University of Virginia — HEP Seminar
September 25, 2013

Evidence for s-channel single top production at DØ

- Electroweak production of top quarks
- Event selection and background estimation
- Multivariate methods
  - Decision Trees, Bayesian NN, Matrix Elements
- Combination
- Cross sections and significance
- Direct measurement of $|V_{tb}|$
- Summary

UNIVERSITY of ROCHESTER

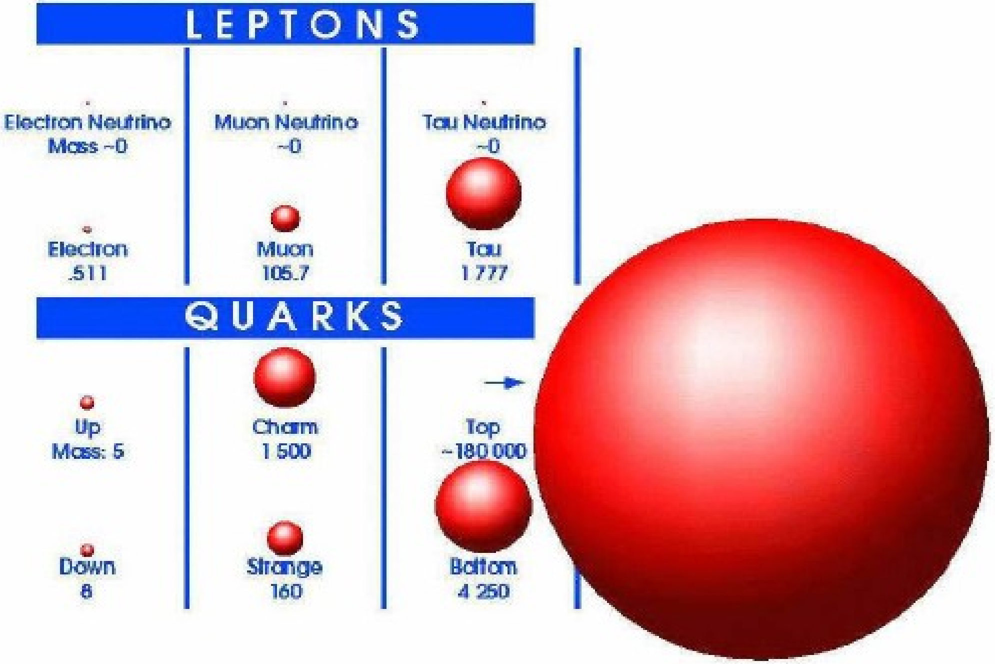
Arán García-Bellido

Top quark physics

The top quark is a very special fermion:

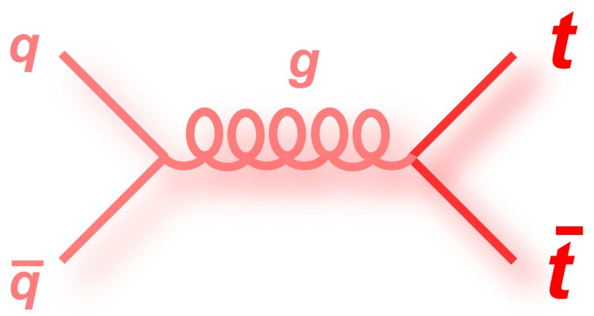
- Heaviest known particle: 173.2±0.9 GeV (arXiv:1305.3929)
  - $m_t \sim v/\sqrt{2}$, $\lambda_t \sim 1$ → Related to EWSB!
  - Sensitive probe for new physics, FCNCs, ...
- Decays as a free quark: $\tau_t = 5\times10^{-25}$ s $\ll \Lambda_{QCD}^{-1}$
  - Spin information is passed to its decay products
  - Test V-A structure of the SM

Wealth of measurements:
Cross section, charge, mass, angular properties, width, lifetime, asymmetries, decays...
LHC is a top factory!

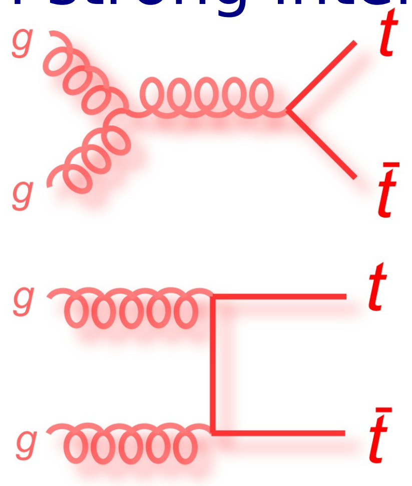


Top quark production

- Dominant mode is through strong interaction: top pairs

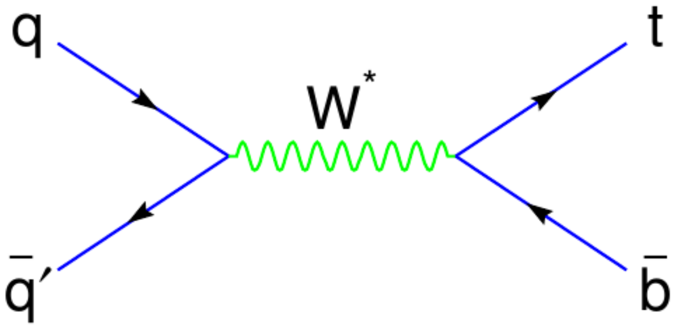


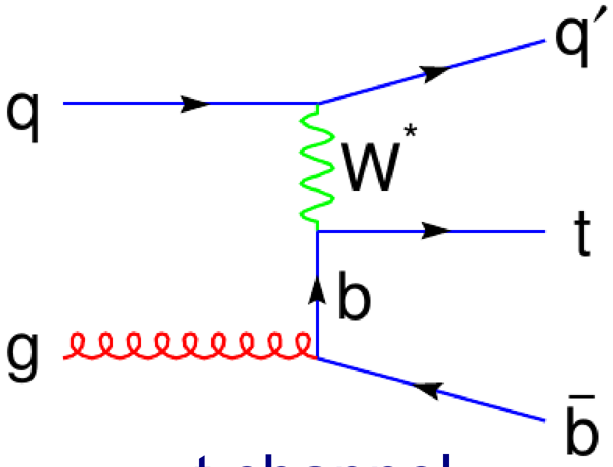
q$\bar{q}$ annihilation
$\sigma(t\bar{t}) = 7.5 \pm 0.7$ pb

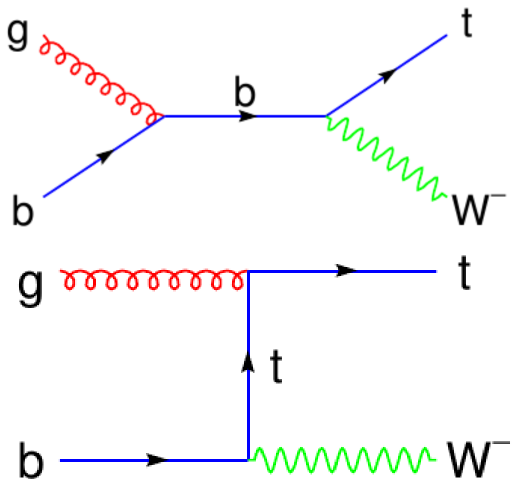


gg fusion

- Electroweak interaction: single top







s channel
$\sigma(tb) = 1.04 \pm 0.08$ pb

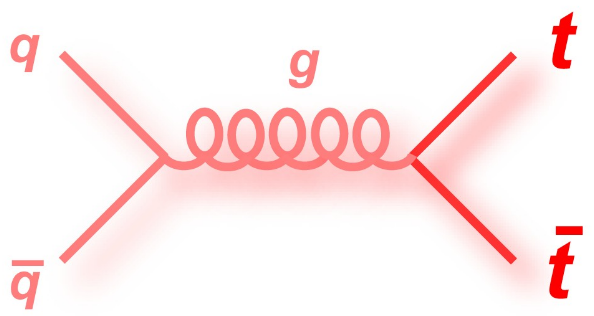
t channel
$\sigma(tqb) = 2.26 \pm 0.12$ pb

tW production
$\sigma(tW)=0.28 \pm 0.06$ pb

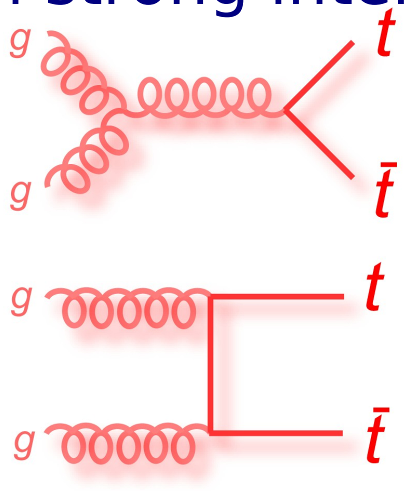
Kidonakis, m_t=172.5 GeV, Phys. Rev. D 74, 114012

Top quark production

- Dominant mode is through strong interaction: top pairs



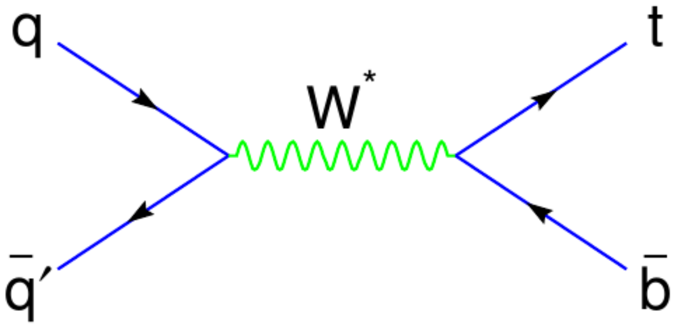
$q\bar{q}$ annihilation
$\sigma(t\bar{t})$ = 7.5±0.7 pb



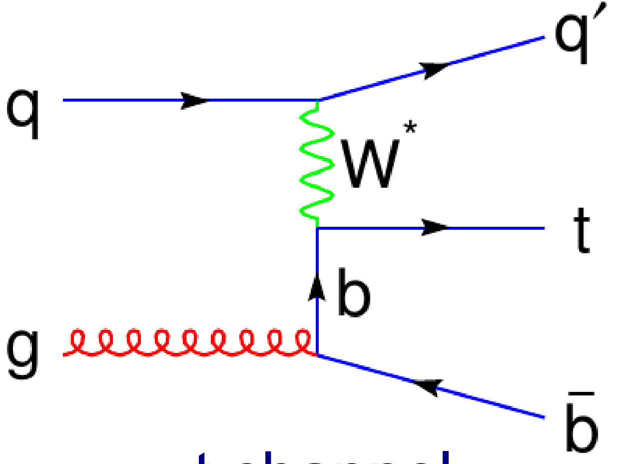
gg fusion

Main modes at the Tevatron

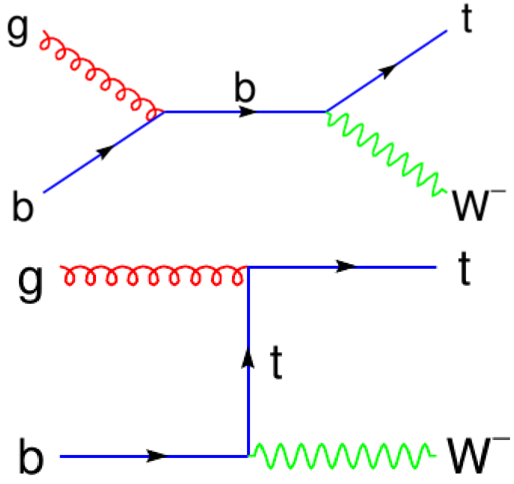
- Electroweak interaction: single top



s channel
$\sigma(tb)$ = 1.04±0.08 pb



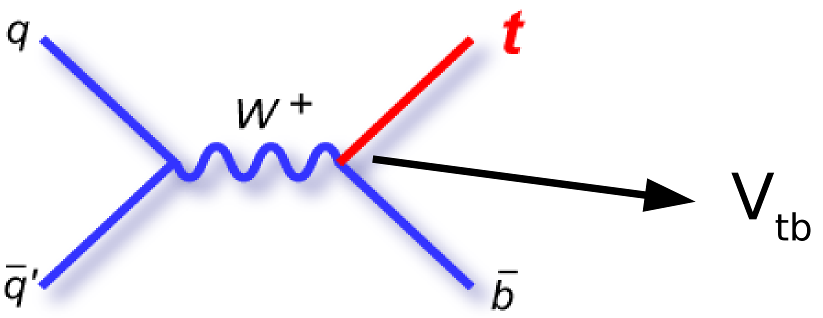
t channel
$\sigma(tqb)$ = 2.26±0.12 pb



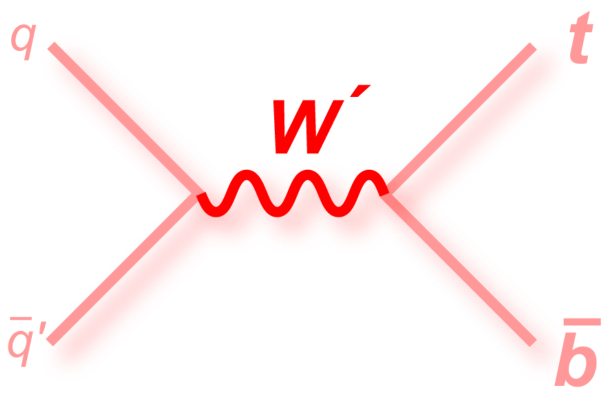
tW production
$\sigma(tW)$=0.28±0.06 pb

Kidonakis, m_t=172.5 GeV, Phys. Rev. D 74, 114012

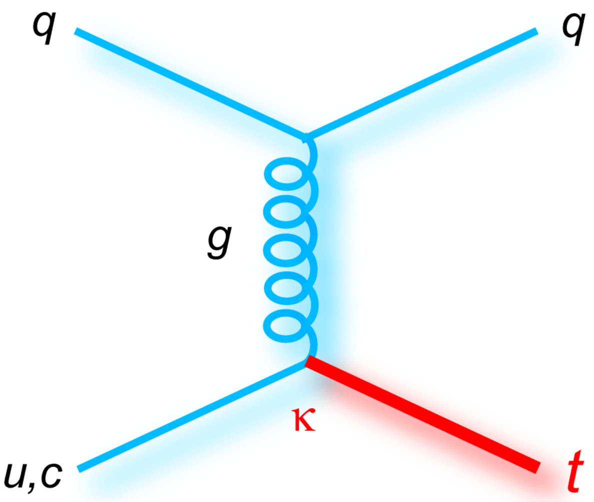
Why do we care?



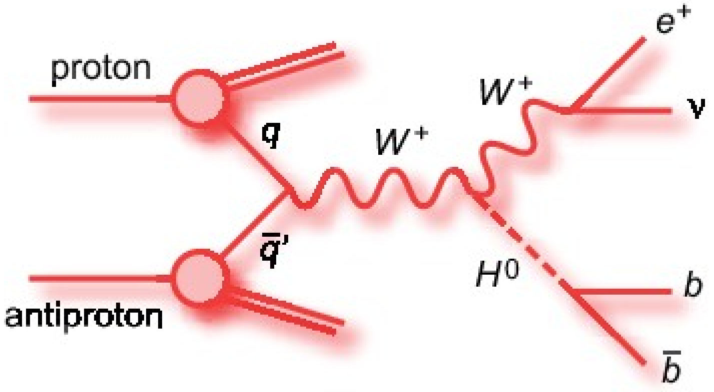
- Access W-t-b coupling
  - measure V_{tb} directly
  - test unitarity of CKM
- New physics:
  - s-channel sensitive to resonances: W’, H^+, top pions, etc...
  - t-channel sensitive to FCNCs, anomalous couplings
- Source of polarized top quarks
- Extract small signal out of a large background



DØ search: 1101.0806



DØ search: 1006.3575



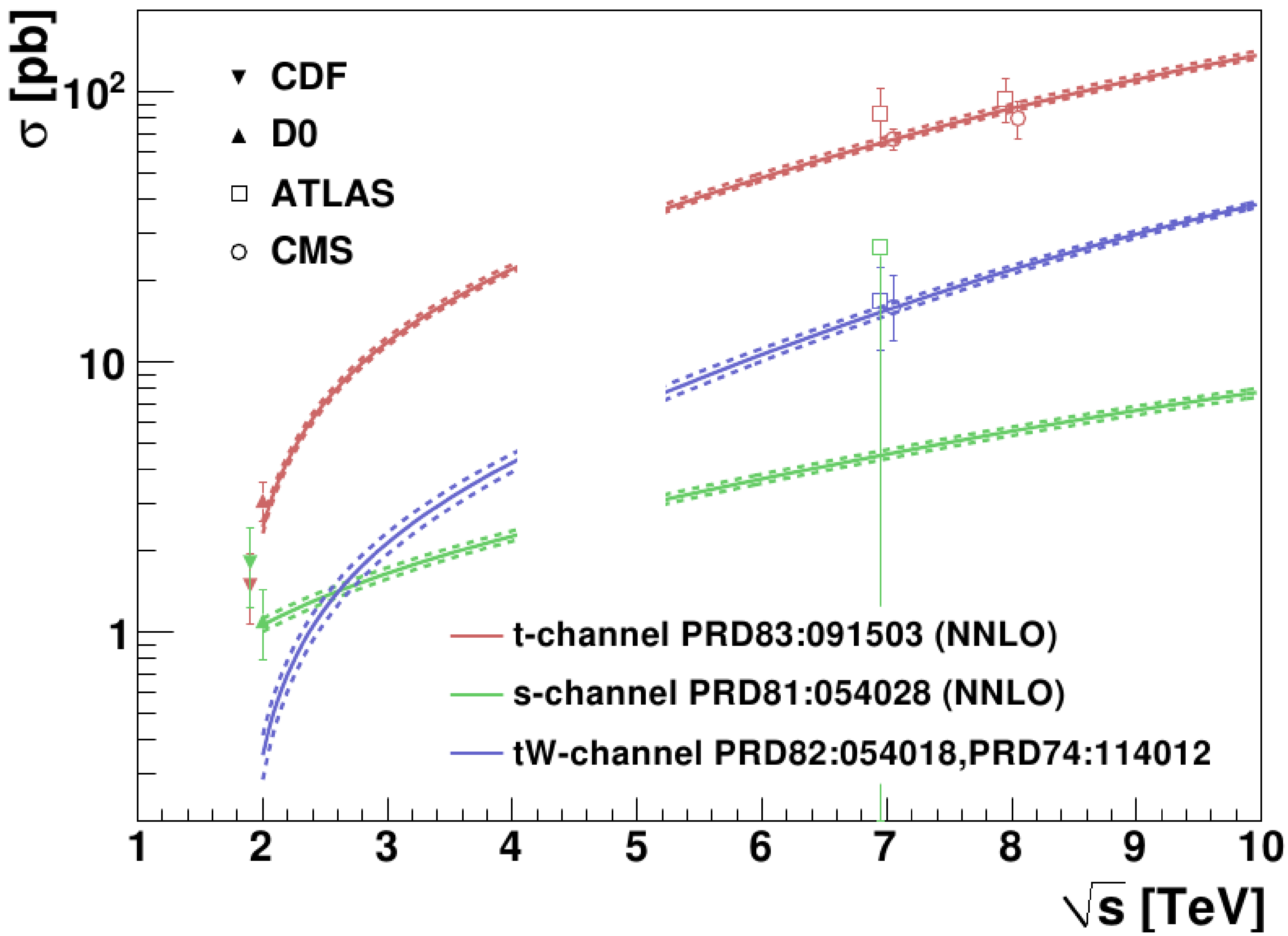
Experimental status

(before this analysis)

σ (NNLO) [pb]	**tb**	**tqb**	**tW**
TeV prediction	1.04	2.26	0.28
CDF (7.5 fb^{-1})	1.8 ± 0.6	1.49 ± 0.45	-
DØ (5.4 fb^{-1})	0.98 ± 0.63	**2.90 ± 0.59**	-

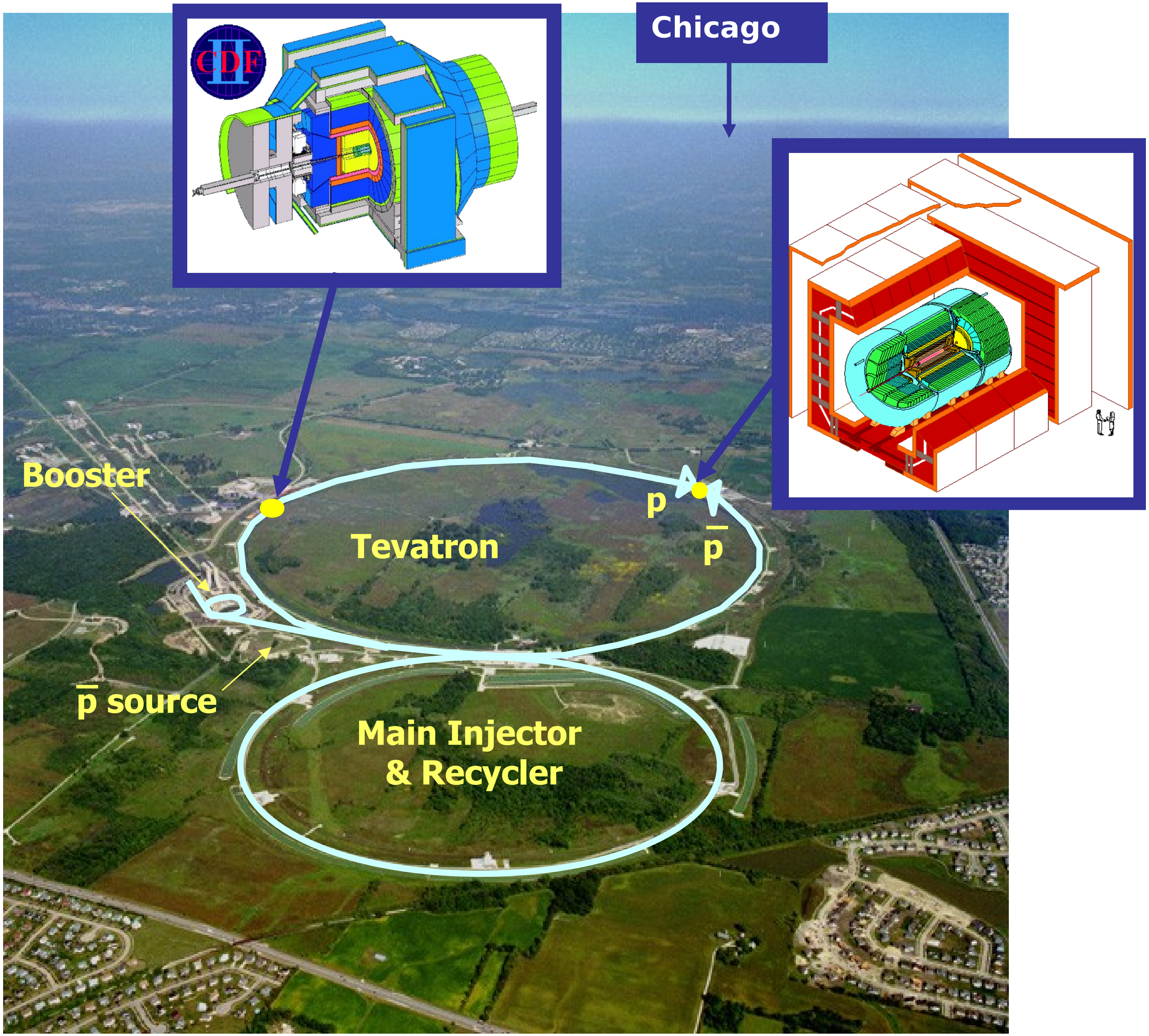
σ (NNLO) [pb]	**tb**	**tqb**	**tW**
LHC prediction (7 TeV)	4.6	64.6	15.7
ATLAS (0.7-2.1 fb^{-1})	<20.5 (95%CL)	**83 ± 20**	17 ± 6
CMS (1.2-4.9 fb^{-1})	-	**67 ± 6**	16 ± 5

Discovery (> 5 SD) Evidence (>3 SD)

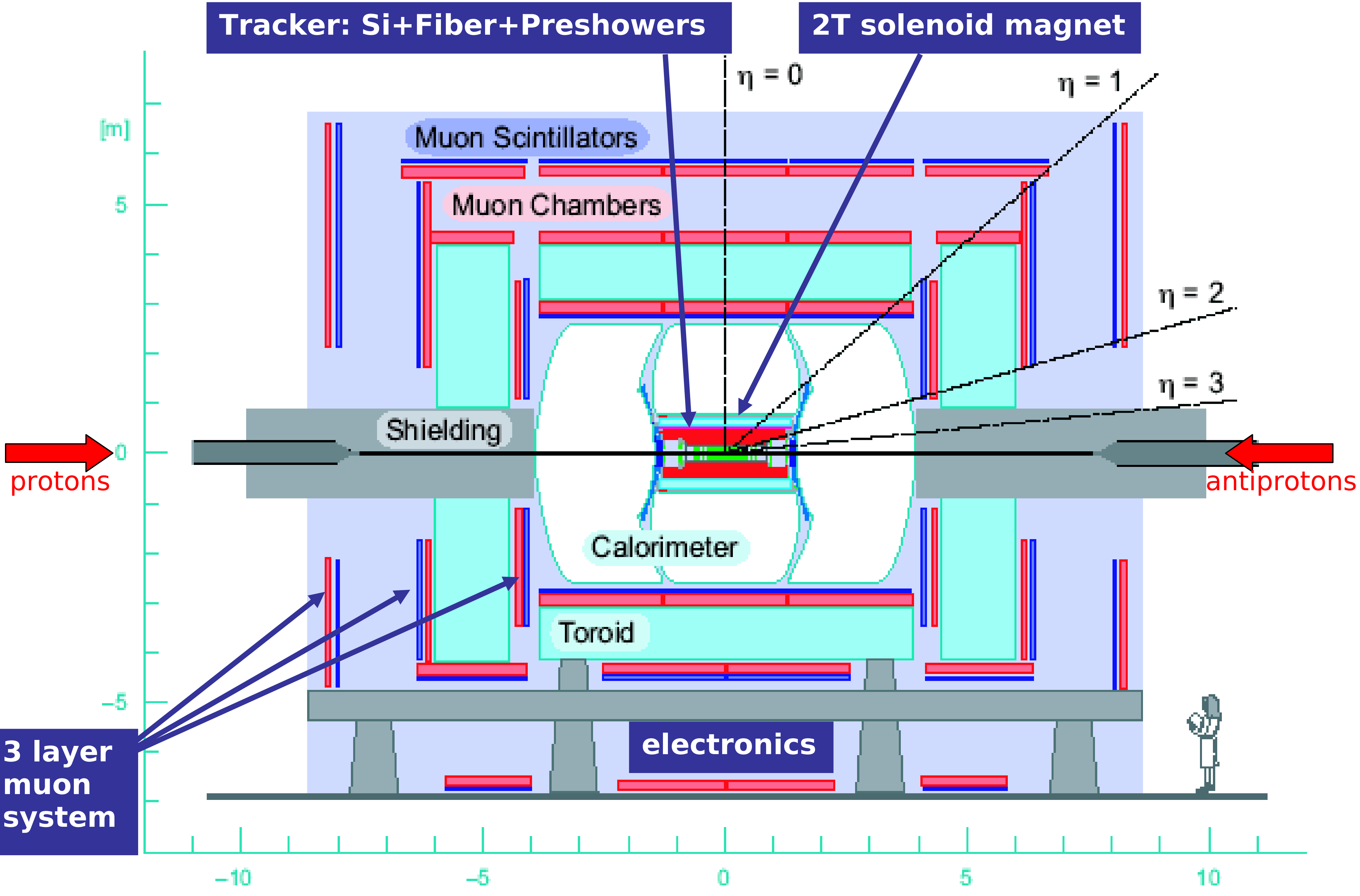

σ [pb]
10^2
10
1
▼ CDF
▲ D0
□ ATLAS
○ CMS
t-channel PRD83:091503 (NNLO)
s-channel PRD81:054028 (NNLO)
tW-channel PRD82:054018,PRD74:114012
1
2
3
4
5
6
7
8
9
10
√s [TeV]

The Tevatron

- 6.3 km $p\bar{p}$ collider
- $\sqrt{s} = 1.96$ TeV
- Run I: 1987-1996
- Run II: 2002-2011
  - 36x36 bunches
  - 10^{11} $\bar{p}$ per bunch
  - 396 ns bunch spacing
  - 1.8 M crossings/s
  - $4.3 \cdot 10^{32}$ $cm^2 s^{-1}$ peak lumi
  - 12 fb^{-1} delivered luminosity
- Detectors recorded data with >90% efficiency



DØ for Run II

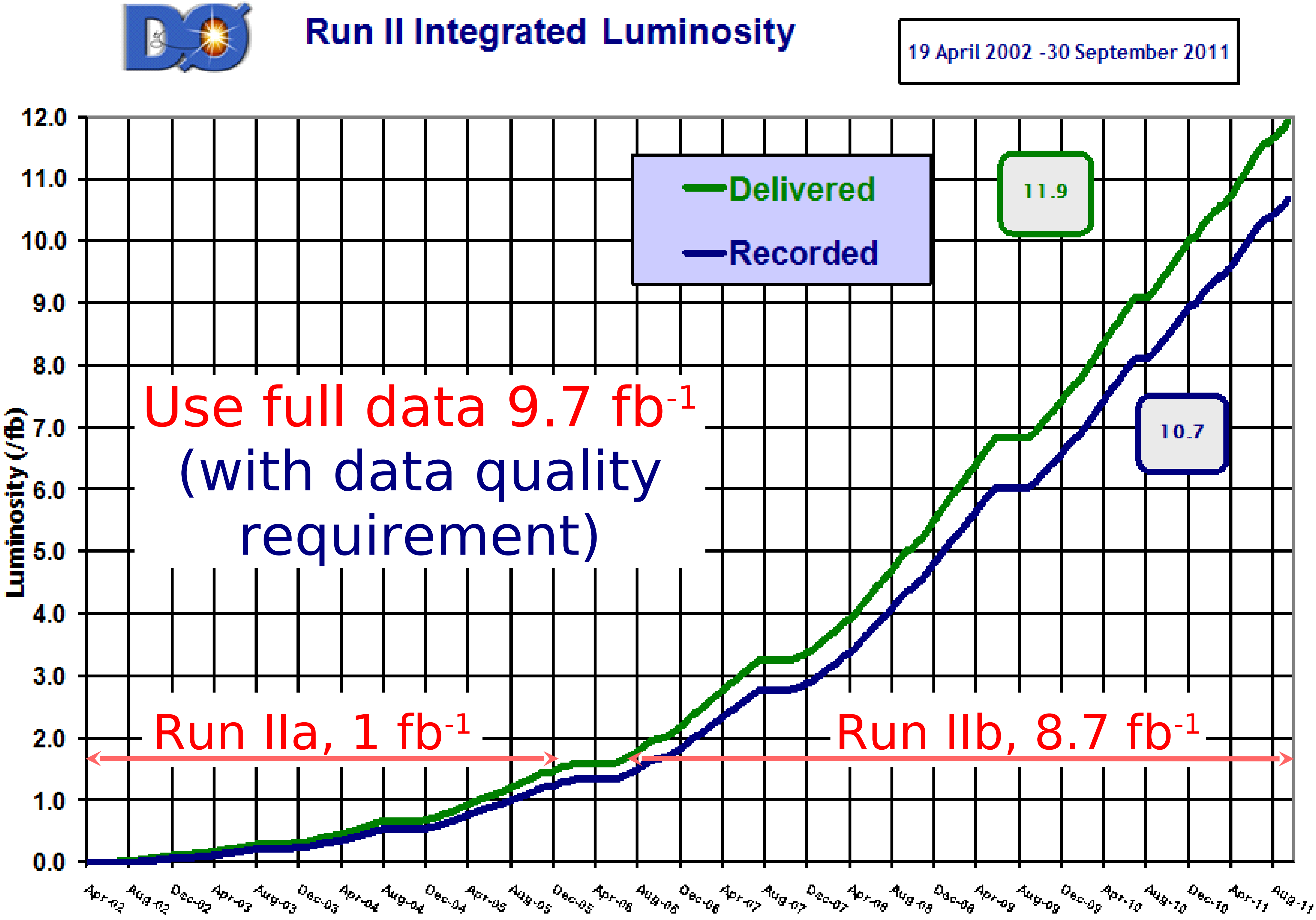


Tracker: Si+Fiber+Preshowers

2T solenoid magnet

3 layer muon system

electronics

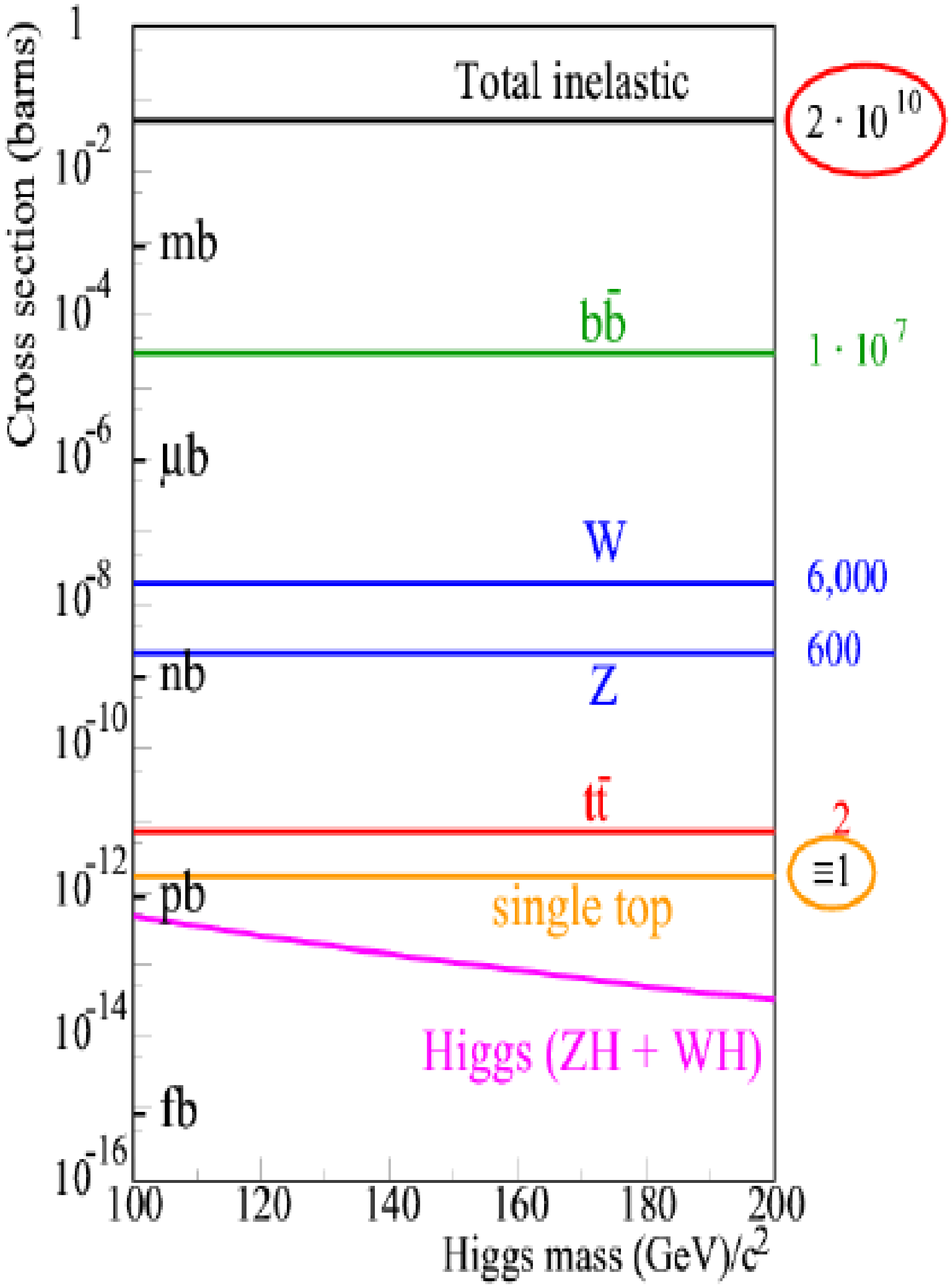


Use full data 9.7 fb^{-1}
(with data quality
requirement)

Run IIa, 1 fb^{-1}

Run IIb, 8.7 fb^{-1}

A big challenge!

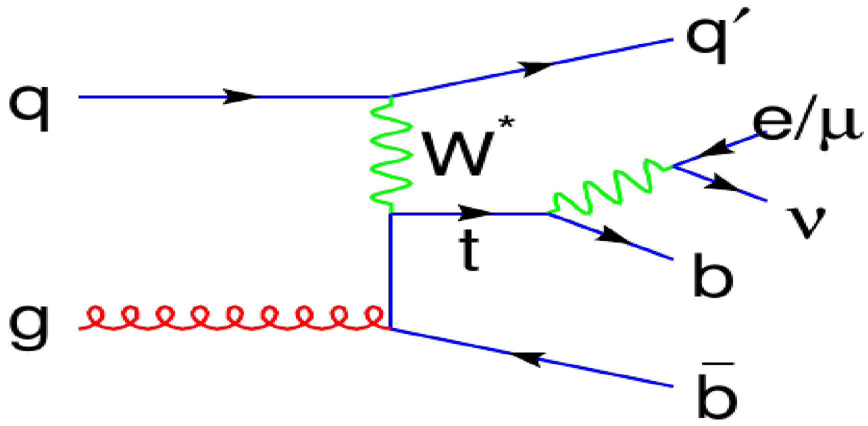


- 32k singletop events produced at the Tevatron
  → Leptonic decays: 6.8k events
- Compare to huge W+jets background
- We needed 50 times more data to discover singletop (s+t) in 2009 than for $t\bar{t}$ in 1995

New for this analysis:

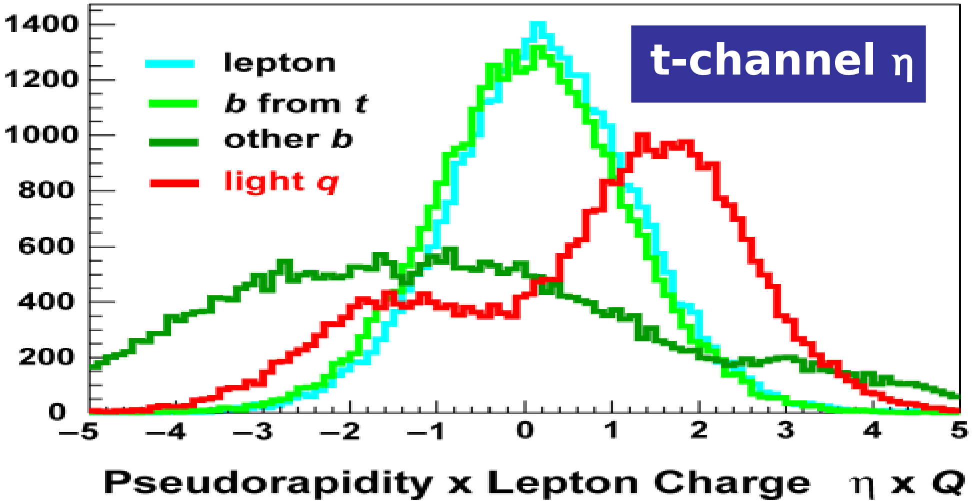
- Optimized for s-channel
- Inclusive trigger
- Better b-tag algorithm
- Use matrix method for W+jets and multijet normalization
- New discriminants with b-tag info, and improved ME method
- Not assume SM ratio σ_{tb}/σ_{tqb} for s+t or V_{tb} measurements
- New joint tb and tqb discriminant

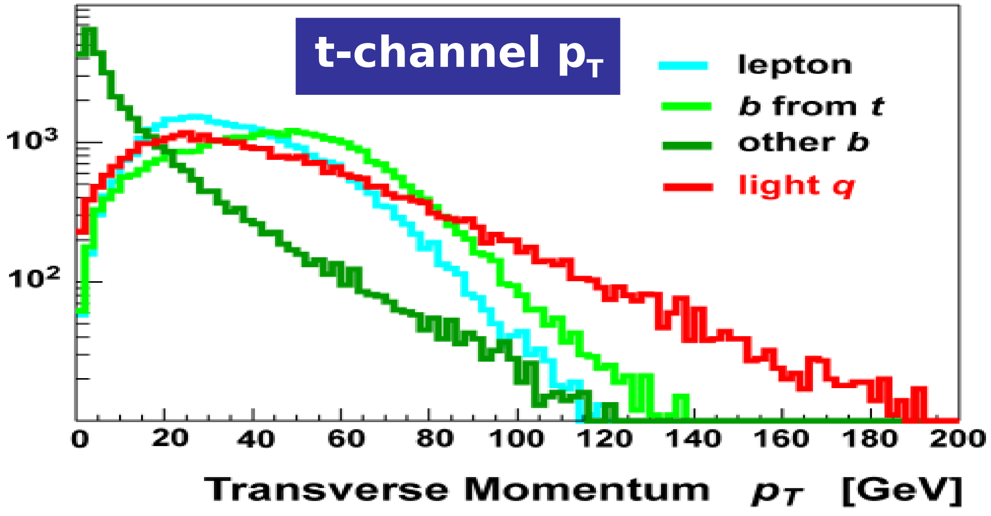
Signal selection



Signature:

- One high p_T isolated lepton (from W)
- MET (ν from W)
- One b-quark jet (from top)
- A light flavor jet and/or another b-jet



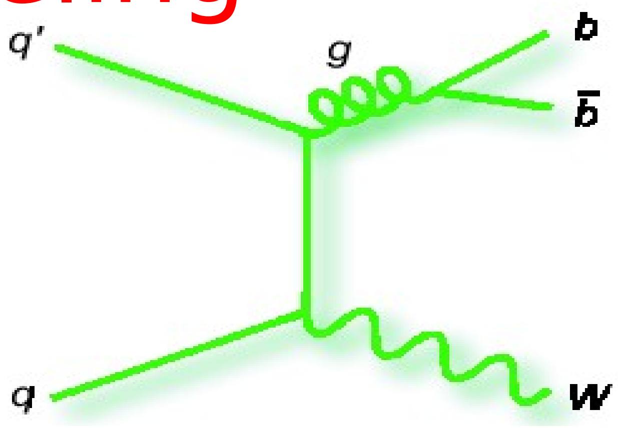


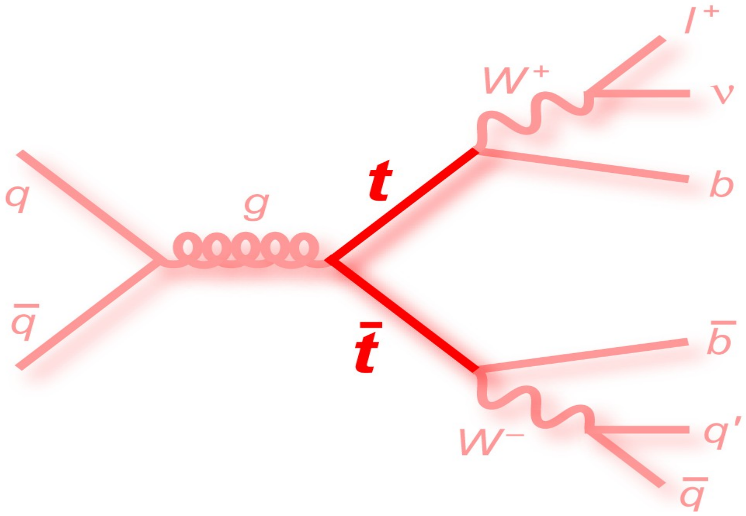
Event selection:

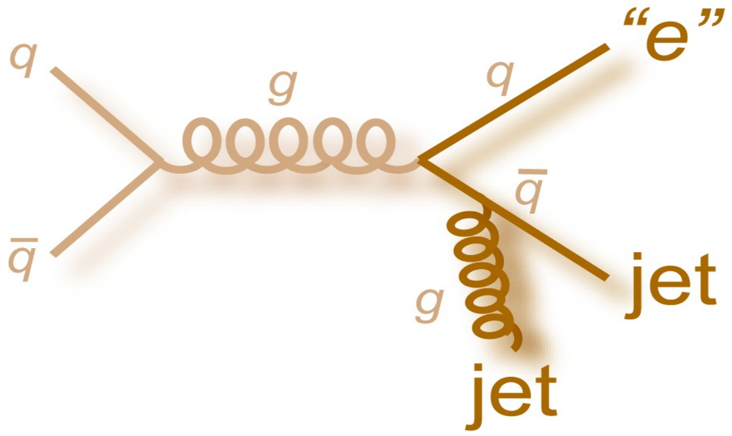
- Only one isolated electron or muon, p_T>20 GeV:
  - Electron: $|\eta|<1.1$
  - Muon: $|\eta|<2.0$
- MET > 20 GeV
- 2-3 jets: p_T >20 GeV and $|\eta|<2.5$
  - Leading jet: p_T>25GeV
  - Second leading jet: p_T >20 GeV
- $H_T(\ell,\text{MET},\text{jets})$ > 120 GeV
- One or two b-tagged jets

Background modeling

- W+jets: ~o(1000) pb
  - Distributions from Alpgen (MLM matching ME↔PS) + Pythia
  - Reweigh η(jet1), η(jet2) from data
  - Normalization from pre-tag data
  - Heavy flavor fraction from NLO
    - Wb$\bar{b}$ k-factor ~ 1.9
- Top pairs: ~7 pb
  - Topologies: dilepton and ℓ+jets
  - Distributions from Alpgen
  - Normalize to NNLO σ
- Multijet events (misidentified lepton)
  - From data with non-isolated lepton
- Z+jets from Alpgen (scaled to NLO)
- Diboson (WW, WZ, ZZ) from Pythia

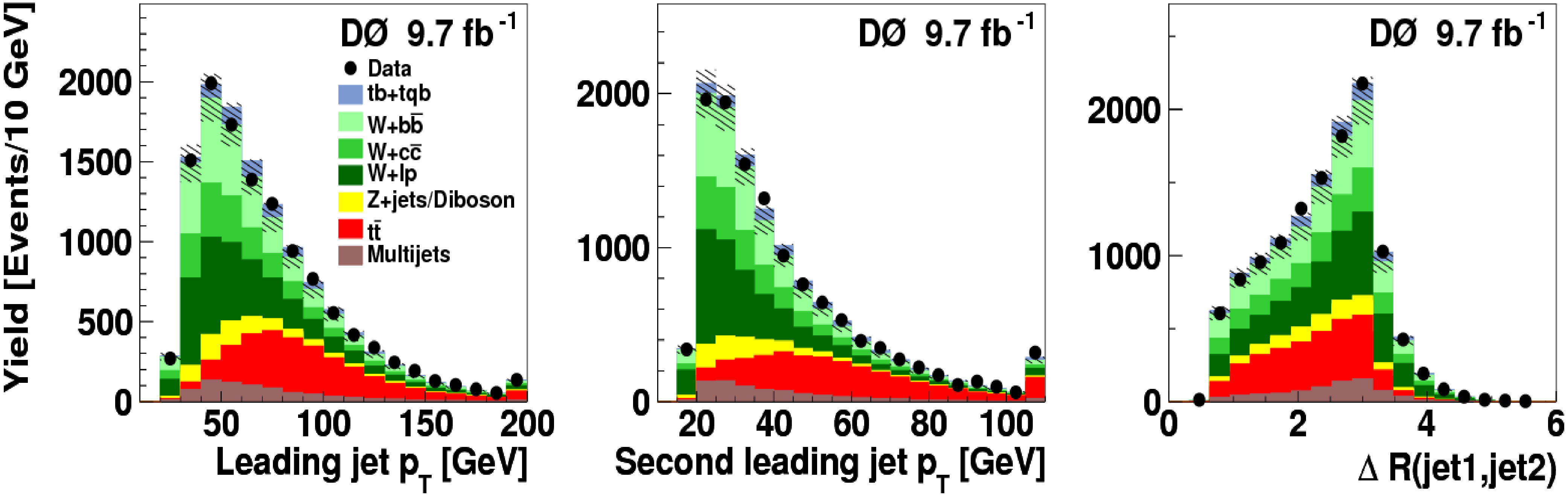






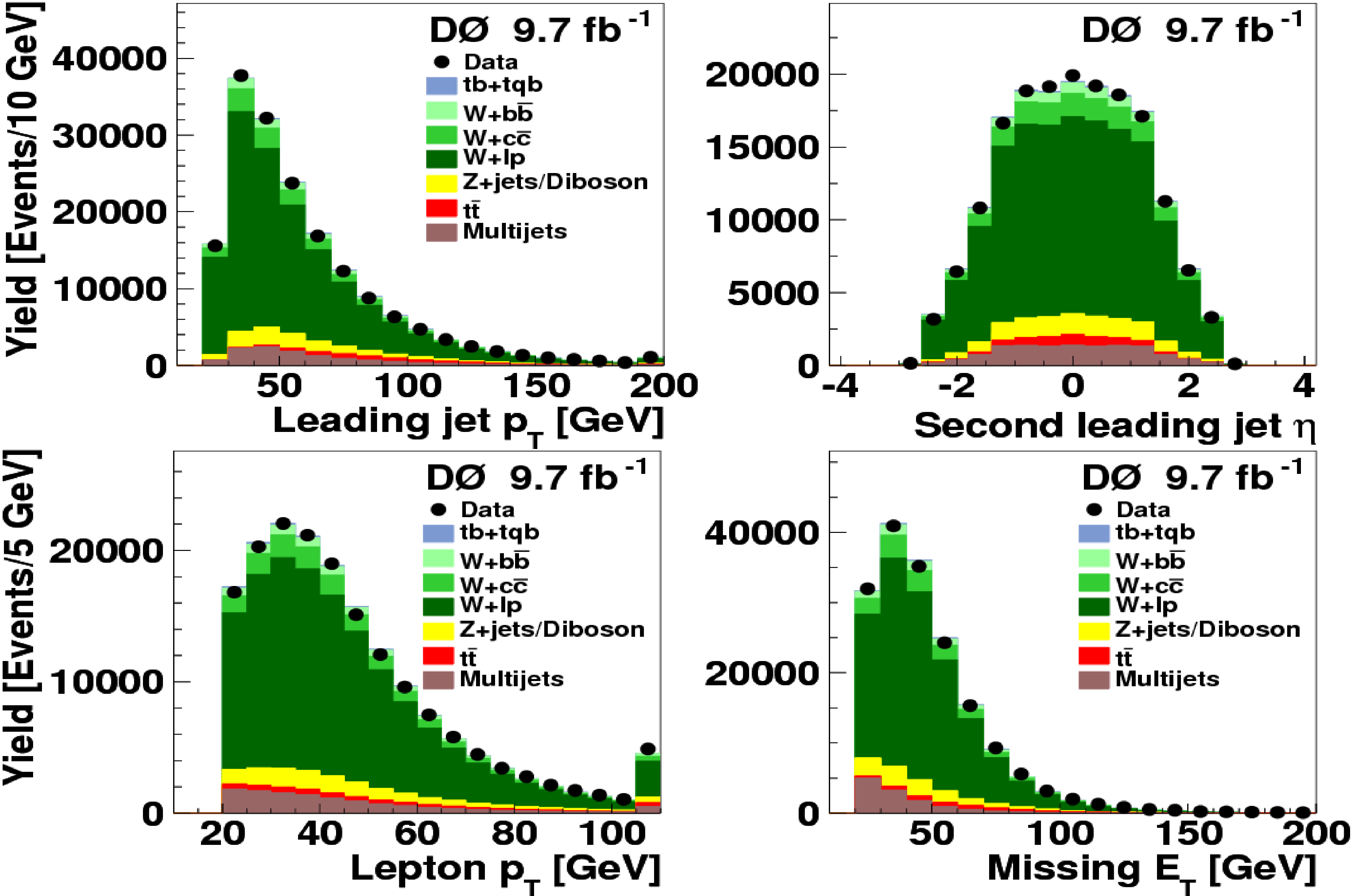
W+jets yield determination

- Normalize W+jets and QCD to data simultaneously before b-tagging (Matrix Method)
  - Split data sample in events with real and fake isolated lepton
  - Measure the probability to have an isolated lepton in each sample
- There are large k-factors for Wbb, Wcc and Wcj
- Determine W$b\bar{b}$, W$c\bar{c}$, Wcj relative fraction from MCFM
  - Cross check with 0-tag sample, and by fitting the b-ID output distribution
  - Source of largest single uncertainty: 20% relative error on HF content

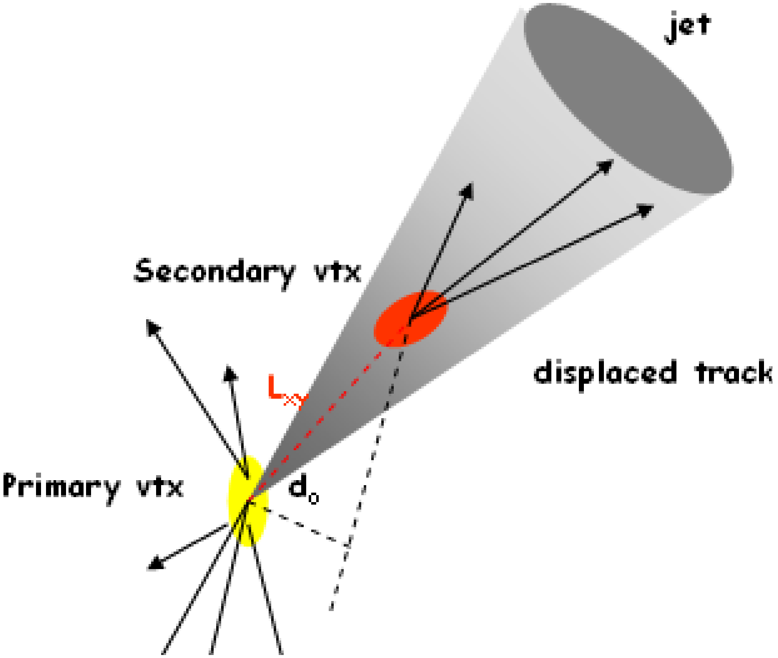


Agreement before b-tagging

- Normalize W+jets and QCD yields to data before b-tagging

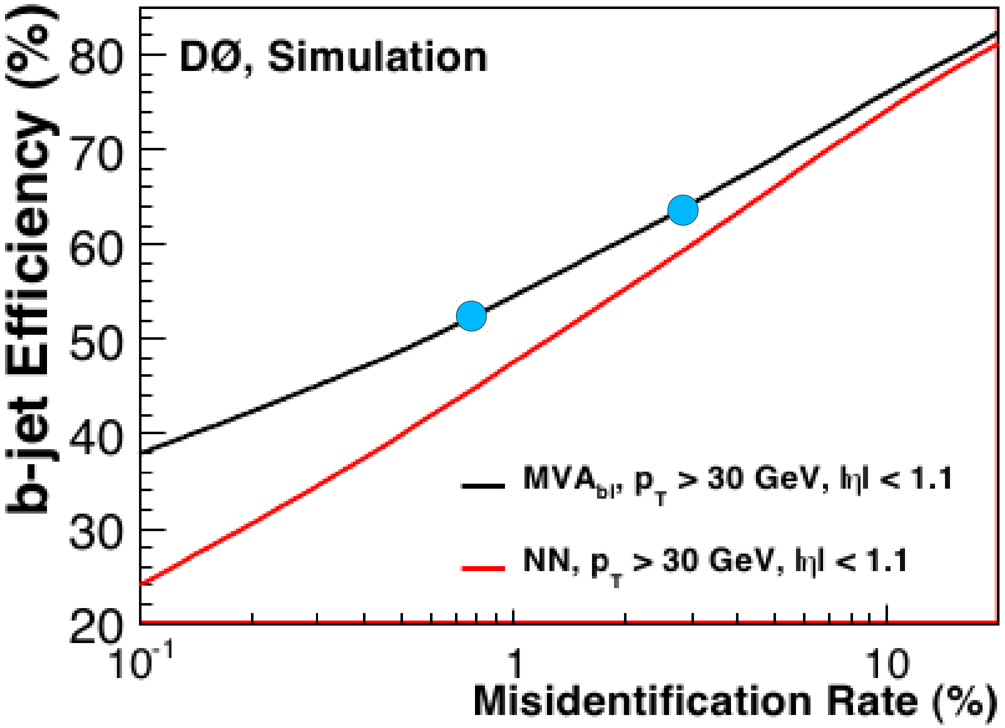


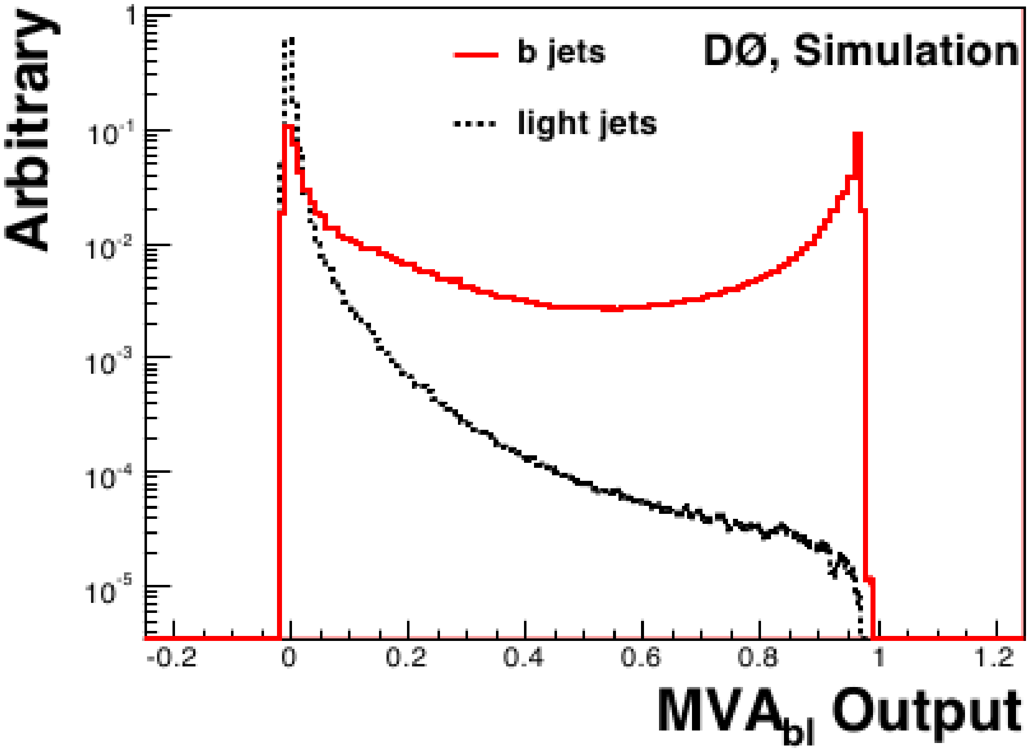
Tagging b-jets



Three different algorithms for b-jet identification at DØ:

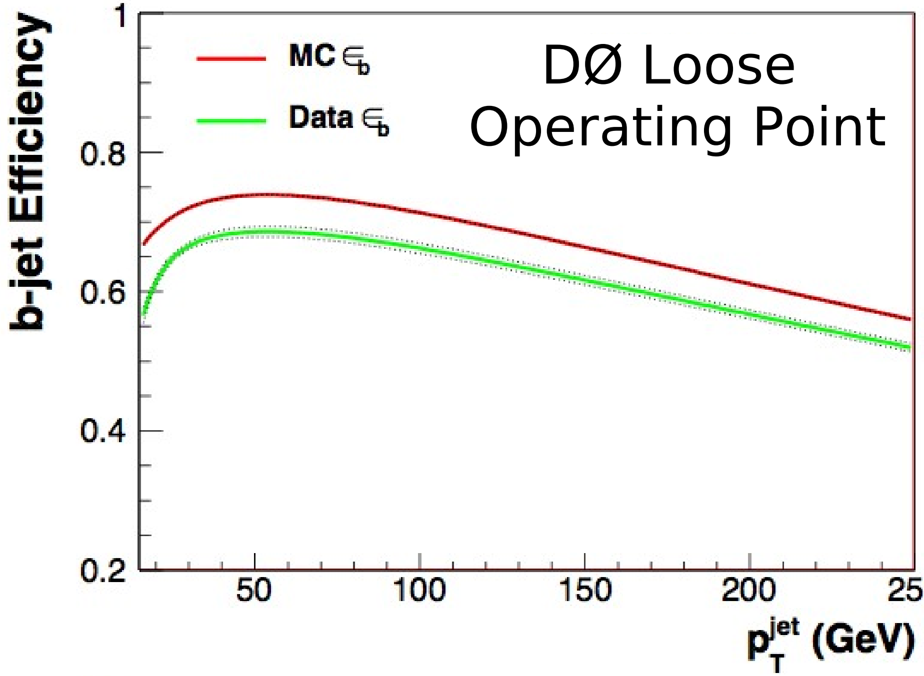
- Two based on tracks with large IP (JLIP, CSIP)
- One based on secondary vertex reconstruction (SVT)
- Combine with MVA NEW

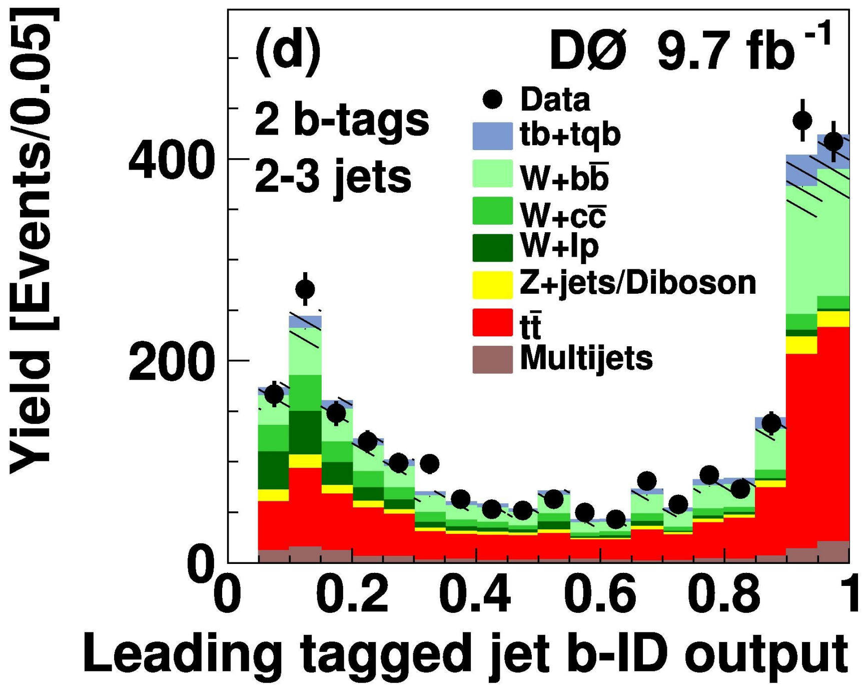


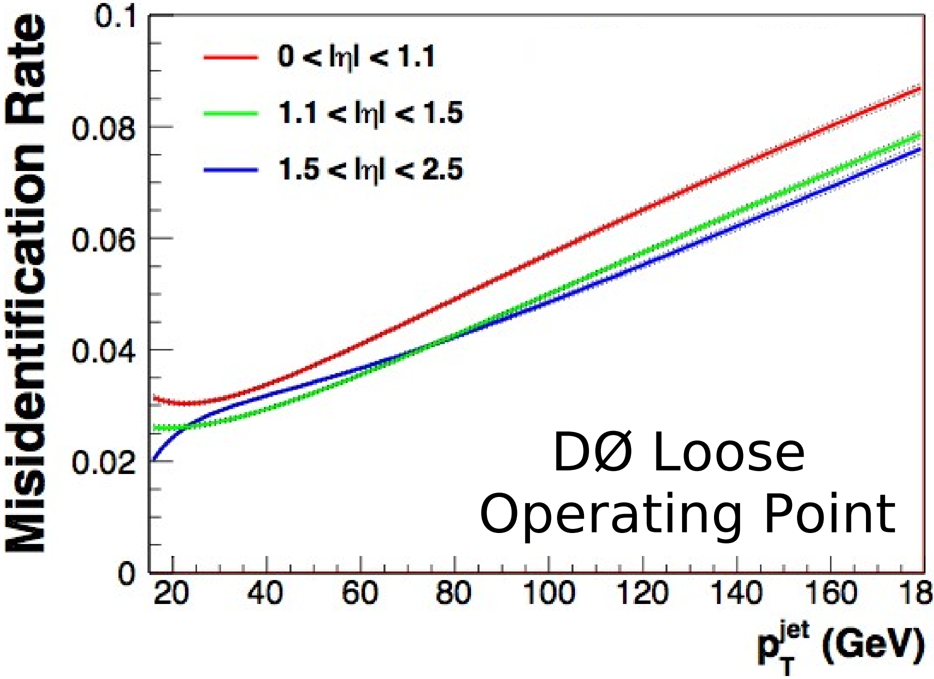


MVA b-jet tagger

- Scale factors derived from data are applied to MC as a function of η, p_T, and z-PV
- 1 tag category: Tight operating point
- 2 tag category: Loose operating point
- Efficiencies in 2 b-tag channel:
  - b-jet efficiency: ~65% per jet
  - c-jet efficiency: ~30% per jet
  - Light efficiency: ~2.9% per jet

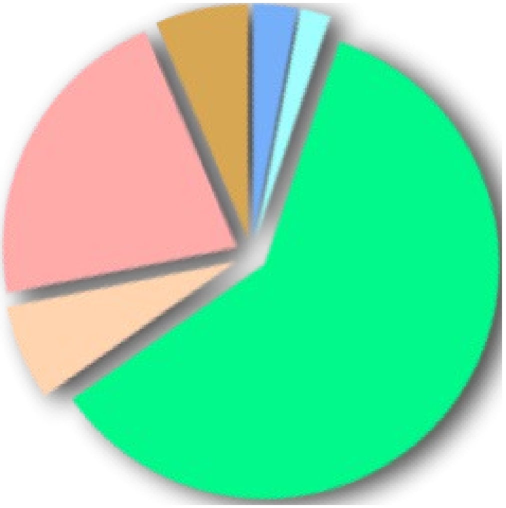






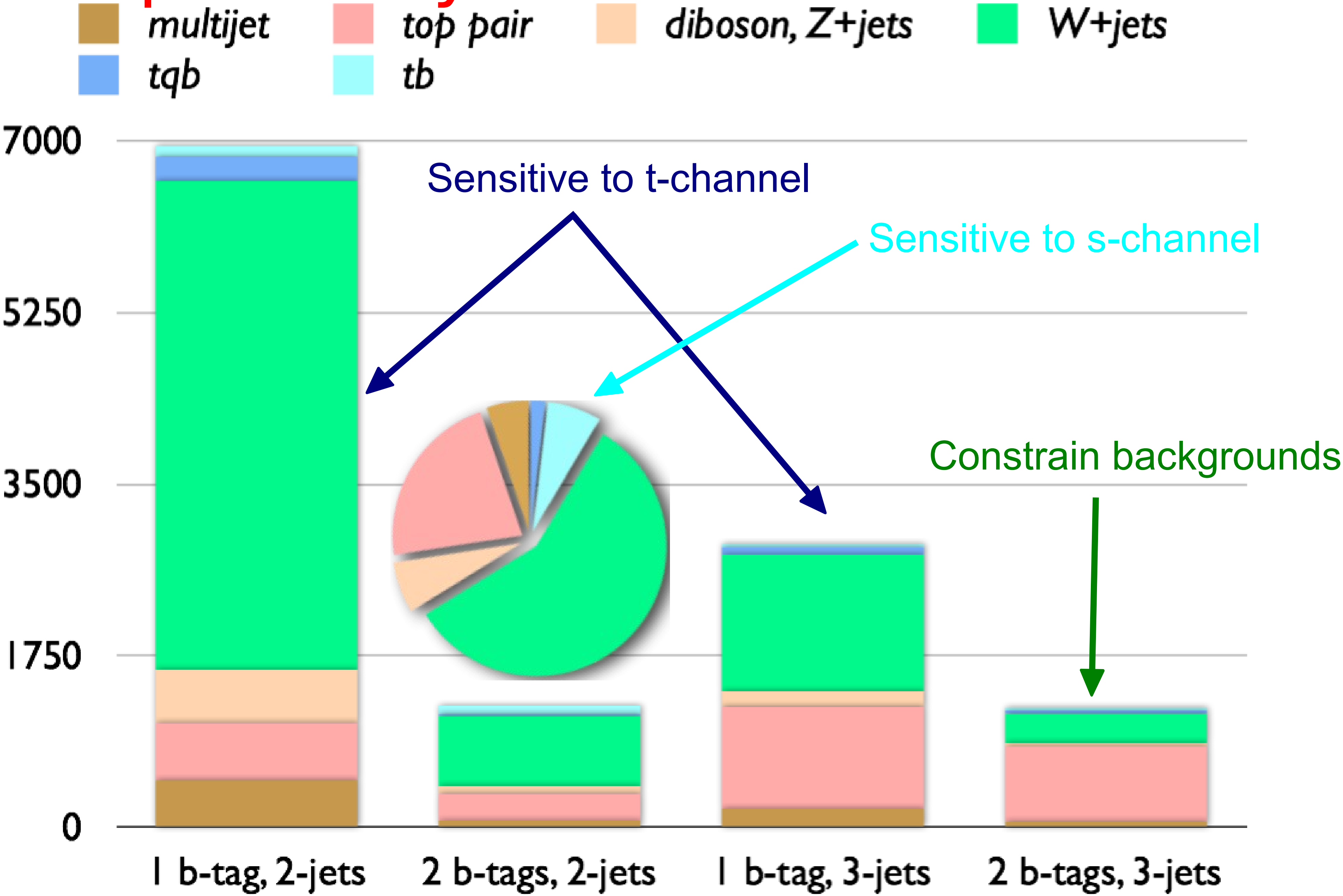
Yields after event selection

e, μ 2, 3-jets 1, 2 *b*-tags combined	
tb	257 ± 31
tqb	378 ± 53
W+jets	7394 ± 401
diboson, *Z*+jets	815 ± 71
top pair	2672 ± 284
multijet	789 ± 81
Background Sum	11669 ± 503
Data	12103



- Optimized the selection to maximize acceptance
  tb = 2.6%   *tqb* = 1.8%
- Allow a lot of background at this stage!
- Then use multiple distributions to separate signal-background

Split analysis in four channels



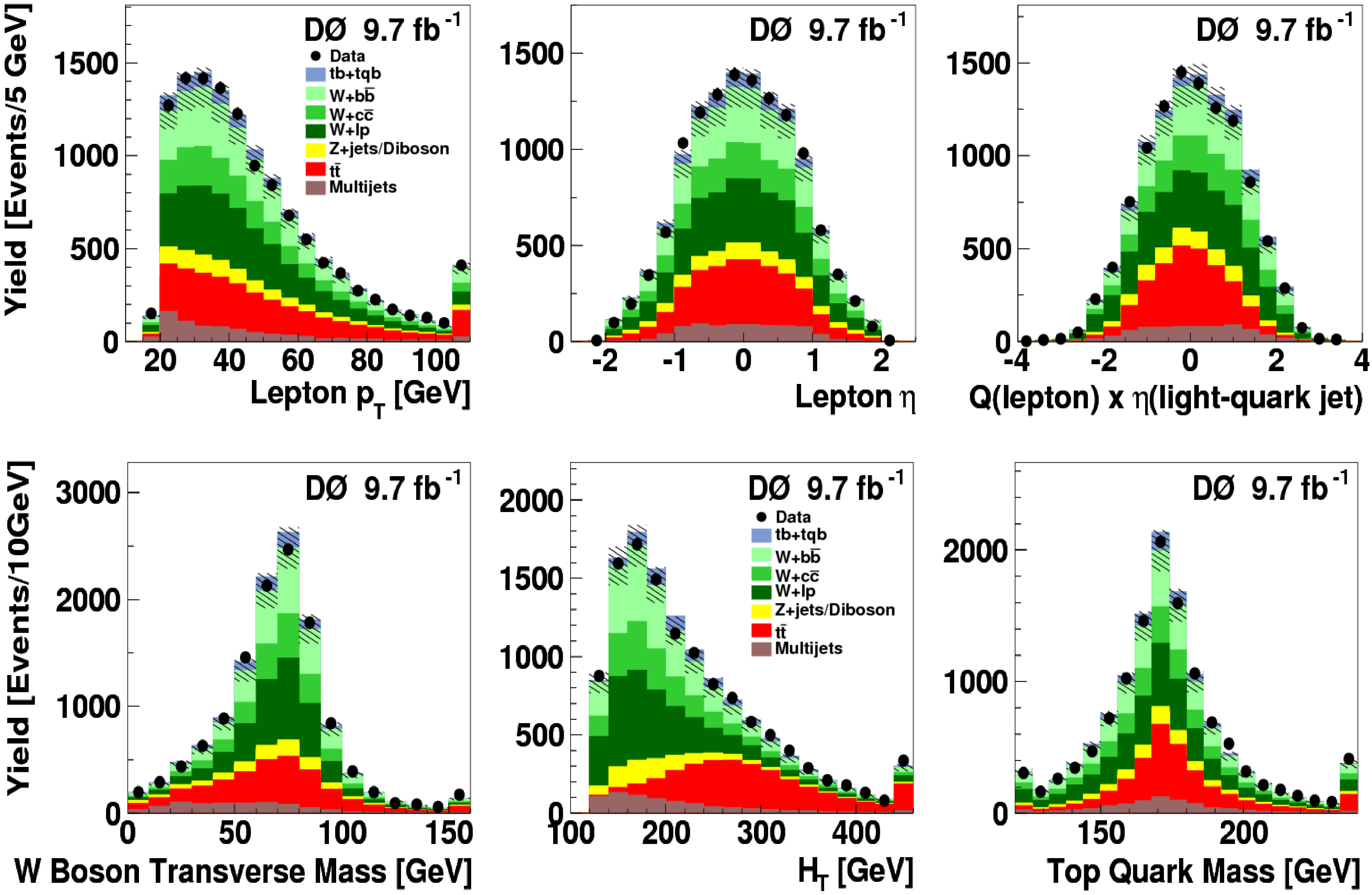
Sensitive to t-channel

Sensitive to s-channel

Constrain backgrounds

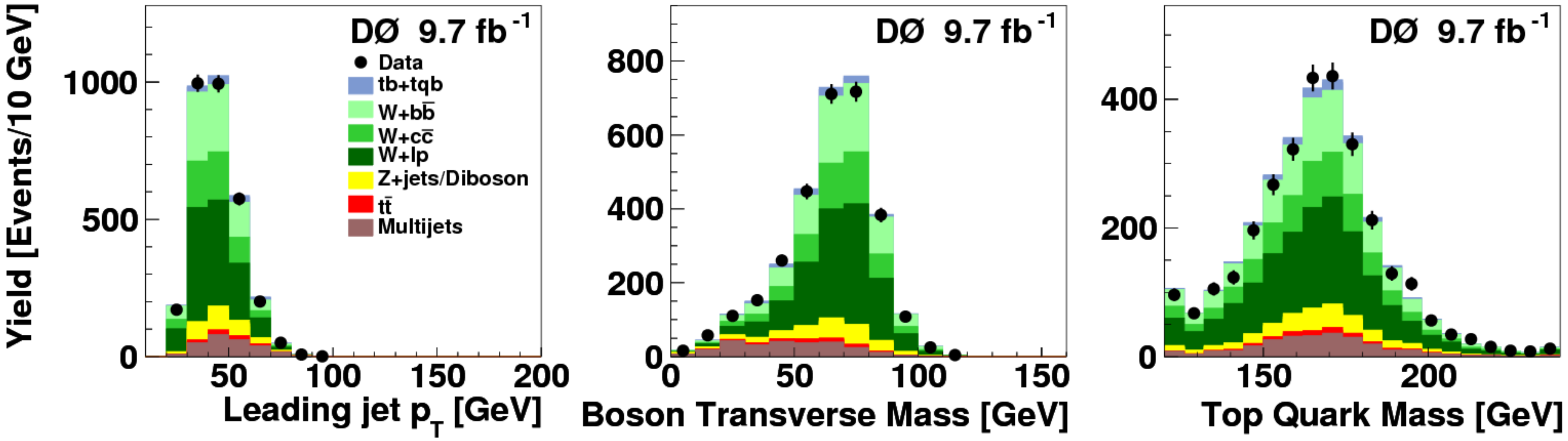
tb:*tqb*:B 1: 2: 59 1: 0.3: 14 1: 2: 85 1: 1: 40

Data-Background comparisons

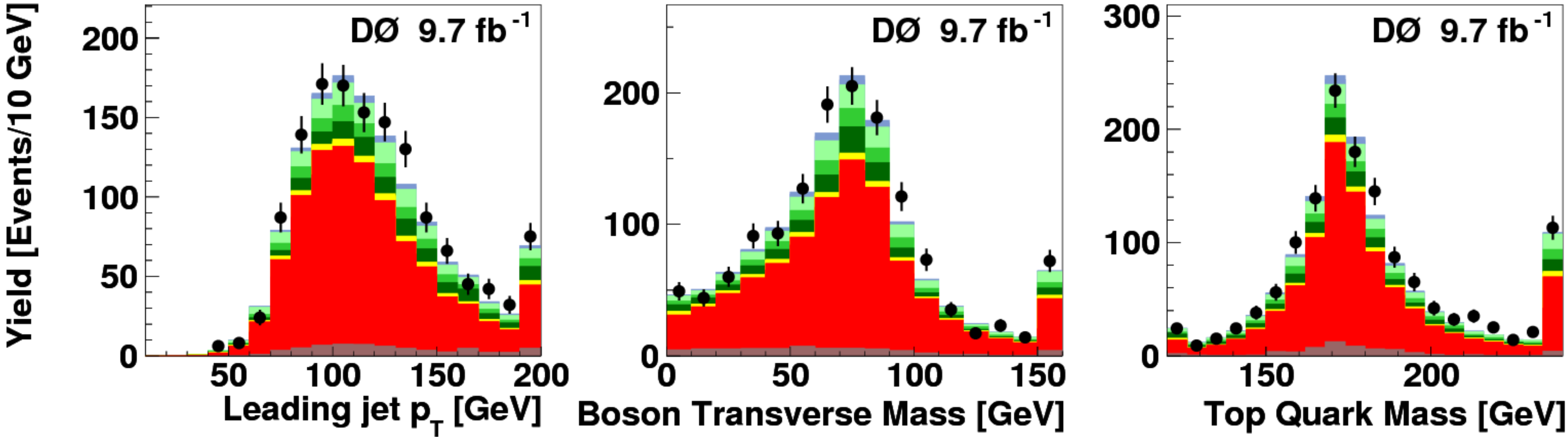


Cross-check samples

- W+jets enriched sample: 2 jets, 1 b-tag, $H_T<175$ GeV



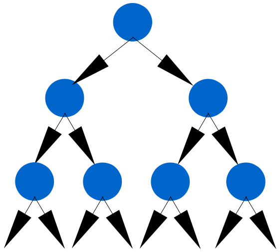
- $t\bar{t}$ enriched sample: 3 jets, ≥1 b-tag, $H_T>300$ GeV



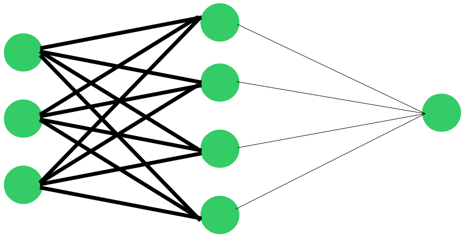
Analysis methods

- Once we understand our data, need to measure the signal
- We cannot use simple cuts to extract the signal: use **multivariate techniques**
- DØ has implemented three analysis methods to extract the signal from the **same dataset**:

Decision Trees



Bayesian NNs



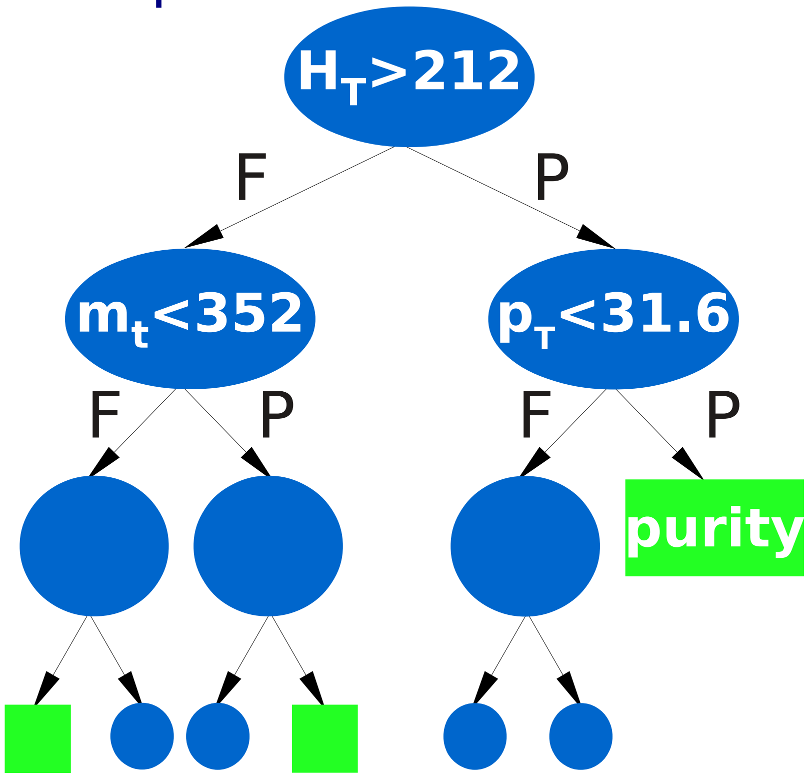
Matrix Elements

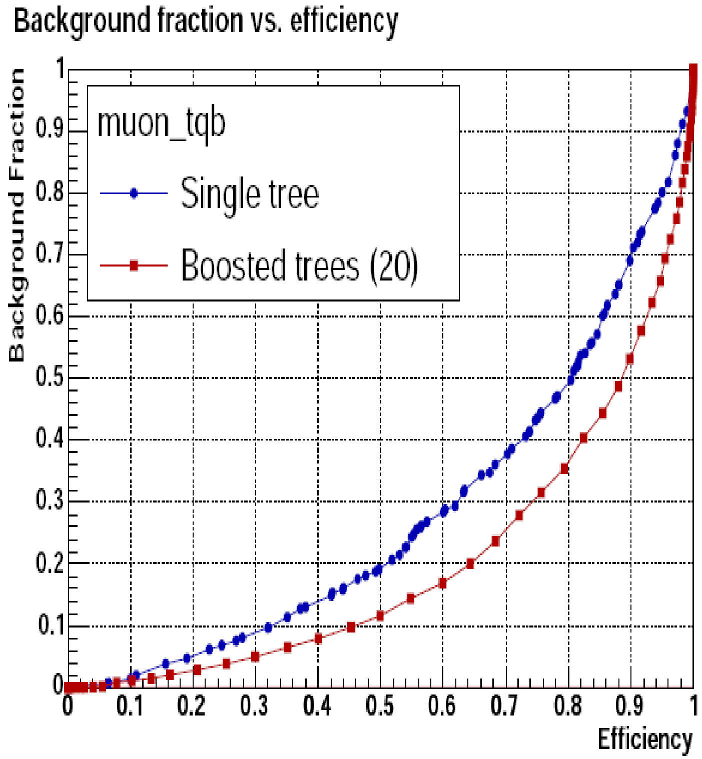
$$\int M$$

- DT, BNN use 1/4 of events for training, 1/2 for measurement
- DT, BNN use well described input variables (KS>0.25)
- DT, BNN are the same used for tqb discovery (5.4 fb^{-1})
- ME method uses 4-vectors of reconstructed objects
- Optimized separately for s-channel and t-channel

Boosted Decision Trees

- Apply sequential cuts but keep the failing events
  - List of 30 variables optimized for s-channel
  - Same list used for all 4 analysis channels
  - Trained with tb and tqb
- Train another tree produced by enhancing misclassified events (boosting)
- Repeat boosting → smoother, more discriminant output

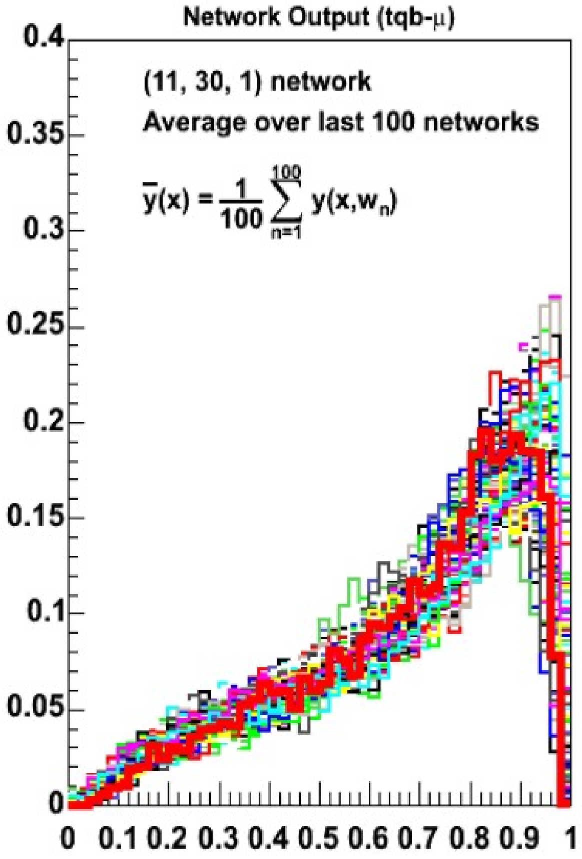


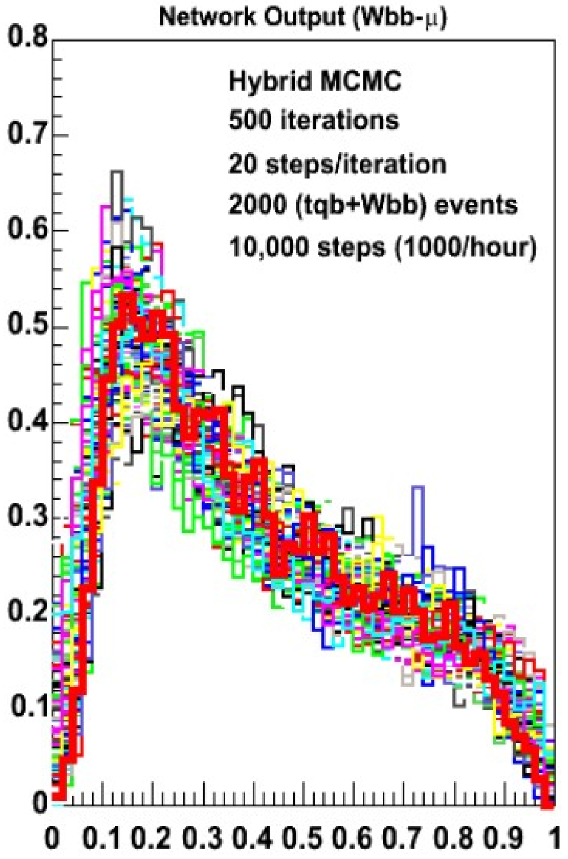


#	BDT input variables
1	$\not{E}_T$
2	$p_T(\ell)$
3	$\eta(\ell)$
4	$M(\text{jet1})$
5	$p_T(\text{untag1})$
6	$E(untag1)$
7	$M(untag1)$
8	$b_{\text{ID}}(\text{untag1})$
9	$p_T(\text{jet2})$
10	$b_{\text{ID}}(\text{tag1})$
11	$\Delta R(\text{jet1},\text{jet2})$
12	$\Delta R_{\text{min}}(\ell,\text{jet})$
13	$\Delta\Phi(\ell,\not{E}_T)$
14	$\Delta\Phi(\text{jet2},\not{E}_T)$
15	$\Delta\Phi(\text{jet1},\not{E}_T)$
16	$Q(\ell)\times\eta(\text{untag1})$
17	$Q(\ell)\times\eta(\text{jet2})$
18	$Q(\ell)\times\eta(\ell)$
19	$Q(\ell)\times\eta(\text{tag1})$
20	$Q(\ell)\times\eta(\text{jet1})$
21	$\cos(\ell,\text{jet2})_{\text{lab}}$
22	$\cos(\ell,\text{jet1})_{\text{lab}}$
23	$H_T(\text{alljets})$
24	$H_T(\ell,\not{E}_T,\text{alljets})$
25	$H_T(\ell,\not{E}_T)$
26	Centrality(alljets)
27	$M\text{jet1},\text{jet2})$
28	$p_T(\text{jet1},\text{jet2})$
29	$M_T(W)$
30	$p_T(W)$

Bayesian Neural Networks

- Uses 4-vectors of objects + Q(ℓ)xη(untag1) + M_T(W) + b-ID output for jets
- Instead of choosing one set of weights, find posterior probability density over all possible weights
- Averages over many networks weighted by the probability of each network given the training data





#	BNN input variables
1	$p_T(\text{tag1})$
2	$\eta(\text{tag1})$
3	$\Delta\Phi(\ell, \text{tag1})$
4	$b_{\text{ID}}(\text{tag1})$
5	$p_T(\text{untag1})$
6	$\eta(\text{untag1})$
7	$\Delta\Phi(\ell, \text{untag1})$
8	$b_{\text{ID}}(\text{untag1})$
9	$p_T(\ell)$
10	$\eta(\ell)$
11	$\not{E}_T$
12	$\Delta\Phi(\ell, \not{E}_T)$
13	$p_T(\text{tag2})$
14	$\eta(\text{tag2})$
15	$\Delta\Phi(\ell, \text{tag2})$
16	$b_{\text{ID}}(\text{tag2})$
17	$p_T(\text{untag2})$
18	$\eta(\text{untag2})$
19	$b_{\text{ID}}(\text{untag2})$
20	$M_T(W)$
21	$Q(\ell) \times \eta(\text{untag1})$

Matrix Elements method

- The idea is to use all available kinematic information from a **fully differential cross-section calculation**

- Calculate an event probability for signal and background hypothesis

$$P(\vec{x})=\frac{1}{\sigma}\Sigma_{x,y}\int f(q_1;Q)dq_1 f(q_2;Q)dq_2\times|M(\vec{y})|^2\phi(\vec{y})dy\times W(\vec{x},\vec{y})$$

Parton distribution functions CTEQ6

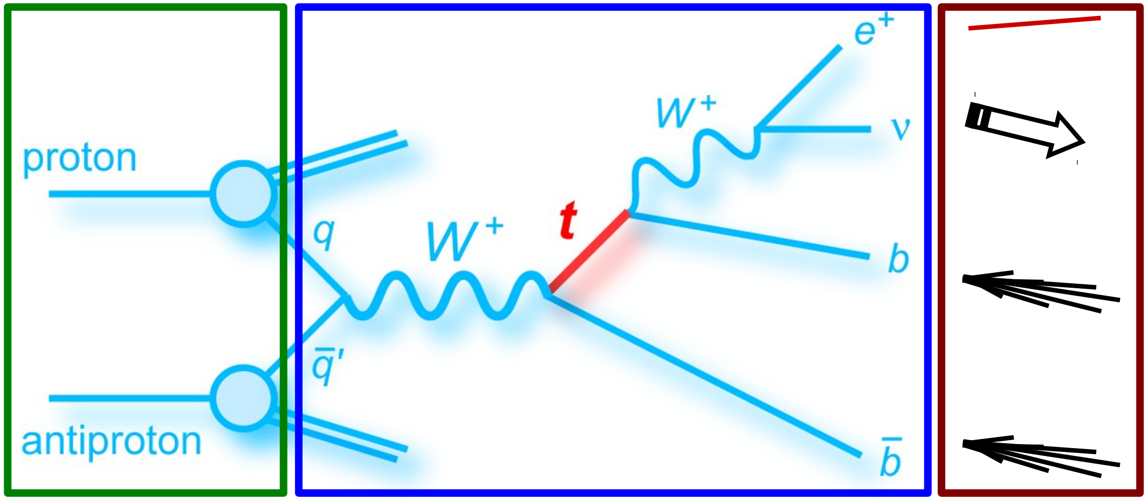
Differential cross section (LO ME from Madgraph)

Transfer Function: maps parton level (y) to reconstructed variables (x)

- Uses the 4-vectors of reconstructed ℓ and jets

- Jet-parton assignment: use b-tag information

- TF for e, μ, jets, and jets misreconstructed as e



- Integrate over 4 (2jet) or 5 (3jet) independent variables: assume angles well measured, known masses, momentum and energy conservation

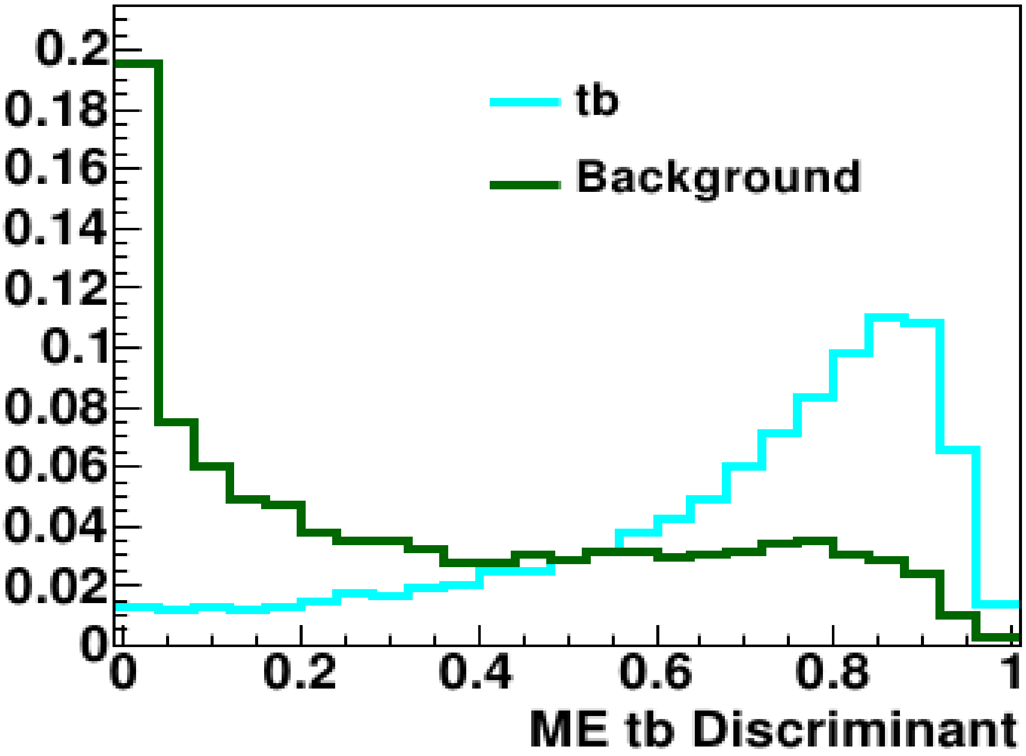
ME discriminants

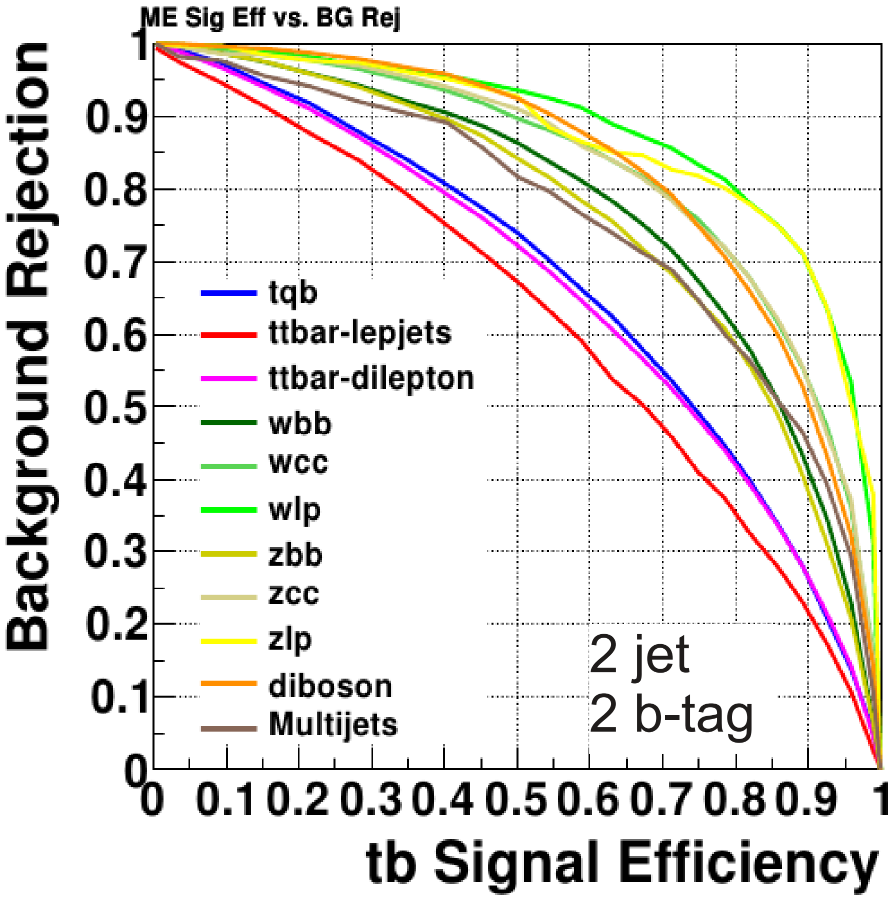
- We use these ME processes
- Define discriminant from probabilities for signal and background

$$D(x) = \frac{P_{signal}(x)}{P_{signal}(x) + P_{background}(x)}$$

	2 Jet	3 Jet
Single Top	*tb, tq*	*tbg, tqb, tqg*
Background	*Wbb, Wcg, Wgg,* *top pair, WW, WZ, ggg*	*Wbbg, Wugg,* *top pair*

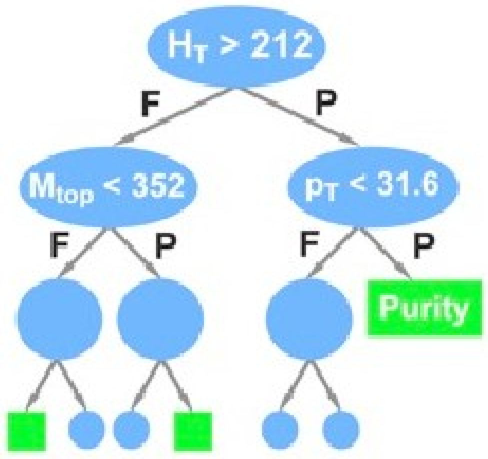
- New: use b-ID weights in Disc.



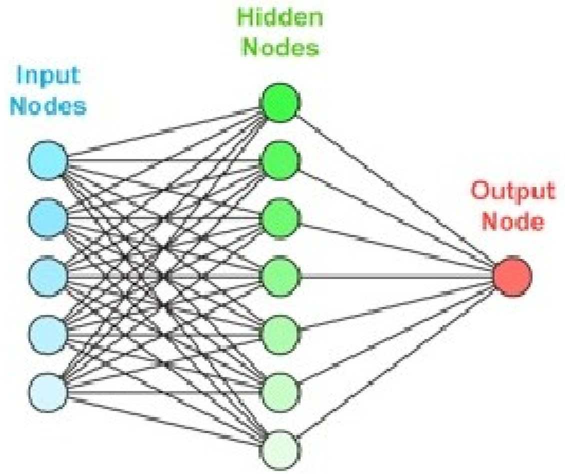


BNN Combination

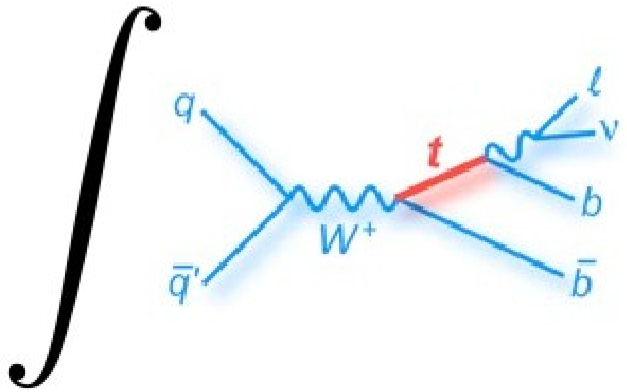
BDT



BNN

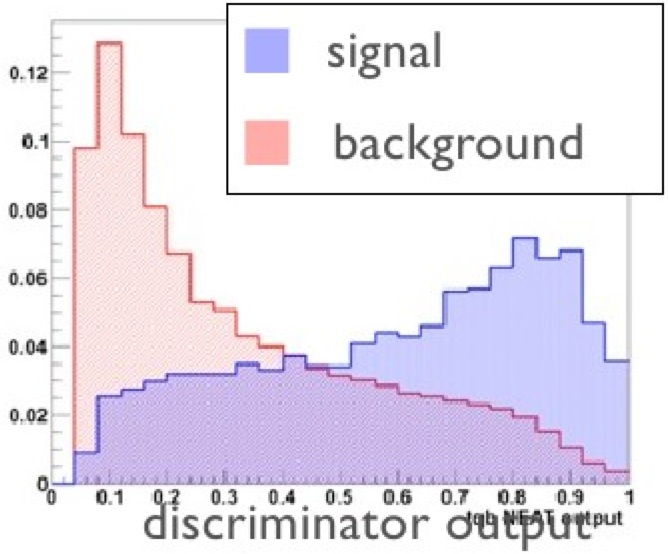


ME

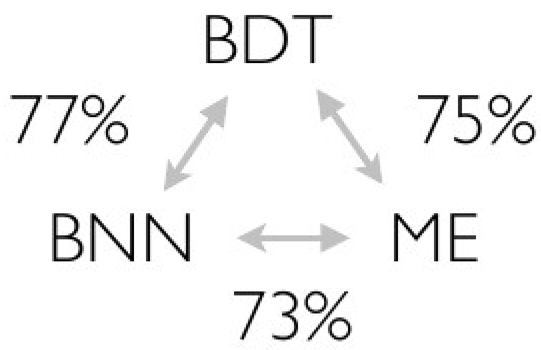


Use BNN to combine the 3 methods

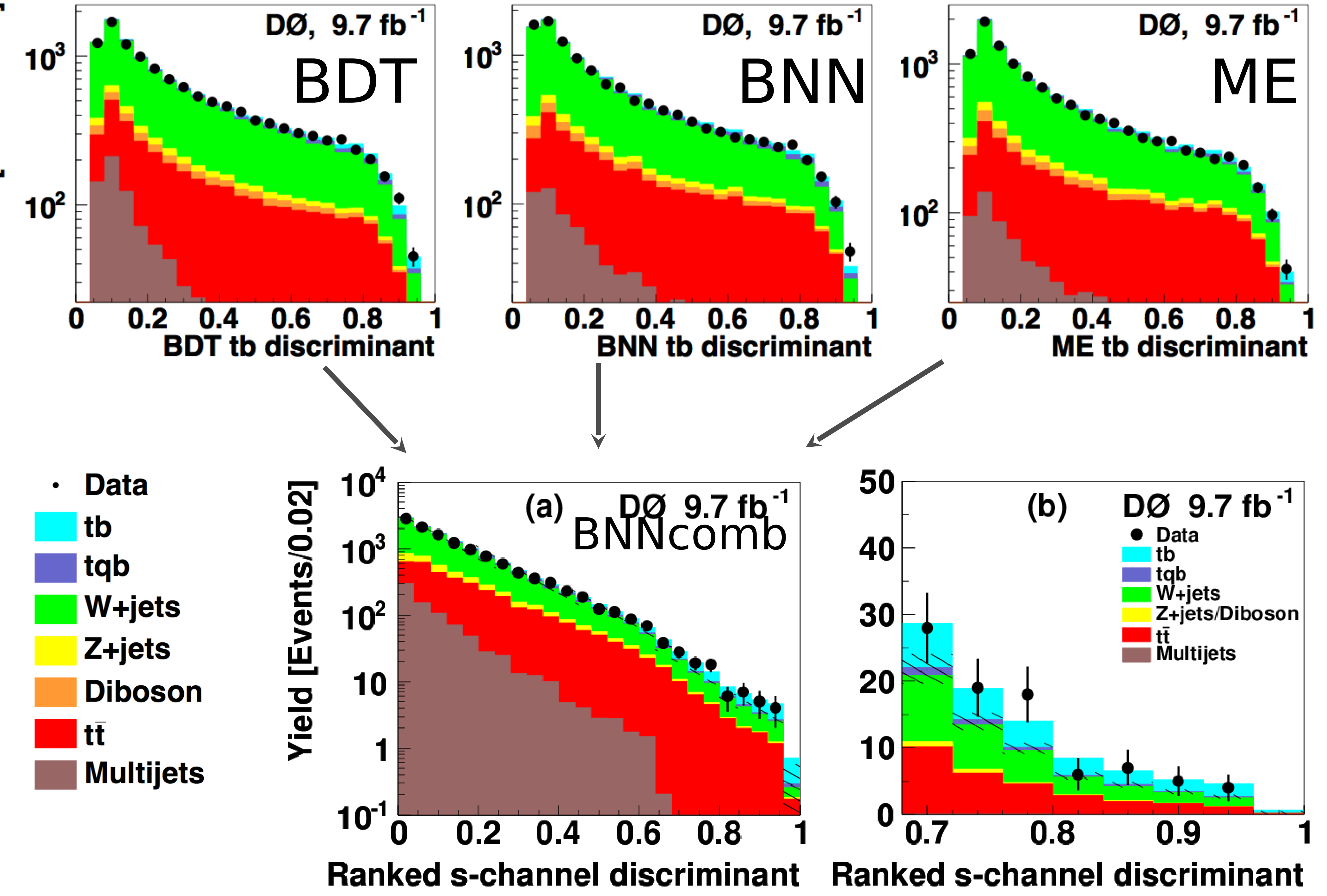
Use 1/4 of sample for training



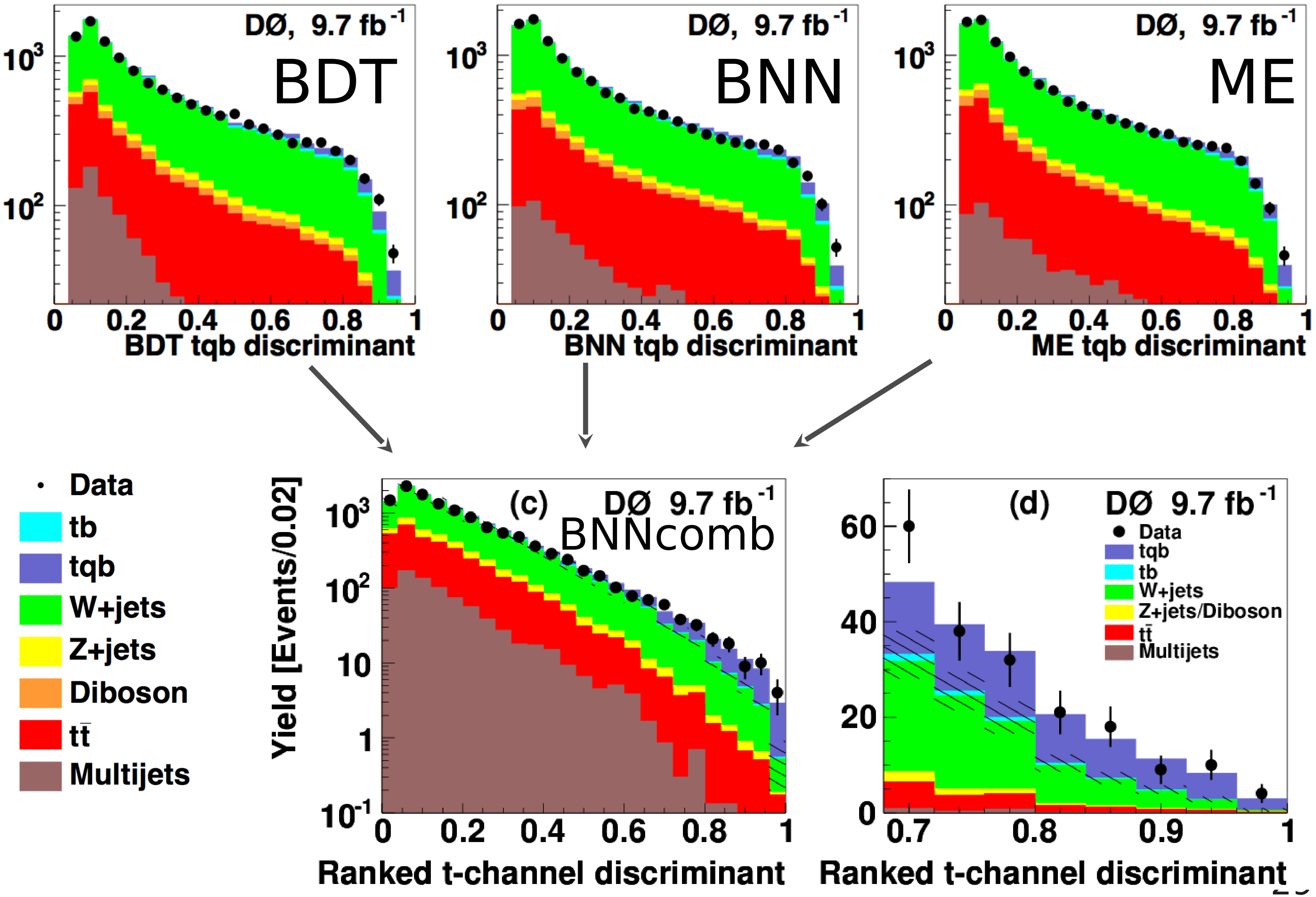
Correlations



BNN tb combination

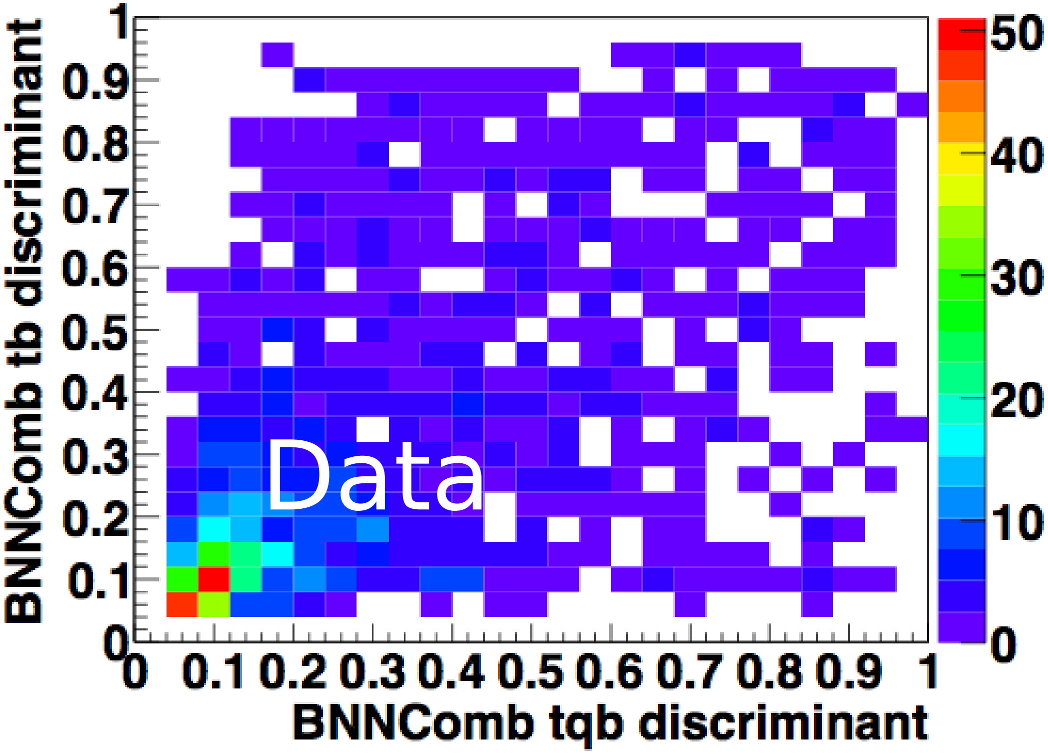


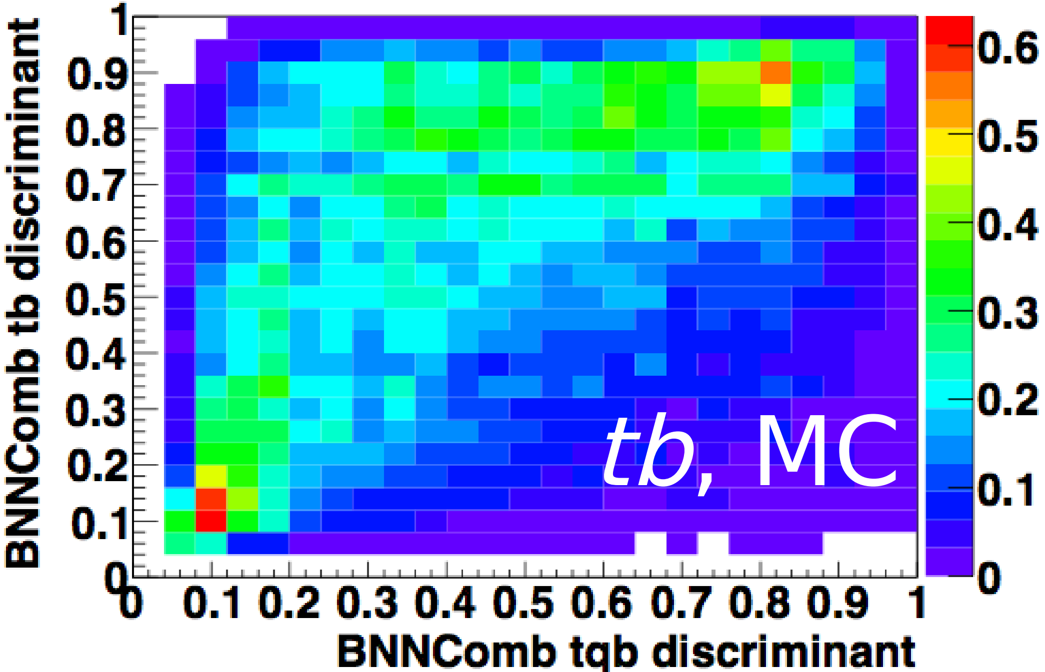
BNN tqb combination



New discriminant

- Aim to simultaneously measure tb and tqb signals without assuming the SM prediction for either
- Need discriminant sensitive to both signals
- Ensure each bin contains enough statistics to have a stable measurement
- Avoid complex binning in 2D
- Split every event based on whether $D_{tb}>D_{tqb}$ or $D_{tb}<D_{tqb}$

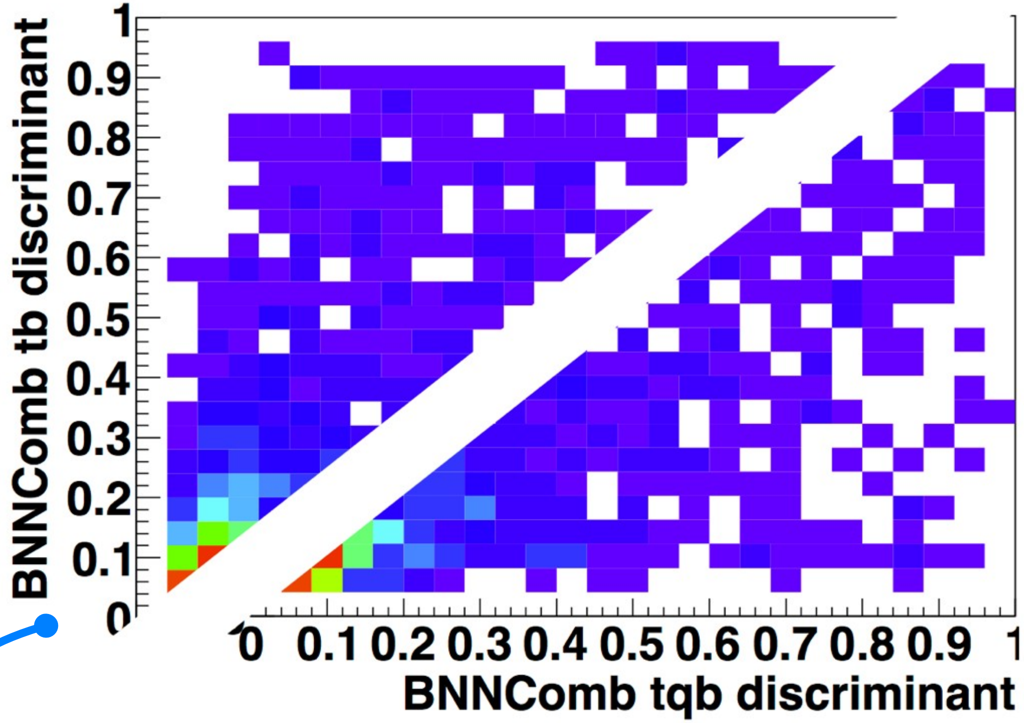




New discriminant

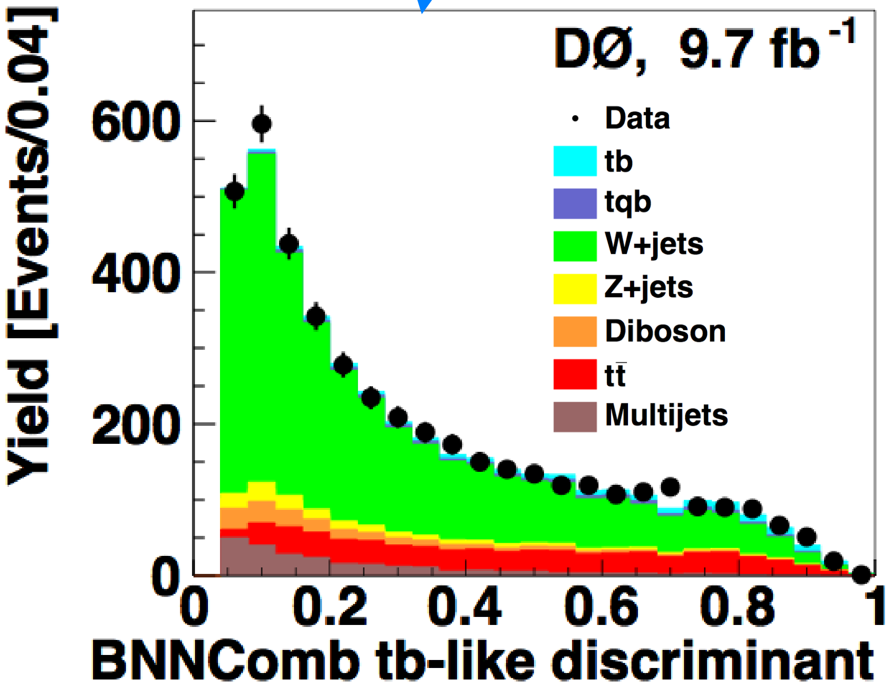
If $D_{tb} > D_{tqb}$:

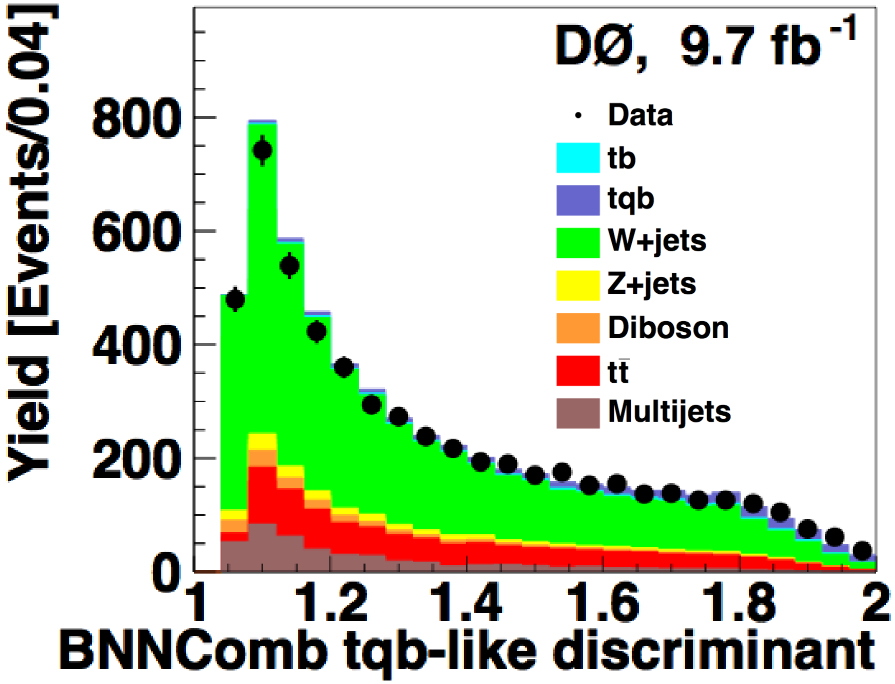
- *tb* category
- Use D_{tb}
- Plot in the range [0, 1]



If $D_{tqb} > D_{tb}$:

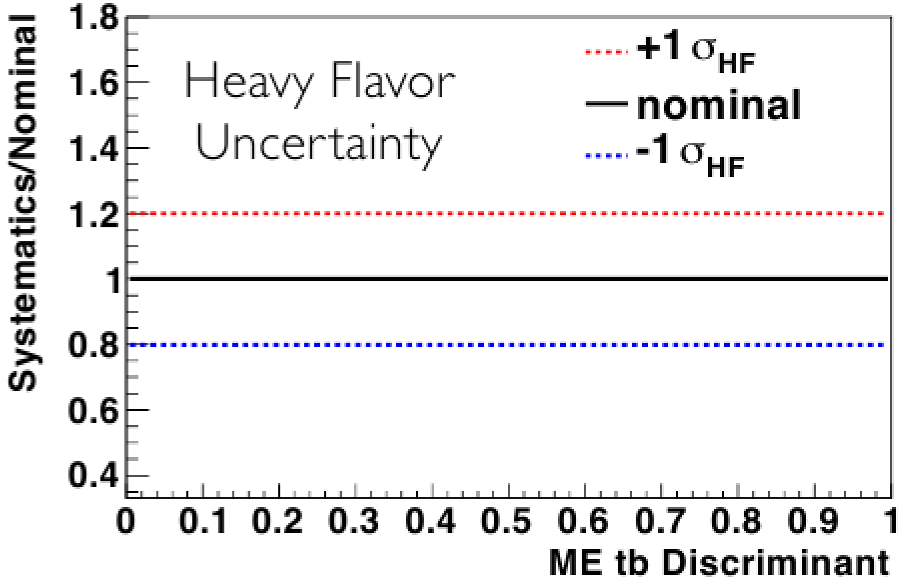
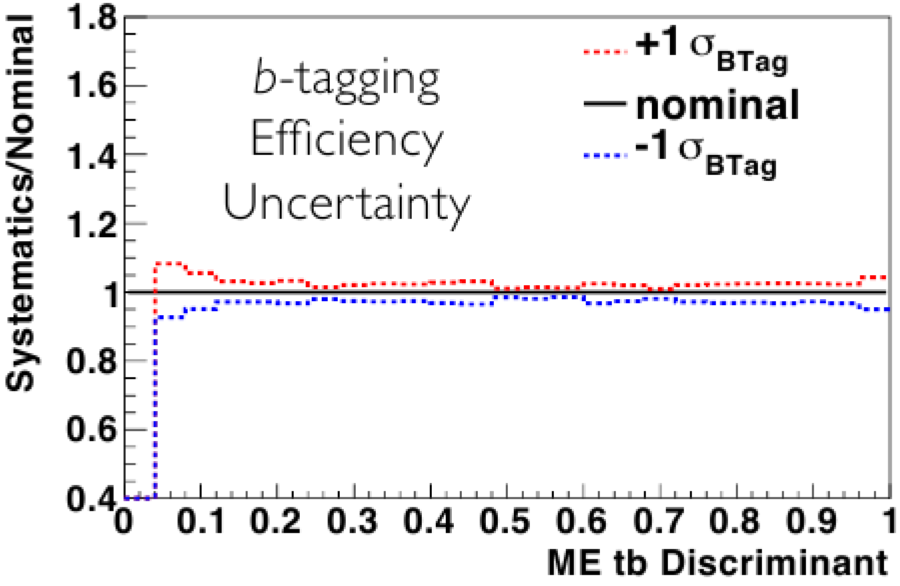
- *tqb* category
- Use D_{tqb}
- Plot in the range [1, 2]





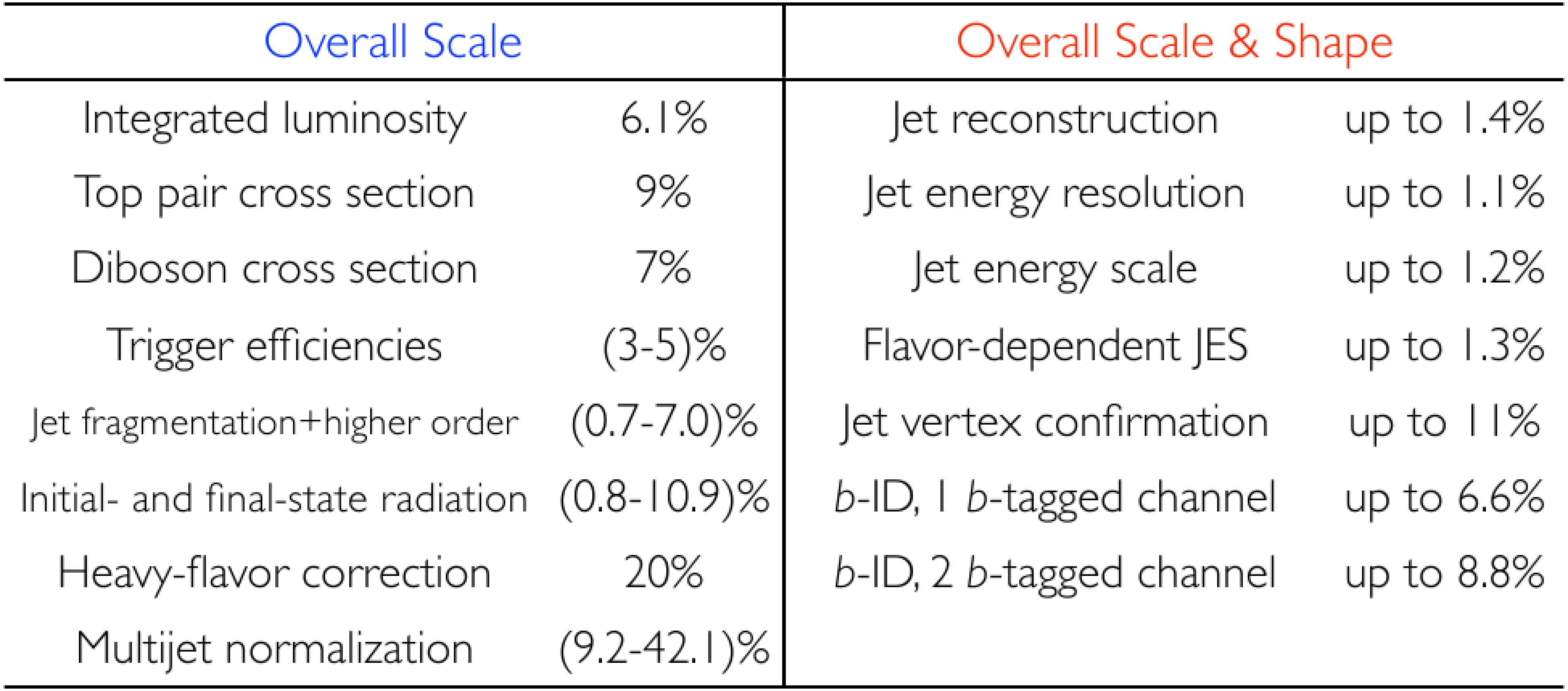
Systematic uncertainties

- Assign to each background and each analysis channel
- Some affect only the overall scale, and others affect also the discriminant outputs bin-by-bin (shape-changing)

Overall Scale	Overall Scale & Shape
	

Systematic uncertainties

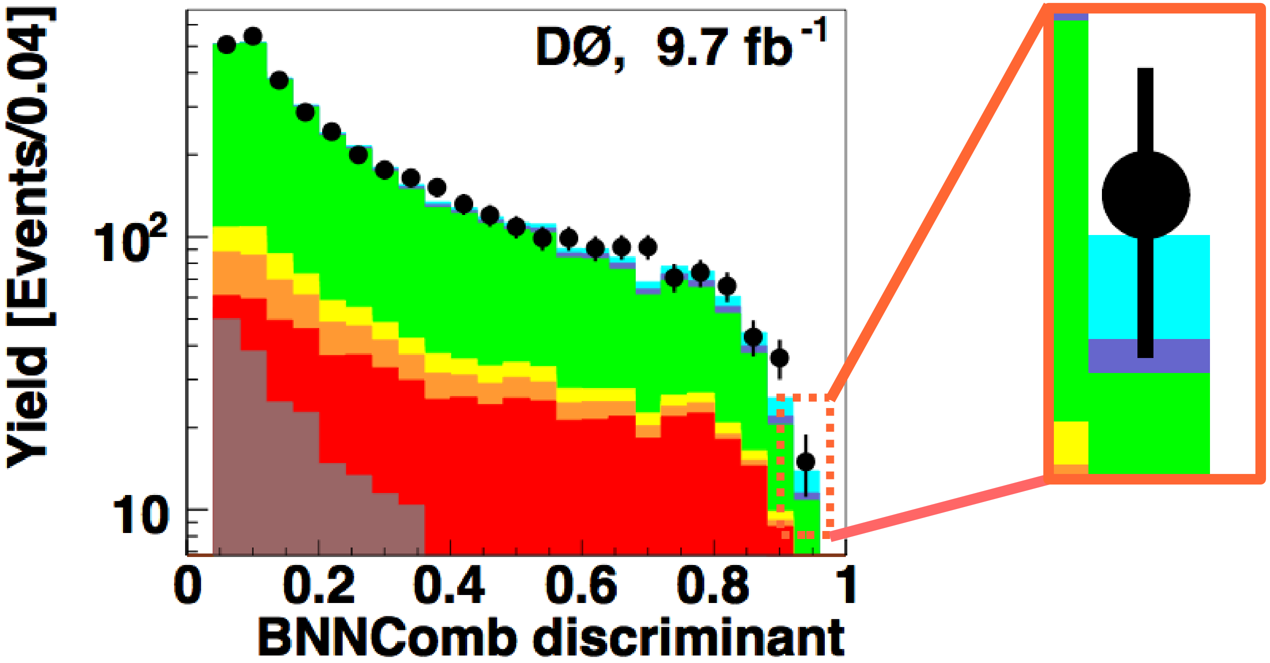
- Assign to each background and each analysis channel
- Some affect only the overall scale, and others affect also the discriminant outputs bin-by-bin (shape-changing)
- Main relative uncertainties are listed here



Overall Scale		Overall Scale & Shape	
Integrated luminosity	6.1%	Jet reconstruction	up to 1.4%
Top pair cross section	9%	Jet energy resolution	up to 1.1%
Diboson cross section	7%	Jet energy scale	up to 1.2%
Trigger efficiencies	(3-5)%	Flavor-dependent JES	up to 1.3%
Jet fragmentation+higher order	(0.7-7.0)%	Jet vertex confirmation	up to 11%
Initial- and final-state radiation	(0.8-10.9)%	*b*-ID, 1 *b*-tagged channel	up to 6.6%
Heavy-flavor correction	20%	*b*-ID, 2 *b*-tagged channel	up to 8.8%
Multijet normalization	(9.2-42.1)%		

Extracting the cross section

- Use the BNN combination discriminant in 25 bins
- Use all bins (we don’t cut on the discriminant)
- For each bin, the likelihood L to observe D data events with a known mean h is modeled by the Poisson distribution



To be measured:
signal cross section

← D

$$\leftarrow h = a_1\sigma_1+a_2\sigma_2+b$$

signal acceptance
tb: 2.6% *tqb*: 1.8%

backgrounds
(known)

Bayesian approach

- Likelihood of observing data distribution D, when h is expected:

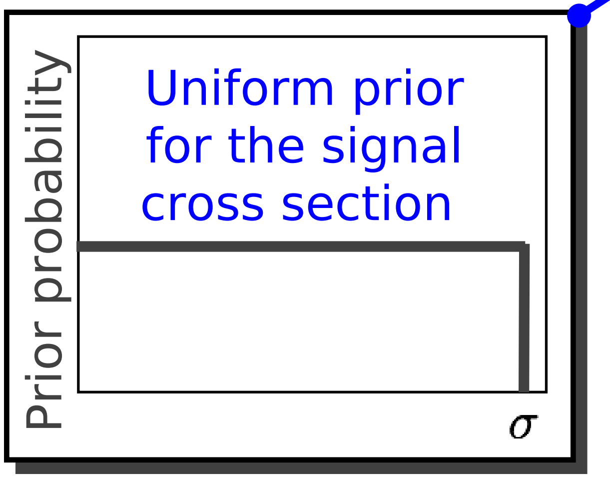
$$L(D|h) \equiv L(D|\sigma_1,\sigma_2,a_1,a_2,b) = \prod_{i=1}^{nbins} L(D_i|h_i)$$

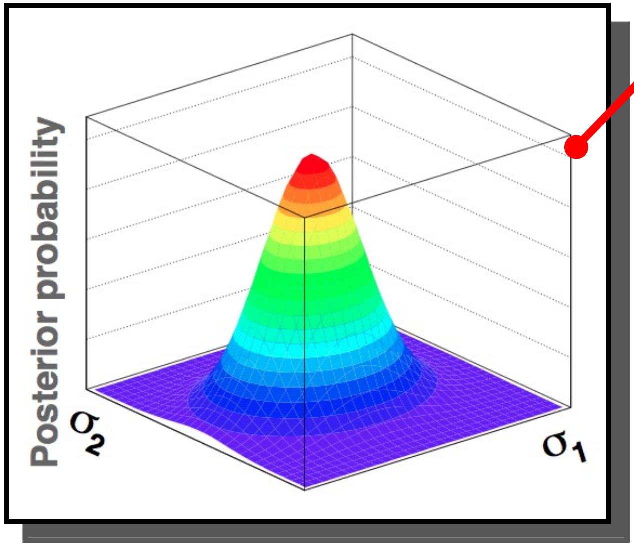
- Obtain Bayesian posterior probability as a function of σ_1, σ_2:

Poisson likelihood

Our state of knowledge, *a*, *b* with systematic uncertainties

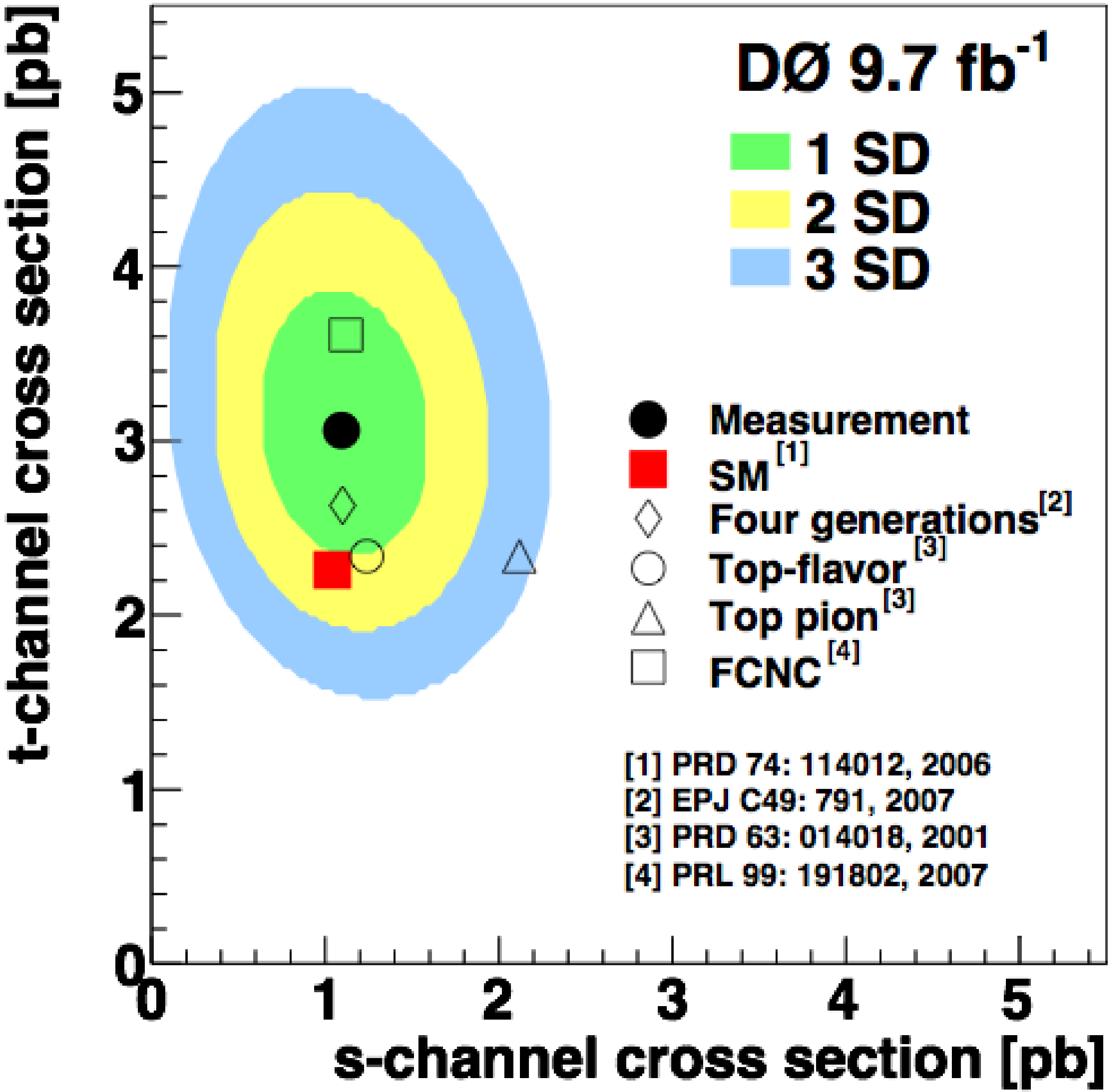
$$\int L(\mathbf{D}|\sigma_1,\sigma_2,\mathbf{a}_1,\mathbf{a}_2,\mathbf{b})\pi(\sigma_1,\sigma_2)\pi(\mathbf{a}_1,\mathbf{a}_2,b)d\mathbf{a}_1 d\mathbf{a}_2 db \propto p(\sigma_1,\sigma_2|D)$$



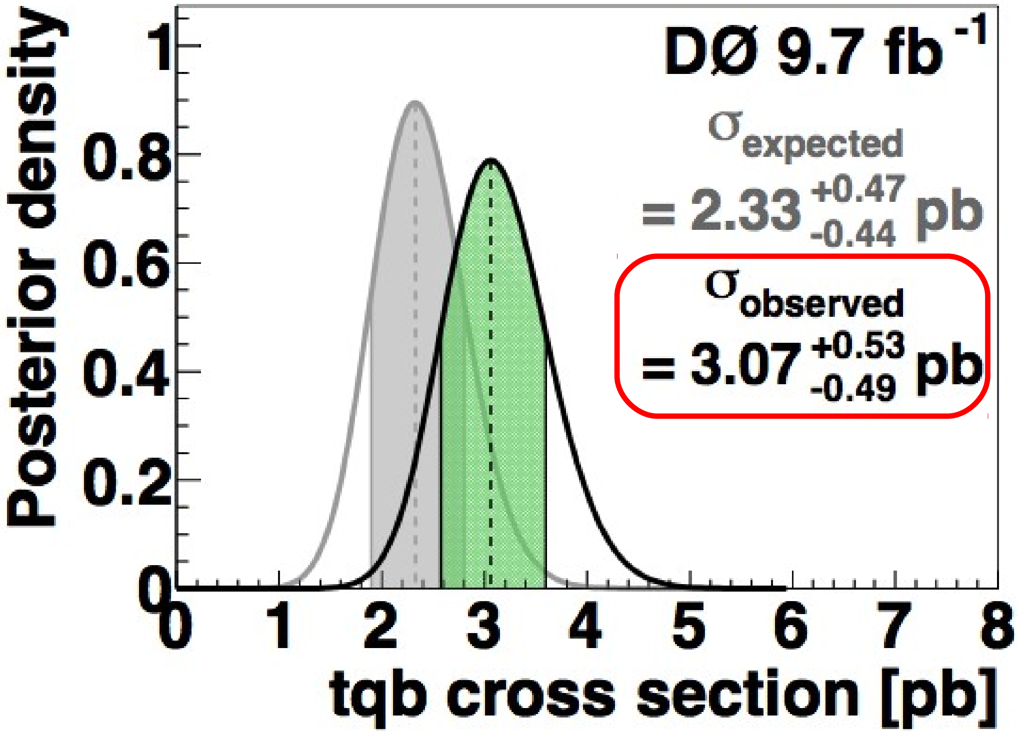


- Shape & normalization systematics treated as nuisance parameters
- Correlations between uncertainties properly accounted for

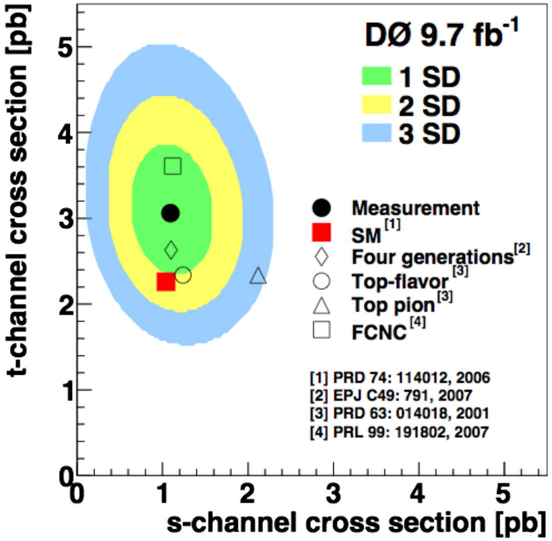
Two dimensional posterior



Measured cross section

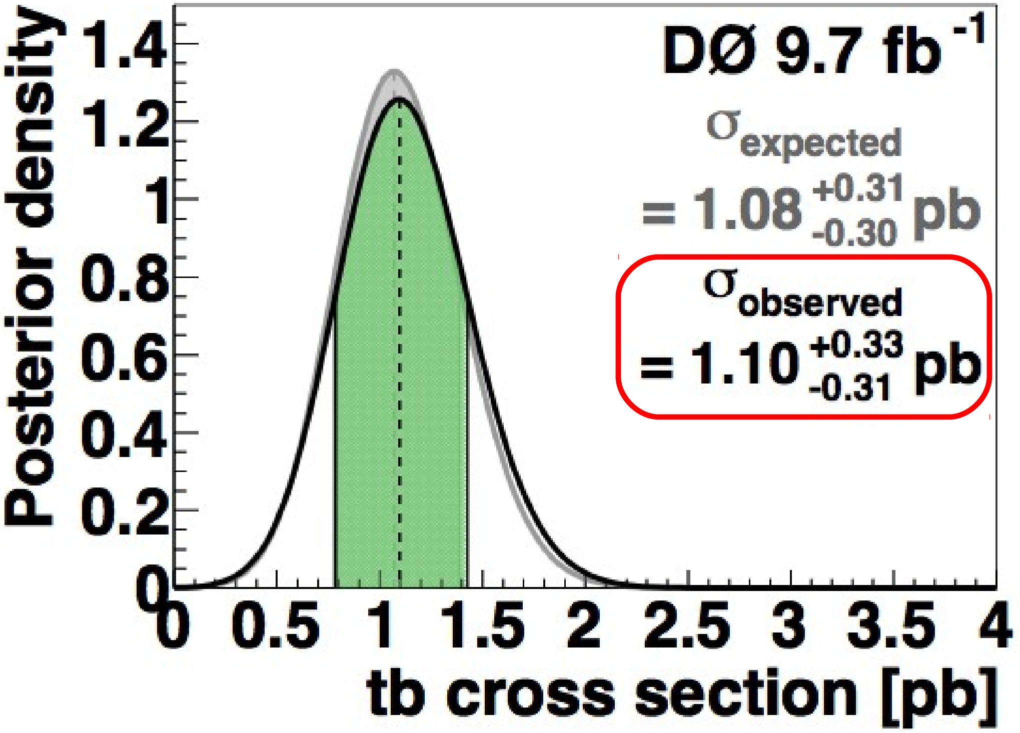


(Integrate over σ_{tb})



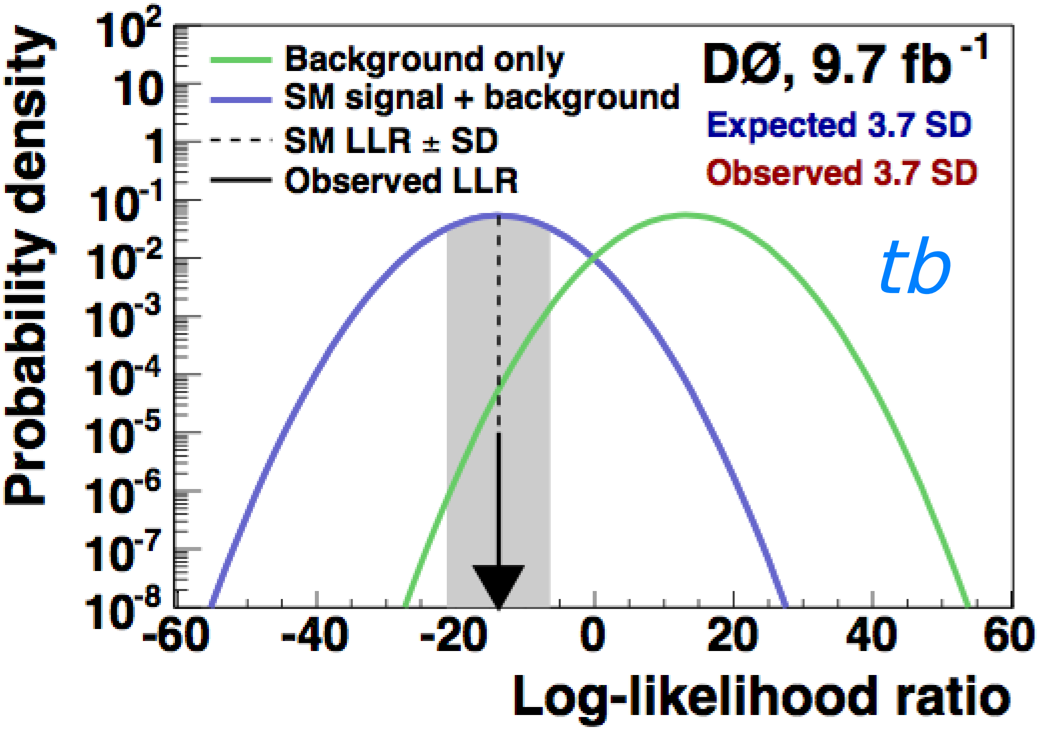
(Integrate over σ_{tqb})

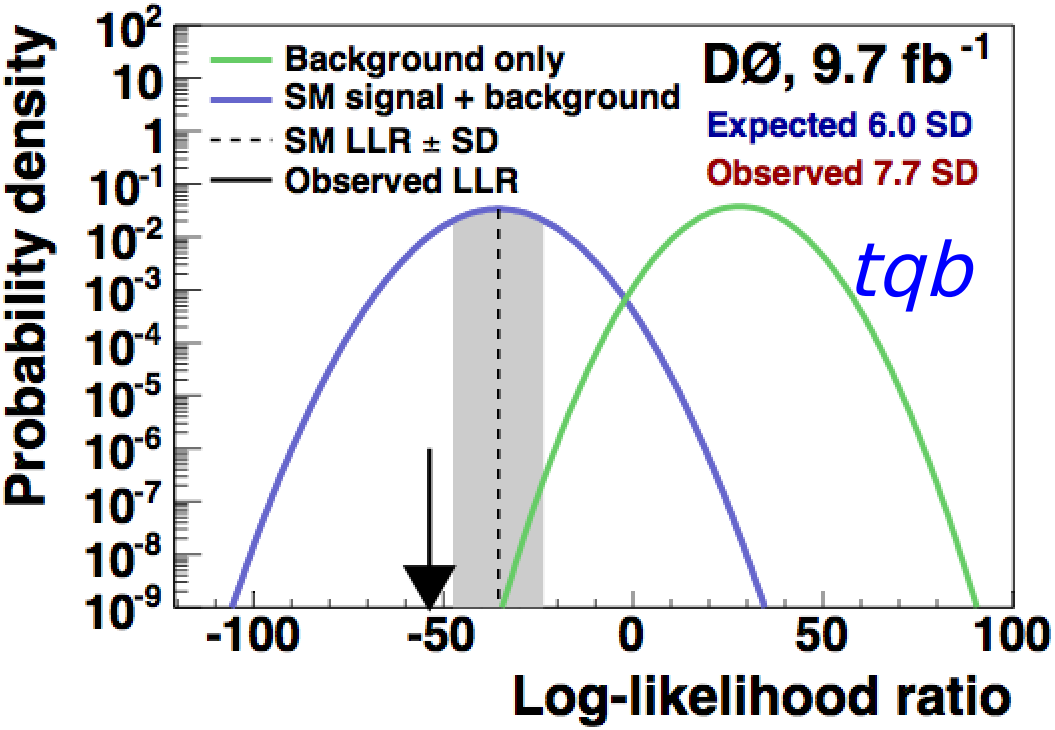
No assumption on SM σ_{tb}/σ_{tqb}



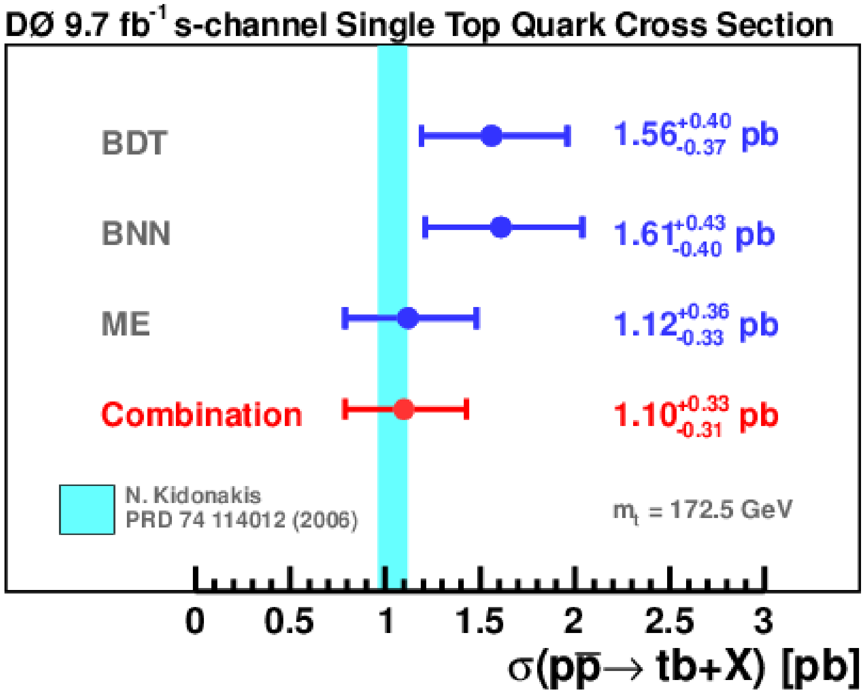
Significance

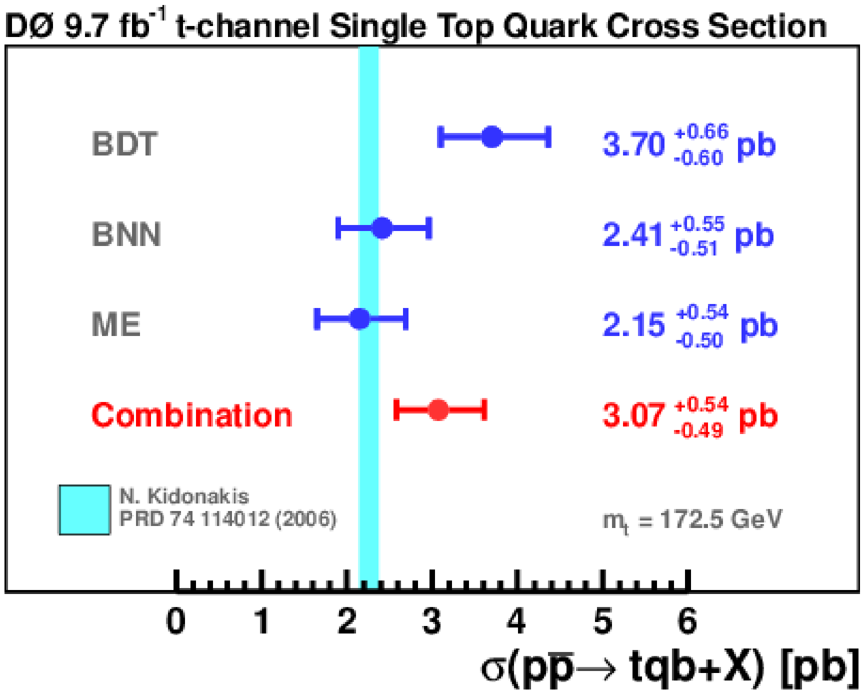
- Asymptotic approximation of the log-likelihood ratio
- Tests how likely the data is to fluctuate to the measured σ value, in the absence of the signals
- Expected p-values:
  - tb: 1.0×10^{-4} (3.7 SD)
  - tqb: 9.9×10^{-10} (6.0 SD)
- Observed p-values:
  - tb: 1.0×10^{-4} (3.7 SD)
  - tqb: 6.1×10^{-15} (7.7 SD)
- All BDT, BNN and ME methods have more than 3 SD significance alone





Individual results

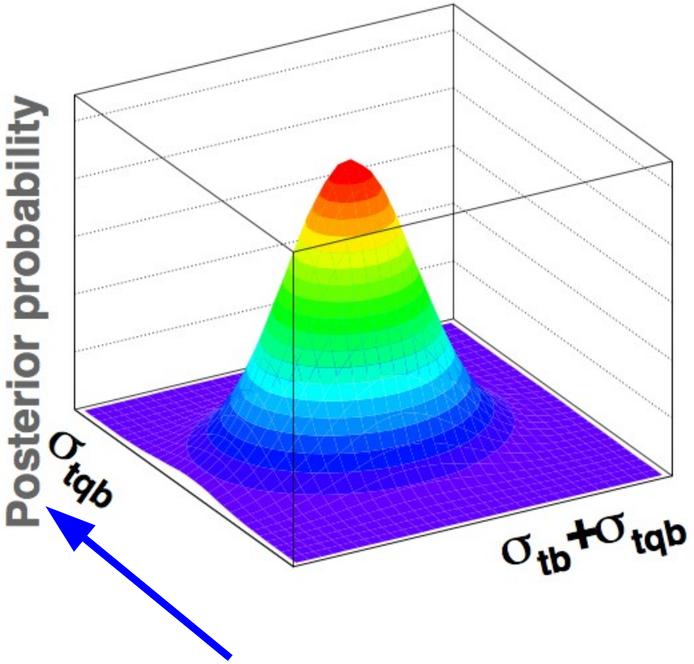

DØ 9.7 fb^{-1} s-channel Single Top Quark Cross Section

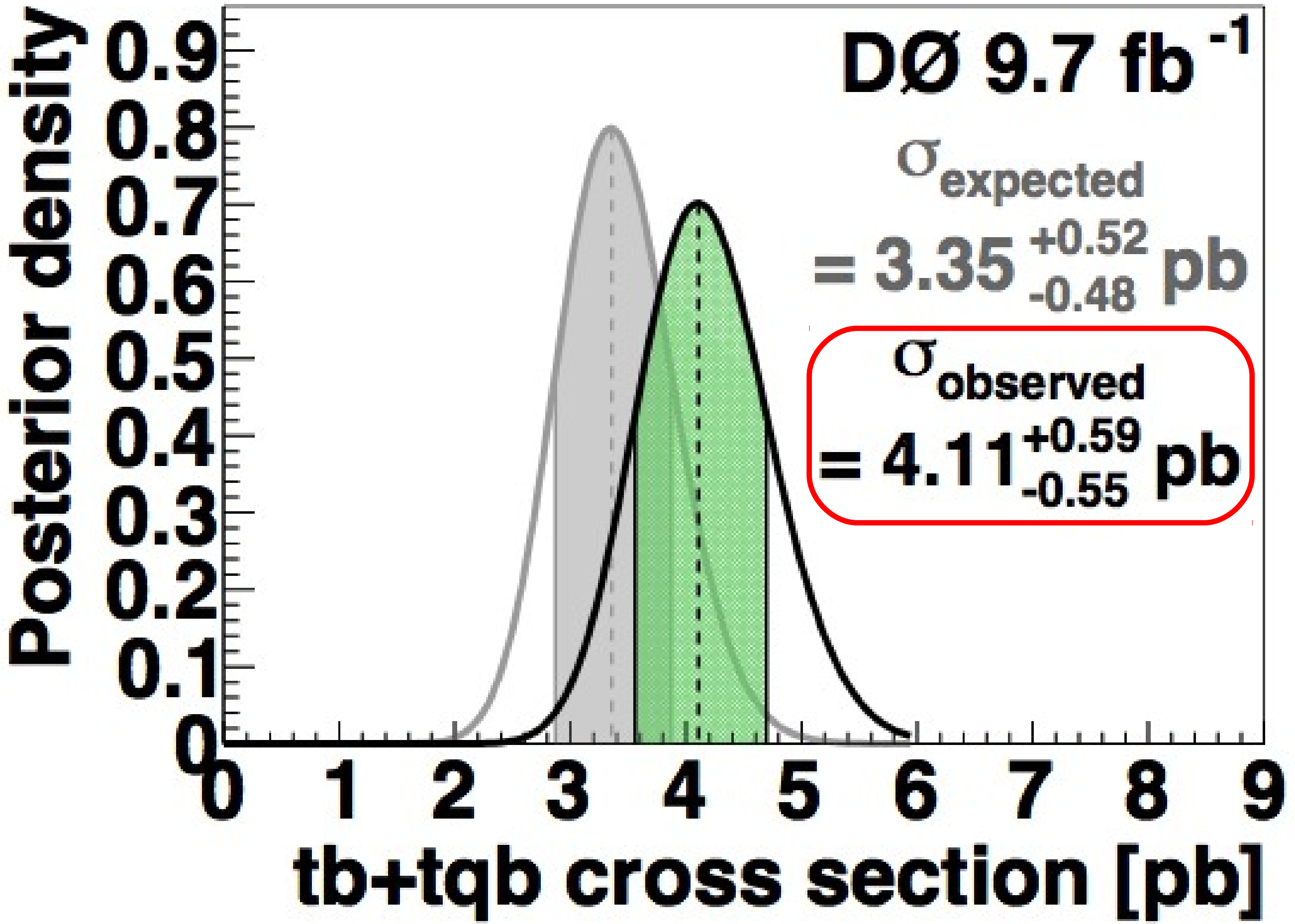

DØ 9.7 fb^{-1} t-channel Single Top Quark Cross Section

Channel	Expected σ (pb)	Observed σ (pb)	Expected p value	Observed p value	Expected Z	Observed Z
ME_s	$1.05^{+0.36}_{-0.34}$	$1.12^{+0.36}_{-0.33}$	8.1×10^{-4}	3.7×10^{-4}	3.2	3.4
BNN_s	$1.06^{+0.41}_{-0.39}$	$1.61^{+0.43}_{-0.40}$	3.3×10^{-3}	1.5×10^{-5}	2.7	4.2
BDT_s	$1.06^{+0.35}_{-0.33}$	$1.56^{+0.40}_{-0.37}$	5.4×10^{-4}	2.3×10^{-6}	3.3	4.6
D_s^{comb}	$1.07^{+0.32}_{-0.30}$	$1.10^{+0.33}_{-0.31}$	1.0×10^{-4}	1.0×10^{-4}	3.7	3.7
ME_t	$2.27^{+0.55}_{-0.51}$	$2.15^{+0.54}_{-0.50}$	6.6×10^{-7}	2.8×10^{-6}	4.8	4.5
BNN_t	$2.31^{+0.54}_{-0.50}$	$2.41^{+0.55}_{-0.51}$	2.4×10^{-7}	1.4×10^{-7}	5.0	5.1
BDT_t	$2.36^{+0.53}_{-0.50}$	$3.70^{+0.66}_{-0.60}$	5.4×10^{-8}	3.4×10^{-15}	5.3	7.8
D_t^{comb}	$2.33^{+0.47}_{-0.44}$	$3.07^{+0.54}_{-0.49}$	1.0×10^{-9}	7.1×10^{-15}	6.0	7.7
D_{s+t}^{comb}	$3.34^{+0.53}_{-0.49}$	$4.11^{+0.60}_{-0.55}$				

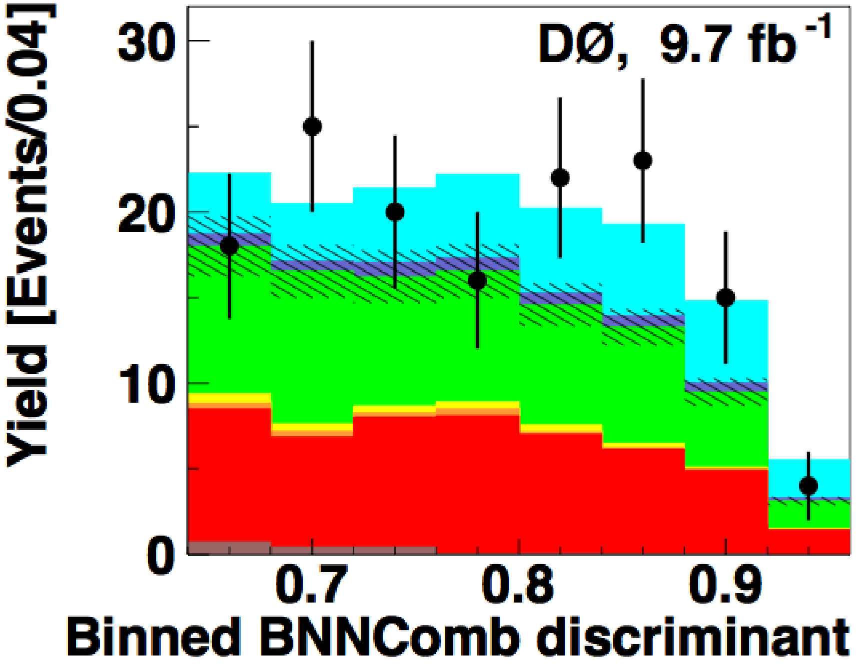
Measure combined σ_{tb+tqb}

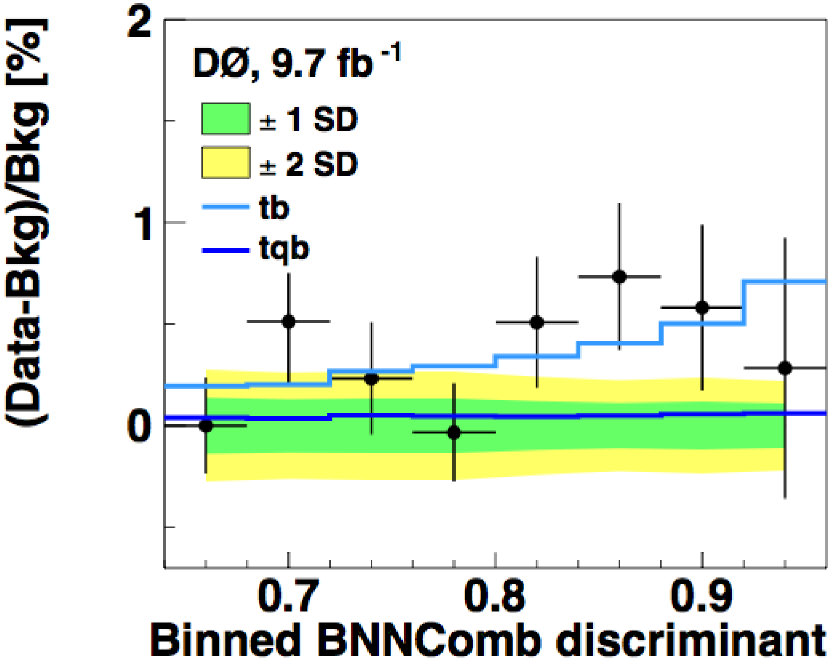
- Measure σ_{tb+tqb} without assuming the SM σ_{tb}/σ_{tqb} NEW
- Use 2D posterior p.d.f.
- Integrate over σ_{tqb} and obtain 1D posterior p.d.f of σ_{tb+tab}



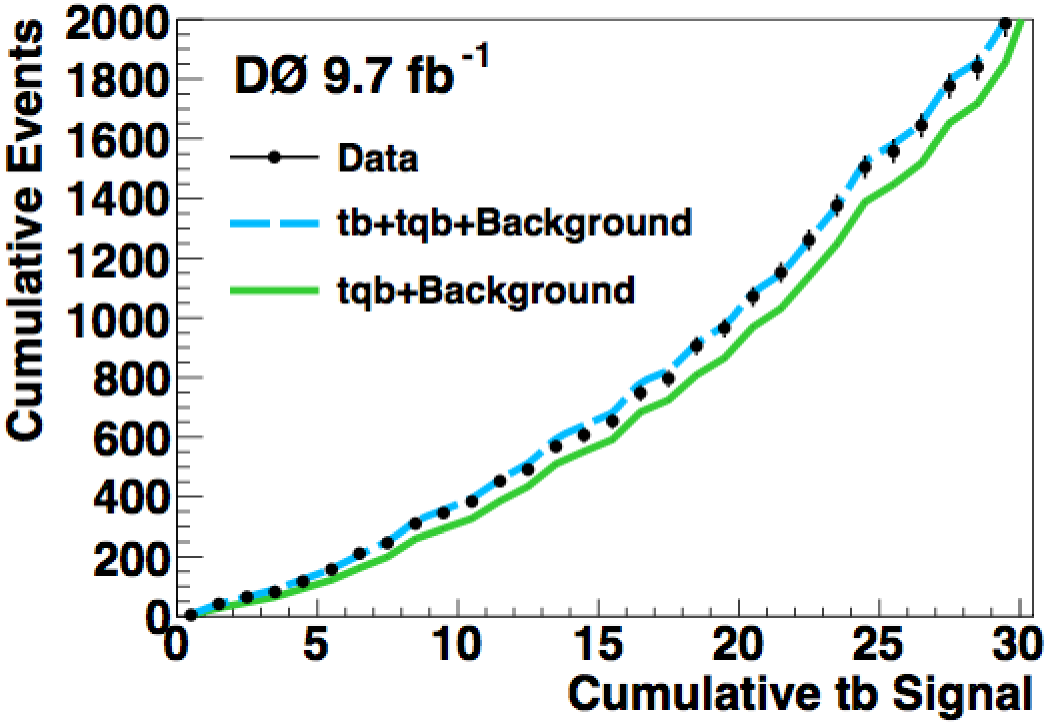


tb or not tb



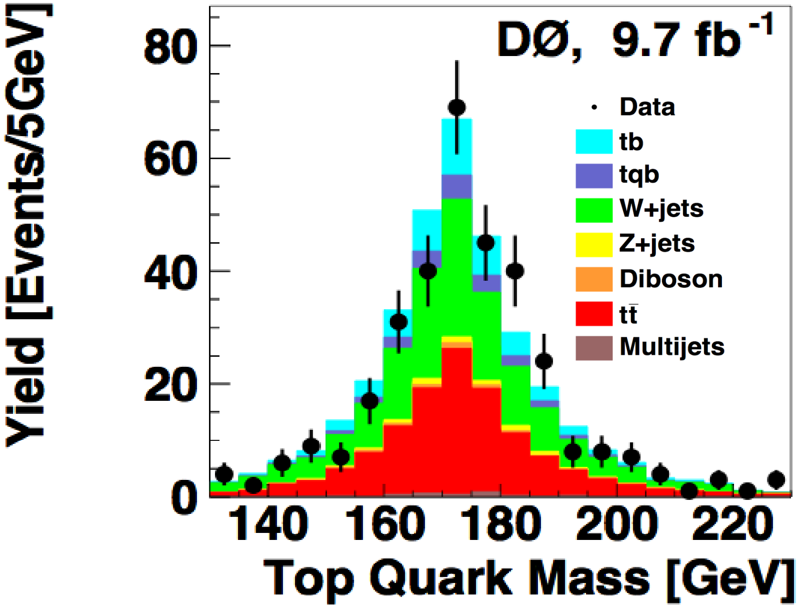


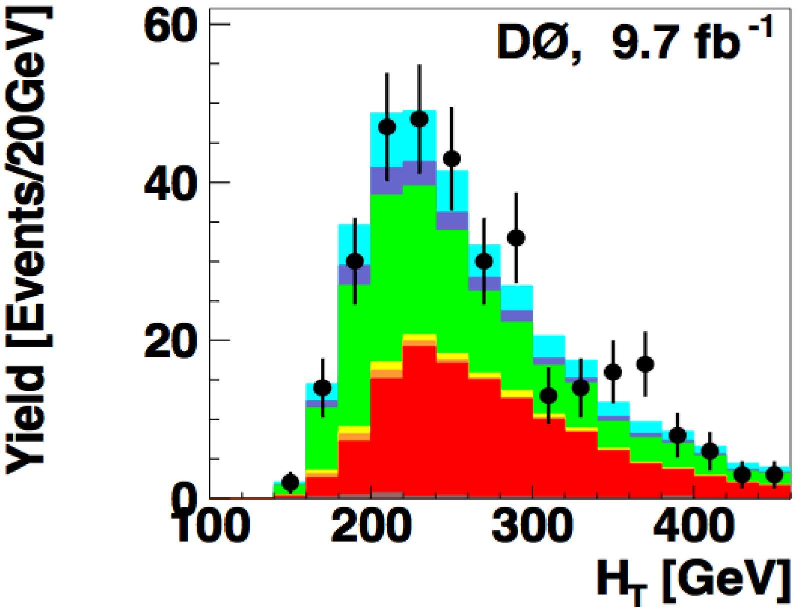
- Examine highest BNNcomb bins, with post-fit uncertainty and measured σ_{tb}
- Integrate from right to left, and plot cumulative signal vs all events



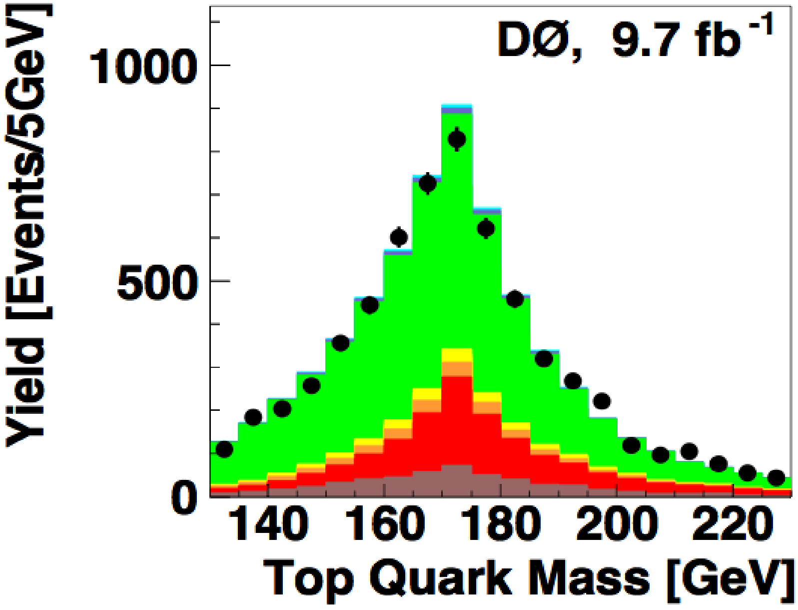
Event characteristics

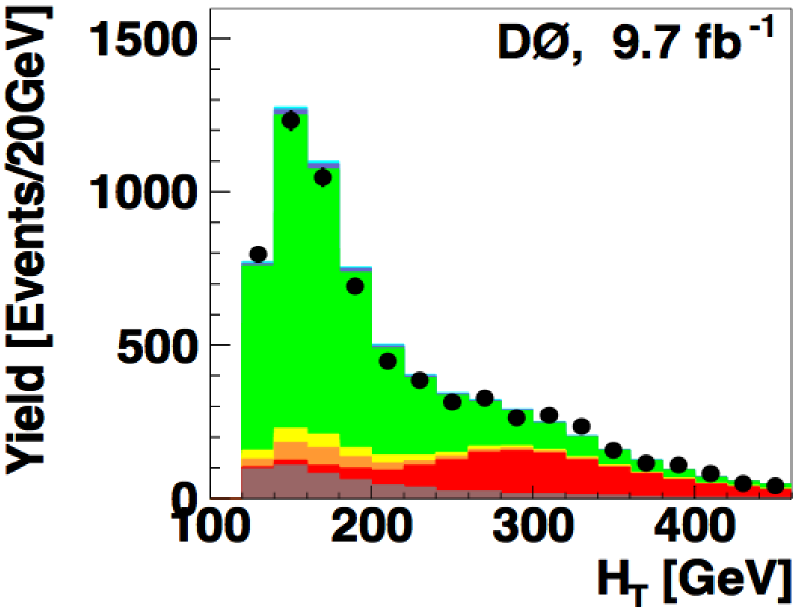
tb Category: $D_{tb} > 0.8$





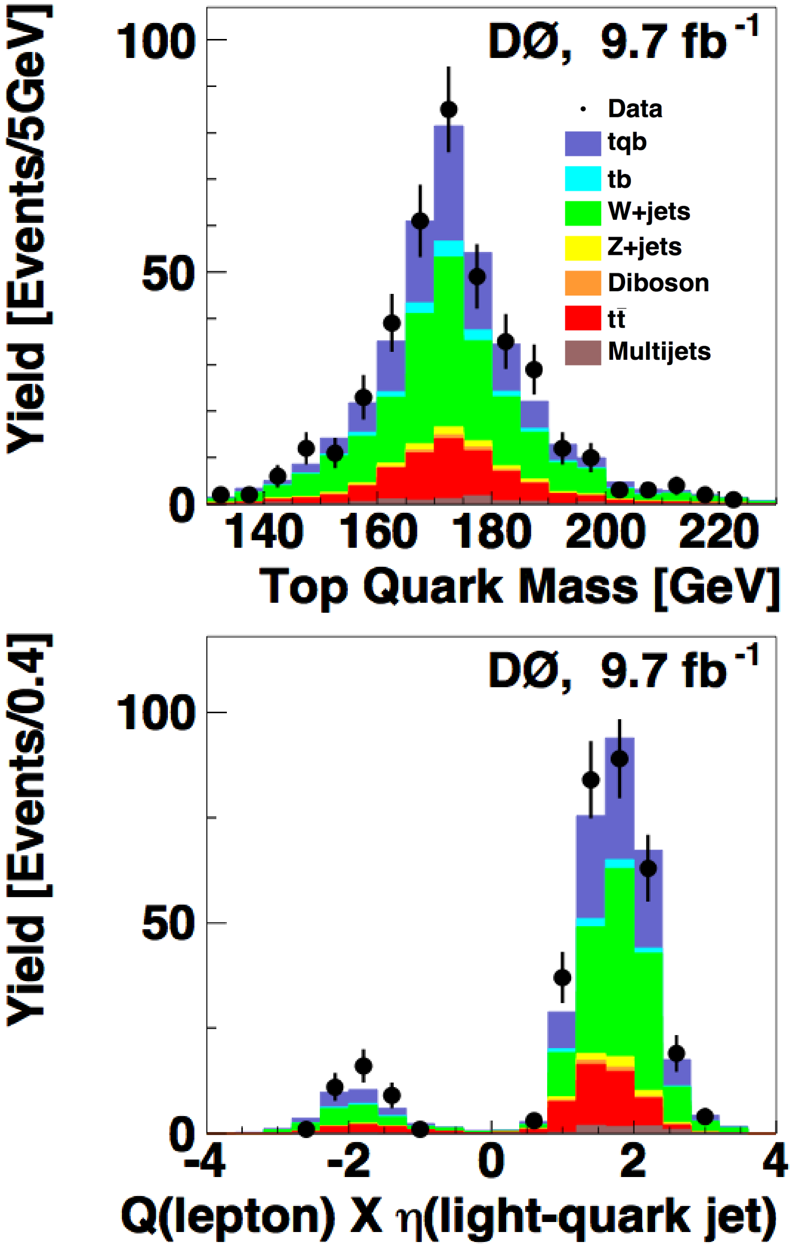
tb & *tqb* Depleted Region



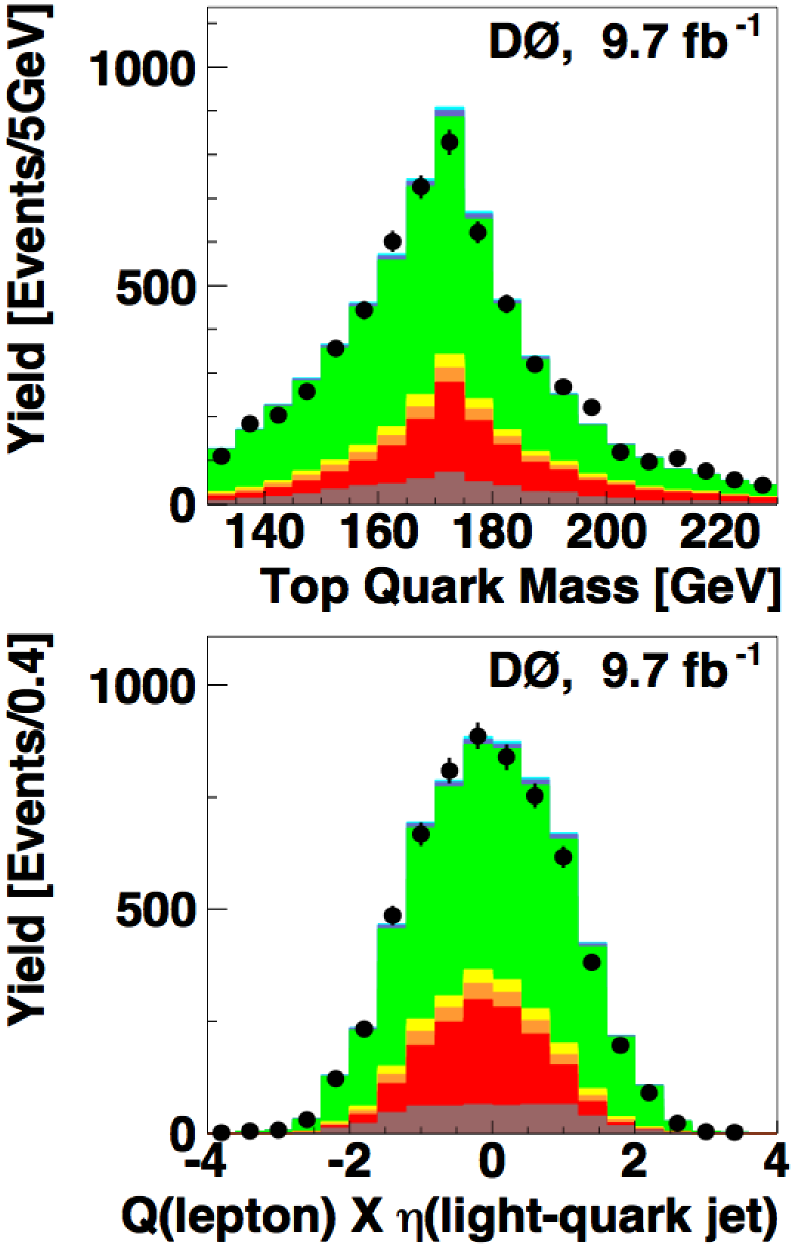


Event characteristics

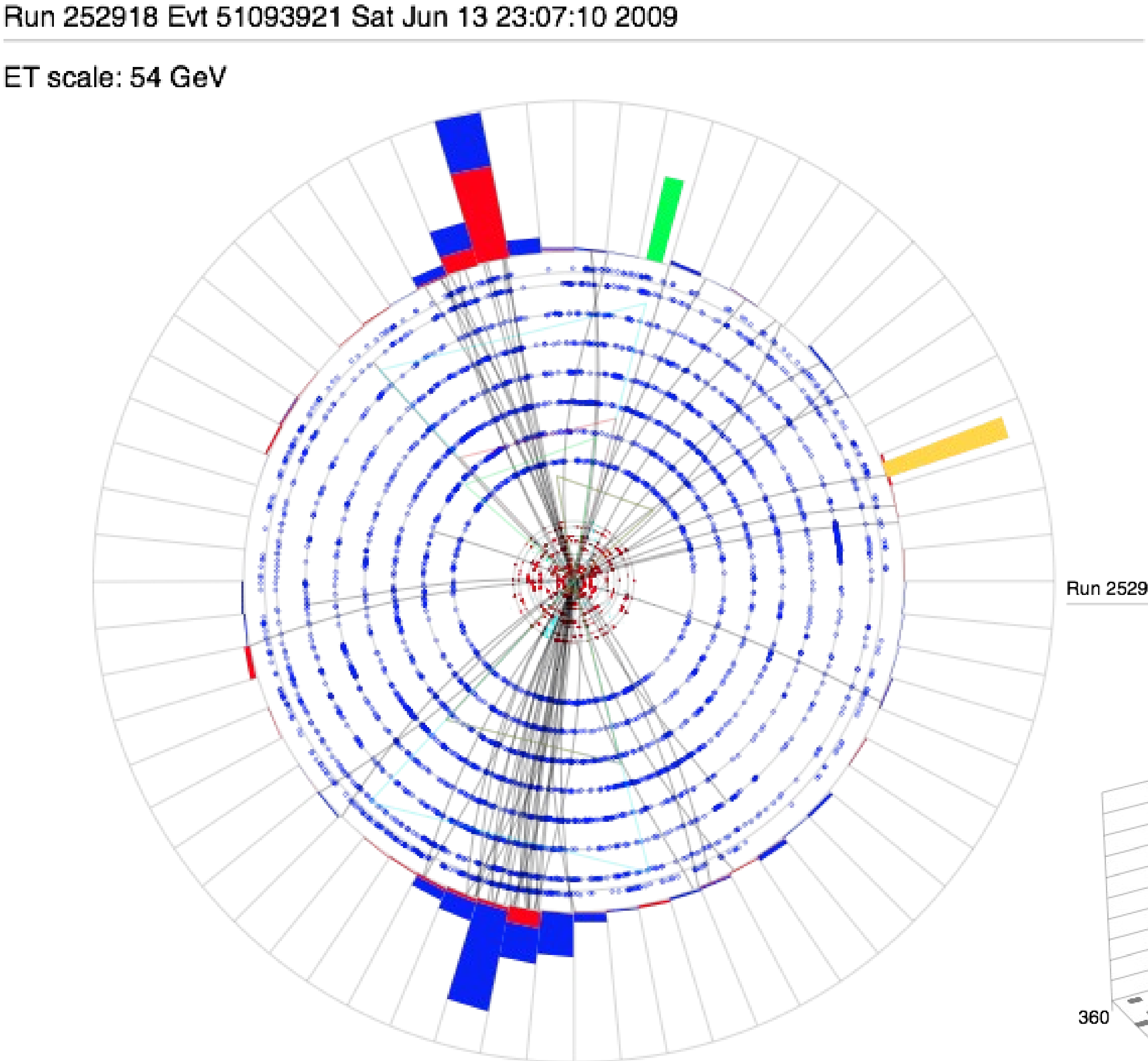
tqb Category: $D_{tqb} > 0.8$



tb & *tqb* Depleted Region



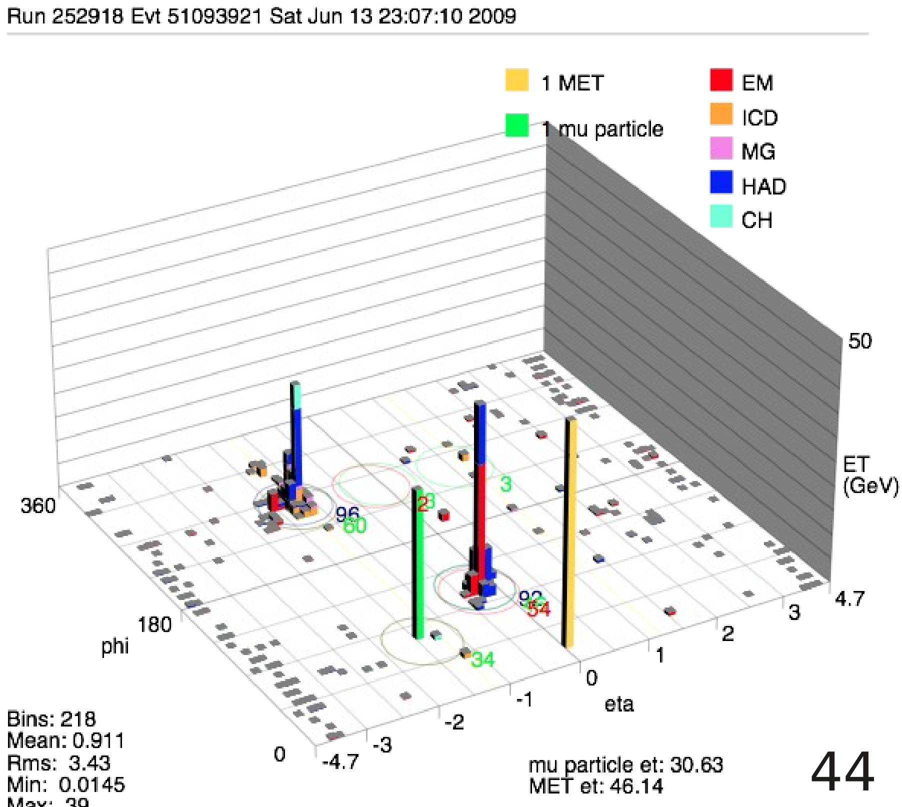
A tb event candidate



$m_t = 171$ GeV

Jet1 b-tag: 0.95

Jet2 b-tag: 0.84



Run 252918

Event 51093921

Sat. June 13 23:07:10 2009

Experimental status

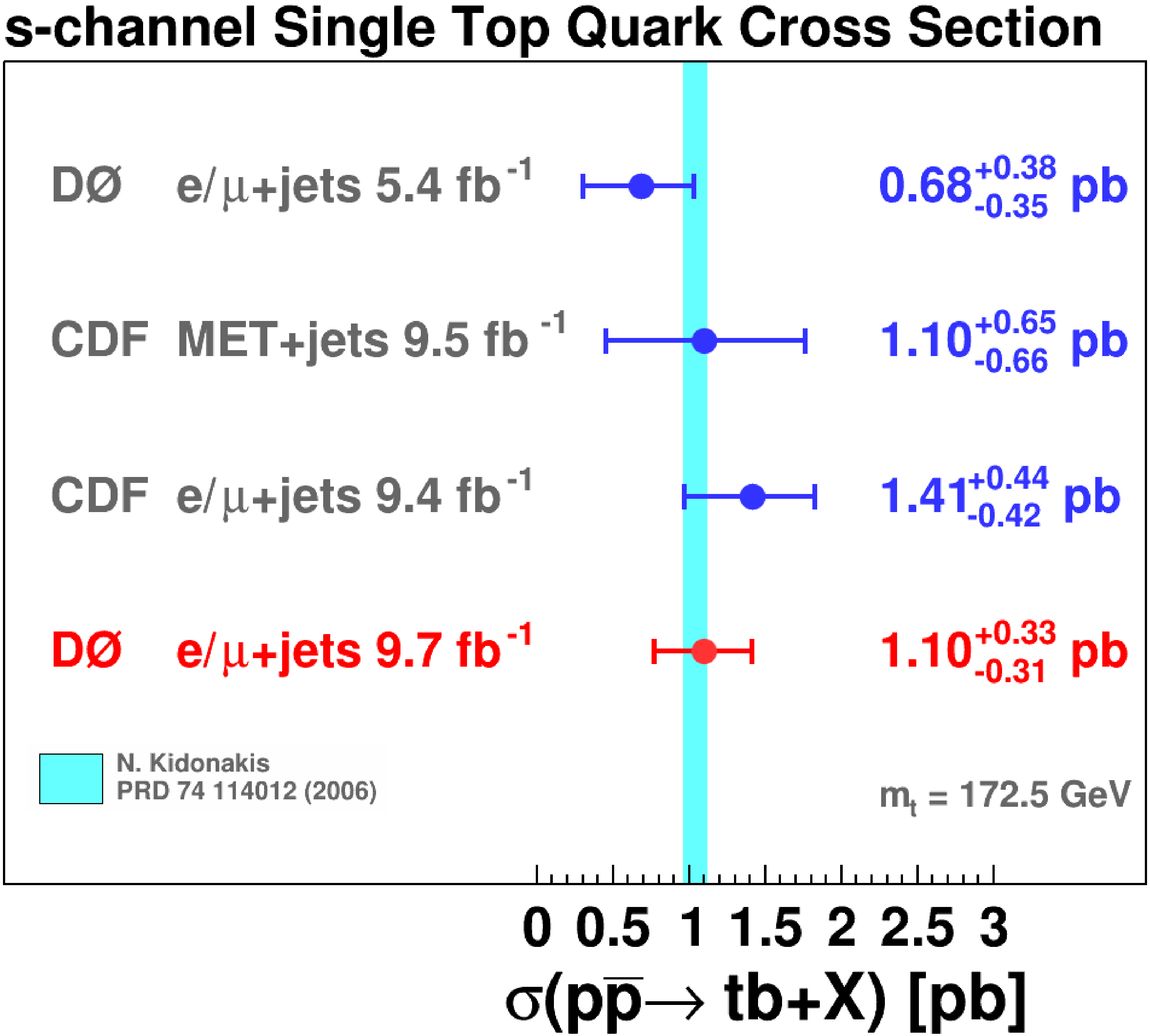
(after this analysis)

σ (NNLO) [pb]	**tb**	**tqb**	**tW**
TeV prediction	1.04	2.26	0.28
CDF (9.4-7.5 fb^{-1})	1.41 ± 0.44	1.49 ± 0.45	-
DØ (9.7 fb^{-1})	1.10 ± 0.33	**3.07 ± 0.53**	-

σ (NNLO) [pb]	**tb**	**tqb**	**tW**
LHC prediction (7 TeV)	4.6	64.6	15.7
ATLAS (0.7-20 fb^{-1})	<20.5 (95%CL)	**83 ± 20**	27 ± 6
CMS (1.2-12.2 fb^{-1})	-	**67 ± 6**	**23 ± 5**

Discovery (> 5 SD) Evidence (>3 SD)

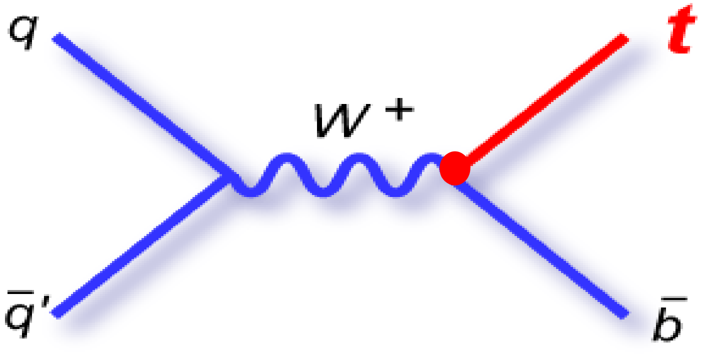
Tevatron latest measurements



Measuring $|V_{tb}|$

- Once we have a cross section measurement, we can make a direct measurement of $|V_{tb}|$, since $\sigma_{tb+tqb} \propto |V_{tb}|^2$

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



- Most general Wtb vertex [PLB 713, 165 (2012)]:

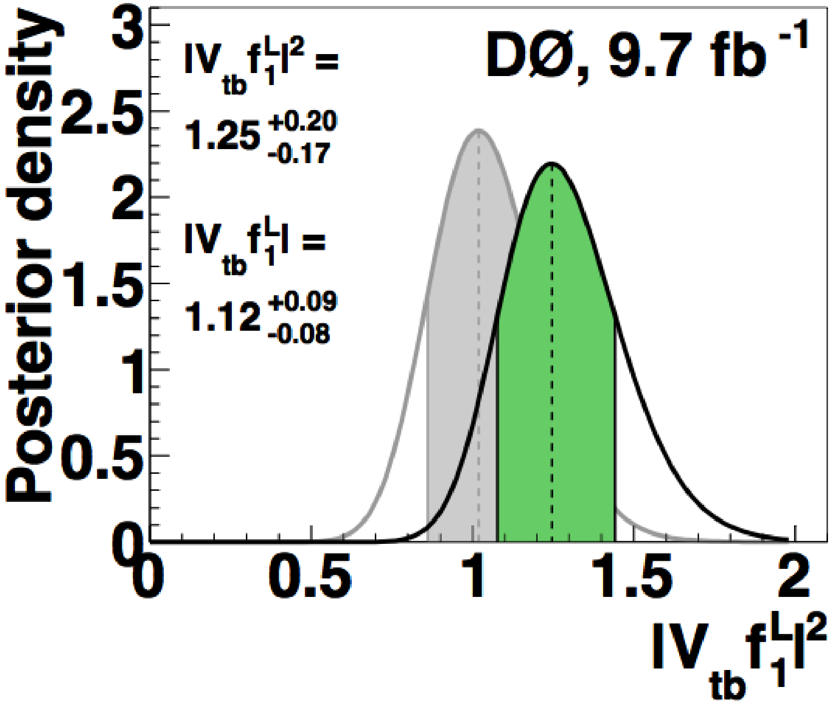
$$L = \frac{g}{\sqrt{2}} \bar{b} \gamma^{\mu} V_{tb} (f_1^L P_L + f_1^R P_R) t W_{\mu}^{-} - \frac{g}{\sqrt{2}} \bar{b} \frac{i \sigma^{\mu\nu} q_{\nu} V_{tb}}{M_W} (f_2^L P_L + f_2^R P_R) t W_{\mu}^{-}$$

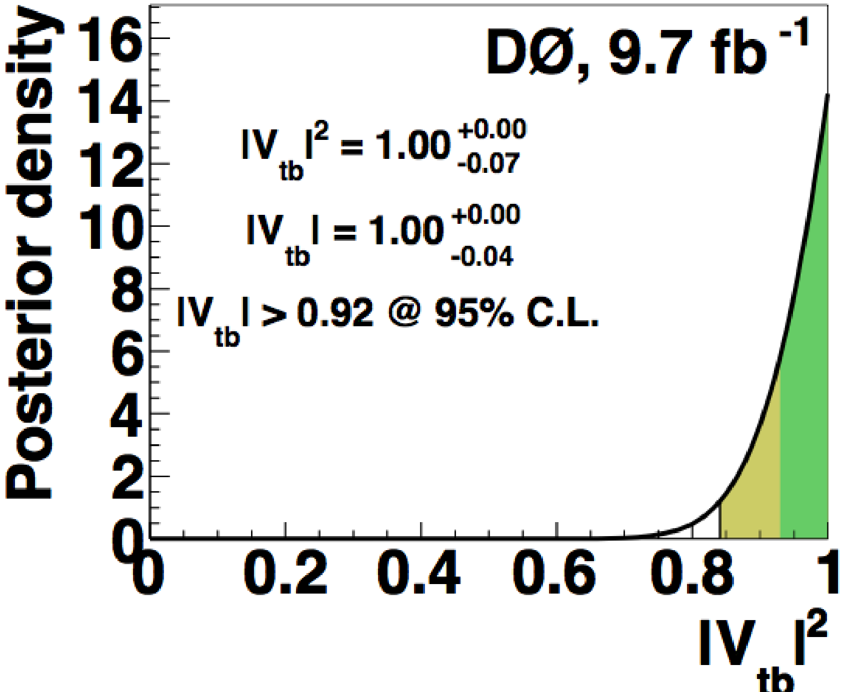
- Assume:
  - **SM top decay**: $V_{td}^2 + V_{ts}^2 \ll V_{tb}^2$
  - Pure V-A interaction: $\mathbf{f_1^R = 0}$
  - CP conservation: $\mathbf{f_2^L = f_2^R = 0}$
- Do not assume:
  - 3 generations
  - Unitarity of CKM
  - New: σ_{tb}/σ_{tqb} NEW

We are effectively measuring the **strength of the V-A coupling:** $\mathbf{|V_{tb} f_1^L|}$, which can be >1

CKM matrix element $|V_{tb}|$

- Allow $|V_{tb}f_1^L|^2 > 1$
  - $|V_{tb}f_1^L| = 1.12^{+0.09}_{-0.08}$
- Assume $0 \leq |V_{tb}|^2 \leq 1$
  - $|V_{tb}| > 0.92$ @ 95% C.L.
- Additional systematic uncertainties
  - Theoretical uncertainty on single top cross sections
- Complementary to $R_{Wb/Wq}$ measurement in top decays [PRL 107, 121802 (2011)]
- Current limits @ 95% C.L.:
  - CDF (7.5 fb^{-1}): $0.78<|V_{tb}|\leq1$
  - ATLAS (6 fb^{-1} 8TeV): $0.80<|V_{tb}|\leq1$
  - CMS (5 fb^{-1} 8TeV): $0.81<|V_{tb}|\leq1$





Conclusions

- First evidence of s-channel single top quark production

$$\boldsymbol{\sigma_{tb} = 1.10 \pm 0.33\ \mathrm{pb}}$$

Accepted by Phys. Lett. B, arXiv: 1307.0731

- Simultaneously measure σ_{tb} and σ_{tqb}, without assuming the SM prediction for either
- Also measure σ_{tb+tqb} and $|V_{tb}|$ without assuming the SM ratio of σ_{tb}/σ_{tqb}

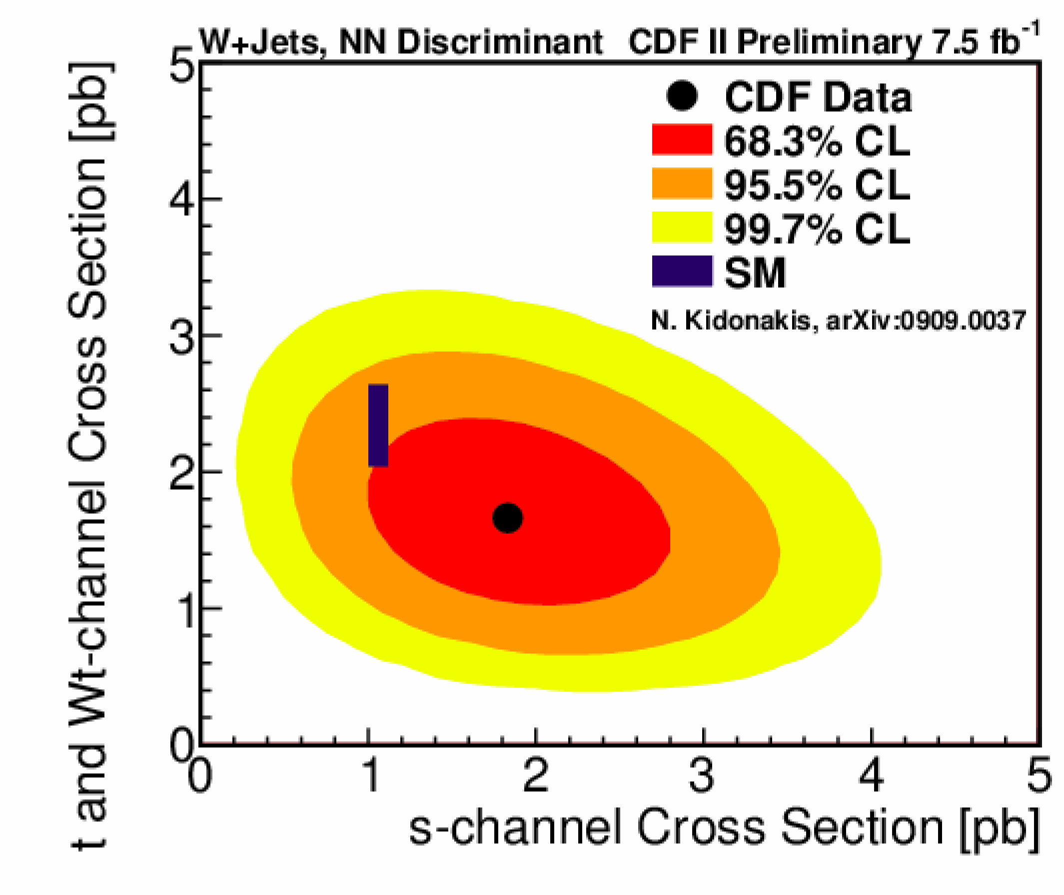
$$\boldsymbol{|V_{tb}| > 0.92\ @\ 95\%\ \mathrm{C.L.}}$$

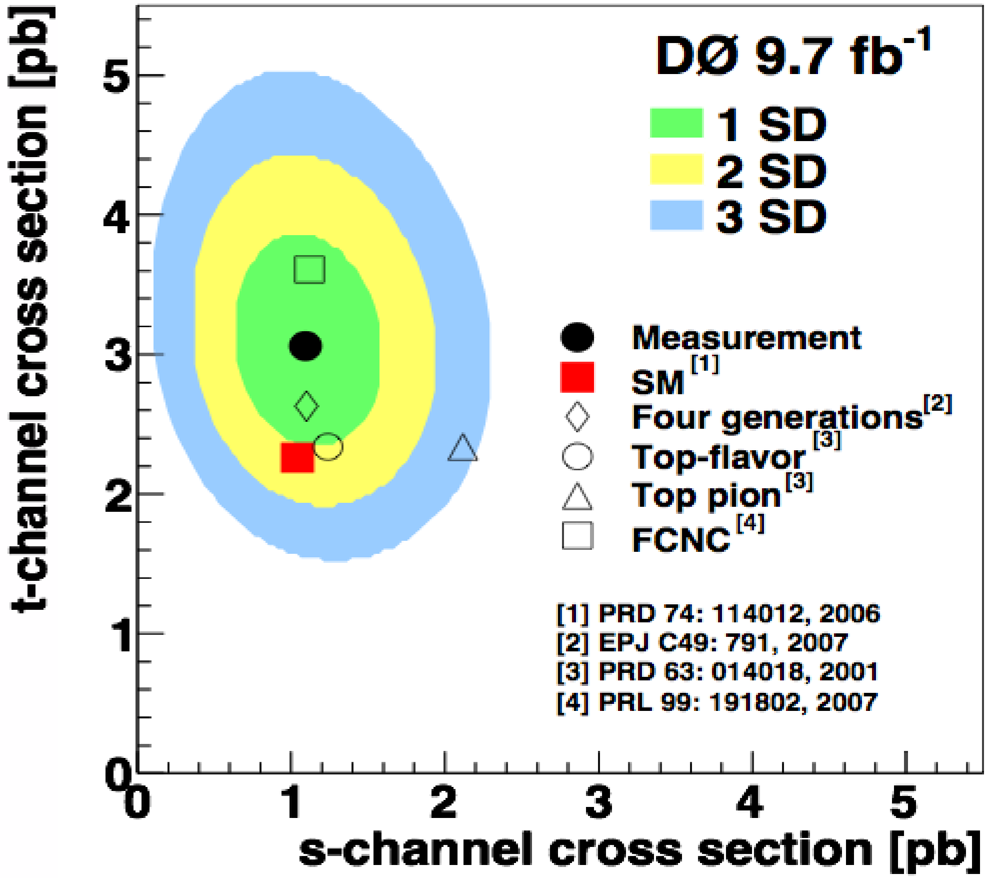
- Results are consistent with the SM predictions
- A legacy measurement at the Tevatron
- Looking forward to combination with CDF

Extra slides

For more information:
http://www-d0.fnal.gov/Run2Physics/top/top_public_web_pages/top_public.htm

- CDF Conference note 10793





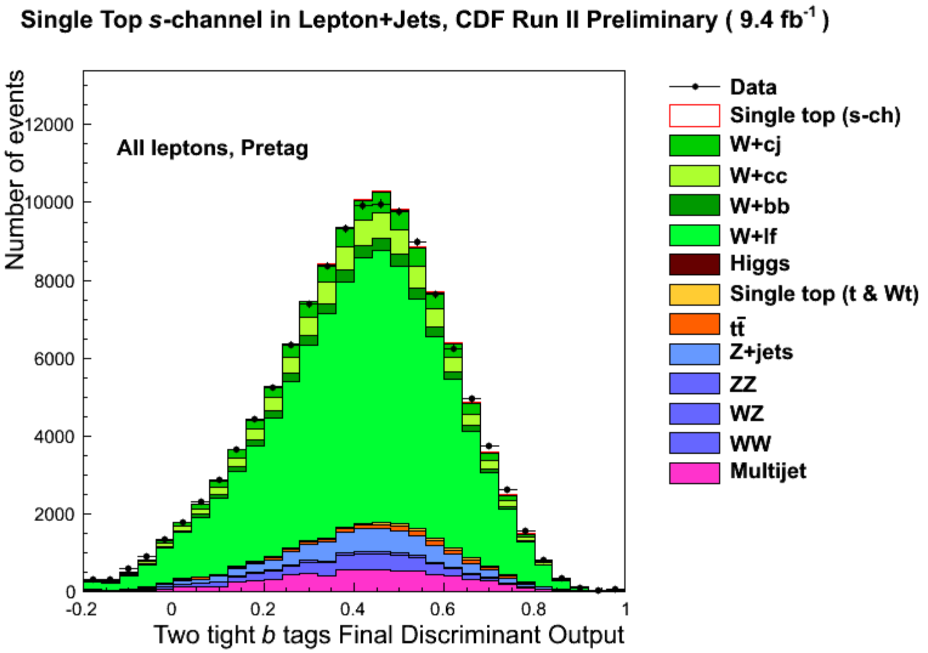
CDF ℓ+jets result 9.4 fb^{-1}

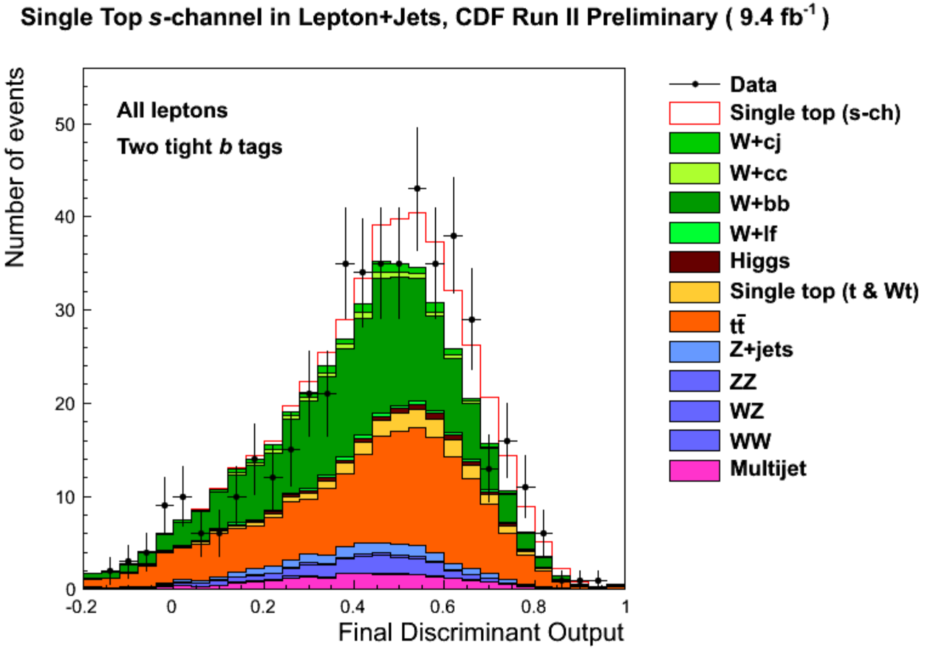
- Isolated e/μ p_T>20 GeV
- MET>10 (20) GeV muon (e)
- Two jets: E_T>20 GeV, |η|<2.0, leading jet: E_T>30 GeV
- H_T>125 GeV, M_{jj}>30 GeV
- W+jets normalization: fit to MET in pretag sample
- Train NN with 8 variables for each lepton and tag category

Category	TT	TL	T	LL
WW	1.7±0.4	13.2±2.7	184±23	24.8±3.9
WZ	17.8±2.2	21.2±2.0	52.7±5.4	9.9±0.9
ZZ	2.4±0.3	2.4±0.2	7.1±0.7	0.96±0.08
Z + jets	10.9±1.2	20.7±2.3	163±18	27.1±3.1
$t\bar{t}$	163±21	194±19	502±50	58.1±6.6
Higgs	6.1±0.6	6.4±0.4	10.3±0.7	1.7±0.2
Wbb	246±99	327±130	1166±468	109±44
Wcc	19.0±7.8	120±49	1158±467	164±67
W + Mistag	4.3±1.3	62±13	978±141	242±34
Multijet	29±12	47±19	281±112	45±18
t and Wt-channel	18.1±2.5	35.3±4.2	251±28	13.6±1.5
s-channel	54.5±6.7	61.2±5.6	109±10	17.8±2.1
Total Prediction	573±155	911±248	4860±1320	714±181
Observed	466	765	4620	718

$\sigma_{tb} = 1.41^{+0.44}_{-0.42}$ pb

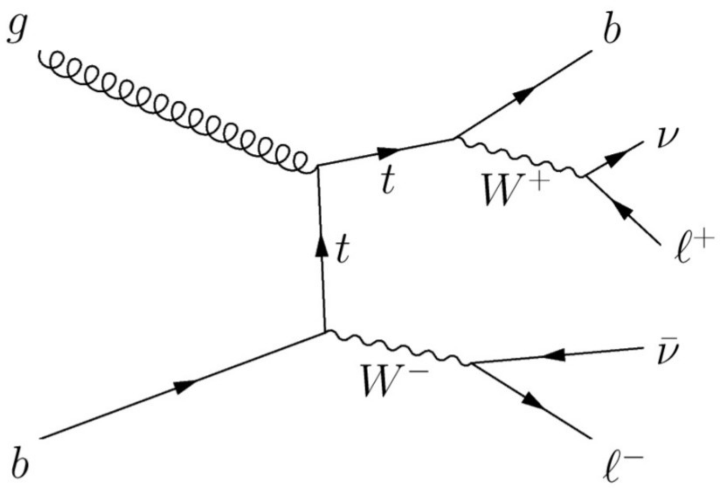
Significance: 3.8σ (2.9σ expected)

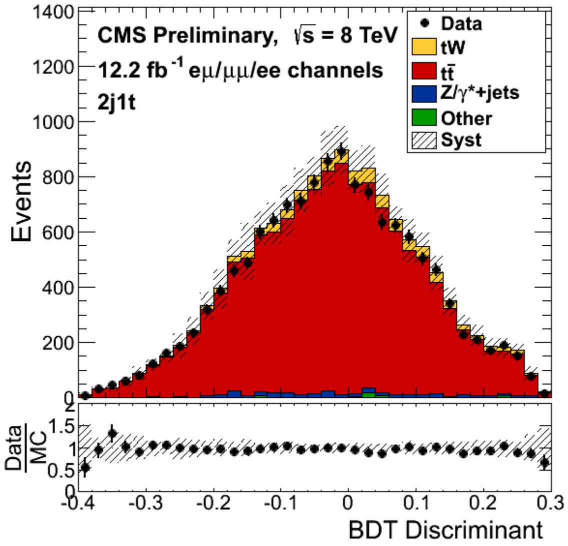


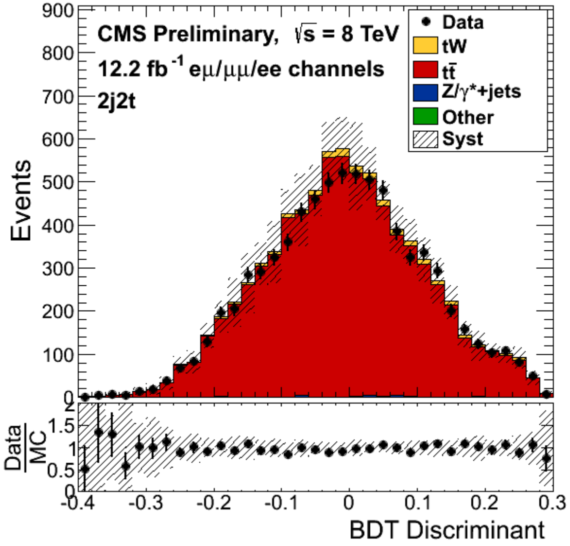


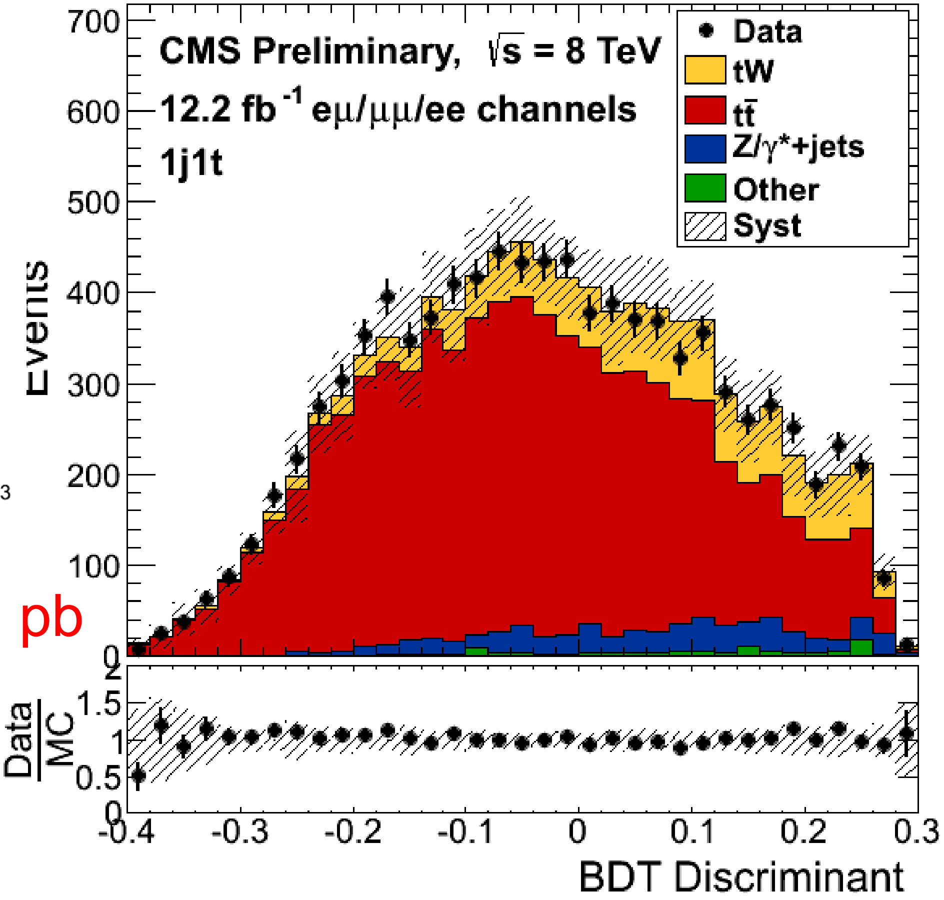
CMS tW observation 12.2 fb^{-1} 8 TeV

- Signal region: 1 tight jet, 1 b-tag
- Control regions, dominated by $t\bar{t}$:
  - 2 tight jets, 1, 2 b-tags
  - Train BDT against $t\bar{t}$ with 13 variables
- Fit done for all channels (ee,eμ,μμ) and regions (1j1t,2j1t,2j2t) simultaneously









σ_{tW} = 23.4+5.5-5.4 pb ; σ_{tW}^{SM} = 22.2±0.6±1.4 pb

Significance: 6.0σ (5.4σ expected)

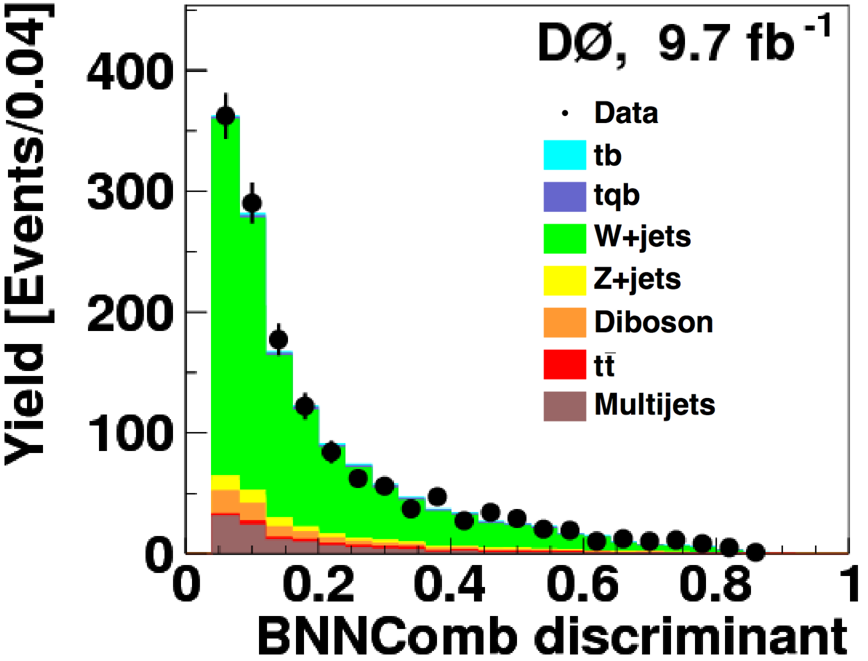
$|V_{tb}|$ = 1.03 ± 0.12(exp) ± 0.04(th)

$|V_{tb}|$>0.78 at 95% C.L.

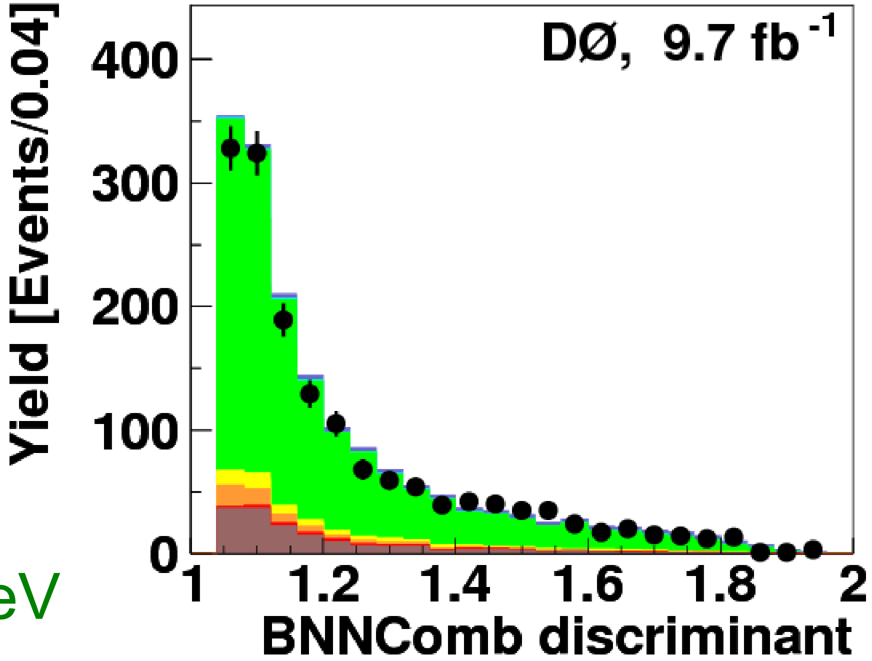
Discriminant on cross-check samples

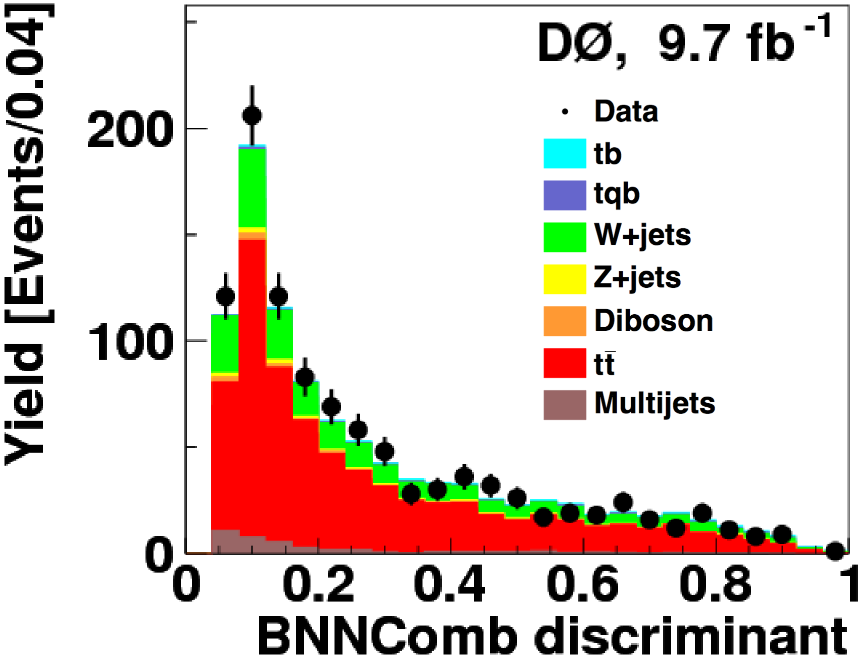
s-like

t-like

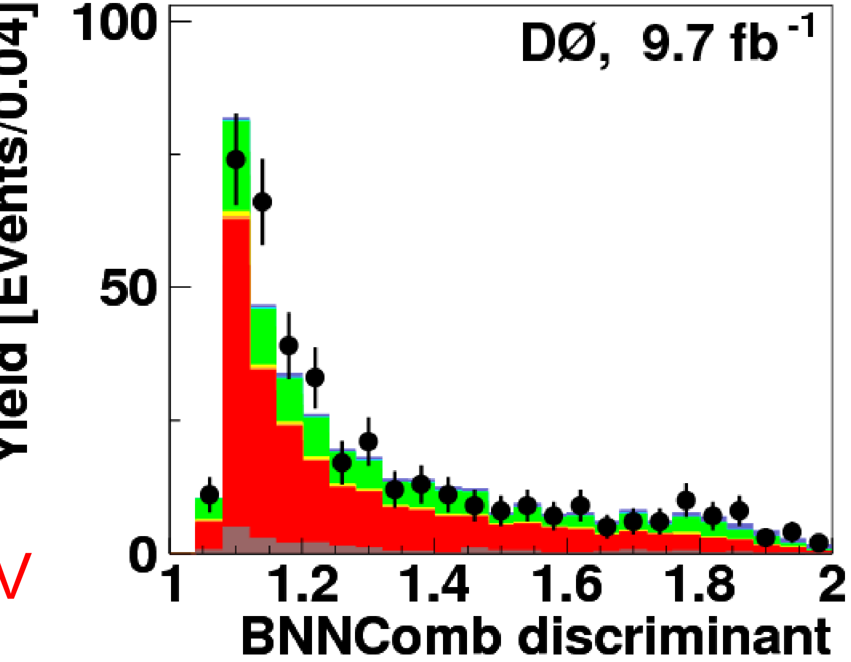


W+jets enriched
1 b-tag
2 jets
H_T<175 GeV



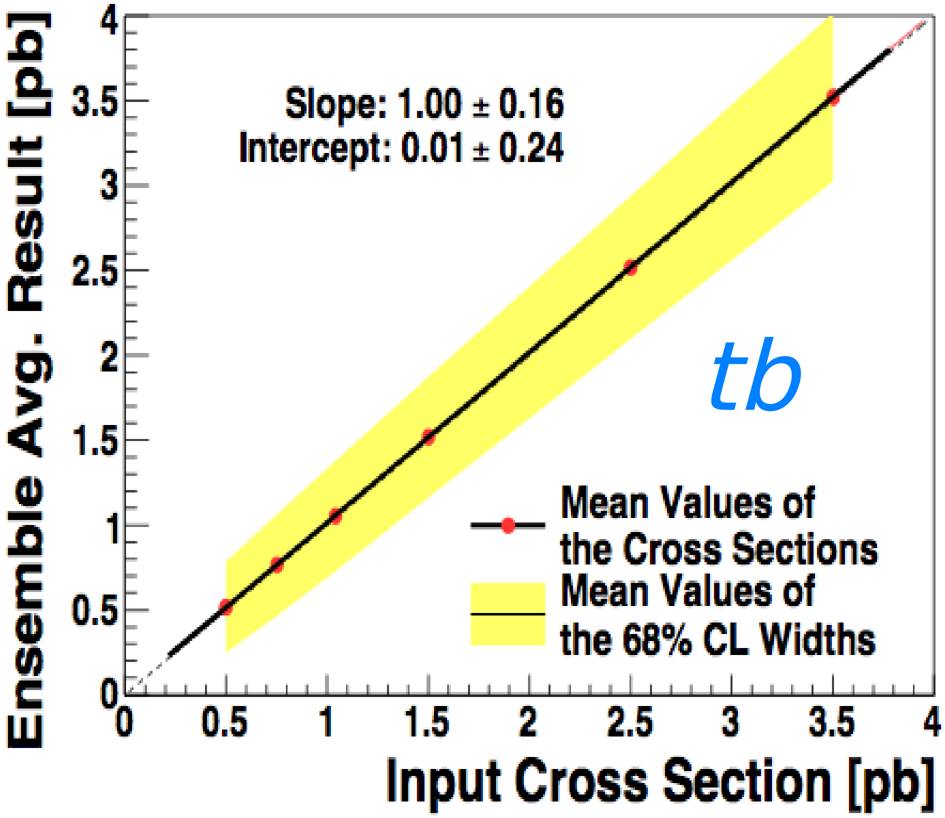


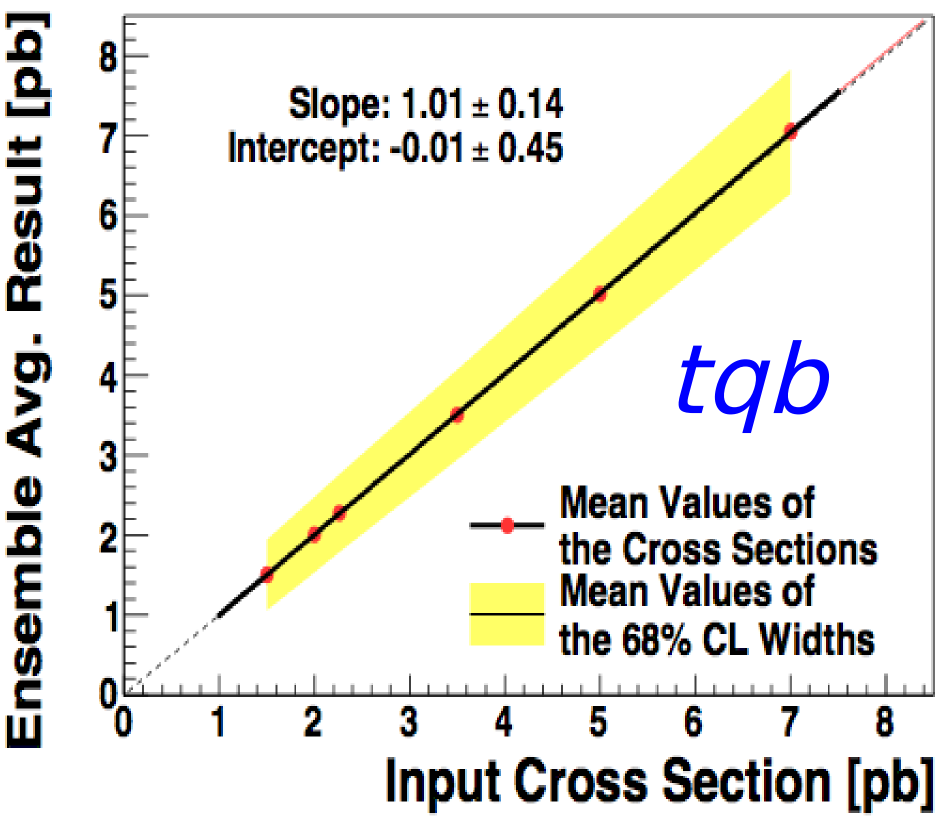
tt enriched
1,2 b-tag
3 jets
H_T>300 GeV



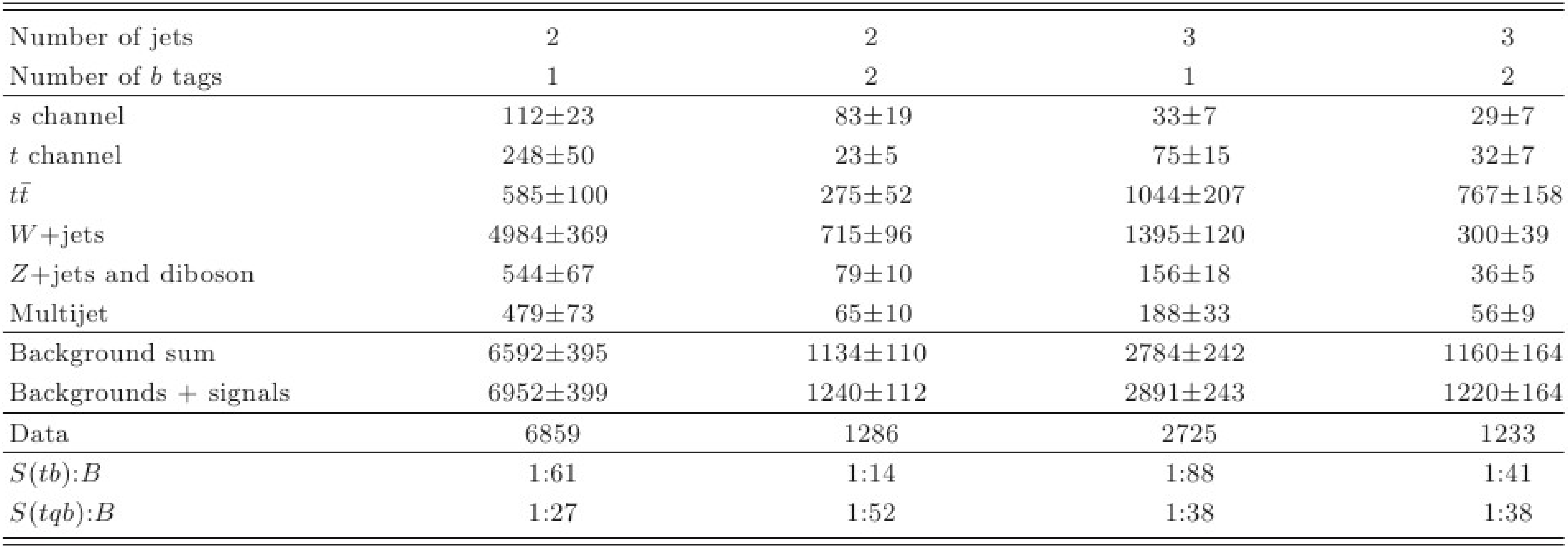
Linearity test

- Generate ensembles of pseudo-data samples
- Each ensemble has a different input signal σ
- All systematics included
- Extract the signal cross section from each pseudo-data sample
- No calibration needed





Yields per channel



Number of jets	2	2	3	3
Number of b tags	1	2	1	2
s channel	112±23	83±19	33±7	29±7
t channel	248±50	23±5	75±15	32±7
$t\bar{t}$	585±100	275±52	1044±207	767±158
W+jets	4984±369	715±96	1395±120	300±39
Z+jets and diboson	544±67	79±10	156±18	36±5
Multijet	479±73	65±10	188±33	56±9
Background sum	6592±395	1134±110	2784±242	1160±164
Backgrounds + signals	6952±399	1240±112	2891±243	1220±164
Data	6859	1286	2725	1233
$S(tb)$:B	1:61	1:14	1:88	1:41
$S(tqb)$:B	1:27	1:52	1:38	1:38

Crash course in Bayesian probability

Bayes’ theorem expresses the degree of belief in a hypothesis A, given another B. “Conditional” probability P(A|B):

$$P(A|B)=\frac{P(B|A)P(A)}{P(B)}$$

In HEP: $B\rightarrow N_{\text{observed}}$, $A\rightarrow n_{\text{predicted}}=n_{\text{signal}}+n_{\text{bkgd}}$, $n_s=Acc*L*\sigma$

P(B|A): “model” density, or likelihood: $L(N_{\text{observed}}|n_{\text{predicted}})=n^N e^{-n}/N!$

P(A): “prior” probability density $\Pi(n_{\text{pred}})=\Pi(Acc*L,n_b)\Pi(\sigma)$

$\Pi(n_s,n_b)$ multivariate gaussian ; $\Pi(\sigma)$ assumed flat

P(B): normalization constant Z: $P(N_{\text{observed}})$

P(A|B): “posterior” probability density $P(n_{\text{predicted}}|N_{\text{observed}})$

$$P(n_{\text{predicted}}|N_{\text{observed}}) = 1/Z\ L(N_{\text{observed}}|n_{\text{predicted}})\Pi(n_{\text{pred}})$$

W+jets normalization

- Find fractions of real and fake isolated ℓ in the data before b-tagging. Split samples in loose and tight isolation:

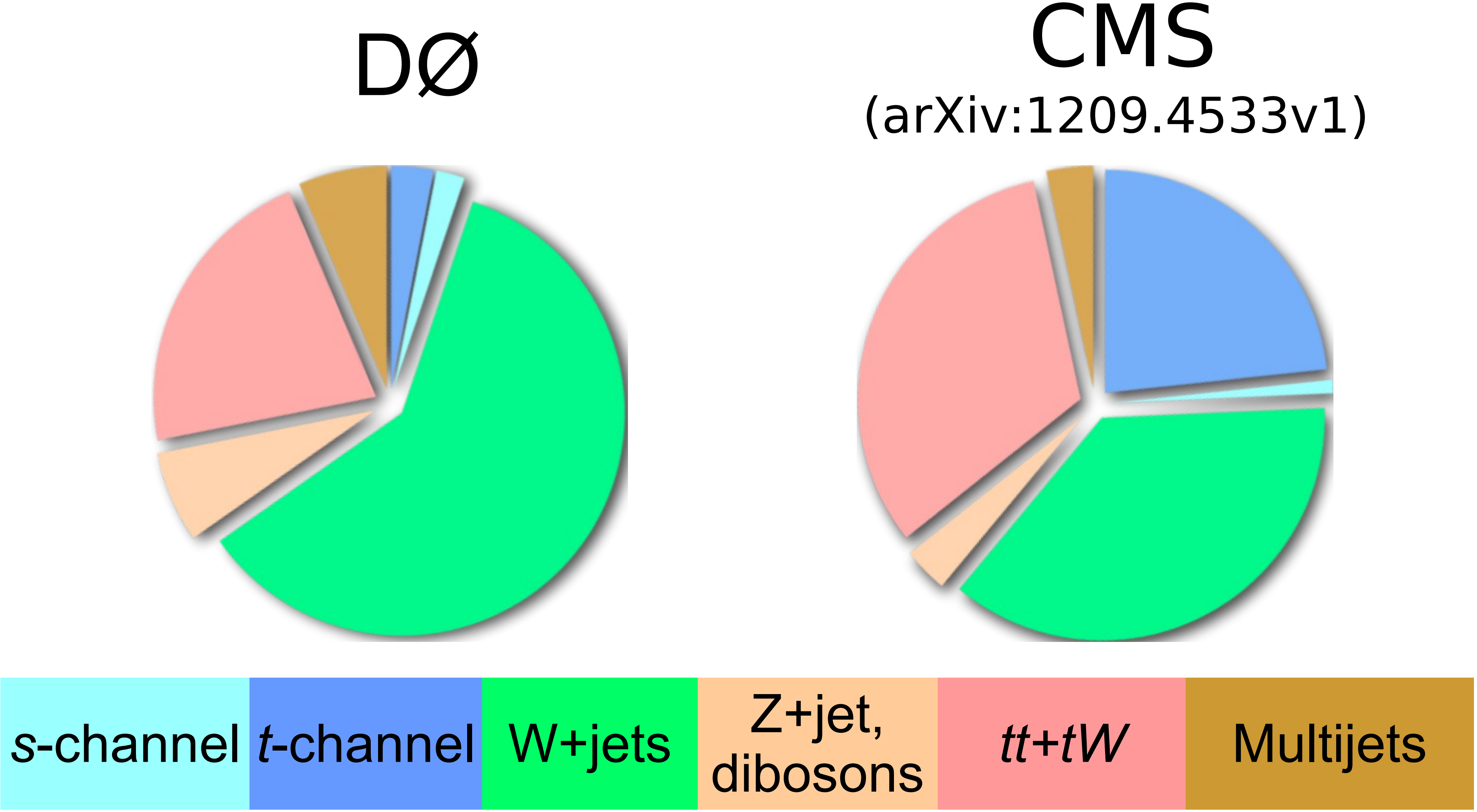
$$N^{loose} = N^{loose}_{fake} + N^{loose}_{real}$$

$$N^{tight} = \varepsilon_{fake} N^{loose}_{fake} + \varepsilon_{real} N^{loose}_{real}$$

Obtain: N^{loose}_{real} and N^{loose}_{fake}

- Obtain ε_{fake} and ε_{real} from MC and data samples
- Then apply b-tagging
  - Greatly reduce W+jets background (Wbb ~5% of Wjj)
  - Shift distributions, changes flavor composition

Background contribution



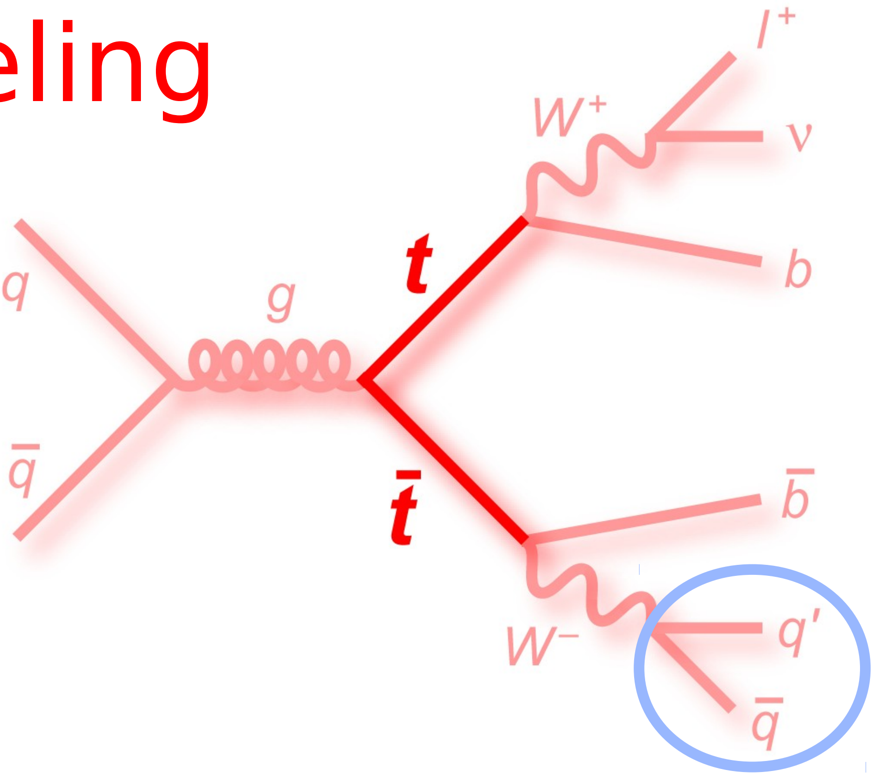
ME processes

Two Jets		Three Jets	
Name	Process	Name	Process
tb	$u\bar{d} \to t\bar{b}$	tbg	$u\bar{d} \to t\bar{b}g$
tq	$ub \to td$	tqb	$ug \to td\bar{b}$
	$\bar{d}b \to t\bar{u}$		$\bar{d}g \to t\bar{u}\bar{b}$
		tqg	$ub \to tdg$
			$\bar{d}b \to t\bar{u}g$
Wbb	$u\bar{d} \to Wb\bar{b}$	$Wbbg$	$u\bar{d} \to Wb\bar{b}g$
Wcg	$sg \to Wcg$	$Wugg$	$\bar{u}g \to W\bar{u}gg$
Wgg	$u\bar{d} \to Wgg$		
WW	$u\bar{u} \to WW$		
WZ	$u\bar{d} \to WZ$		
ggg	$gg \to ggg$		
$t\bar{t}$	$u\bar{u} \to t\bar{t}$	$t\bar{t}$	$u\bar{u} \to t\bar{t}$

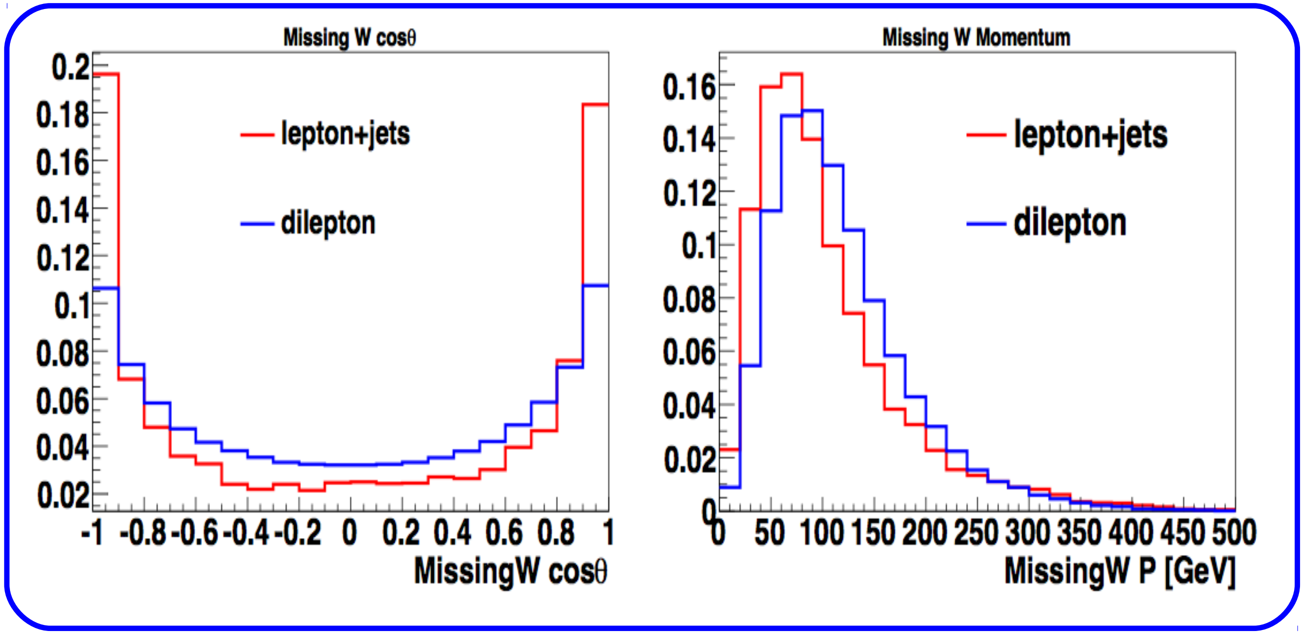
- the more background diagrams, the better discrimination

ME $t\bar{t}$ modeling

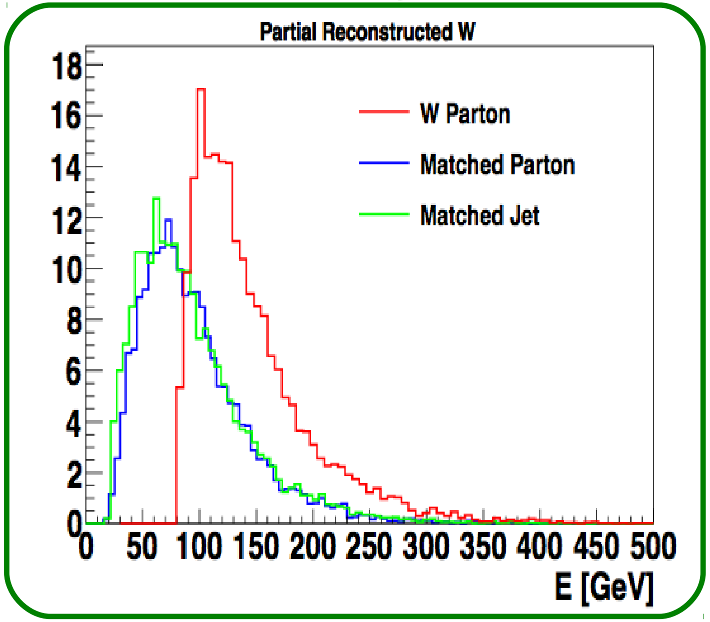
- $t\bar{t} \rightarrow \ell\nu bqq'b$ (4 jets)
- $t\bar{t}$ yields in 2jet & 3jet channels are comparable to single top
- Light-jets are 1.6 times more likely to be lost than b-jets
- Use simulation to derive a prior of missing jet (3jet) or missing W (2jet)



Missing W prior



Missing Jet



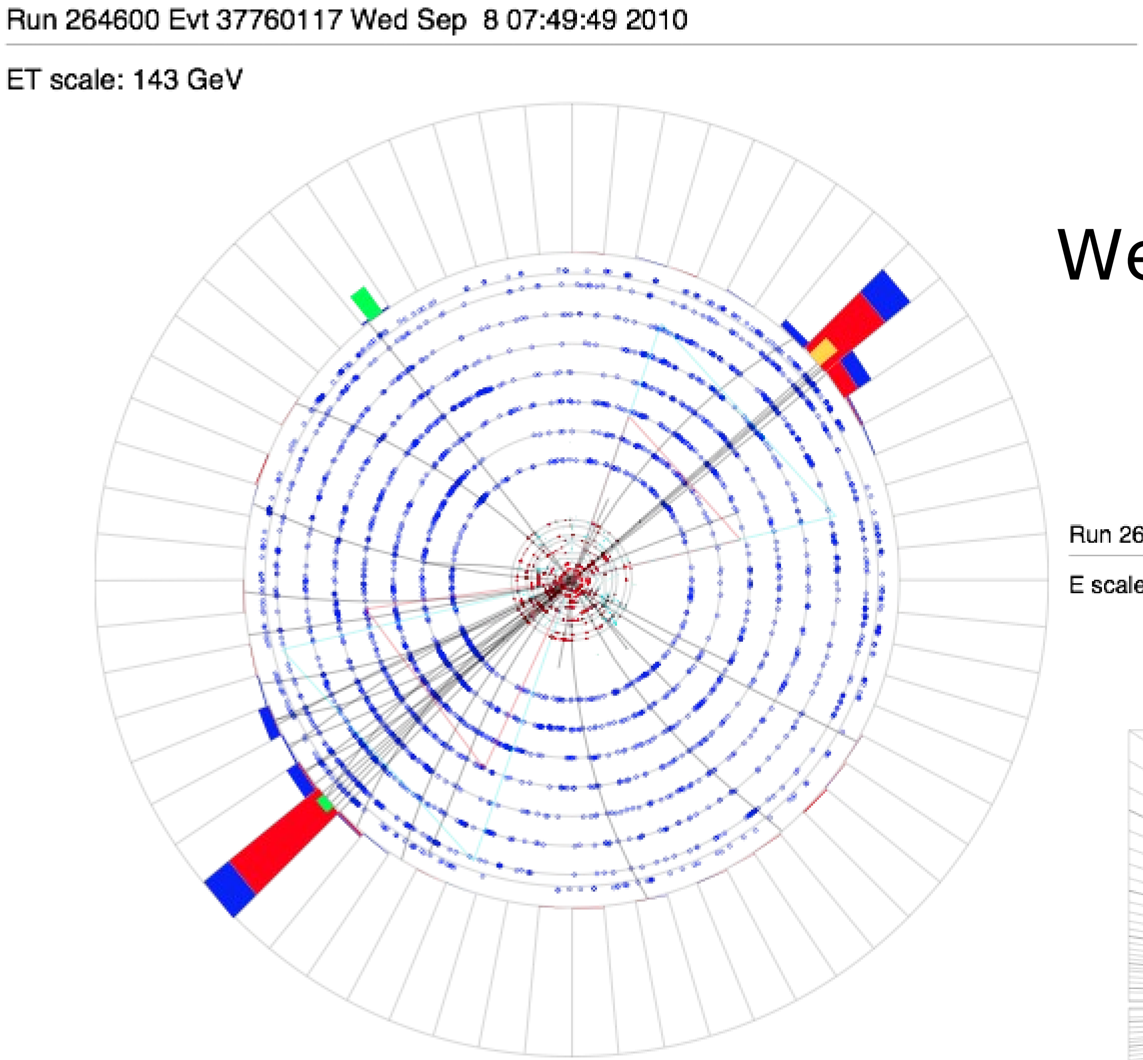
- t-channel gets weights for each probability

$$P=\frac{\Sigma_j w_j d\sigma_j}{\left(\Sigma_j w_j\right)\sigma_j} \qquad w_j = \left\{\begin{matrix} b_1 b_2 \\ b_1\left(1-b_2\right), b_2\left(1-b_1\right) \\ \left(1-b_1\right)\left(1-b_2\right) \end{matrix}\right\}$$

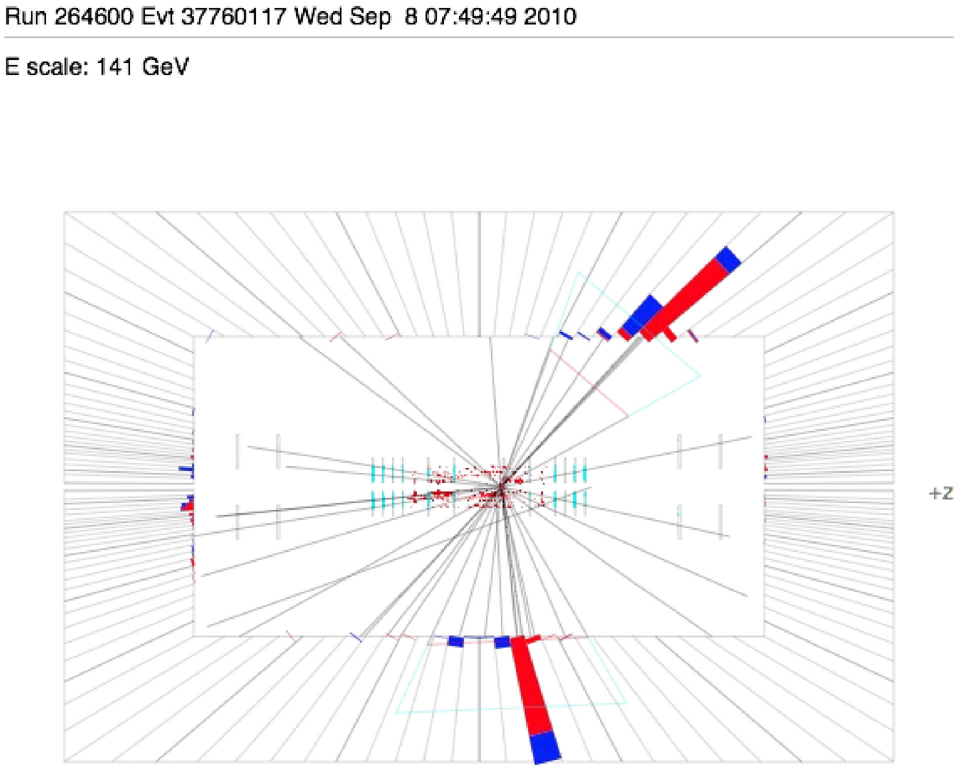
- s-channel gets overall weight for event

$$D=\frac{b_1 b_2 P_{tq}}{b_1 b_2 P_{tb,Wbb,WZ,tt}+b_1\left(1-b_2\right)P_{tq}+\left(1-b_1\right)\left(1-b_2\right)P_{Wcg,Wgg,WW,ggg}}$$

Another candidate event



Run 264600
Event 37760117
Wed. Sep. 8 07:49:49 2010



$m_t = 175$ GeV
Jet1 b-tag: 0.32
Jet2 b-tag: 0.39