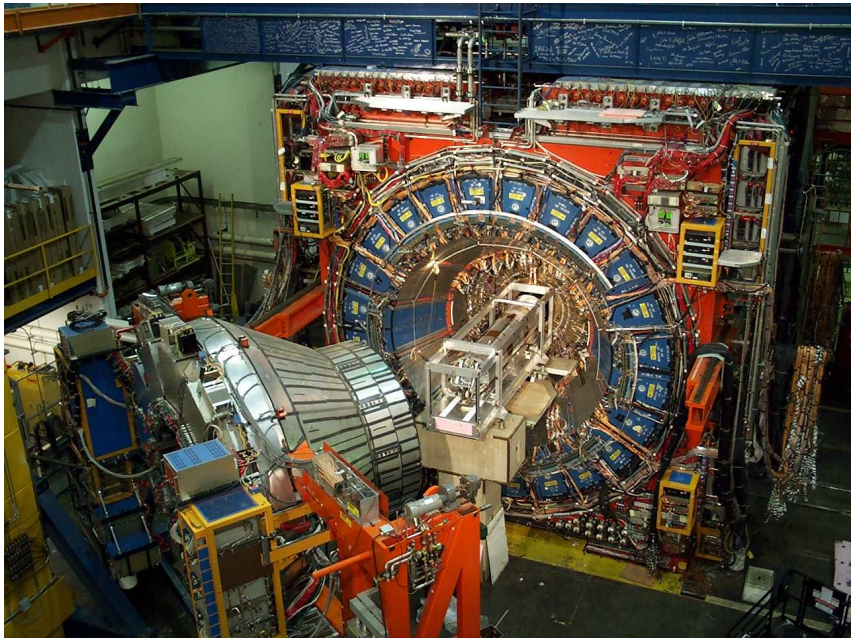
Observation of Single Top at CDF

EPS-HEP2009 • Kraków, Poland • July 16-22, 2009

Bruno Casal Laraña

Instituto de Física de Cantabria
(CSIC-Univ. de Cantabria)

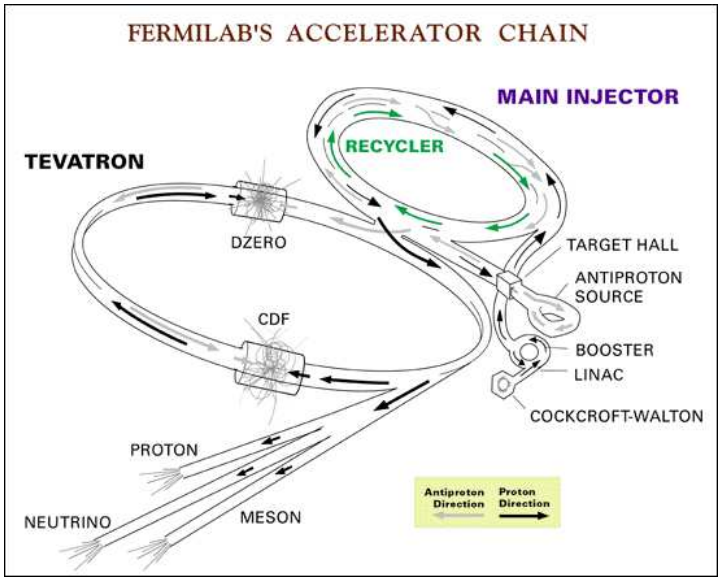
On behalf of the CDF Collaboration

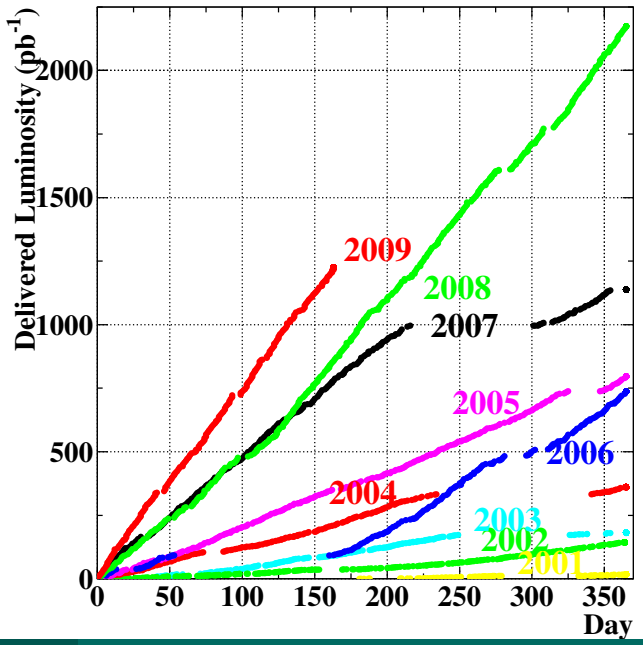


Instituto de Física de Cantabria

THE TEVATRON AT FERMILAB

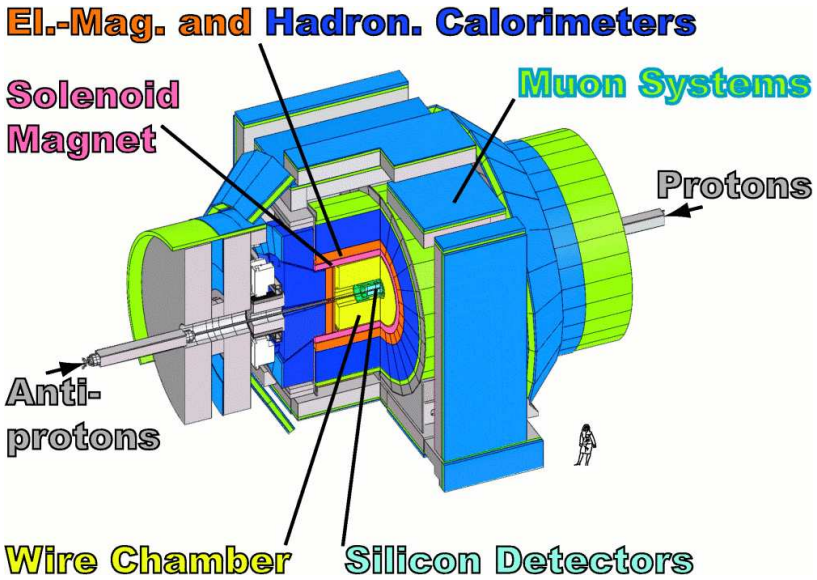
- $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV
- Peak luminosity $\sim 360 \times 10^{30} \mathrm{cm}^{-2}\mathrm{s}^{-1}$
- Tevatron delivered ~ 6.9 fb^{-1}
- CDF collected ~ 5.7 fb^{-1}
  - Analysis shown here use 3.2 fb^{-1}
- Tevatron is performing extremely well
  - Expected >10 fb^{-1} by end of 2010
  - Possibly run in 2011 and beyond

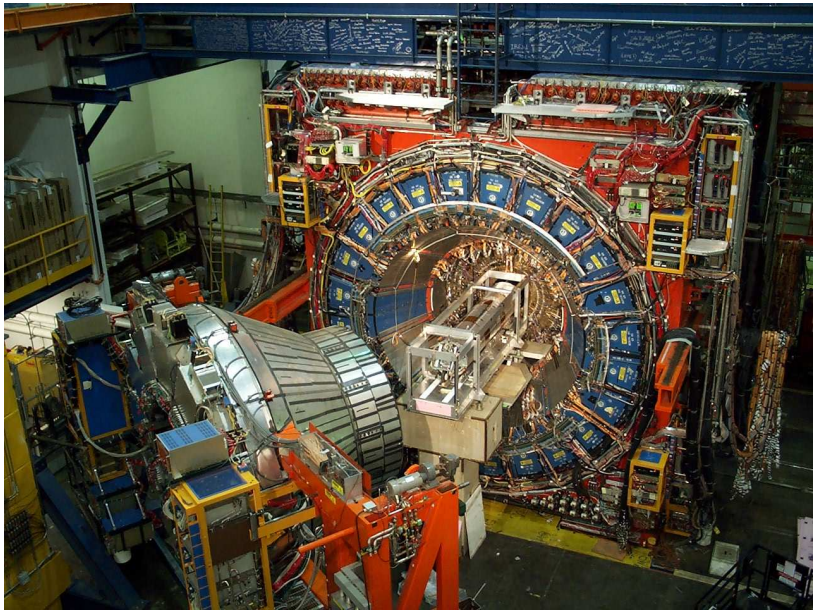




THE CDF II DETECTOR

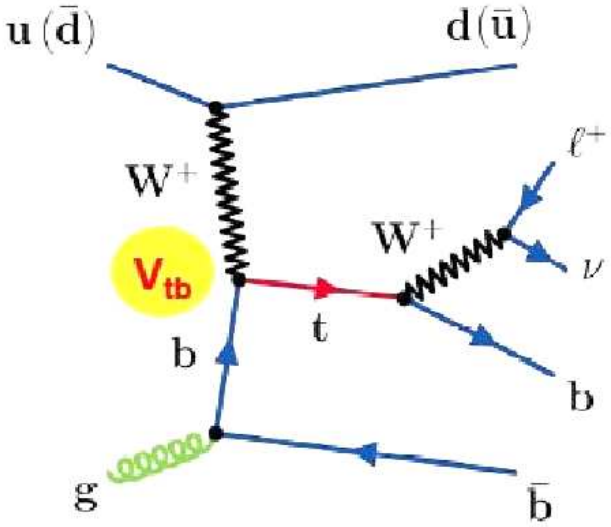
- Multipurpose detector at Tevatron collecting data efficiently
  - Large acceptance and good ID for leptons
    - Tracking and EM calorimeter
    - Muon systems
  - Good calorimetry for jet energy resolution
  - Silicon detectors for b-jet tagging



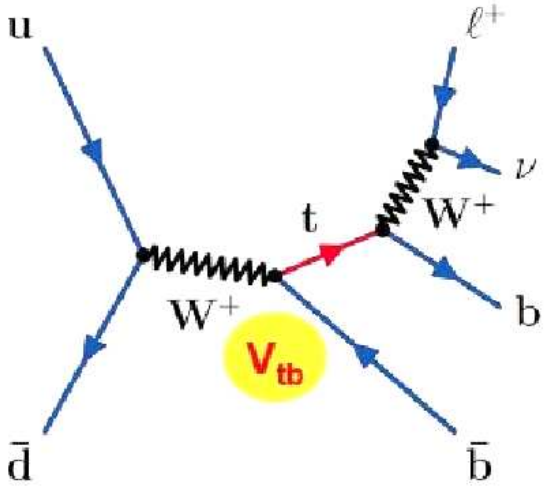


Motivation

- Single top quark produced via electroweak interaction
  - σ_{SM}(t-channel)$= 1.98 \pm 0.21$ pb
  - σ_{SM}(s-channel)$= 0.88 \pm 0.07$ pb
  - $\sigma_{SM}(t\bar{t}) = 6.7 \pm 0.8$ pb (via strong interaction)
  - [Z. Sullivan, Phys. Rev. D70, 114012 (2004)] ($m_t = 175$ GeV)
- Test of the Standard Model
  - Direct measurement of $|V_{tb}|$
  - Top quark properties: polarization, spin, W helicity,...
  - Same final state as WH
- Sensitive to new physics
  - Search for W', H^+ (s-channel signature)
  - Search for FCNC,...



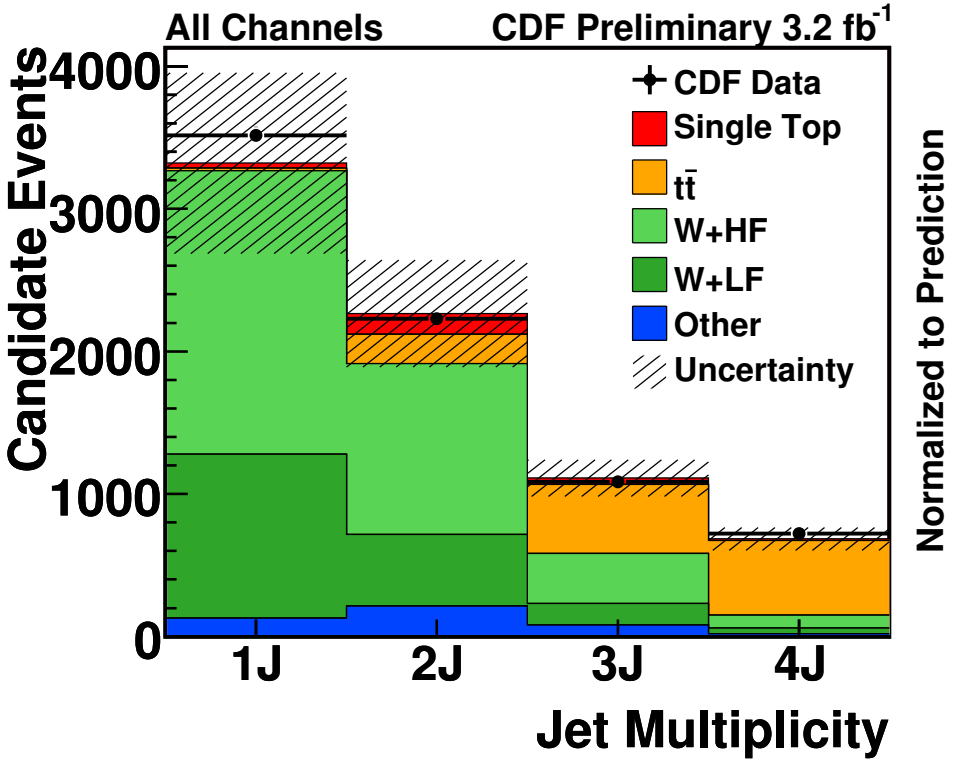
t-channel



s-channel

The Challenge

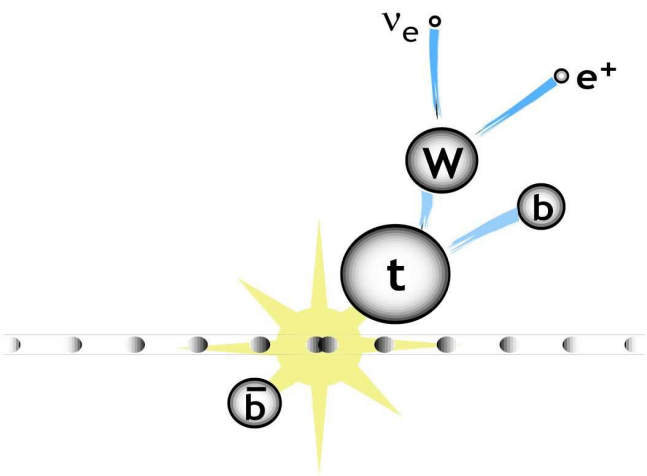
- Rare process at the Tevatron
  - $S : B \sim 1 : 10^9$ before trigger
- Not very distinct signature with many sources of background
  - $S : B \sim 1 : 20$ after final selection
  - No "golden" variable
- Large systematics uncertainties



- Sophisticated analysis methods needed
- Current approaches at CDF: Likelihood Function, Matrix Element, Neural Network, Boosted Decision Trees
  - $S : B > 1 : 1$ in most significant bins

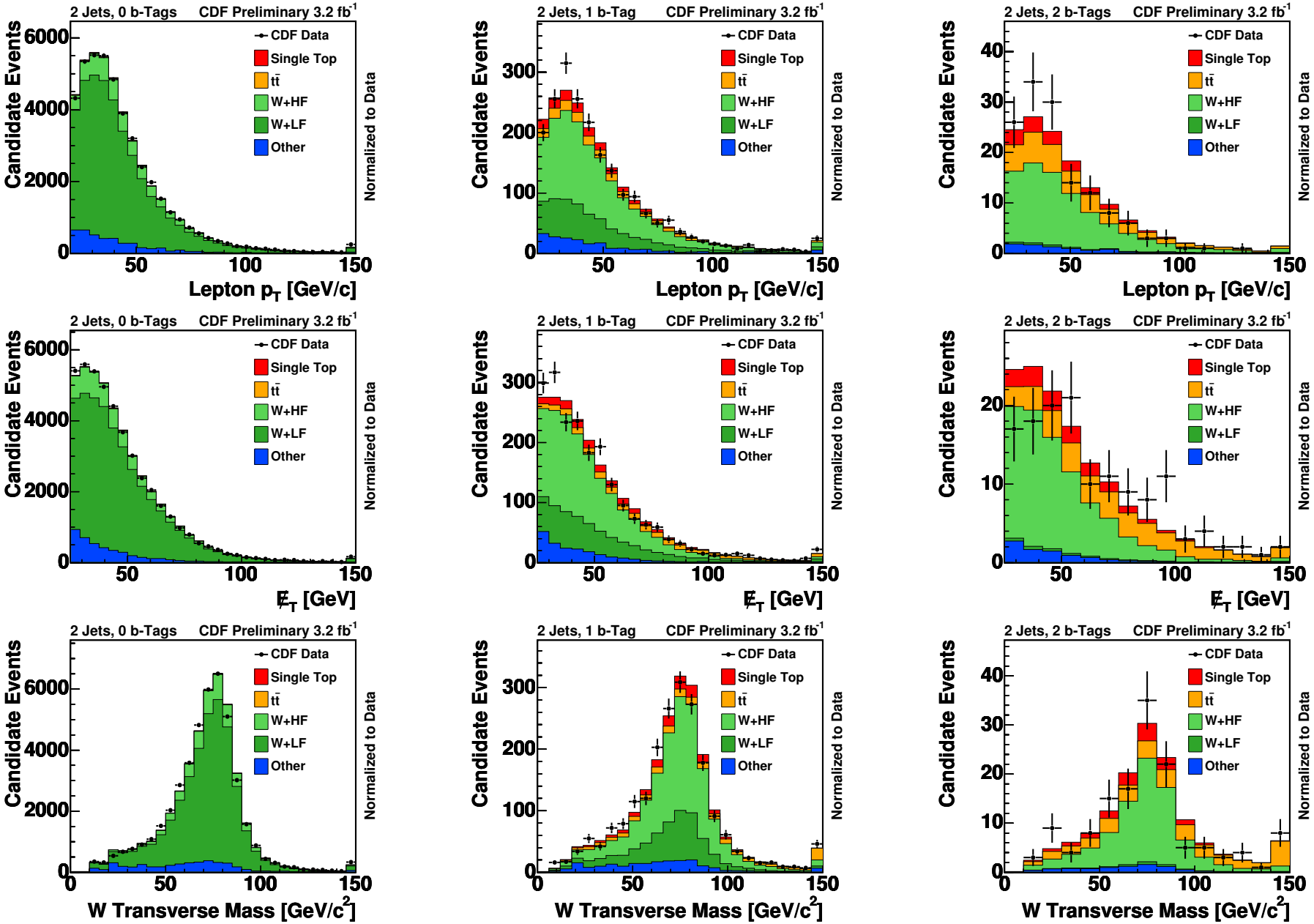
Single Top Selection at CDF

- Top decays most of the times to Wb
- W + 2 or 3 jets with $E_T > 20$ GeV
- One lepton (electron or muon) with $p_T > 20$ GeV
  - Mainly from standard high p_T lepton triggers
  - Extended muon coverage from missing E_T dedicated triggers (30% gain in signal acceptance)
- Missing transverse energy from neutrino, $\not{E}_T > 25$ GeV
- Veto "non-W", Z, dilepton, conversion, cosmics
- At least one b-tagged jet (displaced secondary vertex algorithm)
- Main backgrounds: W+Heavy Flavor, W+Mistags, $t\bar{t}$,...

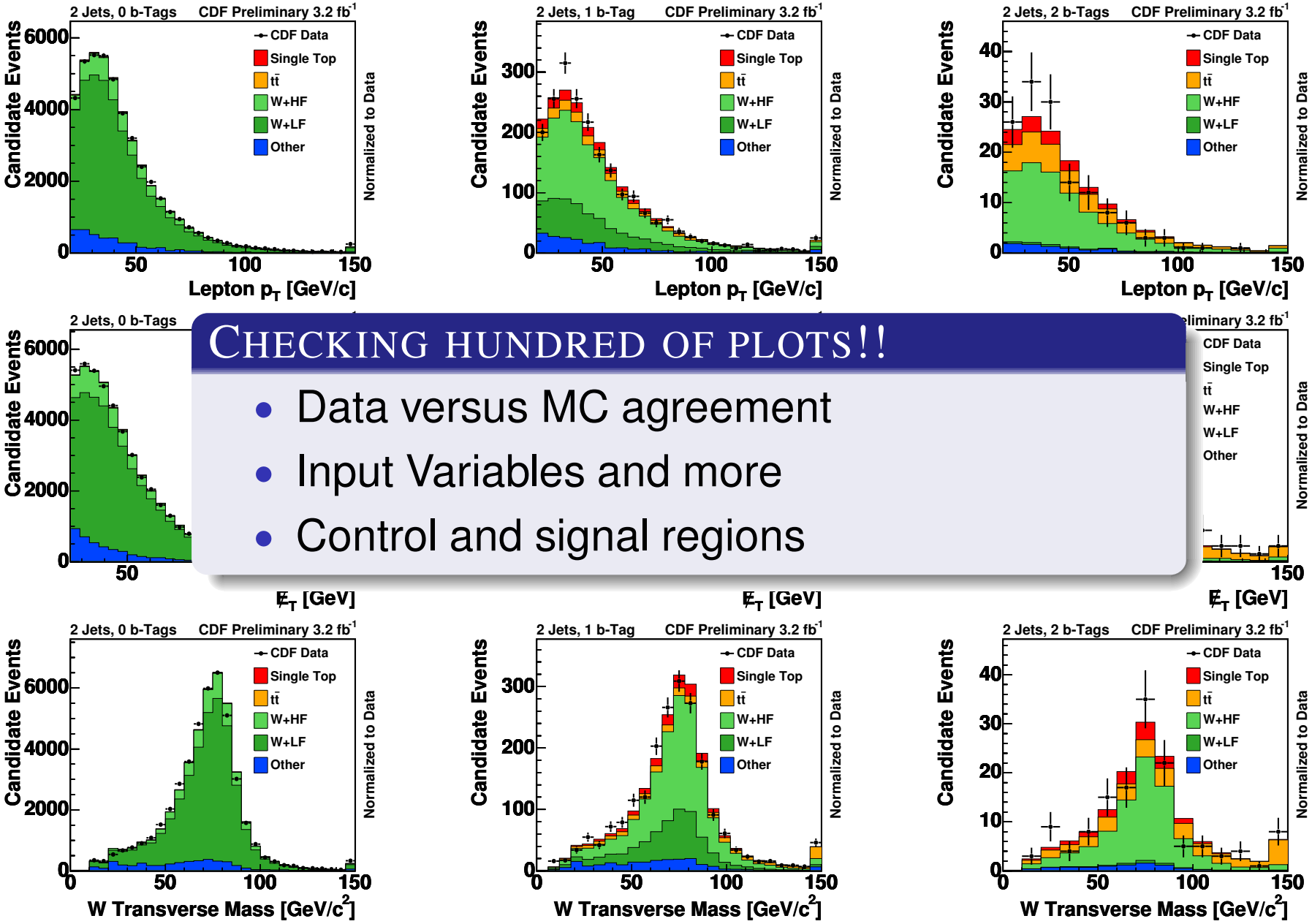


Process	Number of Events in 3.2 fb^{-1}	
	W + 2 jets	W + 3 jets
s-channel	58.1 ± 8.4	19.2 ± 2.8
t-channel	87.6 ± 13.0	26.2 ± 3.9
$Wb\bar{b}$	656.9 ± 198.0	201.3 ± 60.8
$Wc\bar{c}$	292.2 ± 90.1	98.1 ± 30.2
Wcj	250.4 ± 77.2	52.1 ± 16.0
Mistags	501.3 ± 69.6	151.9 ± 21.4
non-W	89.6 ± 35.8	35.1 ± 14.0
WW	58.5 ± 6.6	21.2 ± 2.4
WZ	28.9 ± 2.4	8.5 ± 0.7
ZZ	0.9 ± 0.1	0.4 ± 0.0
$Z + jets$	36.5 ± 5.6	15.6 ± 2.4
$t\bar{t}$ dilepton	69.2 ± 10.0	60.2 ± 8.7
$t\bar{t}$ non-dilepton	134.9 ± 19.6	421.8 ± 61.1
Total signal	145.7 ± 21.4	45.4 ± 6.7
Total prediction	2265.0 ± 375.4	1111.5 ± 129.5
Observed in data	2229	1086

Monte Carlo Modeling



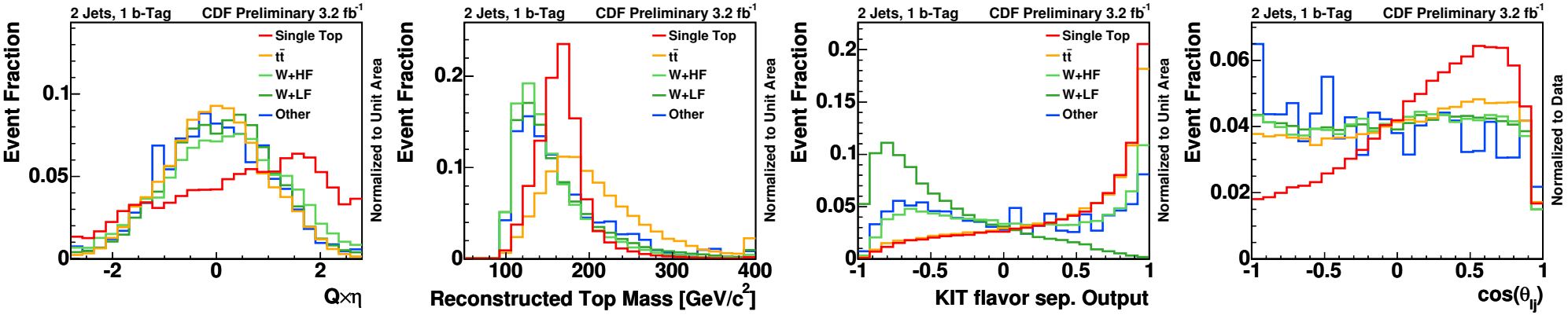
Monte Carlo Modeling



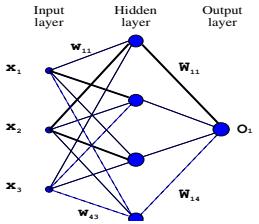
Checking hundred of plots!!

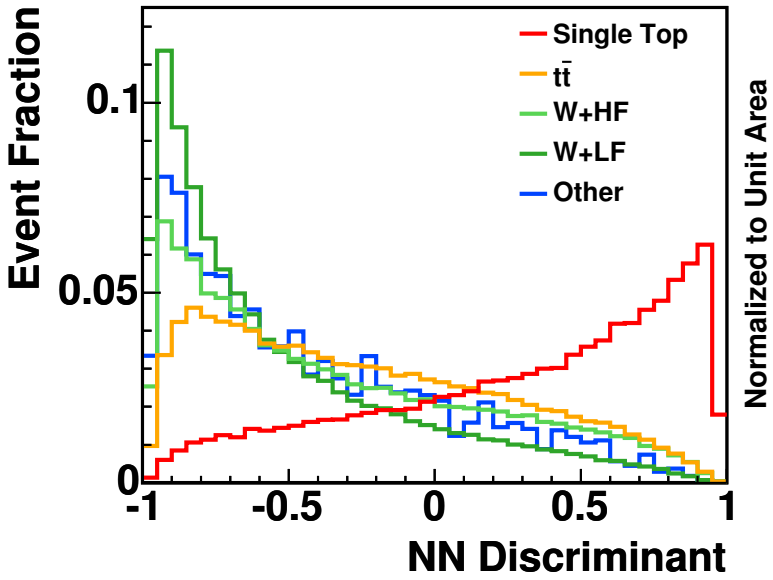
- Data versus MC agreement
- Input Variables and more
- Control and signal regions

Multivariate Analyses

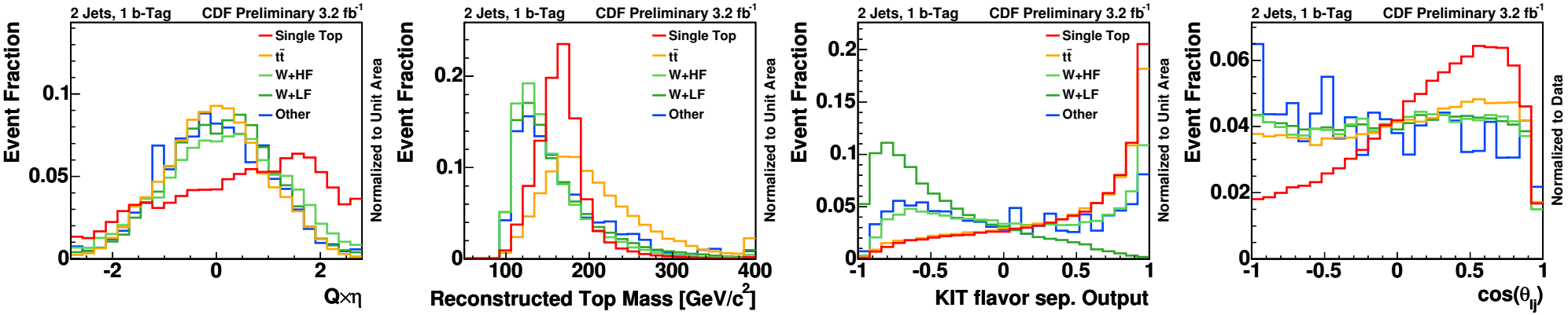


- Once the background estimate is well understood we are ready to perform a multivariate analysis
  - **Neural Network (NN)**
  - Boosted Decision Tree (BDT)
  - Matrix Element Method (ME)
  - Likelihood Function (LF)
    - Dedicated t- and s-channel searches
- Discriminant outputs are thoroughly cross-checked in sideband regions before unblinding the signal region

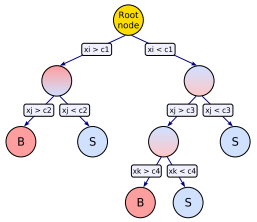


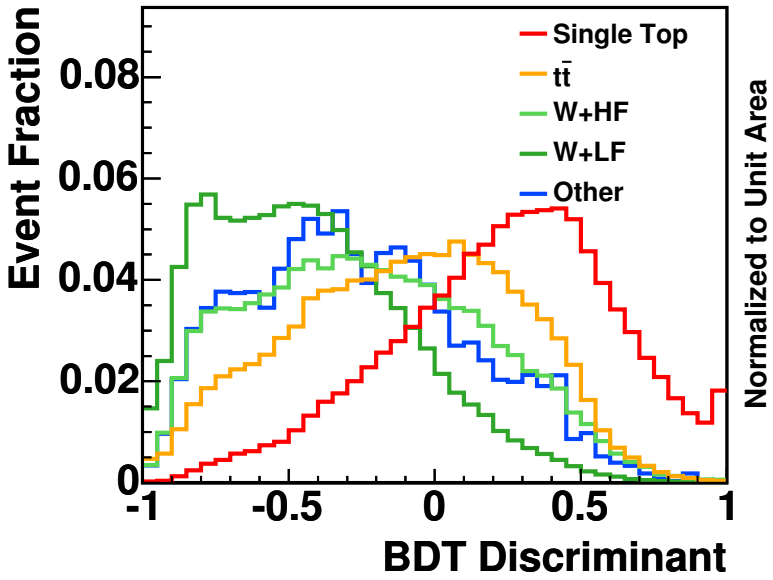


Multivariate Analyses

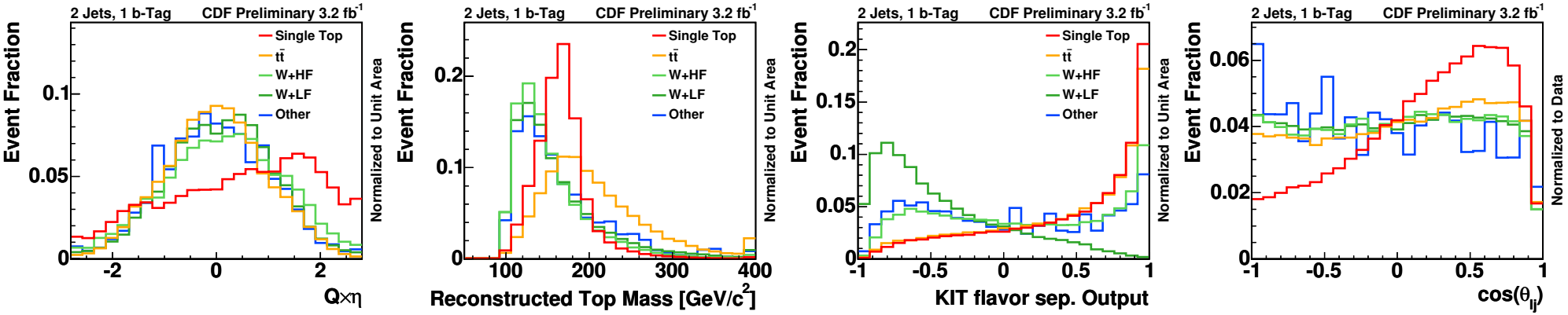


- Once the background estimate is well understood we are ready to perform a multivariate analysis
  - Neural Network (NN)
  - **Boosted Decision Tree (BDT)**
  - Matrix Element Method (ME)
  - Likelihood Function (LF)
    - Dedicated t- and s-channel searches
- Discriminant outputs are thoroughly cross-checked in sideband regions before unblinding the signal region



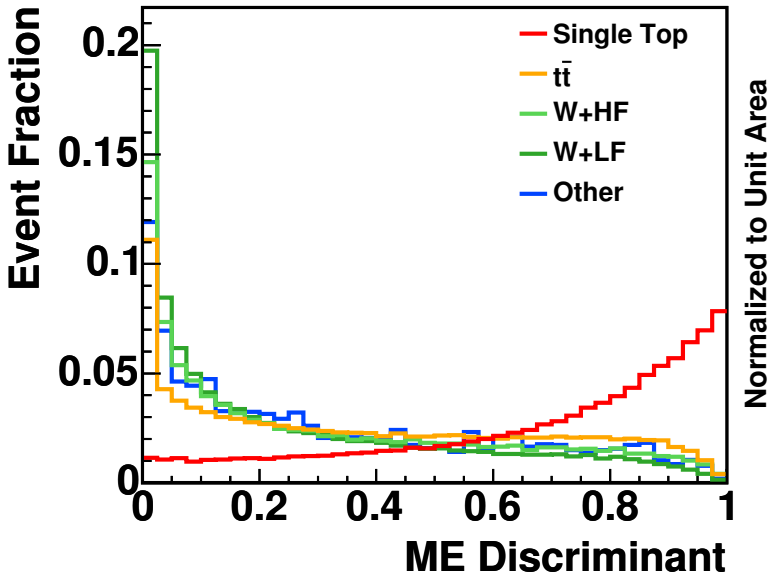


Multivariate Analyses

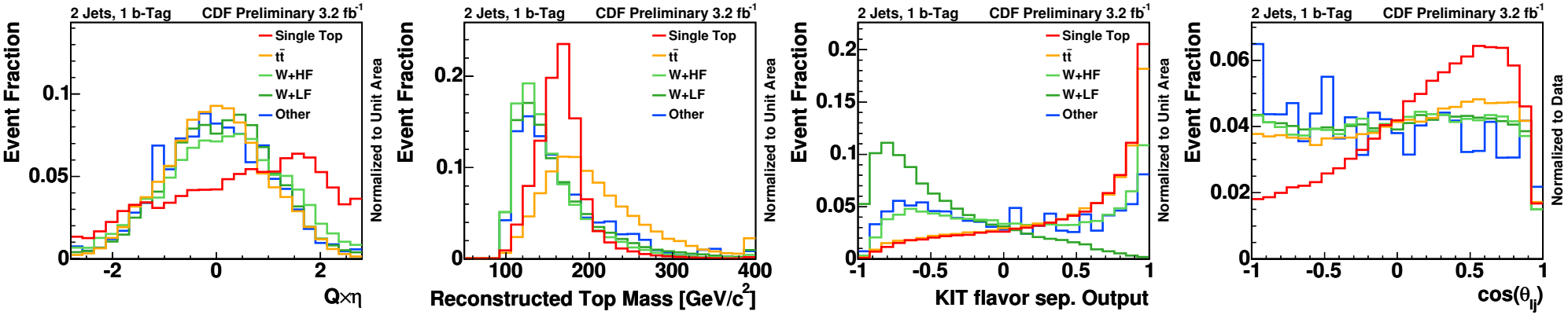


- Once the background estimate is well understood we are ready to perform a multivariate analysis
  - Neural Network (NN)
  - Boosted Decision Tree (BDT)
  - **Matrix Element Method (ME)**
  - Likelihood Function (LF)
    - Dedicated t- and s-channel searches
- Discriminant outputs are thoroughly cross-checked in sideband regions before unblinding the signal region

$$EPD = \frac{b \cdot P_{sig}}{b \cdot P_{sig} + b \cdot P_{b\text{-}bkg} + (1-b) \cdot P_{nonb\text{-}bkg}}$$

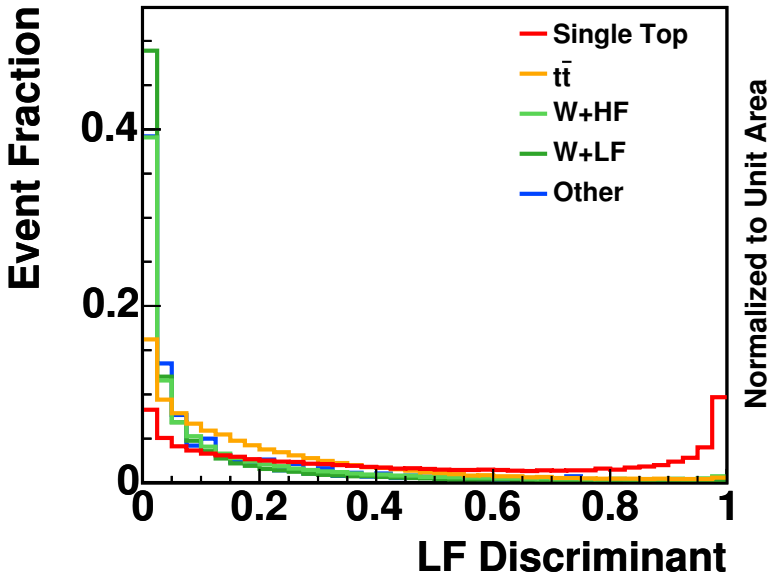


Multivariate Analyses

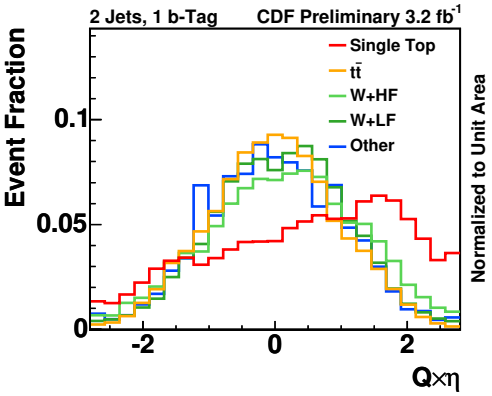


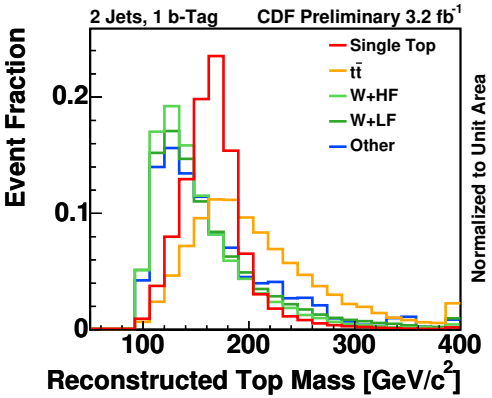
- Once the background estimate is well understood we are ready to perform a multivariate analysis
  - Neural Network (NN)
  - Boosted Decision Tree (BDT)
  - Matrix Element Method (ME)
  - **Likelihood Function (LF)**
    - Dedicated **t-** and s**-channel** searches
- Discriminant outputs are thoroughly cross-checked in sideband regions before unblinding the signal region

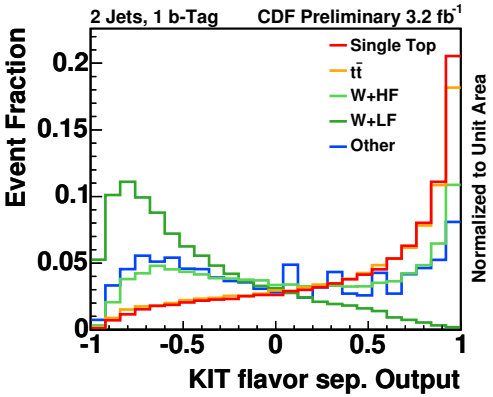
$$L = \frac{\Pi_{i=1}^{nvar} p_i^{sig}(x_i)}{\Pi_{i=1}^{nvar} p_i^{sig}(x_i) + \Pi_{i=1}^{nvar} p_i^{bkg}(x_i)}$$

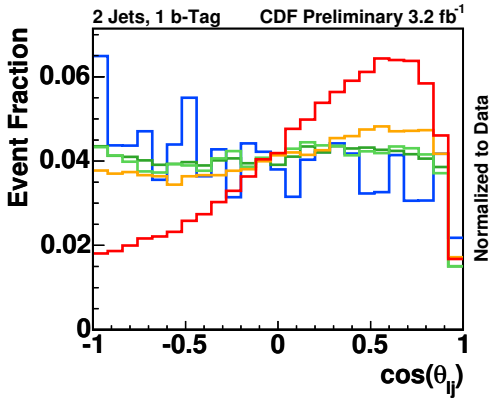


Multivariate Analyses



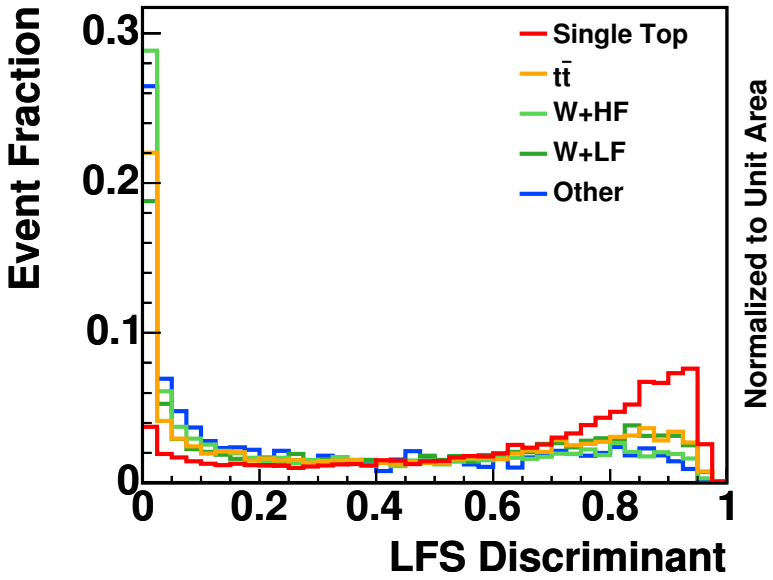




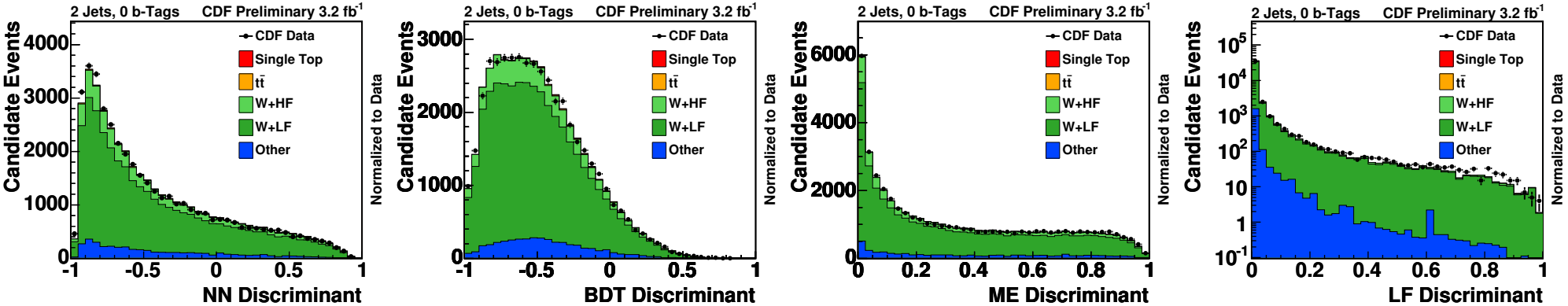


- Once the background estimate is well understood we are ready to perform a multivariate analysis
  - Neural Network (NN)
  - Boosted Decision Tree (BDT)
  - Matrix Element Method (ME)
  - **Likelihood Function (LF)**
    - Dedicated t- and **s-channel** searches
- Discriminant outputs are thoroughly cross-checked in sideband regions before unblinding the signal region

$$L = \frac{\Pi_{i=1}^{nvar} p_i^{sig}(x_i)}{\Pi_{i=1}^{nvar} p_i^{sig}(x_i) + \Pi_{i=1}^{nvar} p_i^{bkg}(x_i)}$$



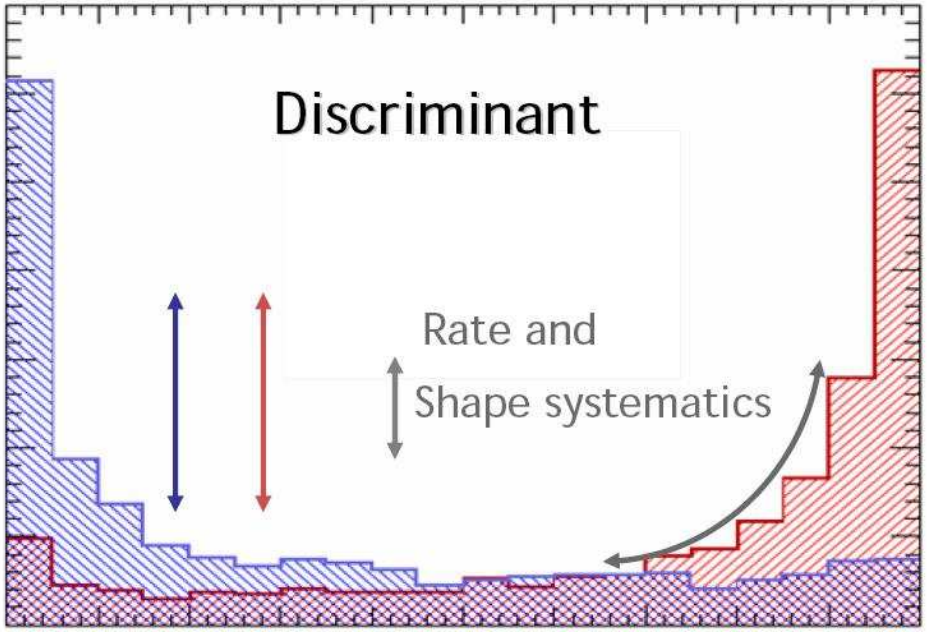
Multivariate Analyses



- Once the background estimate is well understood we are ready to perform a multivariate analysis
  - Neural Network (NN)
  - Boosted Decision Tree (BDT)
  - Matrix Element Method (ME)
  - Likelihood Function (LF)
    - Dedicated t- and s-channel searches
- Discriminant outputs are thoroughly cross-checked in sideband regions before unblinding the signal region

Binned Likelihood Fit and Systematics

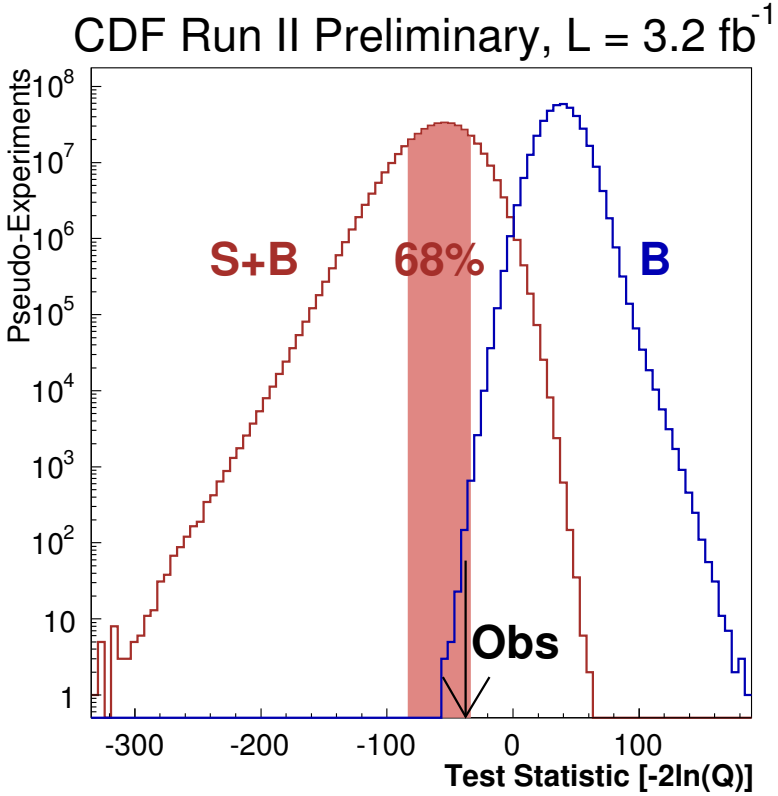
- A binned maximum likelihood fit of the templates to the data is performed to extract the cross-section
- Systematic uncertainties are taken into account as nuisance parameters
  - Rate syst. give freedom to the fit templates to move vertically only
  - Shape syst. allow templates to slide horizontally (bin by bin)
  - MC statistics in each bin treated as a source of systematic uncertainty



Systematic Uncertainty	Rate	Shape
Jet Energy Scale	0...10%	✓
Initial + Final State Radiation	0...15%	✓
Parton Distribution Functions	2...3%	✓
Monte Carlo Generator	1...5%	
Event Detection Efficiency	0...9%	
Luminosity	6%	
Neural Net B-tagger		✓
Mistag Model		✓
Q^2 scale in ALPGEN MC		✓
Input variable mismodeling		✓
Wbb+Wcc normalization	30%	
Wc normalization	30%	
Mistag normalization	17...29%	
ttbar normalization & m_{top}	23%	✓

Significance and Results

- P-value is determined throwing pseudo experiments (PE) for the test hypothesis (S+B) and for the background-only hypothesis (B)
  - Fluctuate all rate and shape systematic uncertainties in PE
  - Test Statistic: $-2\ln(Q)$, $Q = \mathcal{L}(S+B)/\mathcal{L}(B)$
  - Exp. p-value: fraction of B PE with a value smaller than the median of S+B PE
  - Obs. p-value: fraction of B PE with a value smaller than the measured one

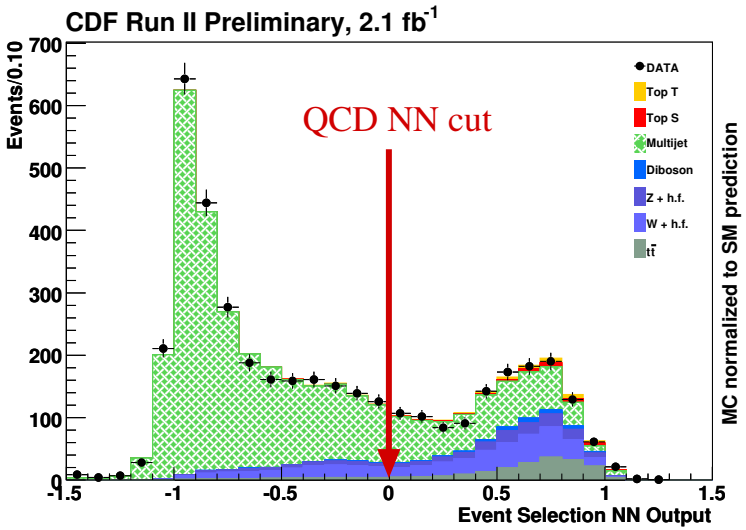


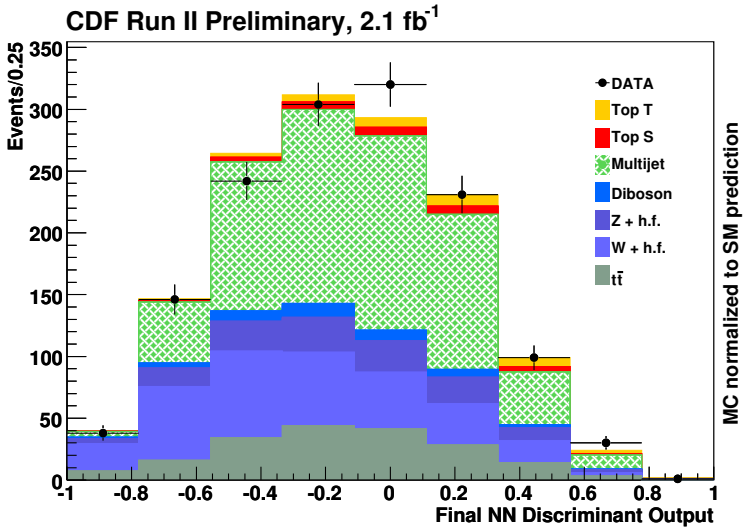
CDF Preliminary, L = 3.2 fb^{-1}			
Analysis	Exp. sig.	Obs. sig.	Cross sect. (pb)
NN	5.2σ	3.5σ	$1.8^{+0.6}_{-0.6}$
BDT	5.2σ	3.5σ	$2.1^{+0.7}_{-0.6}$
ME	4.9σ	4.3σ	$2.5^{+0.7}_{-0.6}$
LF	4.0σ	2.4σ	$1.6^{+0.8}_{-0.7}$
LFS (s-chan)	1.1σ	2.0σ	$1.5^{+0.9}_{-0.8}$

MET + Jets Analysis (MJ)

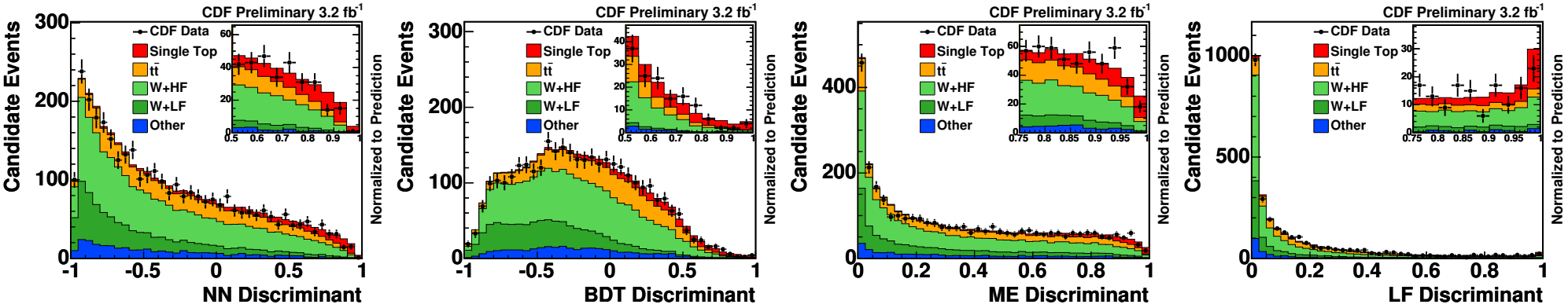
- Another single top search in the orthogonal $\not\!\!E_T$+jets sample
  - Increase acceptance by $\sim$ 30%: events were lepton was not reconstructed and hadronic taus
  - Neural Network to suppress huge instrumental background
  - Final Neural Network to discriminate signal from background
  - Same method to extract signal cross-section and significance

CDF Preliminary, Luminosity = 2.1 fb^{-1}			
Analysis	Exp. sig.	Obs. sig.	Cross section (pb)
$\not\!\!E_T$+jets	1.4	2.1	$4.9^{+2.5}_{-2.2}$

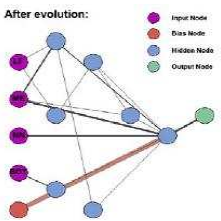


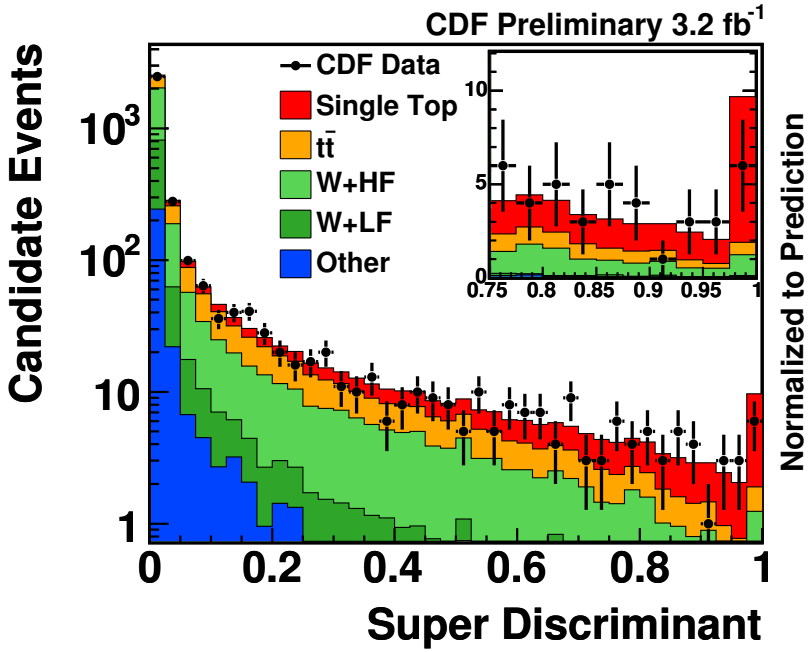


Combination Technique



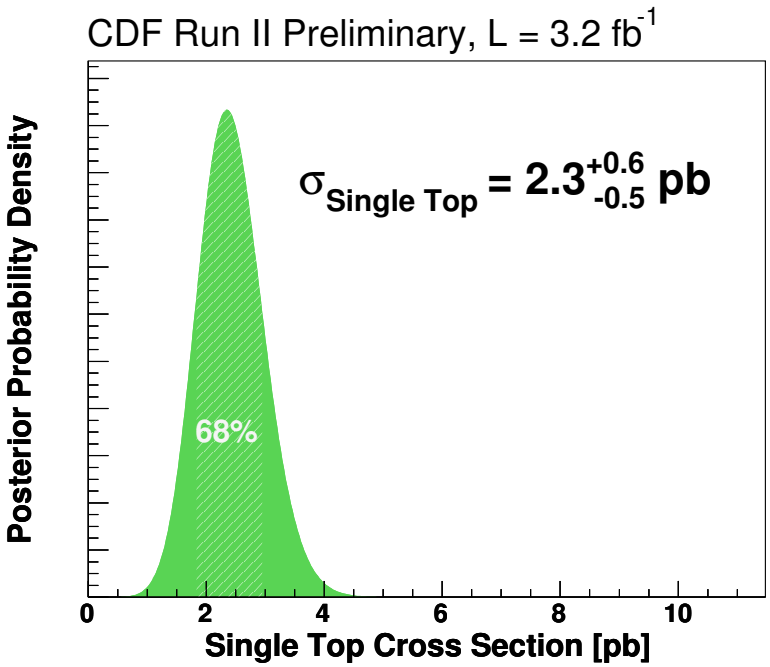
- Combine individual analysis into one "Super Discriminant"
  - Use discriminant outputs (NN, BDT, ME, LF, LFS) as input to NN
  - Candidate networks compete against each other
  - Optimization of network topology, weights, output histogram binning
  - Final network chosen based on best expected p-value
  - Gained 13% over best individual analysis

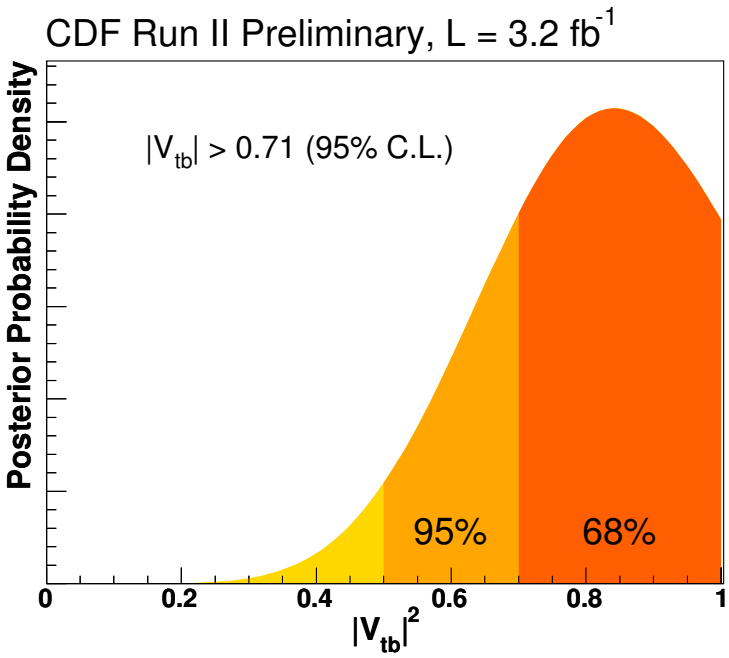




Combination Results

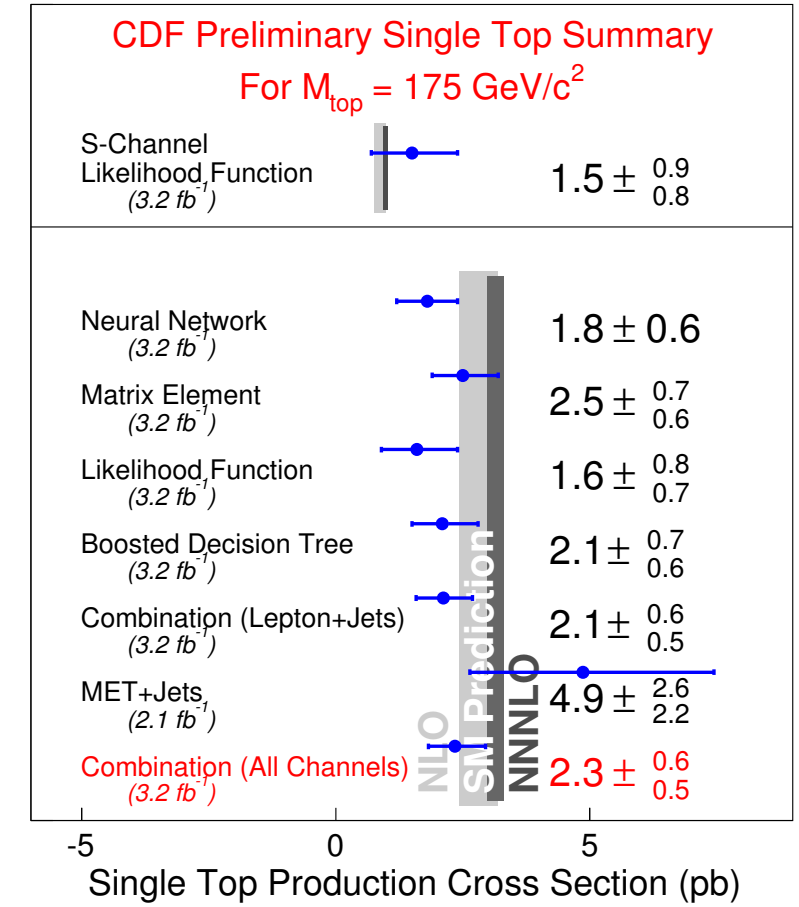
- Perform a simultaneous cross section fit over the two orthogonal analysis (SD + MJ)
- Expected p-value: $> 5.9\sigma$
- Observed p-value: **5.0σ**
- Measured cross-section: $\sigma = 2.3^{+0.6}_{-0.5}$ pb
- Extract $|V_{tb}| = 0.91 \pm 0.11(exp.) \pm 0.07(theory)$
- Submitted to PRL, March 4th, 2009! http://arxiv.org/abs/0903.0885





Conclusions

- Single top searches are challenging
- Advance analysis tools essential to establish small signals buried underneath huge backgrounds
- We report Observation (5σ) for electroweak single top quark at CDF
  - $\sigma_{\text{SingleTop}} = 2.3^{+0.6}_{-0.5}$ pb
  - $|V_{tb}| = 0.91 \pm 0.11(exp.) \pm 0.07(theory)$
- Future: properties in the single top sample
- Important milestone along the way to the Higgs!
  - All SM backgrounds for Higgs searches are now firmly established

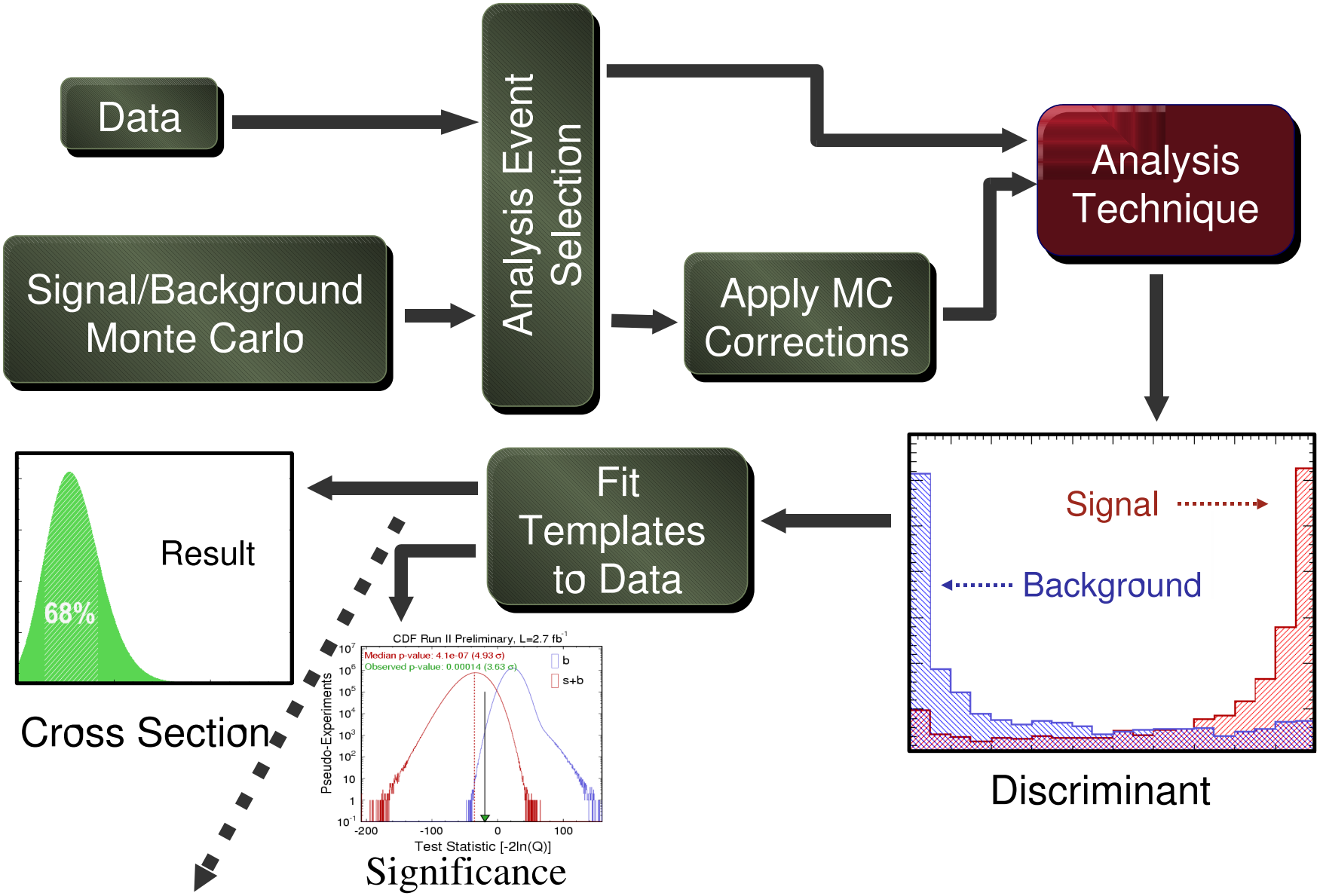


NLO: Z.Sullivan, Phys.Rev.D70,114012 (2004) NNNLO: N.Kidonakis, Phys.Rev.D74,114012 (2006)

- Check all the results in the public web page http://www-cdf.fnal.gov/physics/new/top/public_singletop.html

BACKUP

Search Strategy



Rate/shape systematic uncertainties: implemented as nuisance parameters

Background Estimate

Top/EWK (WW/WZ/Z→ττ, ttbar, single t)

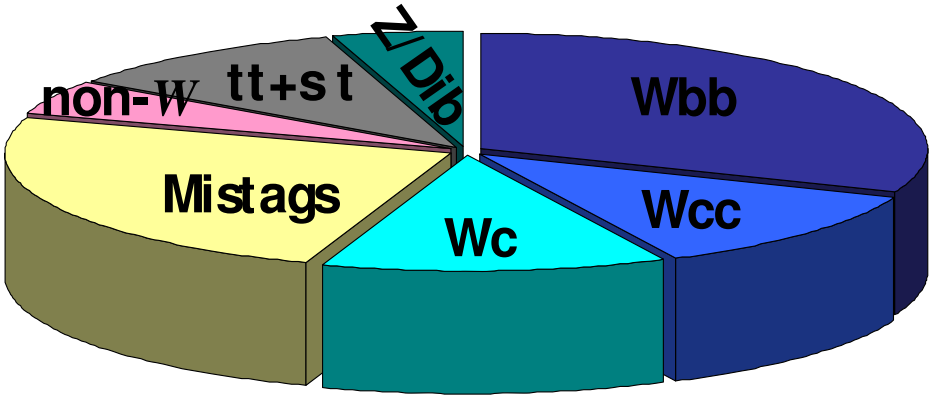
- MC normalized to theoretical cross-section

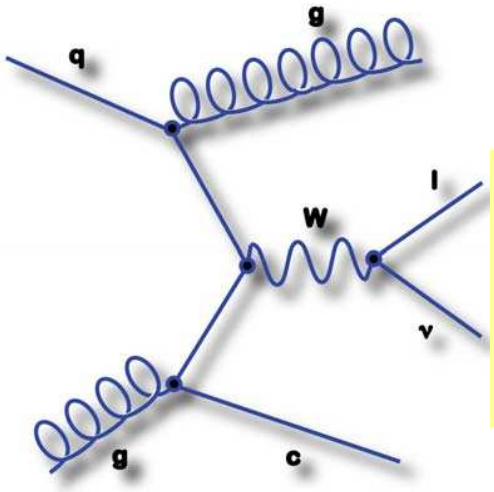
W+HF jets (Wbb/Wcc/Wc)

- W+jets normalization from data and heavy flavor (HF) fractions from ALPGEN Monte Carlo, calibrated in generic multijet data

Non-*W* (QCD)

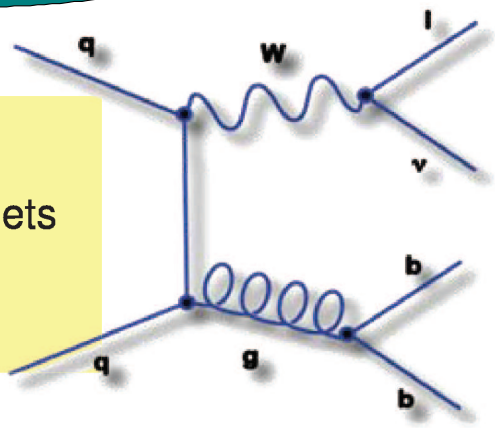
- Multijet events with semileptonic *b*-decays or mismeasured jets
- Fit low missing E_T data and extrapolate into signal region



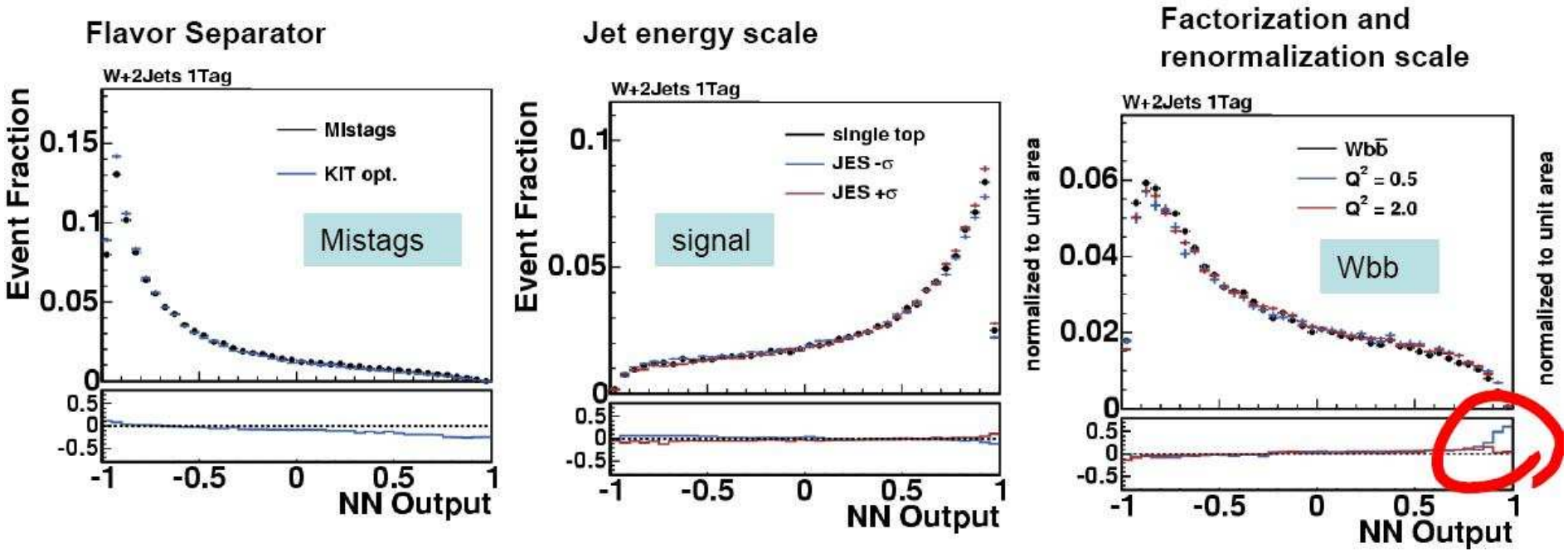


Mistags (W+2jets)

- Falsely tagged light quark or gluon jets
- Mistag probability parameterization obtained from inclusive jet data

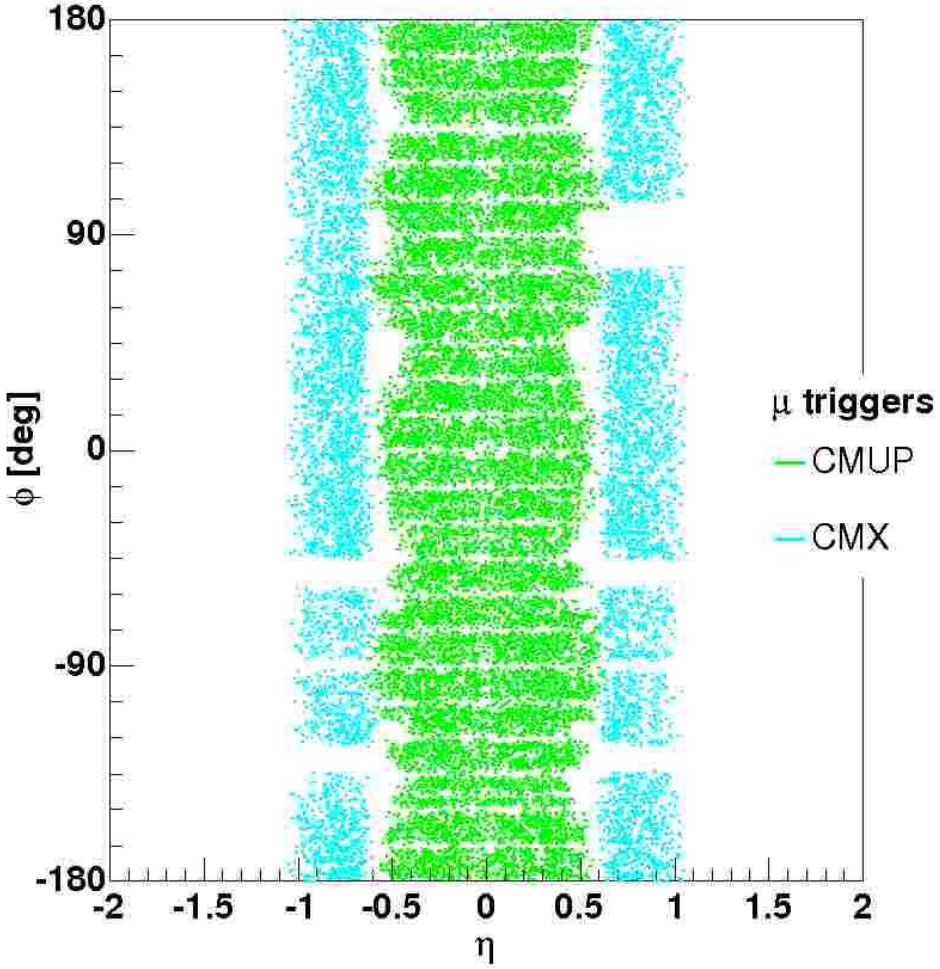


TEMPLATE SHAPE SYSTEMATICS



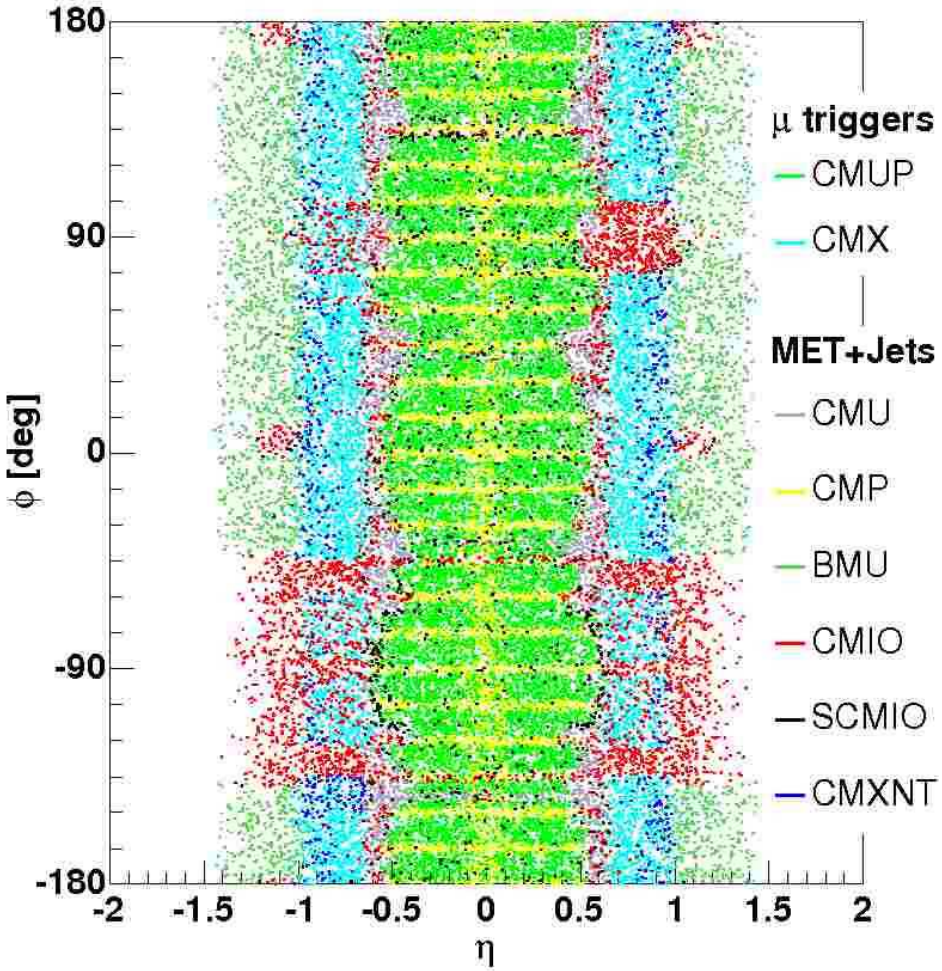
- A total of 370 shape uncertainties evaluated!
- Each template, each source of shape error, each channel (#tags, #jets, extended muon coverage)
- Shape uncertainties affect sensitivity - most are quite small but some appreciable

EXTENDED MUON COVERAGE



- Traditional muons are triggered muons with hits in
  - Both central muon chambers (CMUP, $|\eta| < 0.6$)
  - In the central extension chamber (CMX, $|\eta| < 1.0$)
- Extended muons
  - Events with missing $E_T > 35$ GeV plus two jets are triggered
  - Recover muons in the gaps left by the high p_T muon triggers
  - Acceptance gain of $\sim 30\%$
  - Orthogonal muon definitions

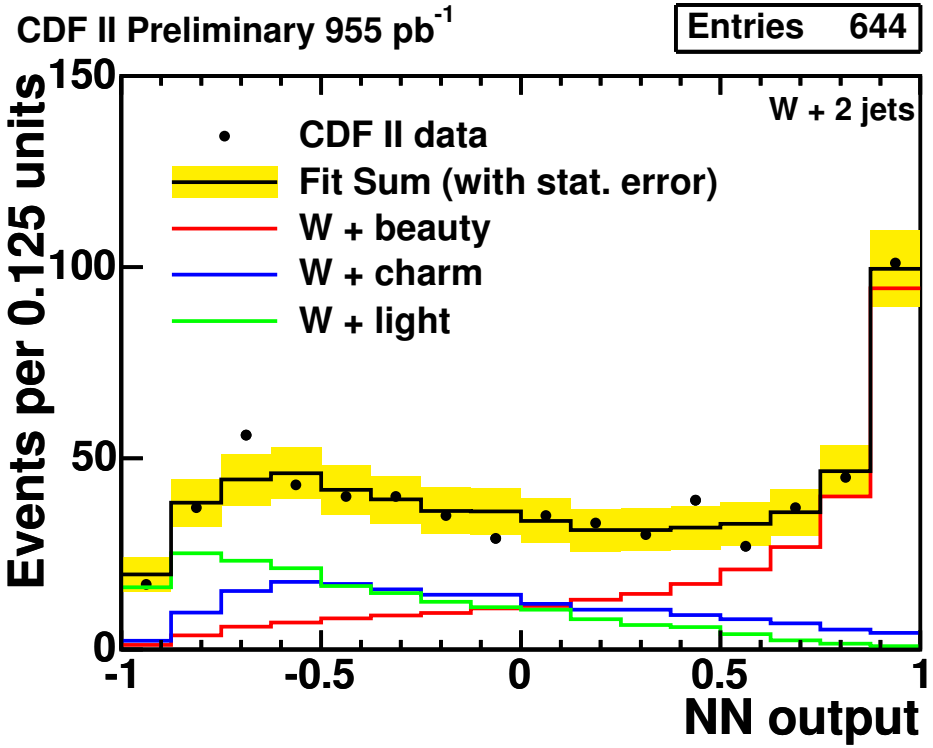
Extended Muon Coverage



- Extended muons
  - Events with missing $E_T > 35$ GeV plus two jets are triggered
  - Recover muons in the gaps left by the high p_T muon triggers
  - Acceptance gain of $\sim 30\%$
  - Orthogonal muon definitions
    - ➜ CMU-only muons (η-Gap)
    - ➜ CMP-only muons (ϕ-Gap)
    - ➜ BMU muons
    - ➜ CMIO muons (stubless and no fiducial to any chamber)
    - ➜ SCMIO (stubbed and no match in other categories)
    - ➜ CMXNT (non-triggered CMX)

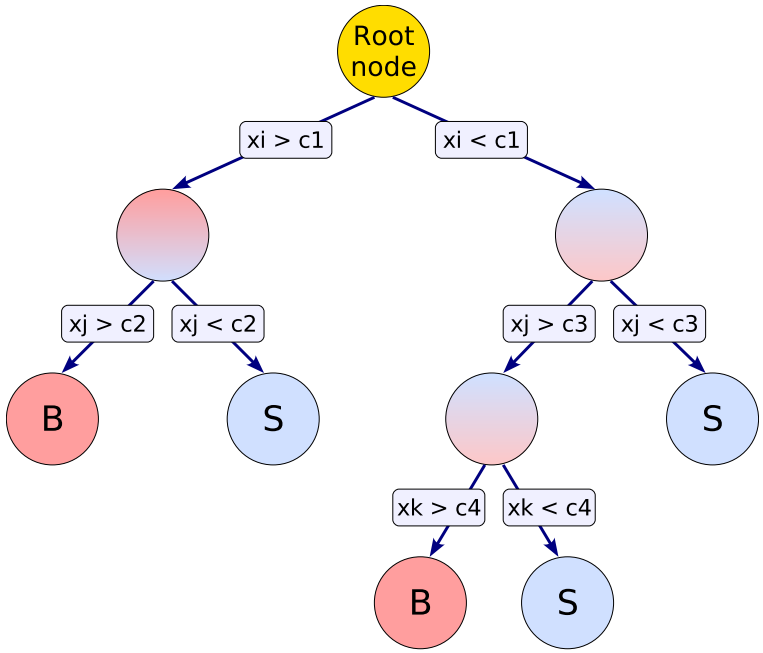
NN Jet Flavor Separator

- About 50% of the background in the W+2 jets sample do NOT contain b quarks (even though a secondary vertex is required)
- Train Neural Network to separate b-quark jets from charm and light-quark jets
- Jet and track variables (vertex mass, decay length, track multiplicity...)
- Replace Yes-No tag decision by a continuous variable
- Improves single top sensitivity by $\sim$10-15%!



Overview of Boosted Decision Trees

- Decision Tree (DT): Sequence of binary splits using the discriminating variable which gives best sig-bkg separation
- Leaf nodes are classified as sig-like or bkg-like depending on majority of events ending up in the respective leaf
- A Boosting algorithm improves the discrimination power and statistical stability

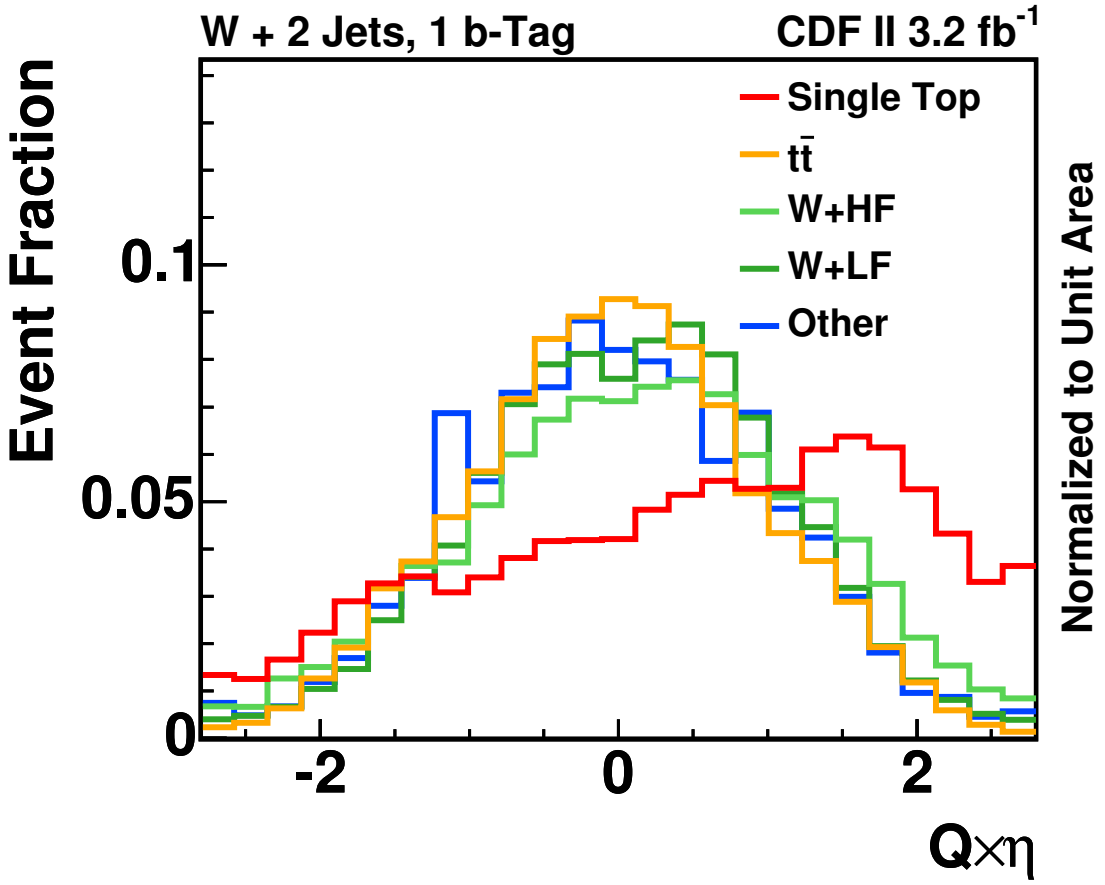


  - Events misclassified during a DT training are given a higher weight in the next DT training

- To avoid overtraining a pruning algorithm (BDT) is performed after the BDT training
  - Removes statistically insignificant nodes by cutting back a tree from the bottom up

Likelihood Function Overview

- Combines several sensitive variables into a single one
- 7 (10) variables used in the 2 (3) jet bin: H_T, $Q\times\eta$, M_{jj}, cos(l, j), $\log(ME_{t-chan})$...



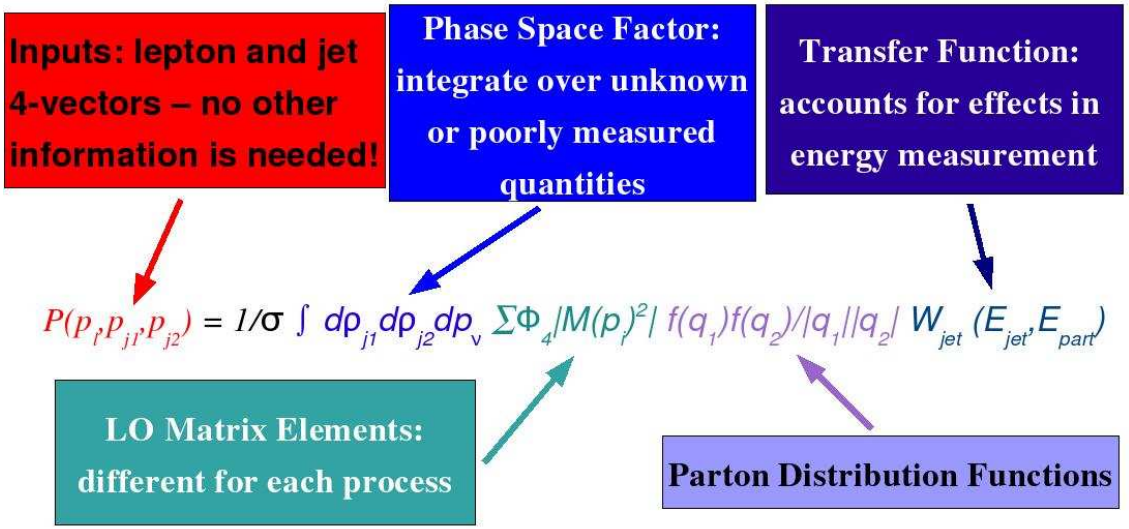
$$p_i^{sig}(x_i) = \frac{N_i^{sig}(x_i)}{N_i^{sig}(x_i)+N_i^{bkg}(x_i)},$$

$$L = \frac{\Pi_{i=1}^{nvar} p_i^{sig}(x_i)}{\Pi_{i=1}^{nvar} p_i^{sig}(x_i)+\Pi_{i=1}^{nvar} p_i^{bkg}(x_i)}$$

Matrix Elements Overview

- Compute, for each event, the probability for signal and background hypotheses

- Use full event kinematic information
- Calculate probabilities for signal and backgrounds

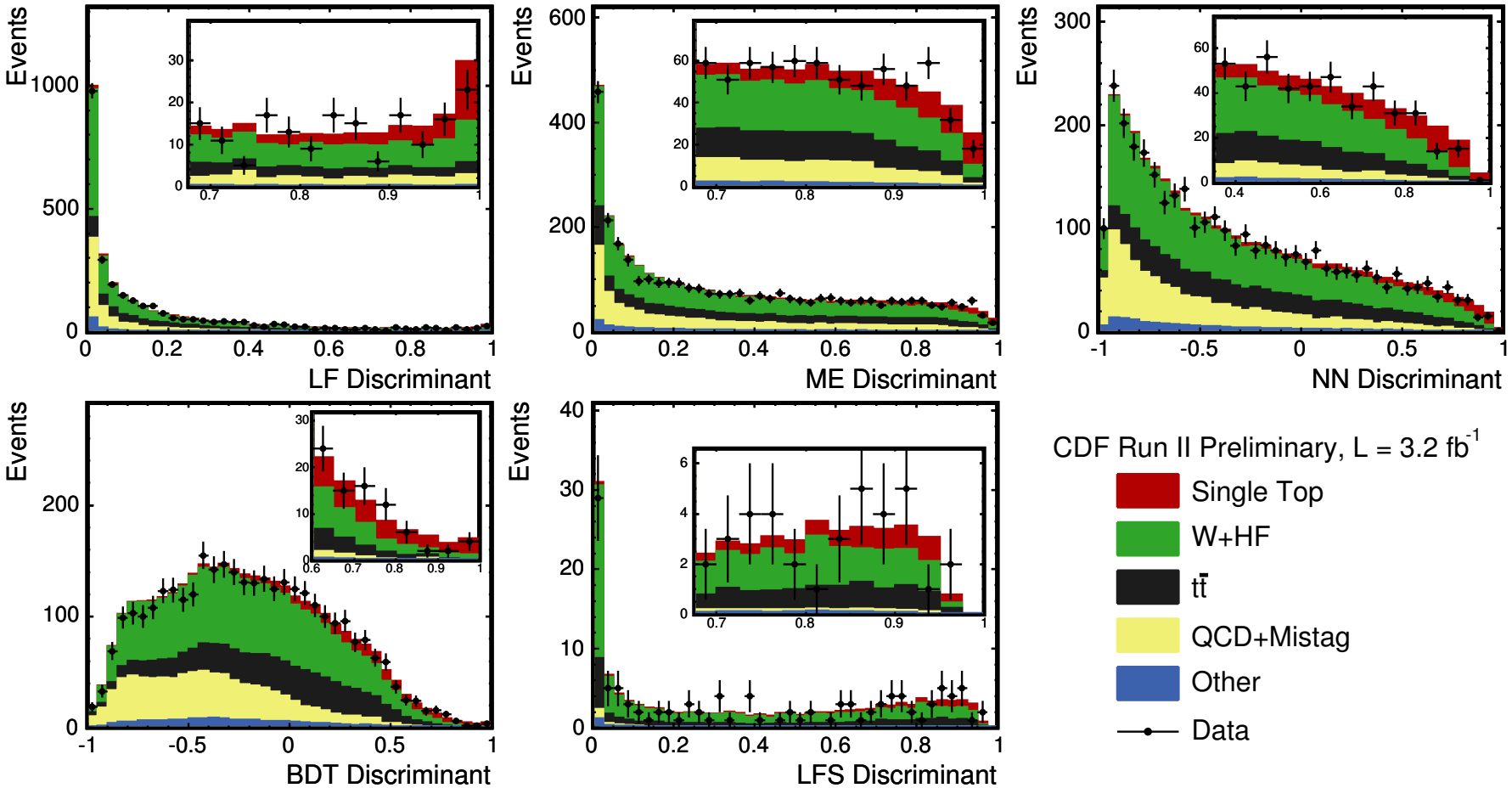


- Build a discriminant

$$EPD = \frac{b \cdot P_{sig}(\vec{x})}{b \cdot P_{sig}(\vec{x}) + b \cdot P_{b\text{-}bkg}(\vec{x}) + (1-b) \cdot P_{nonb\text{-}bkg}(\vec{x})}$$

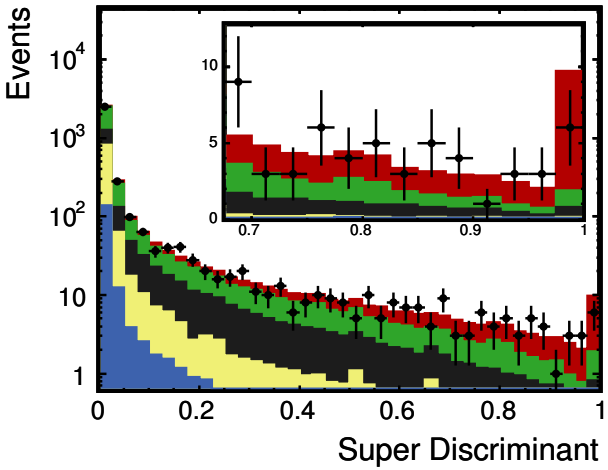
- b: b-likeness form NN Flavor Separator

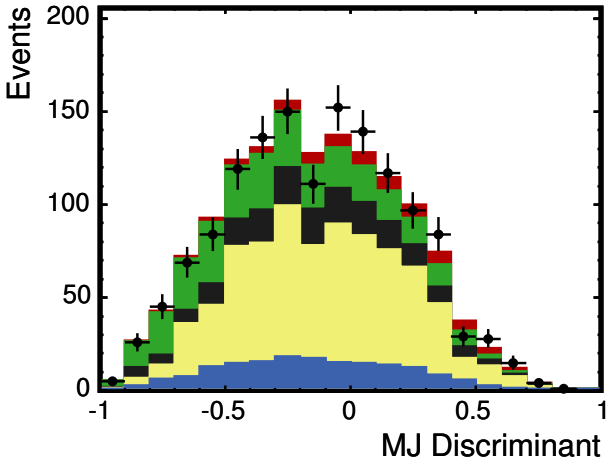
PRL Combination I

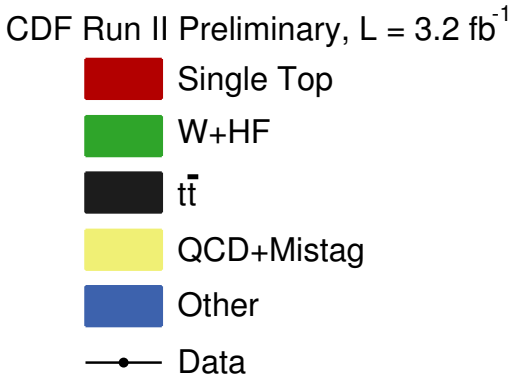


- 5 independent ℓ+jets analyses at CDF sharing single top selection (correlations 60–85%)
- Combined into a super discriminant trained using genetic algorithms to provide optimal sensitivity

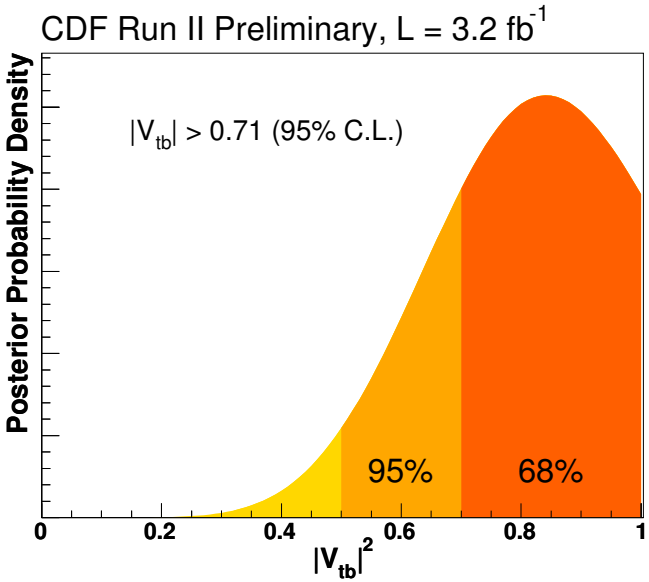
PRL Combination II







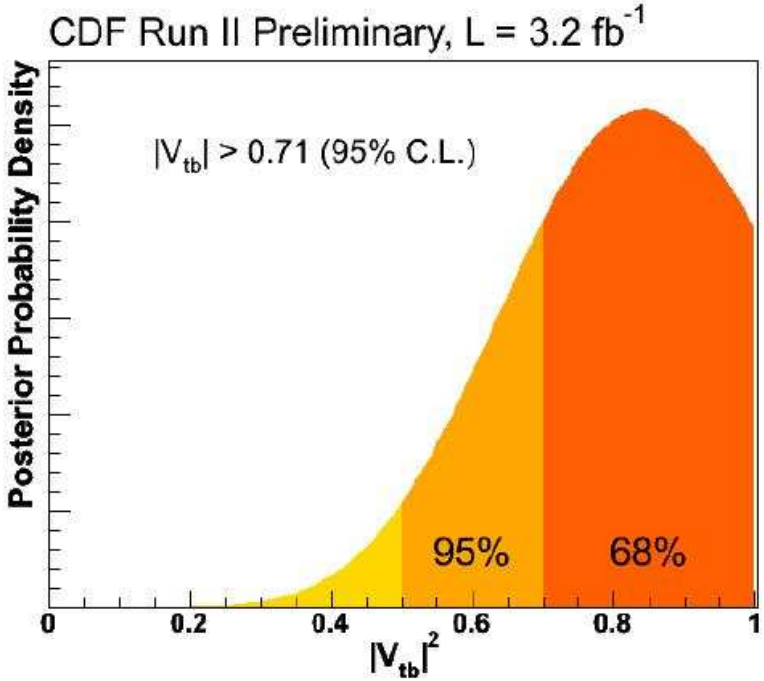
- Another orthogonal MET + jets analysis where the lepton is missed
- A simultaneous fit is performed to the super discriminant and the MET + jets channel
- Expected p-value: $> 5.9\sigma$
- Observed p-value: $\mathbf{5.0\sigma}$
- Measured cross-section: $\sigma = 2.3^{+0.6}_{-0.5}$ pb
- Extract $|V_{tb}| = 0.91 \pm 0.11(exp.) \pm 0.07(the.)$
- Submitted to PRL, March 4th, 2009! http://arxiv.org/abs/0903.0885

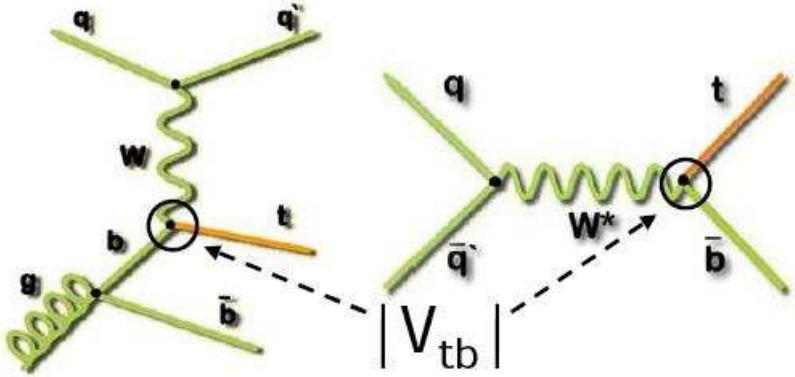


Direct $|V_{tb}|$ Measurement

- Using cross section result measure $|V_{tb}|$
- Assume Standard Model (V-A) coupling and $|V_{tb}| >> |V_{ts}|, |V_{td}|$ (from BR(t →Wb) measurements)

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





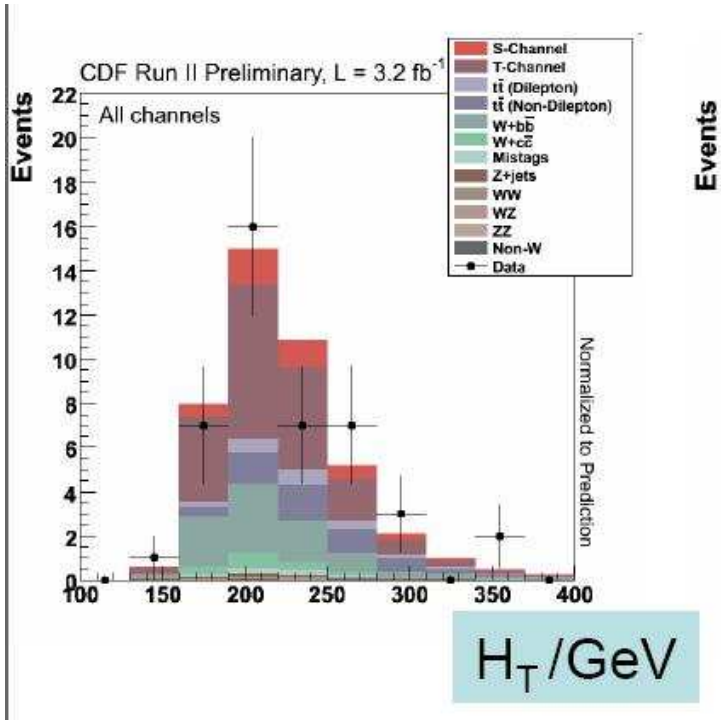
CDF combined fit:

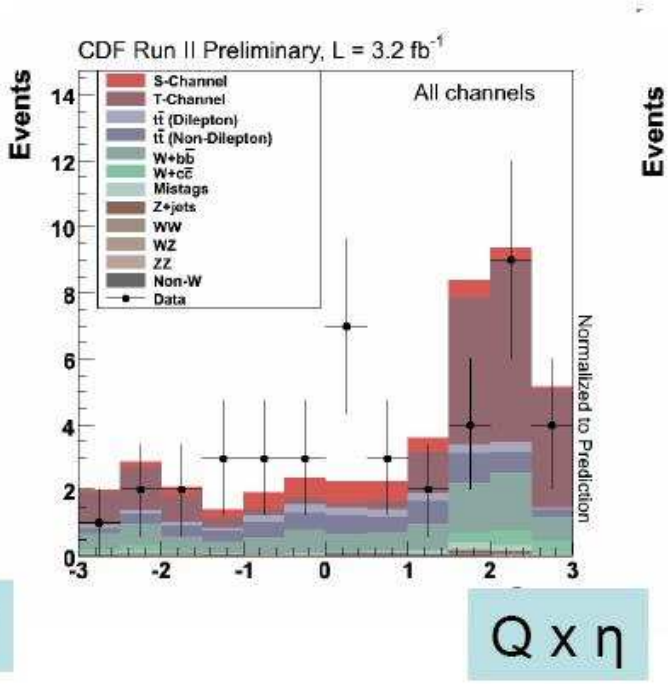
$|V_{tb}|>0.71$ at 95% C.L.

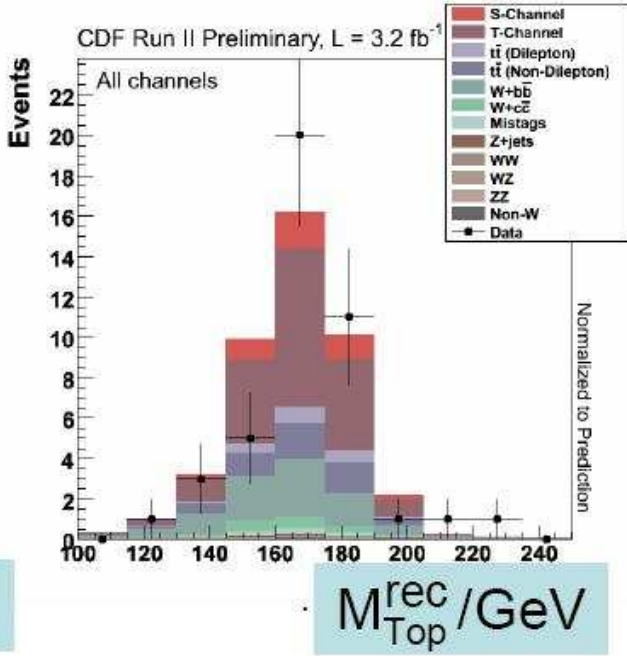
Z. Sullivan, Phys.Rev. D70 (2004) 114012

$|V_{tb}| = 0.91 \pm 0.11$ (stat+syst) ± 0.07 (theory)

Signal Features







- Super Discriminant > 0.72
- Purity S/B $\sim$ 1.2