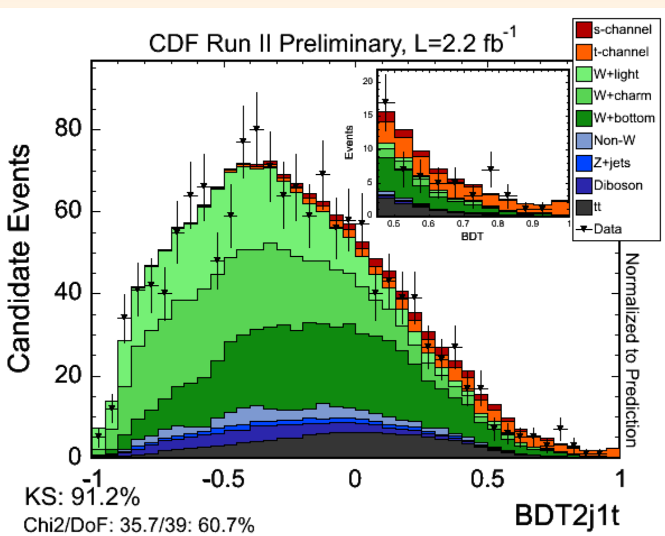


$$\sigma_{s+t} = 2.0^{+0.9}_{-0.8} \text{ pb}$$

Likelihood function



$$\sigma_{s+t} = 1.8^{+0.9}_{-0.8} \text{ pb}$$



Boosted decision tree

$$\sigma_{s+t} = 1.9^{+0.8}_{-0.7} \text{ pb}$$

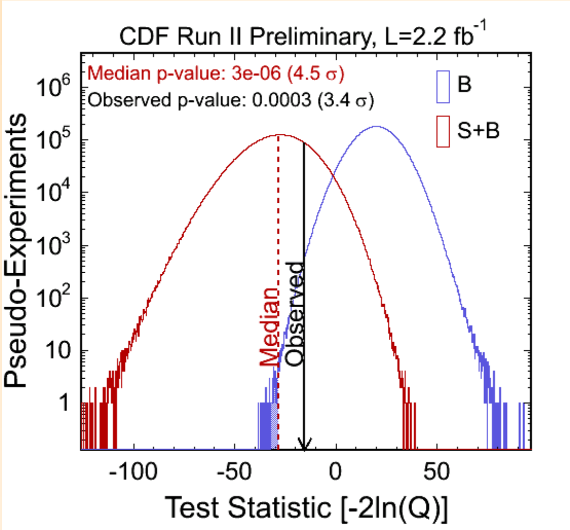
SM prediction:

$\sigma_{s+t} \approx 2.9 \pm 0.4$ pb

Analyses are consistent with each other

Significance

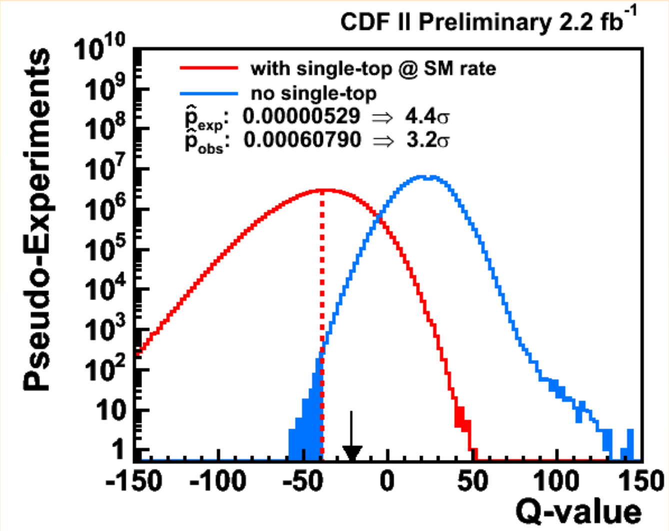
Matrix element method



Exp. p-value: 0.0003% (4.5σ)

Obs. p-value: 0.03% (3.4 σ)

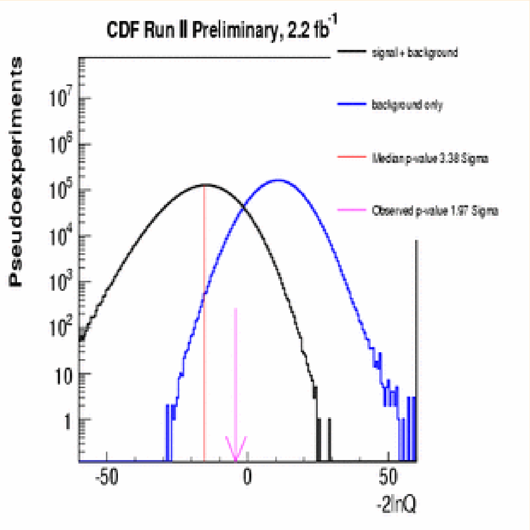
Neural networks



Exp. p-value: 0.0005% (4.4σ)

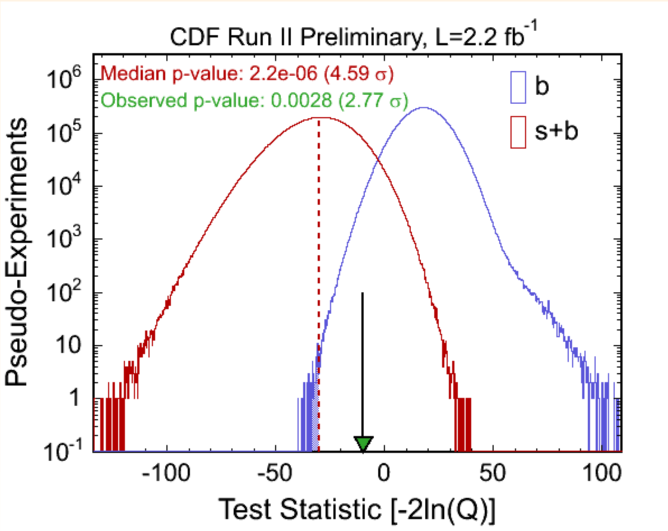
Obs. p-value: 0.06% (3.2 σ)

Likelihood function



Exp. p-value: 0.03% (3.4σ)

Obs. p-value: 2.5% (2.0 σ)



Boosted decision tree

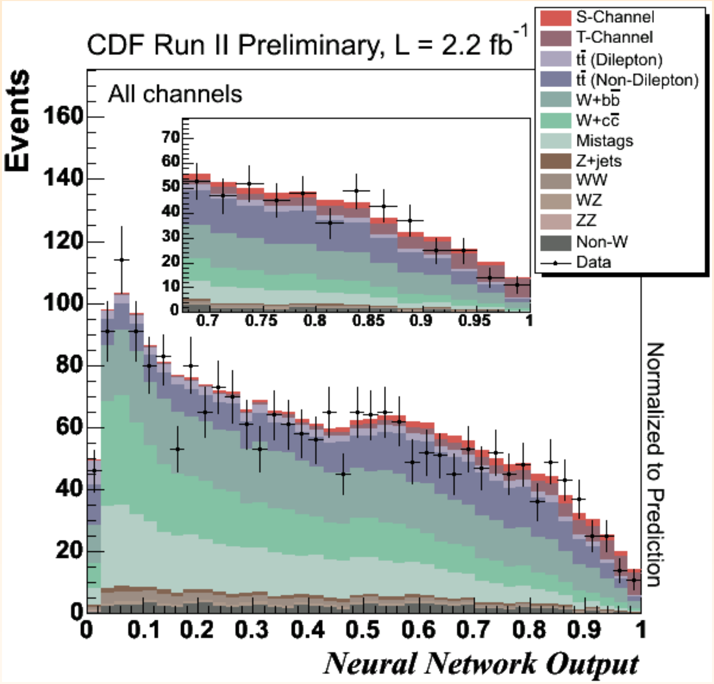
Exp. p-value: 0.0002% (4.6σ)

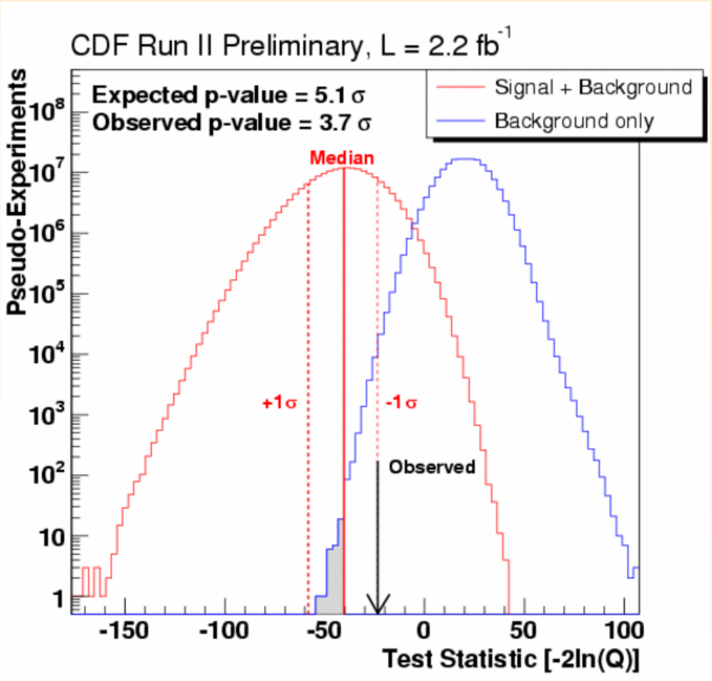
Obs. p-value: 0.28% (2.8 σ)

Combination – Super-Discriminant

Combine ME, NN, LF discriminants as inputs to a neural net (NEAT)

Neuro-**E**volution of **A**ugmenting **T**opologies: *designed to optimize the discovery significance*





Exp. p-value:

0.00002% (5.1σ)

Obs. p-value:

0.0094% (3.7 σ)

Cross section:

$$\sigma_{s+t} = 2.2^{+0.7}_{-0.7} \text{ pb}$$

SM prediction:

$\sigma_{s+t} \approx 2.9 \pm 0.4$ pb

σ_{s+t} lower than SM prediction but consistent with SM prediction

Combination – Checks

Original BLUE:
L. Lyons, NIM A270, 110 (1988)

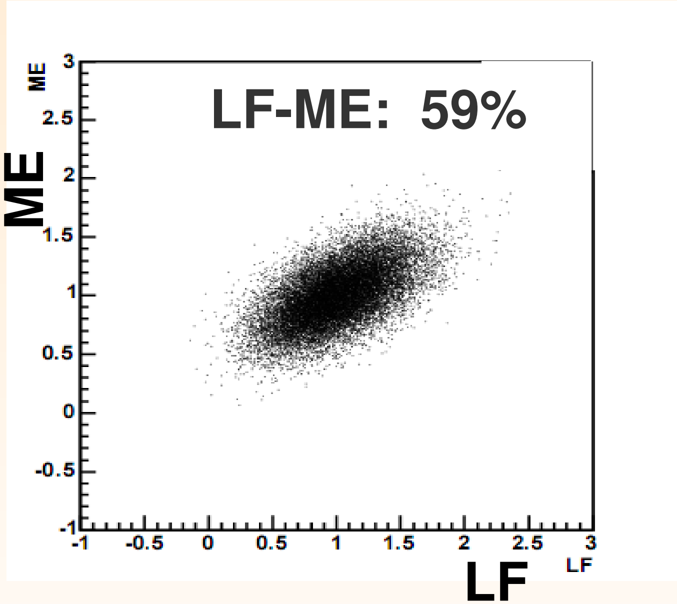
Use **B**est**L**inear**U**nbiased**E**stimator to estimate:

- Compatibility between three analyses and compatibility with SM
- Correlation between single analyses

Compatibility:

- Perform PEs generated assuming a certain single top cross section
- Determine χ^2 of combination for each PE
- Compatibility = fraction of PE with a larger χ^2 than in data

Correlation



61% (ME-NN) , 74% (LF-NN)

Discrepancy $\lesssim 1.1\,\sigma$

Self consistency: 87 %
(σ_{s+t} = 2 pb)

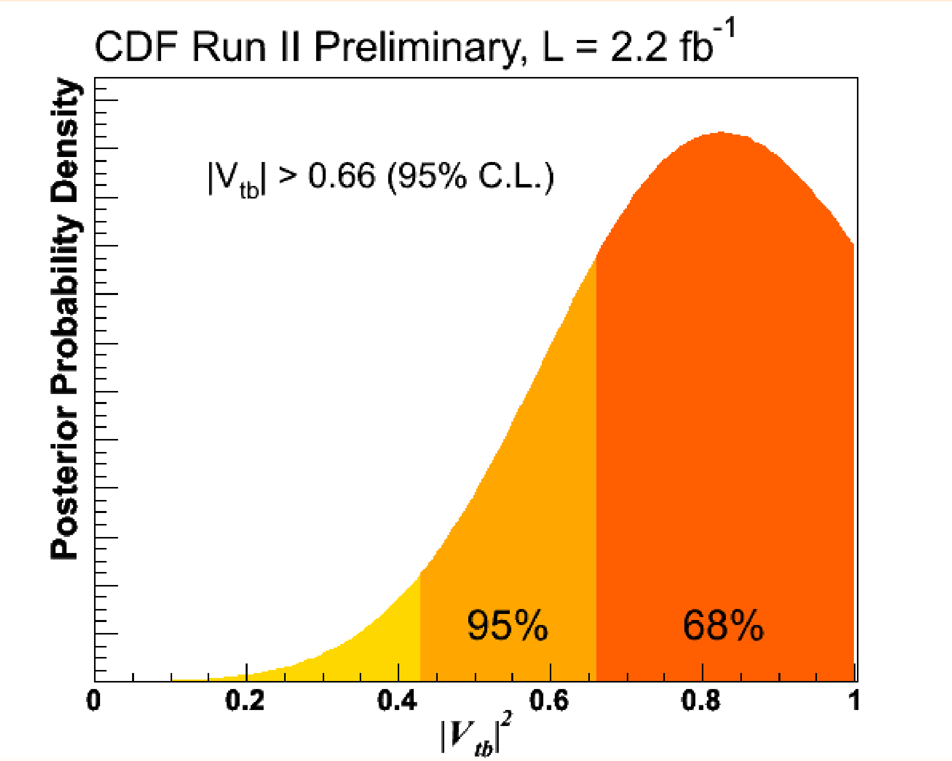
Compatibility with SM: 15 %
(σ_{s+t} = 2.9 pb)

V_{tb} Measurement

Assuming SM production:

- **Pure V-A and CP conserving interaction**
- $|V_{tb}|^2 \gg |V_{td}|^2 + |V_{ts}|^2$ **or** $B(t \to Wb) \sim 100\%$

$$|V_{tb,meas}|^2 = \frac{\sigma_{meas}}{\sigma_{SM}} \cdot |V_{tb,SM}|^2$$



Super-discriminant:

$|V_{tb}| = 0.88 \pm 0.14$ (exp.) ± 0.07 (theory)

$|V_{tb}| > 0.66$ (95% C.L.)

Measurement consistent with SM prediction, $V_{tb} \approx 1$

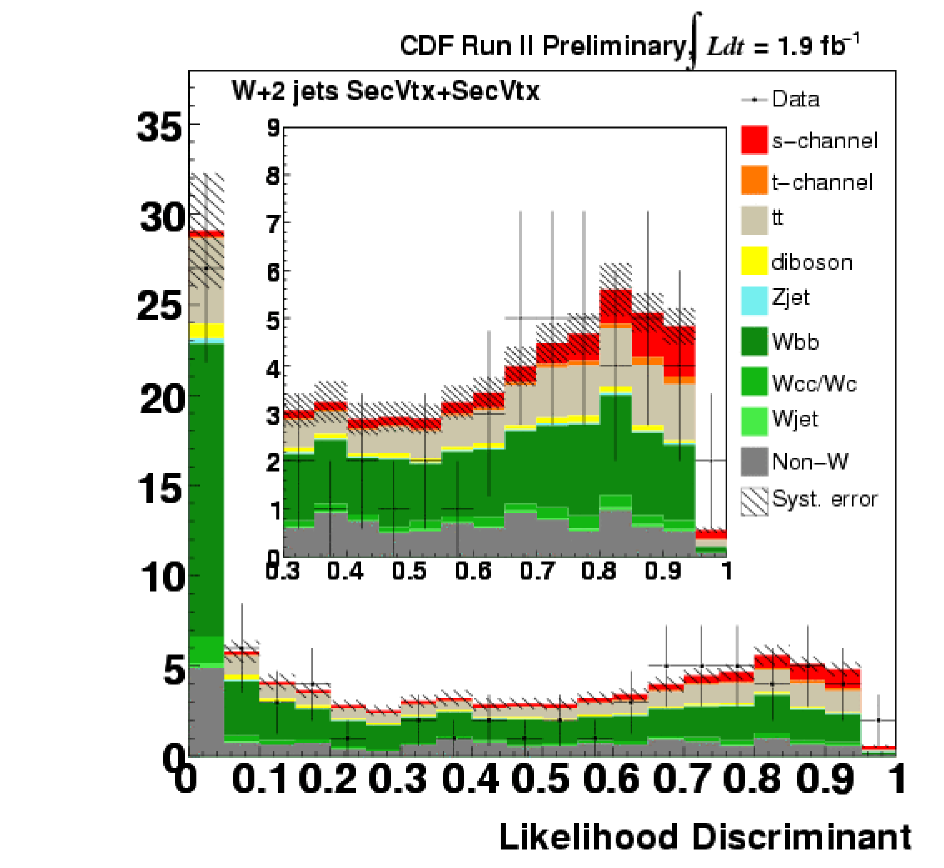
Separate s-Channel Search

No constraint of SM cross section ratio

- Use double b-tagged events with
  - 2 secondary vertices
  - 1 sec. vertex and a "loose" b-tag

    (calc. from d_0 of tracks assigned to jet)
- No QCD veto applied

  (QCD bkg small, acceptance gain)
- Use only W+2jet events
- Use Likelihood function as multi-variate analysis



s-channel:

$$\sigma_s = 0.6^{+0.9}_{-0.6} \text{ pb}$$

(t-channel fixed to SM, treated as bkg)

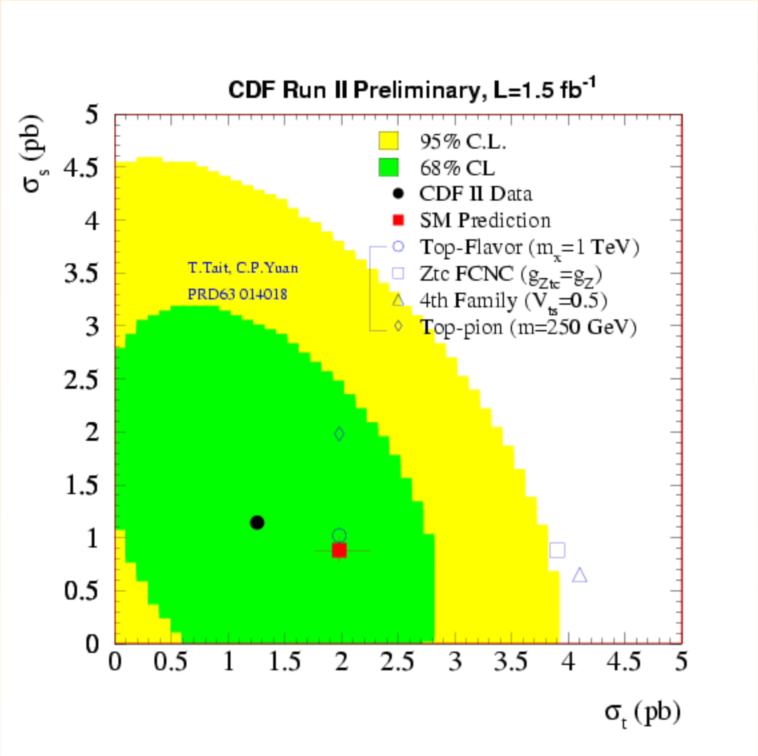
Measured σ_s consistent with SM prediction, σ_s = (0.88 ± 0.11) pb

Separate s- and t-Channel Search

Measure both channels simultaneously

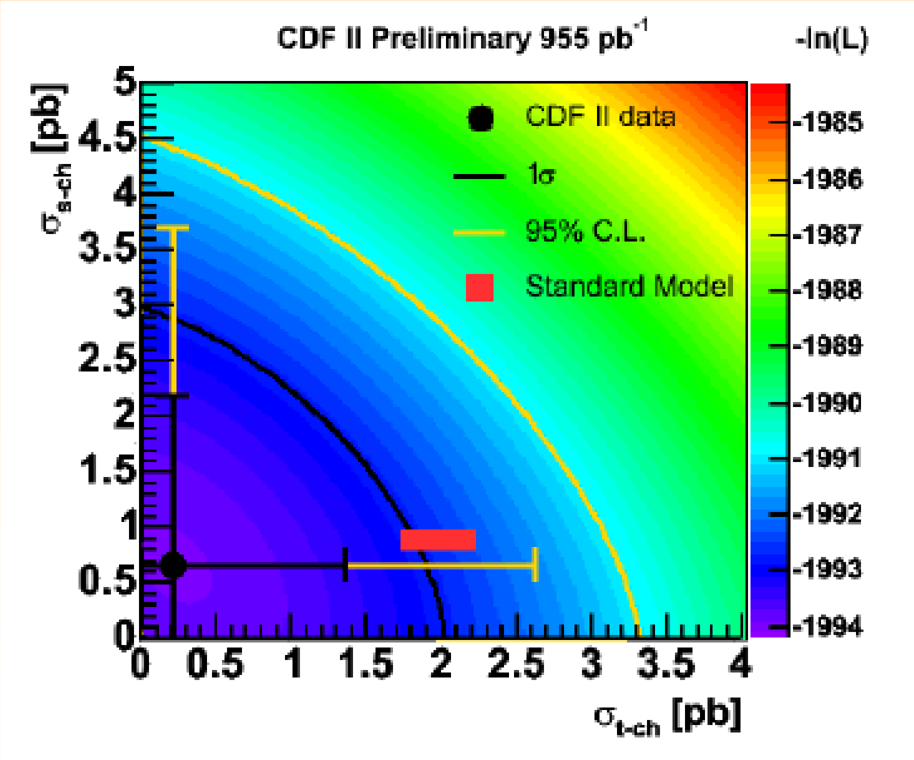
(no constraint on the cross section ratio)

Likelihood function



New results coming soon

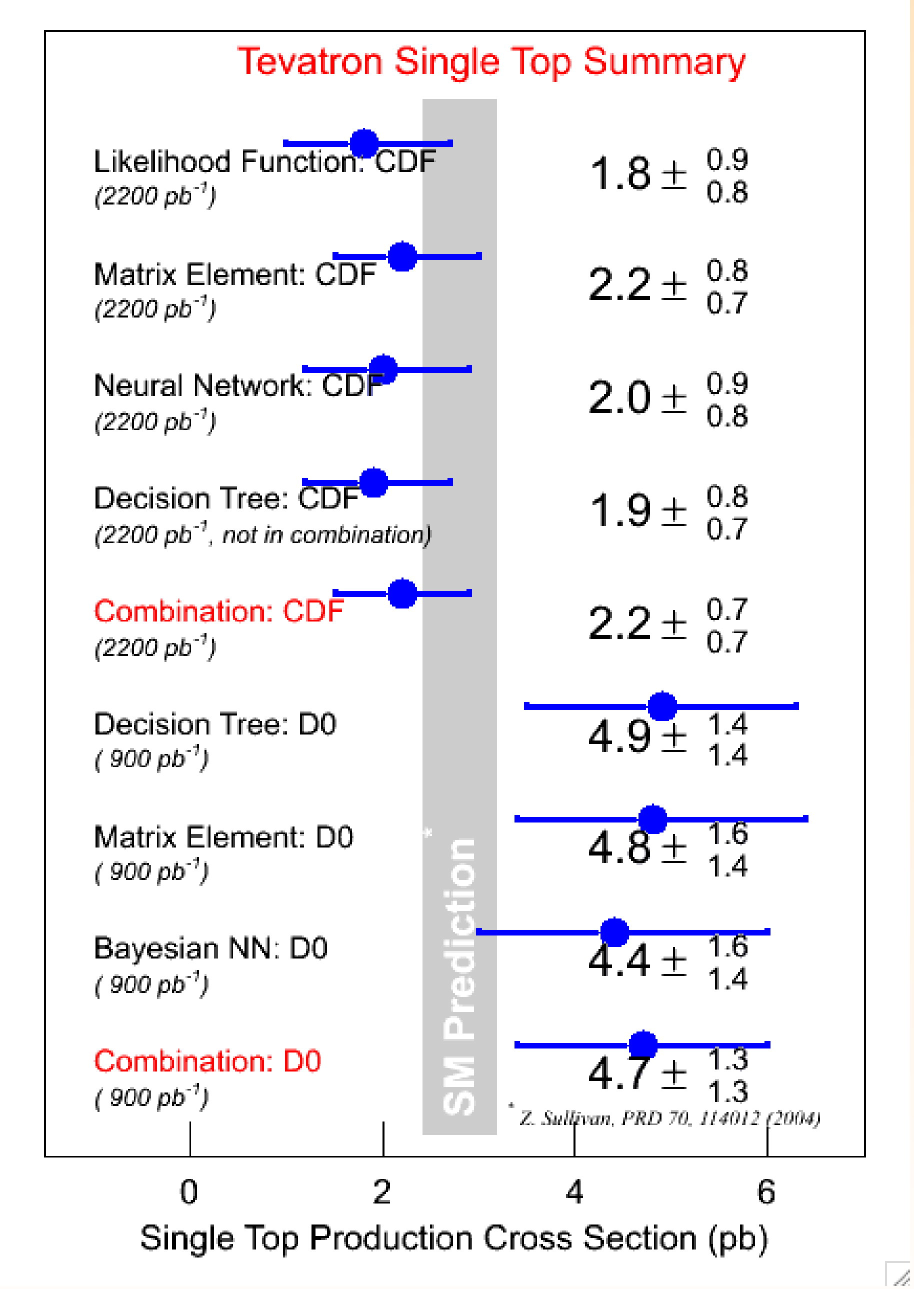
Neural networks



s−channel ≈ SM prediction

t−channel : lower than SM prediction

Summary of CDF Results



Combined search:

Cross section: $\sigma_{s+t} = 2.2^{+0.7}_{-0.7}$ pb

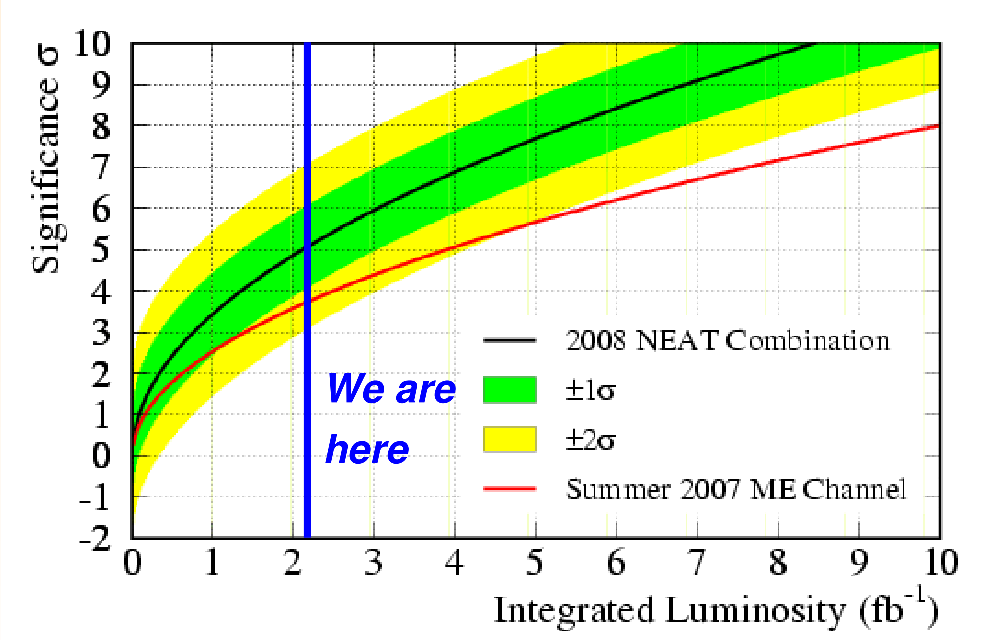
Exp. p-value: 5.1σ Obs.: 3.7 σ

$|V_{tb}| = 0.88 \pm 0.16$

Separate search:

σ_s ≈ SM prediction

σ_t : lower than SM prediction

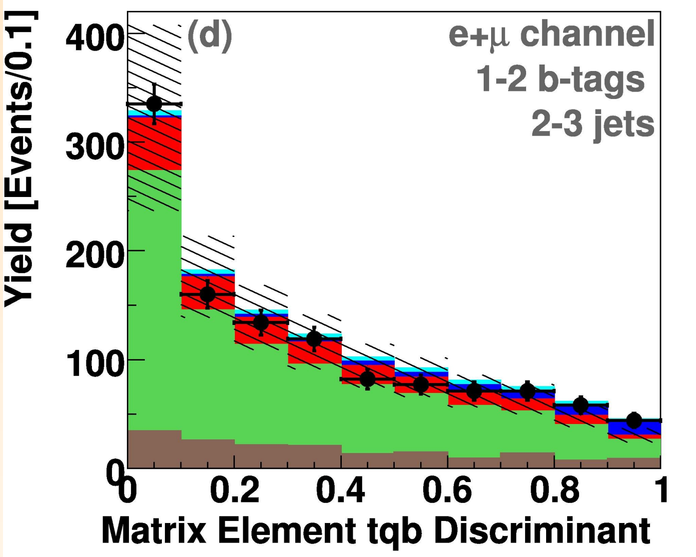


Results from DO

Measured Cross Sections

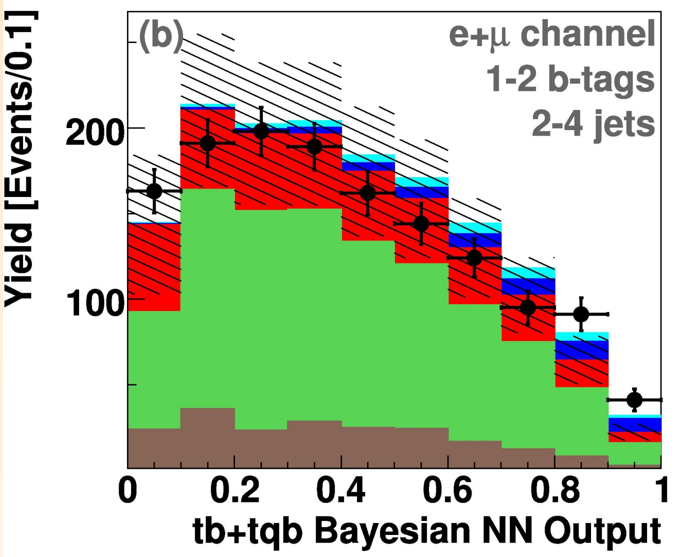
Measured using a Bayesian binned likelihood calculation

Matrix element method



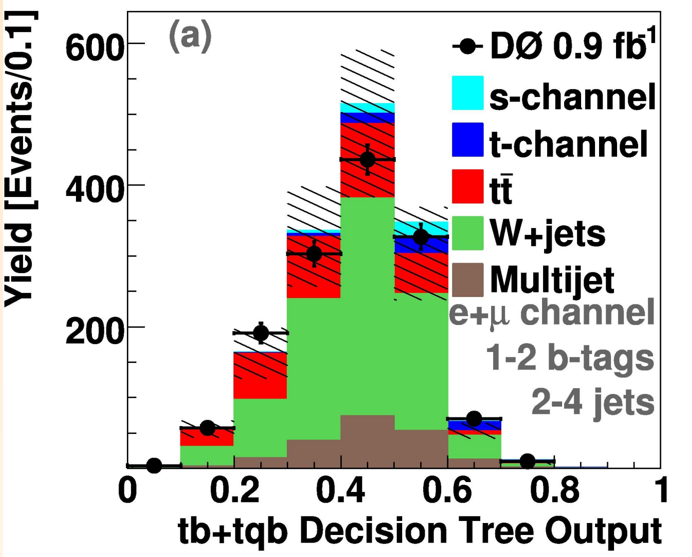
$$\sigma_{s+t} = 4.8^{+1.6}_{-1.4} \text{ pb}$$

Bayesian neural network



$$\sigma_{s+t} = 4.4^{+1.6}_{-1.4} \text{ pb}$$

Boosted decision tree



$$\sigma_{s+t} = 4.9^{+1.4}_{-1.4} \text{ pb}$$

PRL 98, 18102 (2007), PRD article available at arXiv: 0803.0739 [hep-ex]

SM prediction:

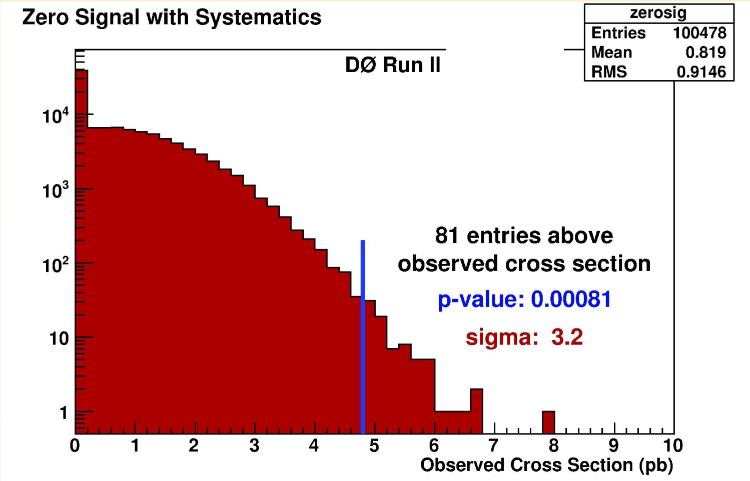
$\sigma_{s+t} \approx 2.9 \pm 0.4$ **pb**

Measurements are consistent with each other

Significance

Determined using ensemble tests without signal contribution

Matrix element method



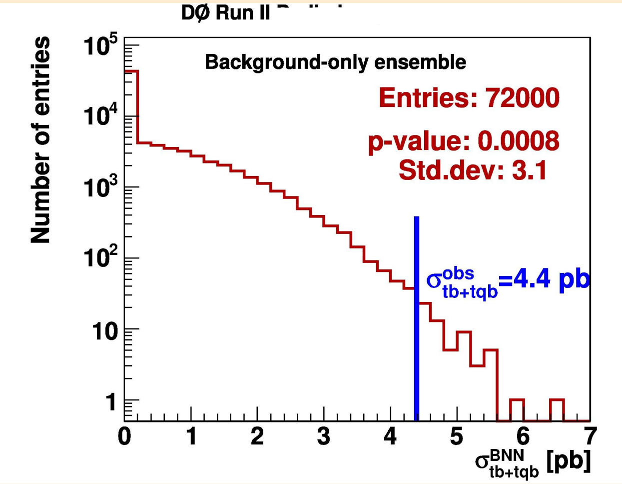
Expected p-value:

3% (1.9 σ)

Observed p-value:

0.081% (3.2 σ)

Bayesian neural network



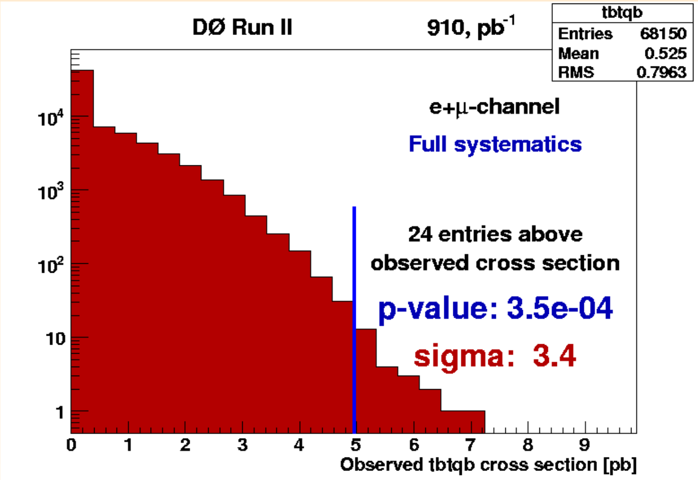
Expected p-value:

1.6% (2.2 σ)

Observed p-value:

0.08% (3.1 σ)

Boosted decision tree



Expected p-value:

1.9% (2.1 σ)

Observed p-value:

0.035% (3.4 σ)

Combination and V_{tb}

Combination:

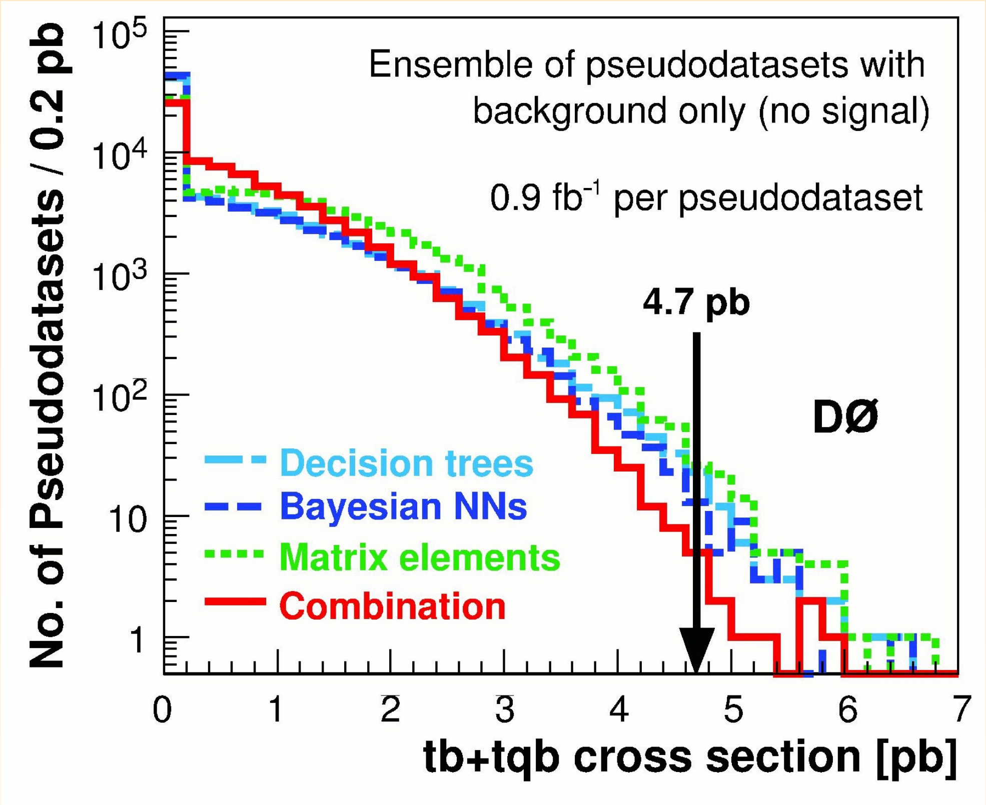
- **Combine results with BLUE**
- **Correlations:**

  59% (ME-BNN) – 66% (BNN-BDT)

Cross section:

$$\sigma_{s+t} = 4.7 \pm 1.3 \text{ pb}$$

Exp. p-value: 2.3 σ , Obs: 3.6 σ



Boosted decision tree analysis used to determine V_{tb} :

$$|V_{tb}| = 1.31^{+0.25}_{-0.21} \text{ pb}$$

- **σ_{s+t} higher than SM prediction but consistent with SM prediction**
- **$|V_{tb}|$ consistent with SM pred.**

Separate Search

No constraint of SM cross section ratio

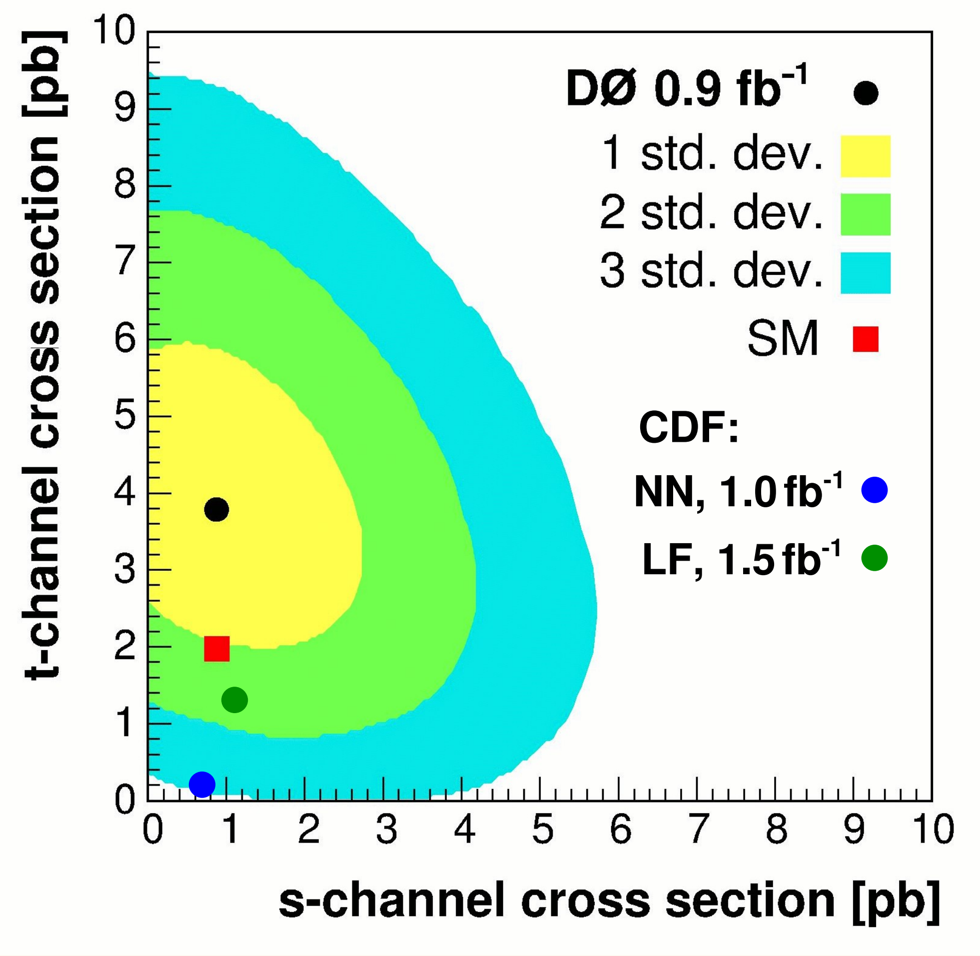
Boosted decision tree analysis

Measure only one channel and assume SM value for non-measured channel *(treated as background)*

$$\sigma_s = 1.0^{+0.9}_{-0.9}\ \text{pb}$$
$$\sigma_t = 4.2^{+1.8}_{-1.4}\ \text{pb}$$

Measure both channels simultaneously

$$\sigma_s = 0.9\ \text{pb}$$
$$\sigma_t = 3.8\ \text{pb}$$



s−channel ≈ SM prediction

t−channel ≈ 2 · SM prediction

Summary

Evidence for single top quark production

Cross section:

$$\sigma_{s+t} = 2.2^{+0.7}_{-0.7} \text{ pb}$$

Exp. p-value: 5.1σ Obs.: 3.7 σ

DO: $\sigma_{s+t} = 4.7 \pm 1.3$ pb
Exp. p-value: 2.3σ Obs.: 3.6 σ

V_{tb} CKM matrix element:

$|V_{tb}| = 0.88 \pm 0.14$

DO: $|V_{tb}| = 1.31^{+0.25}_{-0.21}$

Separate search:

$\sigma_s \approx$ SM prediction

σ_t lower than SM prediction

DO: $\sigma_t \approx 2 \cdot$ SM prediction