Evidence for s-channel single top quark production at DØ

Jyoti Joshi
University of California, Riverside

HEP Seminar, BNL, 07/18/2013

Outline

- Electroweak production of top quarks.
- Event selection and background modeling.
- Multivariate Techniques.
- Cross section and significance.
- Direct measurement of $|V_{tb}|$.

The Top Quark

- Discovered in 1995 by the DØ and CDF collaborations at Fermilab

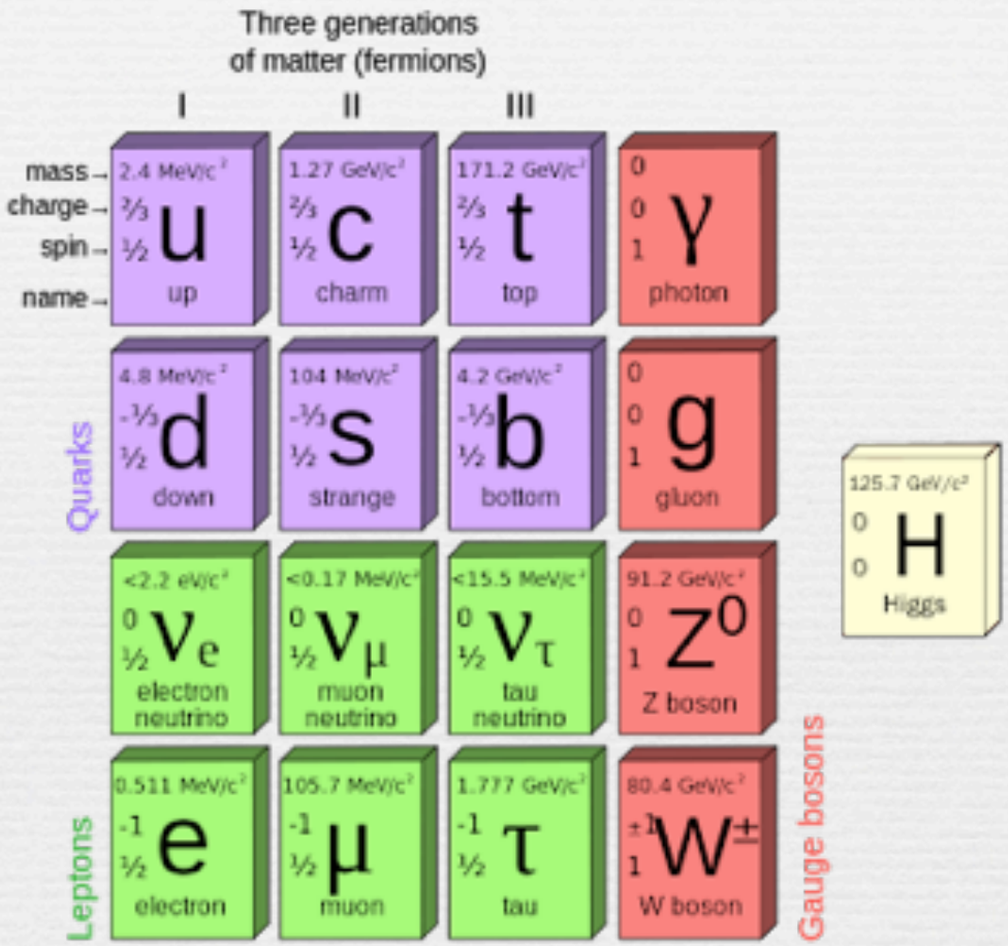
- By far the **heaviest fundamental particle** we know of : ~ **173.2 ± 0.9 GeV**

  (arXiv:1305.3929)

  - couples very strongly to higgs because of its mass
  - sensitive to probe new physics

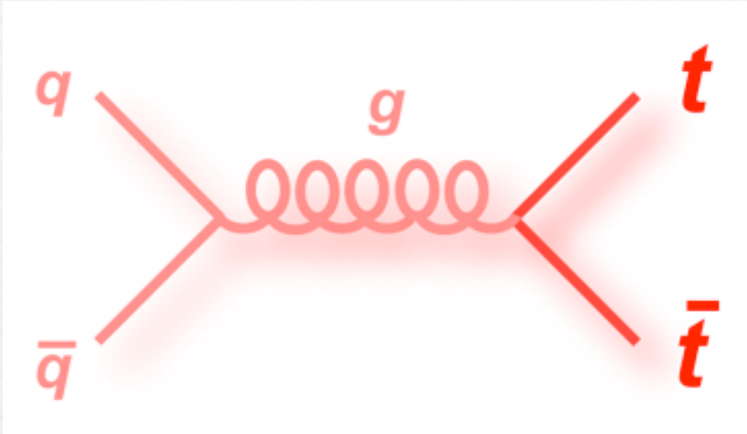
- **Decays before it hadronizes**: $t \longrightarrow \cancel{W}b$

  - very short lifetime
  - imparts its spin properties to the decay products.

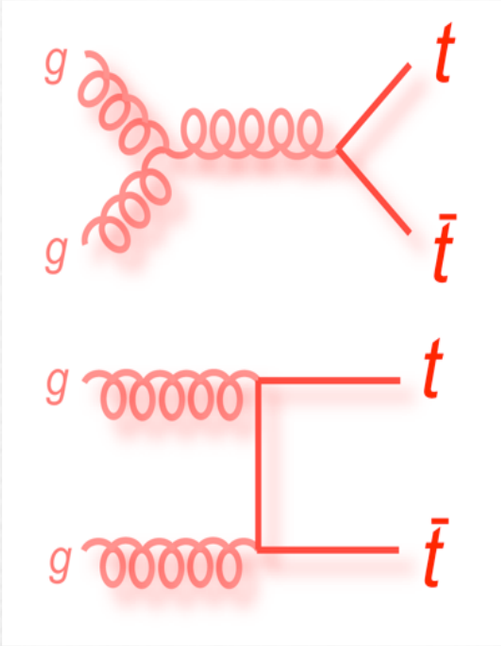


Top Production at the Hadron Colliders

- **Strong interaction : top pair**

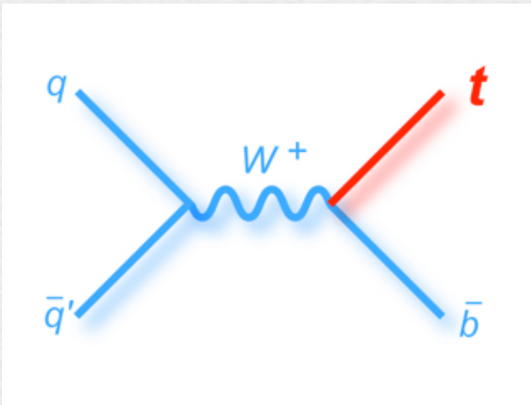


$q\bar{q}$ annihilation

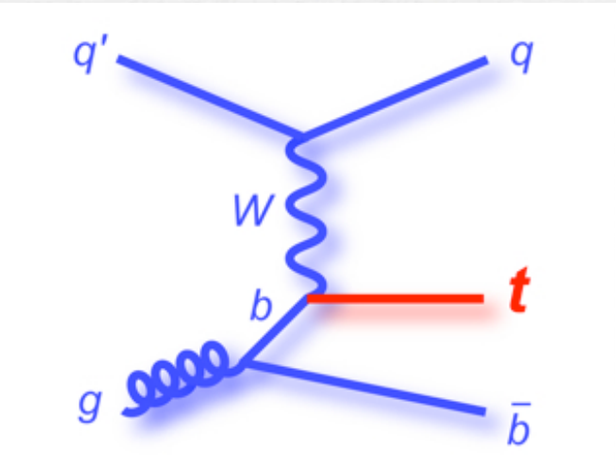


gg fusion

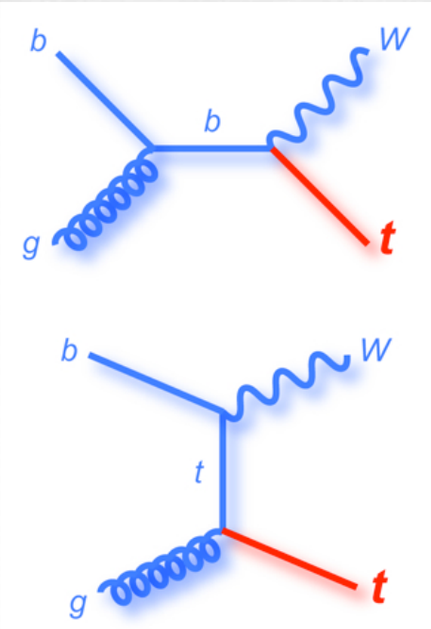
- **Electroweak interaction : single top**



s-channel (tb)



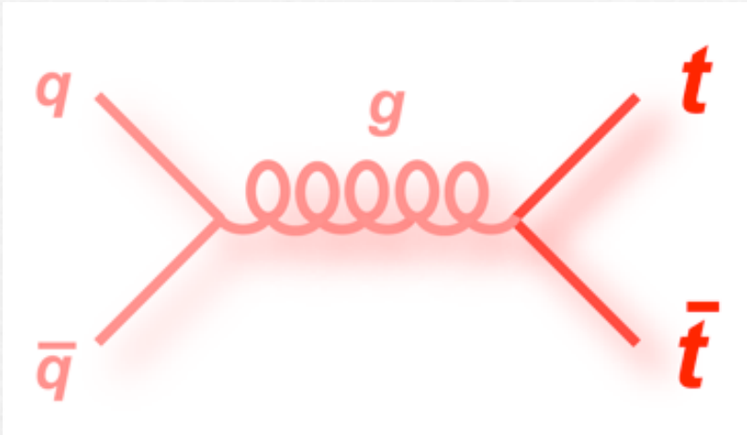
t-channel (tqb)



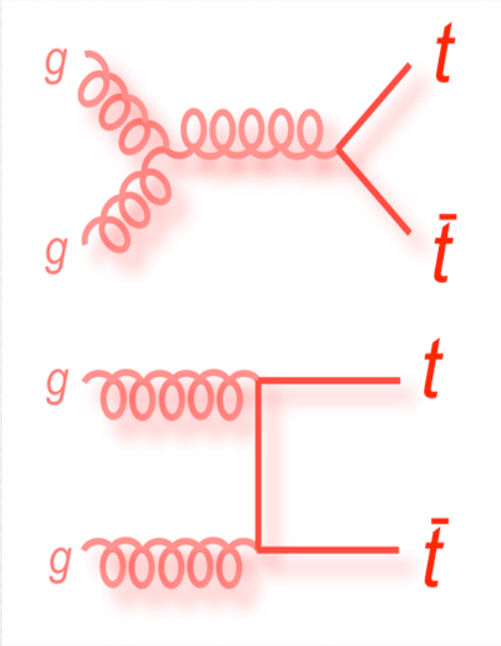
Associated (tW)

Top Production at the Tevatron

- **Strong interaction : top pair**



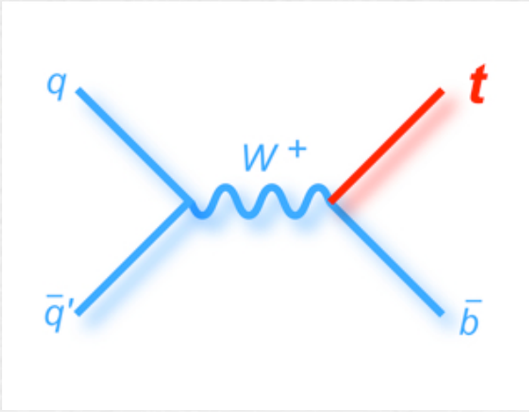
$q\bar{q}$ annihilation



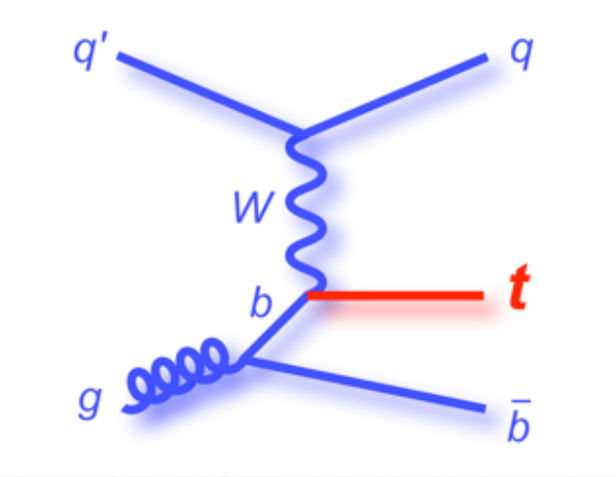
gg fusion

Main production modes at the Tevatron

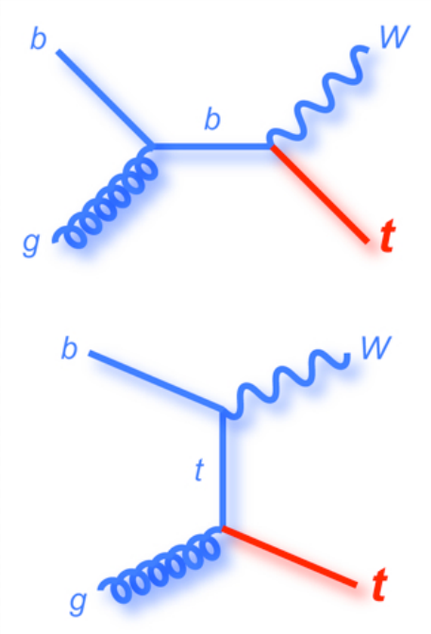
- **Electroweak interaction : single top**



s-channel (tb)



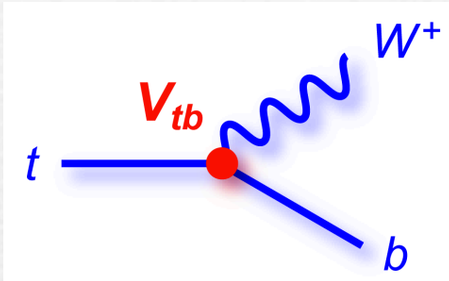
t-channel (tqb)



Associated (tW)

Electroweak Top Production

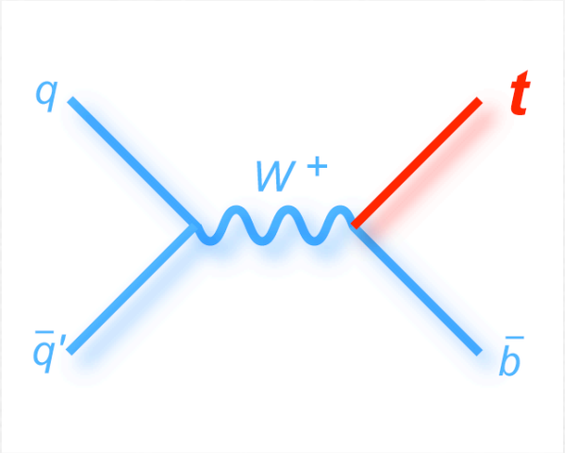
- s-channel and t-channel production cross section measurement
- Direct $|V_{tb}|$ measurement : test CKM unitarity



- New physics can change σ_{tb} and σ_{tqb} differently:
  - σ_{tb} : New bosons, H^+, W'
  - σ_{tqb} : Anomalous couplings, FCNCs

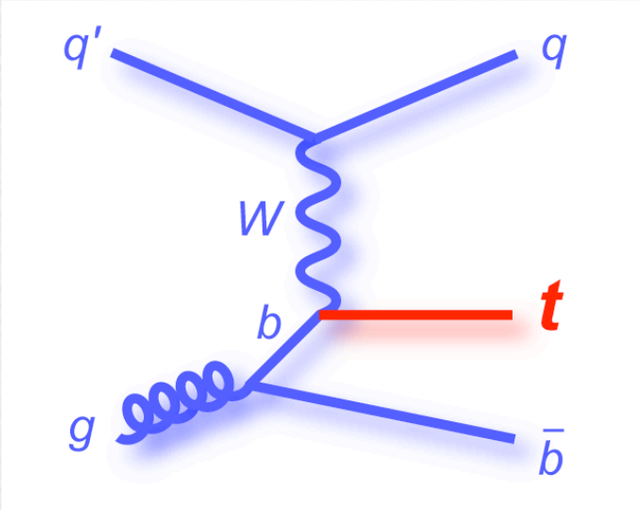
	tb [pb]	tqb [pb]	tW [pb]
Tevatron[1] (1.96 TeV)	1.04	2.26	0.30
	×4	×30	×50
LHC[2] (7 TeV)	4.59	64.6	15.7

s-channel (tb)



σ_{tb} = **1.04 ± 0.04 pb**

t-channel (tqb)



σ_{tqb} = **2.26 ± 0.12 pb**

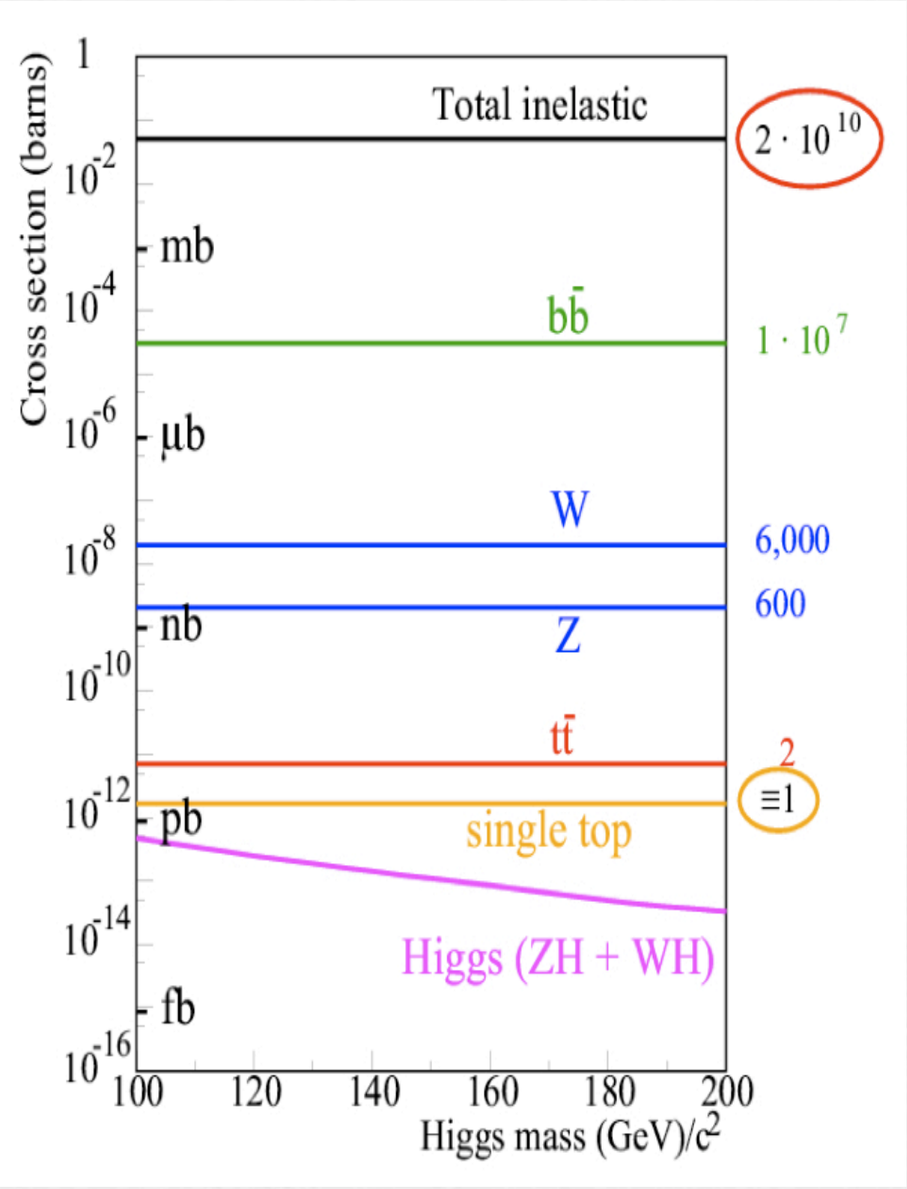
Top Mass = 172.5 GeV; σ = (N)NLO pb

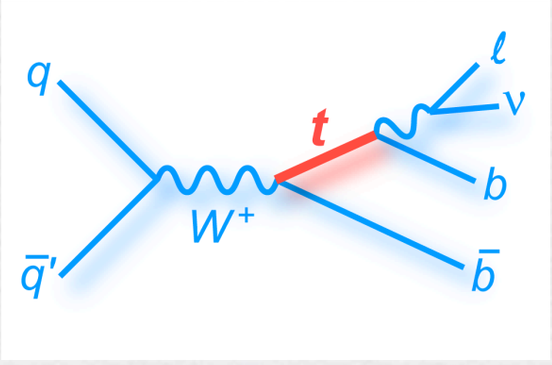
1: PRD 74, 114012 (2006)

2: PRD 81, 054028 (2010)
PRD 83, 091503 (2011)

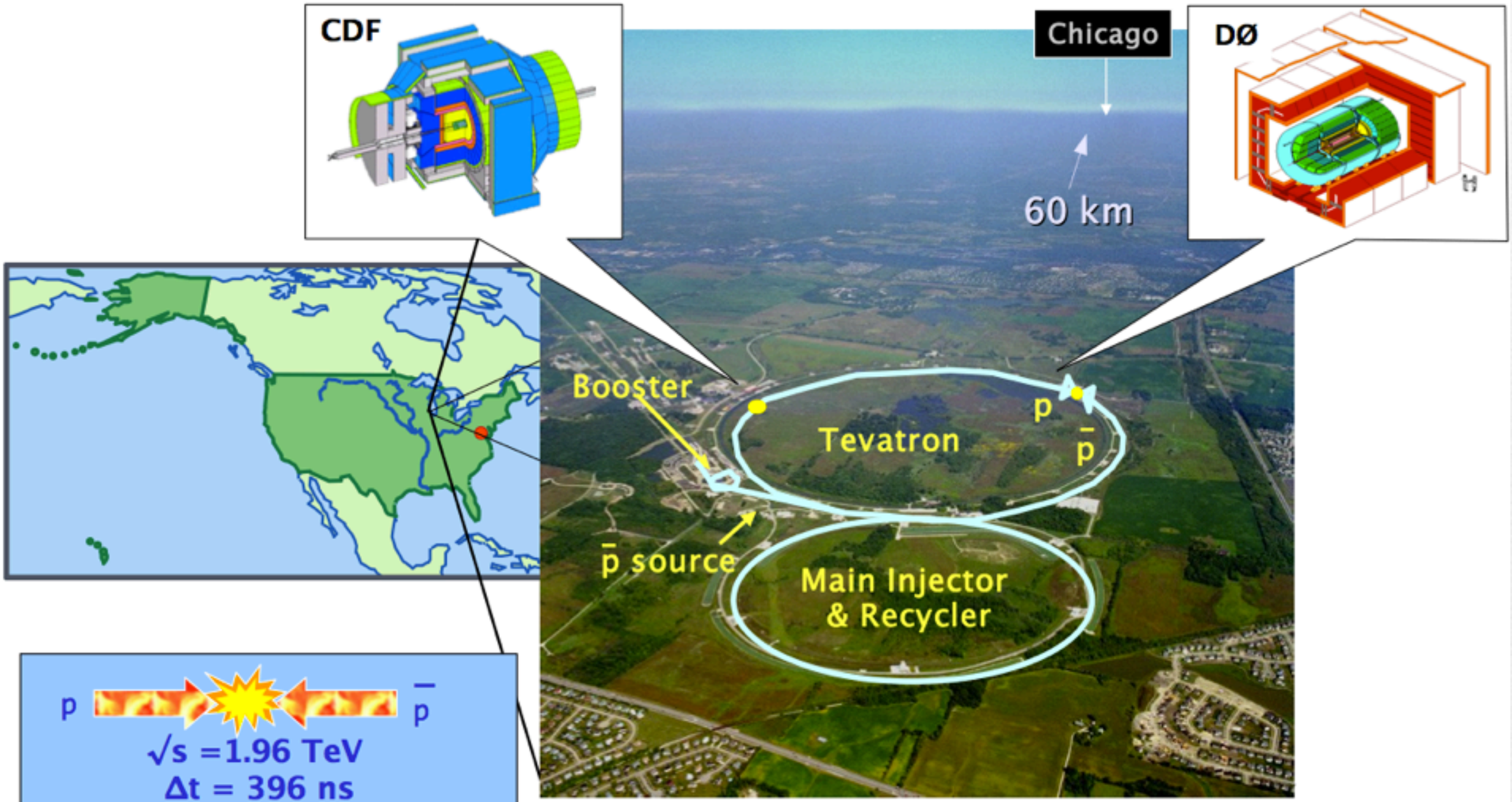
A Challenging Analysis

- Observed at Tevatron 14 years after the observation of top quark produced by strong interaction.
  - σ (top pair) ~ 7.5 pb
- Smaller Cross section ~ 3 pb
  - Tevatron produced 32K single top events
  - Analyze leptonic final states : 6.8K single top events
- Huge Background domination
  - Main background *W*+jets





The Tevatron Collider



p p̄

$\sqrt{s}$ =1.96 TeV

Δt = 396 ns

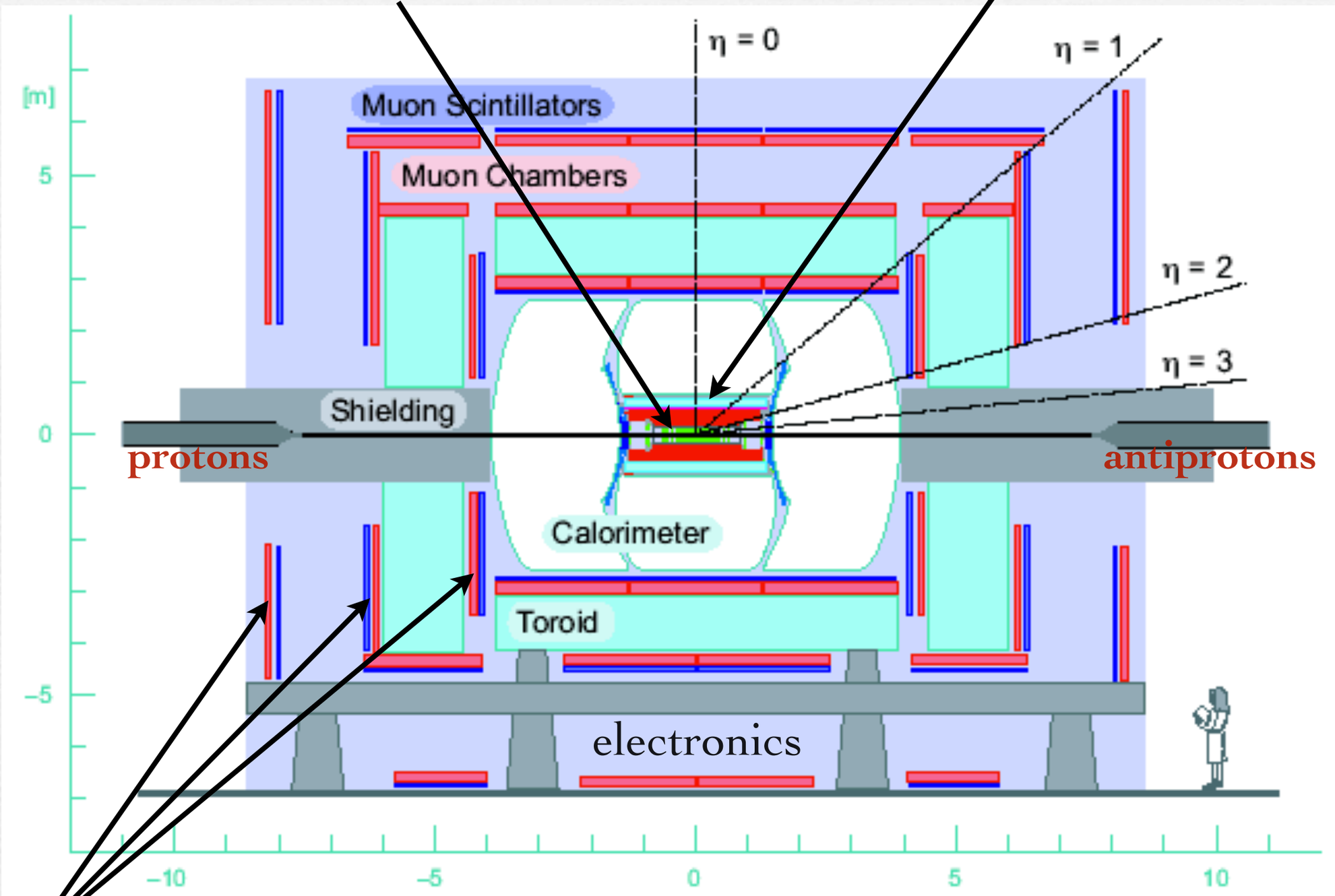
Run I 1987 (92)–95
Run II 2001–11: 100x larger dataset at increased energy

10^{-12}s after big bang!

RunII DØ Detector

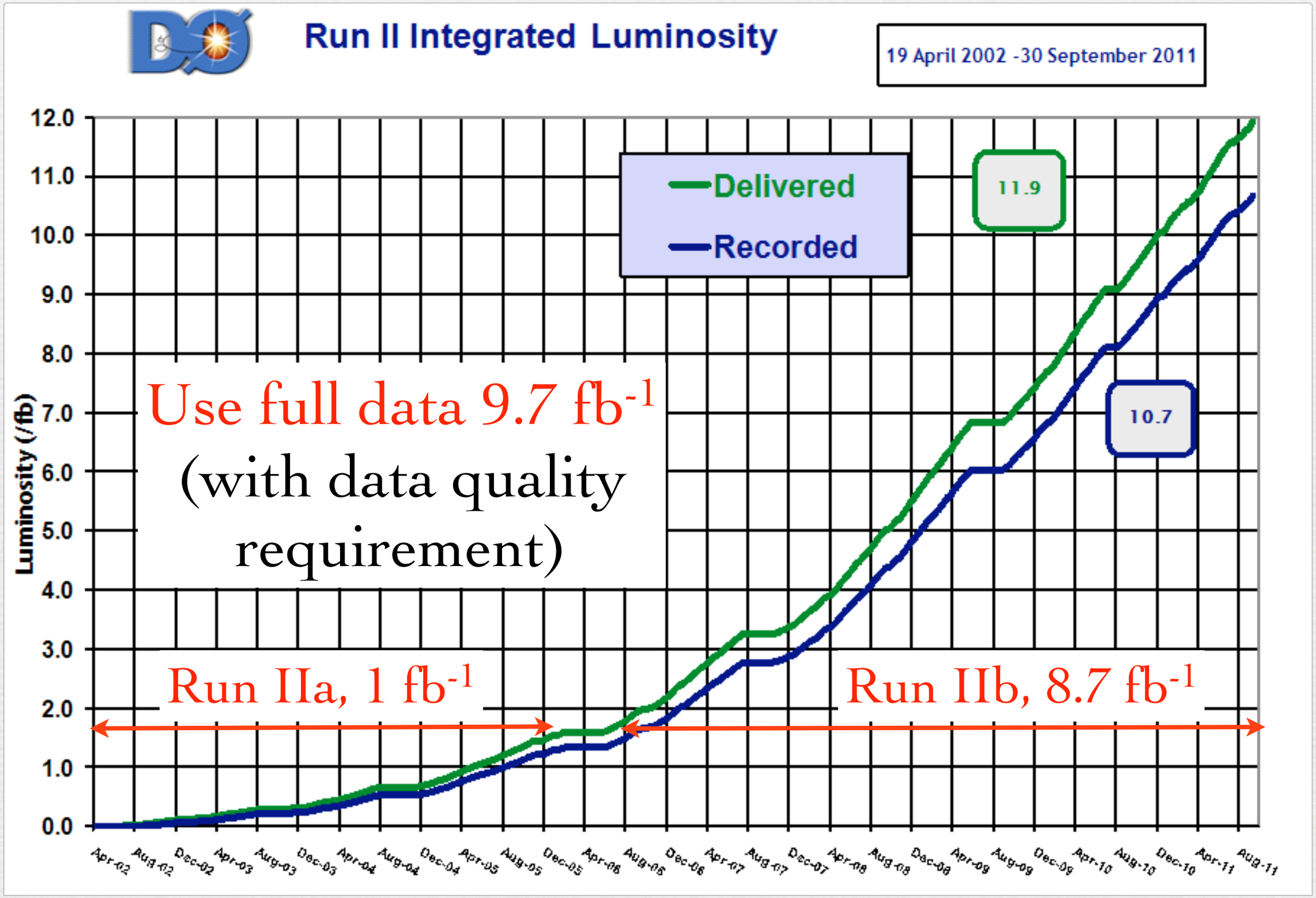
Central Tracking :
Silicon and Fiber tracker

2T Solenoid Magnet



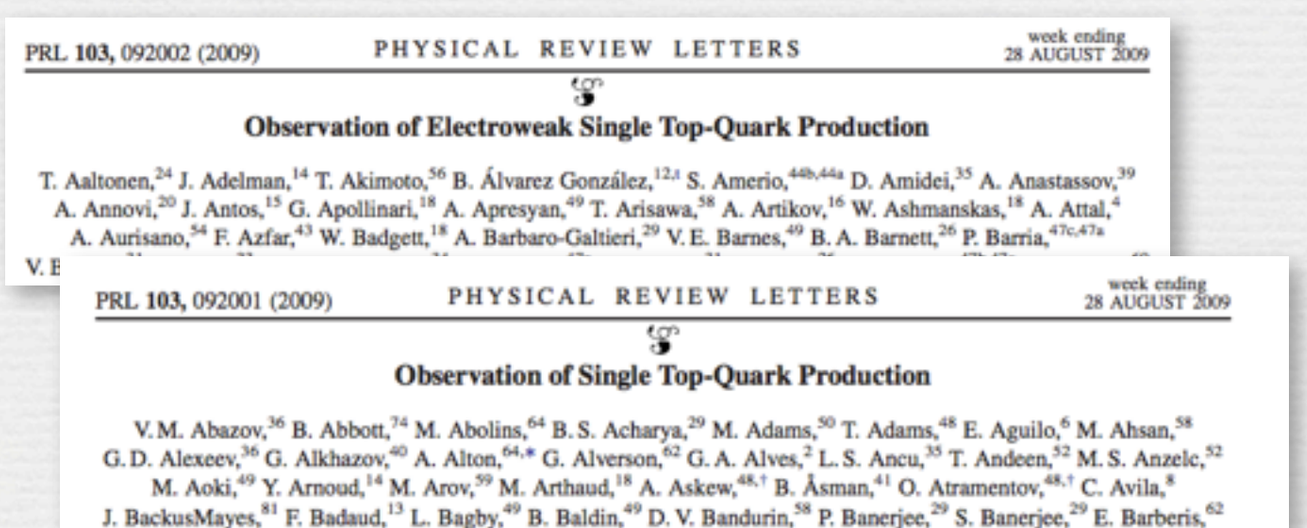
3 layer Muon System

Dataset Used



Single top at the Tevatron

- Search: PRD 65, 091102 (2002)
- W': PRL 90, 081802 (2003)
- Search: PRD 69, 052003 (2004)
- Search: PRD 71, 012005 (2005)
- **Evidence: PRL 101, 252001 (2008)**
- FCNC: PRL 102 151801 (2009)
- W': PRL 103 041801 (2009)
- **Observation: PRL 103 092002 (2009), PRD82 112005 (2009)**
- **Tevatron combination: arxiv:0908.2171 (2009)**
- Search of s-channel: conf. note 9712 (2009)
- Wtb: conf. note 9920 (2009)


PRL 103, 092002 (2009) PHYSICAL REVIEW LETTERS week ending 28 AUGUST 2009

Observation of Electroweak Single Top-Quark Production

T. Aaltonen,[24] J. Adelman,[14] T. Akimoto,[56] B. Álvarez González,[12,l] S. Amerio,[44b,44a] D. Amidei,[35] A. Anastassov,[39] A. Annovi,[20] J. Antos,[15] G. Apollinari,[18] A. Apresyan,[49] T. Arisawa,[58] A. Artikov,[16] W. Ashmanskas,[18] A. Attal,[4] A. Aurisano,[54] F. Azfar,[43] W. Badgett,[18] A. Barbaro-Galtieri,[29] V. E. Barnes,[49] B. A. Barnett,[26] P. Barria,[47c,47a]

PRL 103, 092001 (2009) PHYSICAL REVIEW LETTERS week ending 28 AUGUST 2009

Observation of Single Top-Quark Production

V. M. Abazov,[36] B. Abbott,[74] M. Abolins,[64] B. S. Acharya,[29] M. Adams,[50] T. Adams,[48] E. Aguilo,[6] M. Ahsan,[58] G. D. Alexeev,[36] G. Alkhazov,[40] A. Alton,[64,*] G. Alverson,[62] G. A. Alves,[2] L. S. Ancu,[35] T. Andeen,[52] M. S. Anzelc,[52] M. Aoki,[49] Y. Arnoud,[14] M. Arov,[59] M. Arthaud,[18] A. Askew,[48,†] B. Åsman,[41] O. Atramentov,[48,†] C. Avila,[8] J. BackusMayes,[81] F. Badaud,[13] L. Bagby,[49] B. Baldin,[49] D. V. Bandurin,[58] P. Banerjee,[29] S. Banerjee,[29] E. Barberis,[62]

- Search: PRD 63, 031101 (2000)
- Search: PLB 517, 282 (2001)
- Search: PLB 622, 265 (2005)
- W': PLB 641, 423 (2006)
- Search: PRD 75, 092007 (2007)
- **Evidence: PRL 98, 181802 (2007)**
- FCNC: PRL 99, 191802 (2007)
- W': PRL 100, 211802 (2007)
- Evidence: PRD 78, 012005 (2008)
- Wtb: PRL 101, 221801 (2008)
- Wtb: PRL 102, 092002 (2009)
- Wtb: DØ Note 5838-CONF (2009)
- H+: PRL 102, 191802 (2009)
- **Observation: PRL 103, 092001 (2009)**
- **Tevatron combination: arxiv:0908.2171**
- Search of t-,s-channel: PLB 682, 363 (2010)
- Cross section (tau+jets): PLB 690, 5 (2010)
- FCNC: PLB 693, 81 (2010)
- W': PLB 699, 145 (2011)
- Top width and lifetime: PRL 106, 022001 (2011)
- Observation of t-channel: PLB 705, 313 (2011)
- Cross section of s+t- and s-channel: PRD 84, 112001 (2011)

Next ?

Experimental Status : LHC limited on s-channel

σ (pb) ~ NNLO	tb [pb]	tqb [pb]	tW [pb]
Tevatron Prediction[1] (1.96 TeV)	1.04	2.26	0.3
CDF	1.81 ± 0.6	1.49 ± 0.45	
DØ	0.68 ± 0.36	★ 2.9 ± 0.59	
LHC Prediction[2] (7 TeV)	4.59	64.6	15.7
ATLAS	< 20.5 (95% C.L.)	★ 83 ± 20	● 17 ± 6
CMS		★ 67 ± 6	● 16 ± 5

★ **Observation (5 SD)**

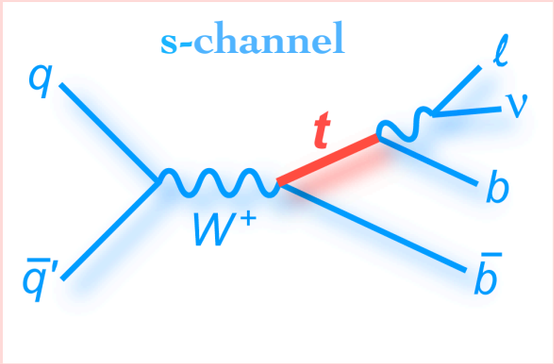
● **Evidence (3 SD)**

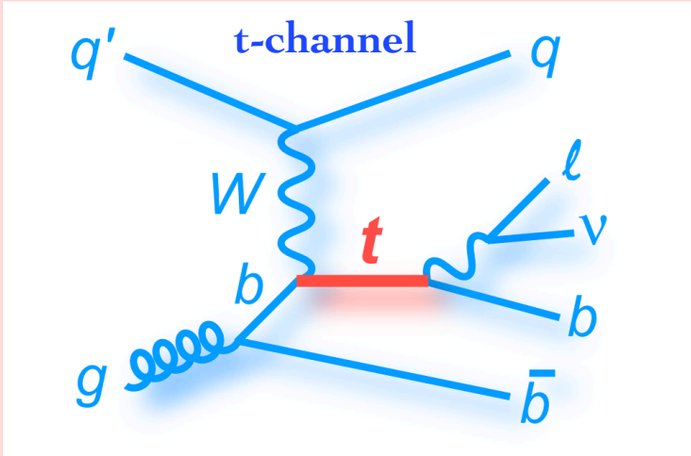
1: PRD 74, 114012 (2006)
2: PRD 81, 054028 (2010)
PRD 83, 091503 (2011)

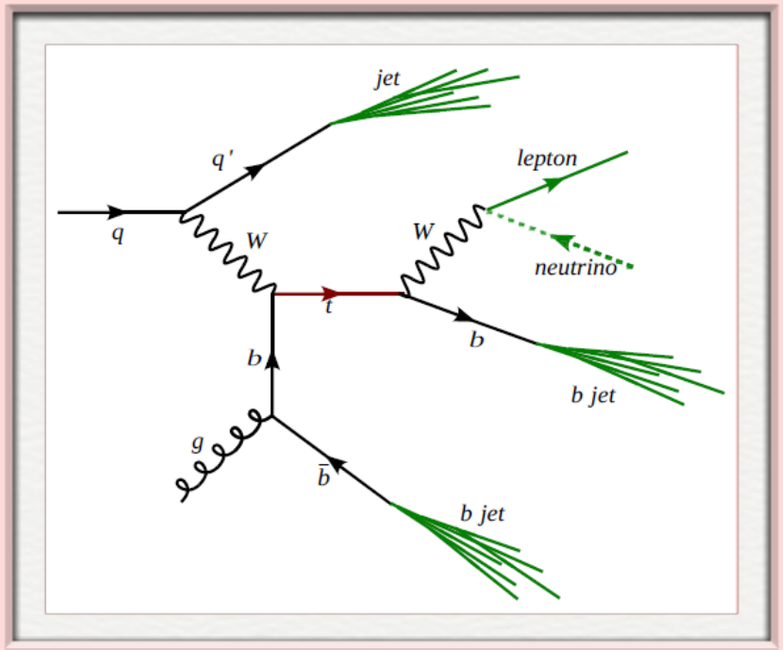
Event Signature and Selection

- **One High p_T isolated Lepton from W boson**
  - **Electron Selection -** $p_T > 20$ GeV, $|\eta| < 1.1$
  - **Muon Selection -** $p_T > 20$ GeV, $|\eta| < 2.0$
- **Large Missing transverse energy from W boson**
  - MET > 20 GeV
- **Two or three jets**
  - $p_T > 25$ GeV (jet1), $p_T > 20$ GeV (other jets)
  - $|\eta| < 2.5$
- **Total Transverse Energy**
  - H_T (l,MET,jets) > 120 GeV
- **B-Tagging Selection**
  - One or Two jets originating from fragmentation of b quarks.

After Selection, samples are splitted into 4 analysis channels based on Tag bin (1Tag/2Tag) and Jet bin (2Jet/3Jet). This helps as different channel has different composition of background.







Signal and Background Simulation

Single Top signal, MC

- COMPHEP (NLO)-PYTHIA generator

W+jets, MC

- Largest Background
- ALPGEN-PYTHIA generator
- Normalization and heavy flavor fraction

ttbar, MC

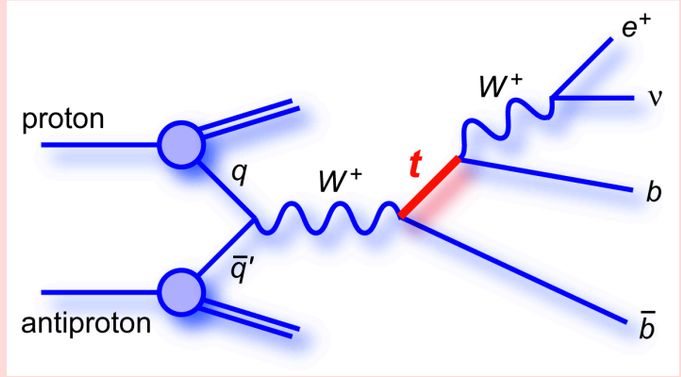
- ALPGEN-PYTHIA generator
- dilepton & lepton+jets
- Normalized to σ_{NNLO} = 7.27 pb

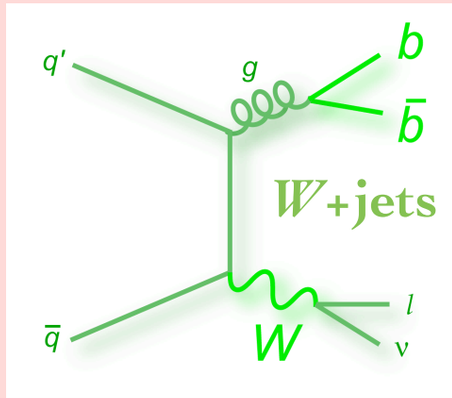
Mulitjets, MC

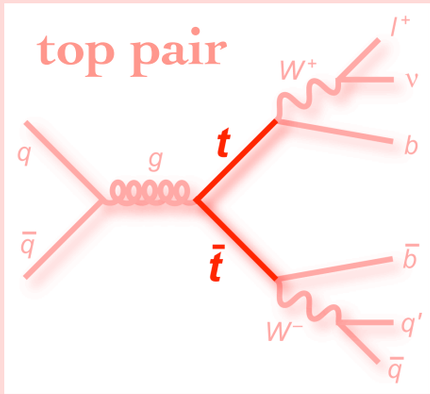
- Orthogonal sample for data
- non-isolated lepton

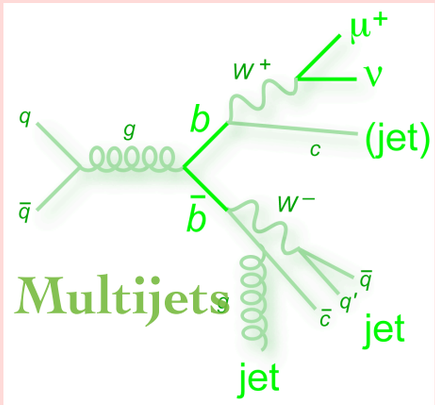
Z+jets and diboson, MC

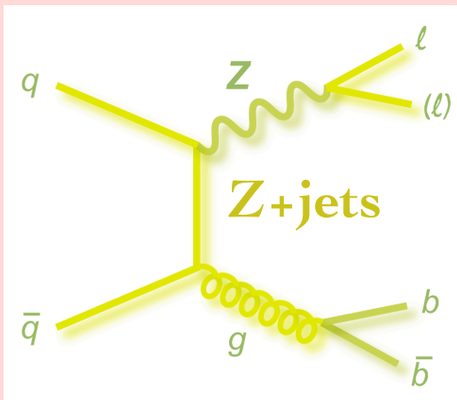
- ALPGEN-PYTHIA generator

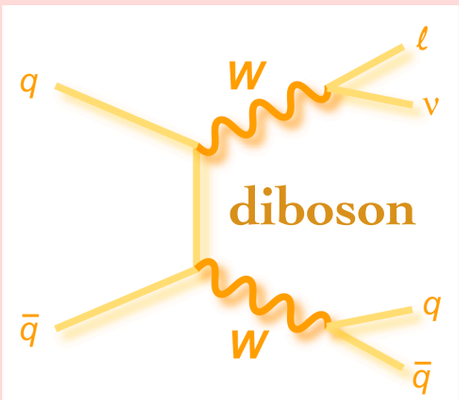






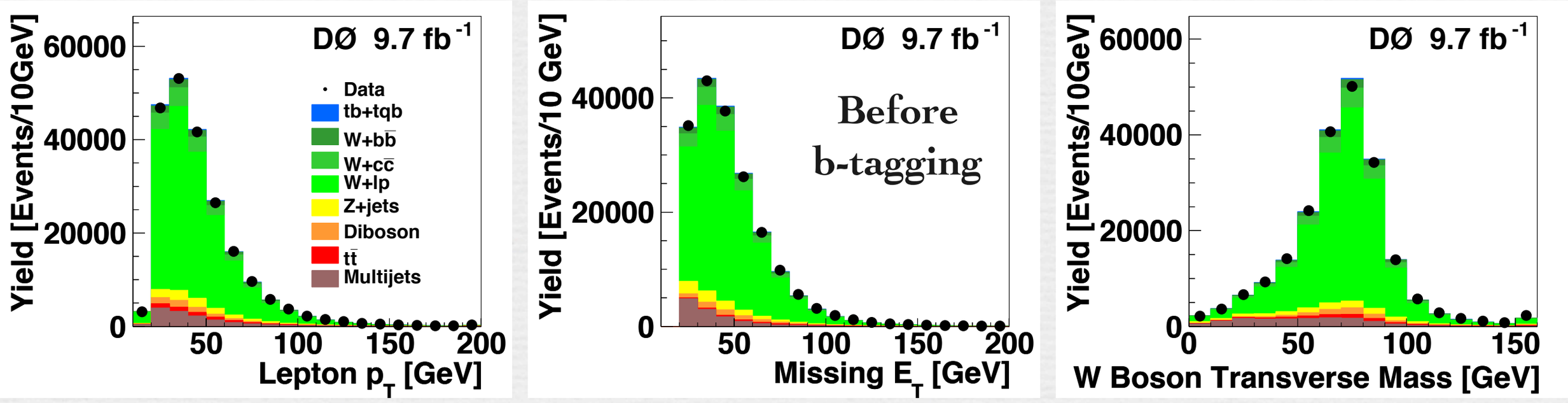






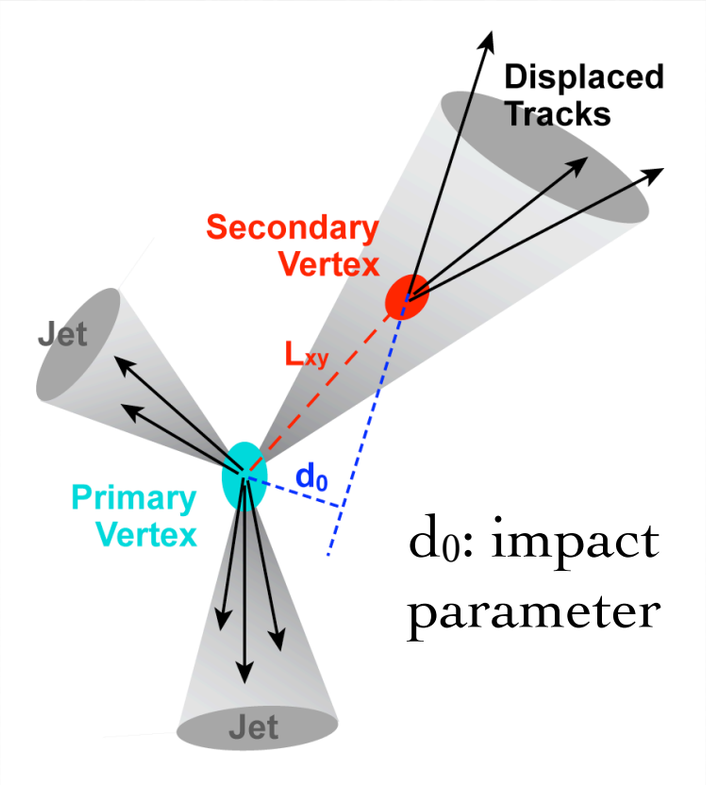
Background Modeling

- Dominated by backgrounds
- Normalize *W*+jets and multijets to data simultaneously before b-tagging using Matrix Method.
  - Split data sample in events with real and fake isolated lepton
  - Measure the probability to have an isolated lepton in each sample.
- *W/Z* Heavy flavor normalized to theory (MCFM-NLO)
- Correct the efficiency of the simulated samples to that of data.



b-jet Identification

- Identify b-jet from light-quark and gluon jets to reduce a lot of backgrounds
- Features of b-hadron in b-jet
  - Long lifetime (~1 ps, L_{xy}~3 mm)
  - Large invariant mass



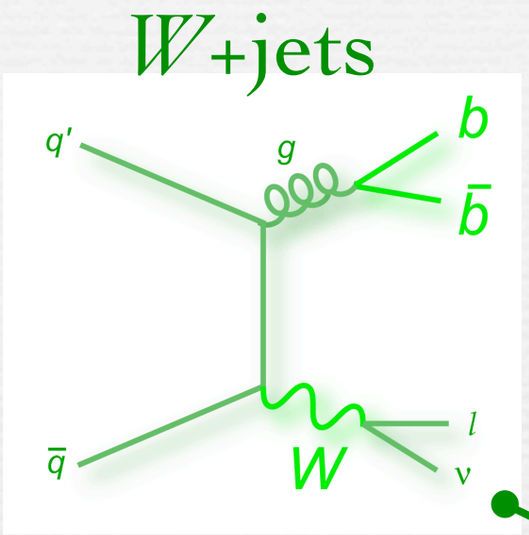
d_0: impact parameter

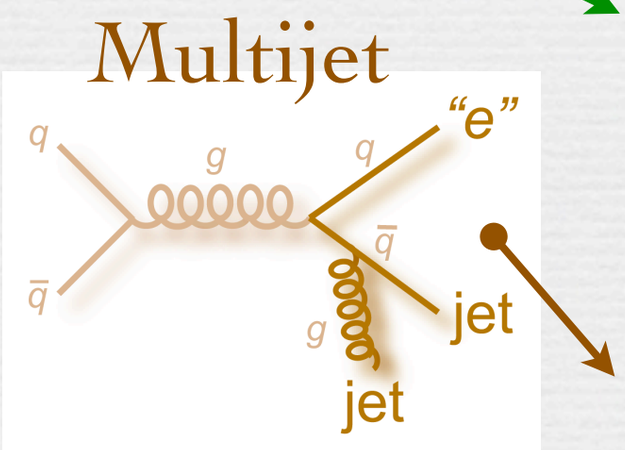
Three different algorithms to identify b-jets :

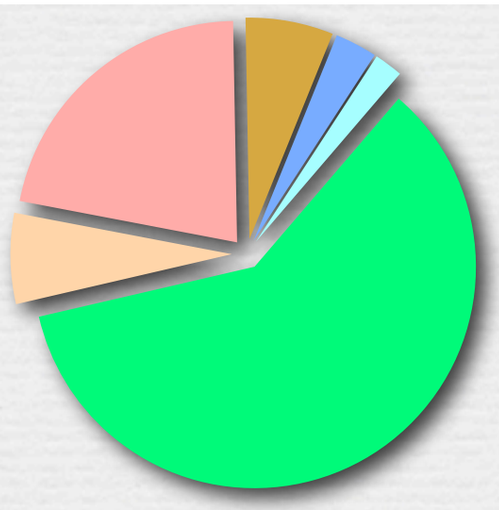
- Two makes use of the displaced tracks (with large impact parameter)
- One based on secondary vertex reconstruction
- Multivariate technique to combine these methods

- b-identification efficiency: (50-70)%
- Misidentification rate: (3-8)%
- Obtain scale factors to correct the MC samples

Event Yield with full DØ dataset

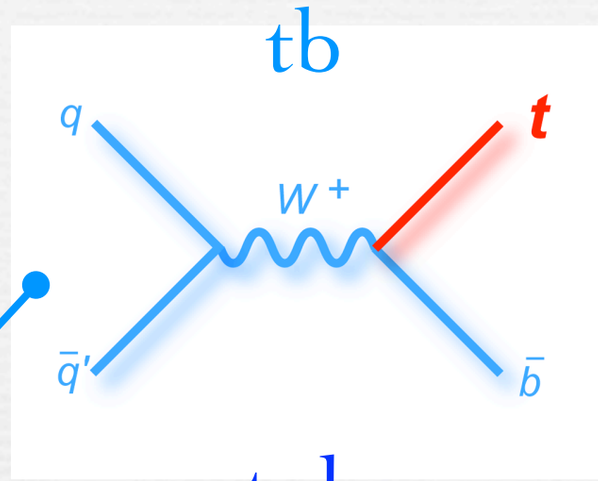


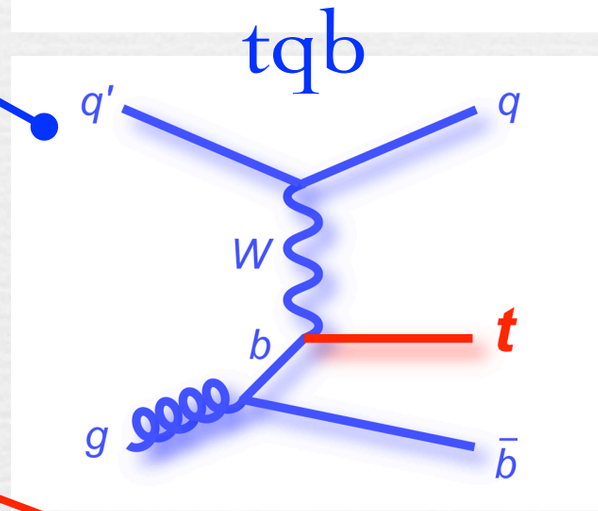


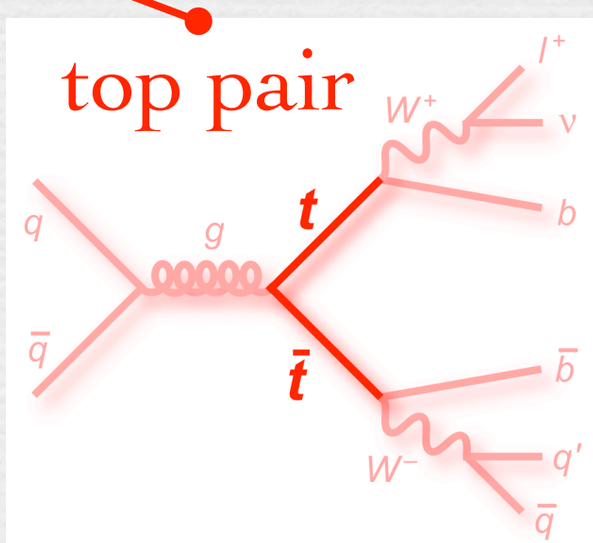


Event yields in 9.7/fb DØ data	
e,μ, 2,3,4-jets 1,2-tags combined	
s-channel	257 ± 31
t-channel	378 ± 53
W+jets	7394 ± 401
Z+jet, dibosons	815 ± 71
ttbar	2672 ± 284
Multijets	789 ± 81
Total prediction	11669 ± 503
Data	12103

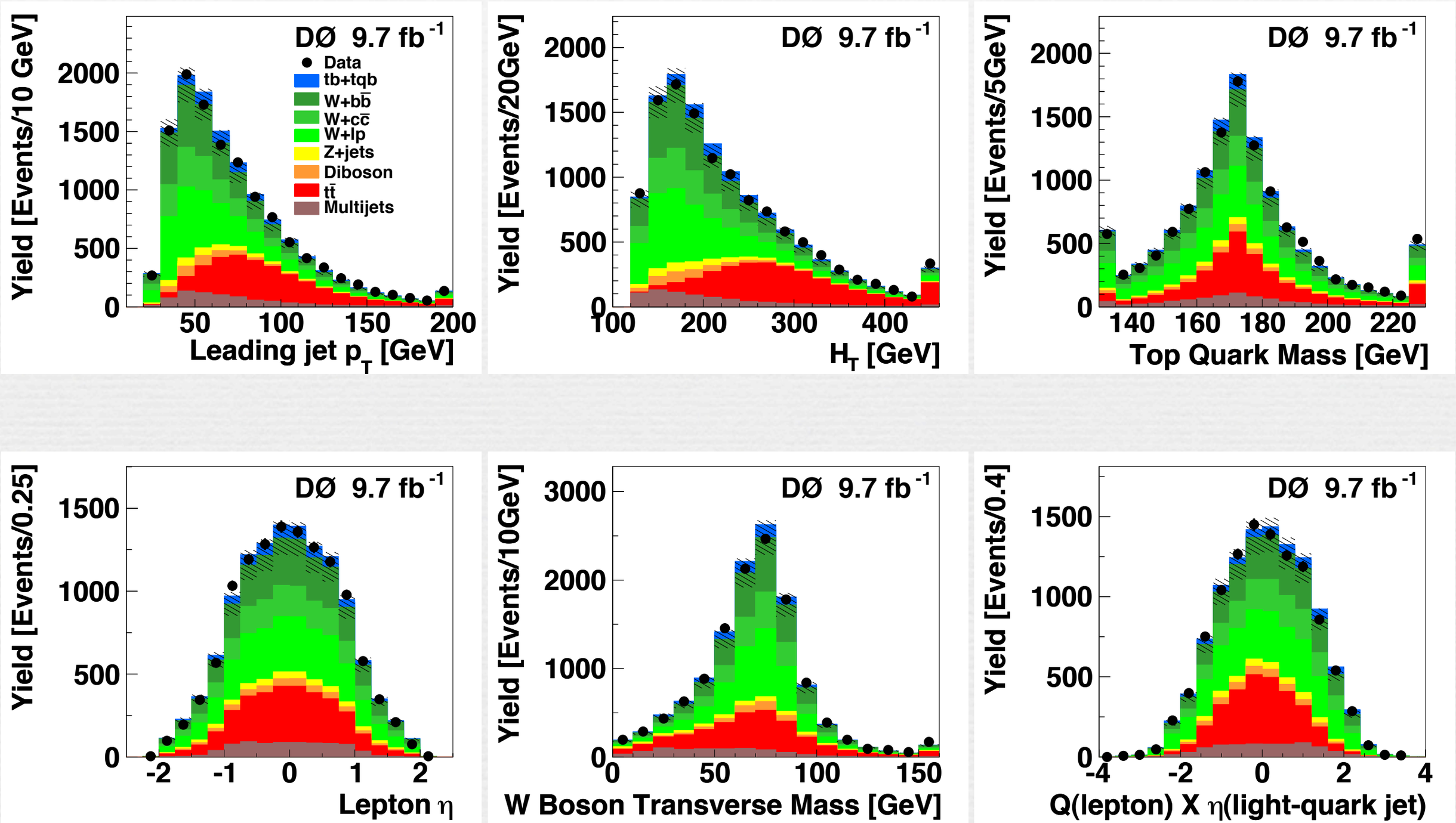
tb: tqb: B = 1: 1.5: 45







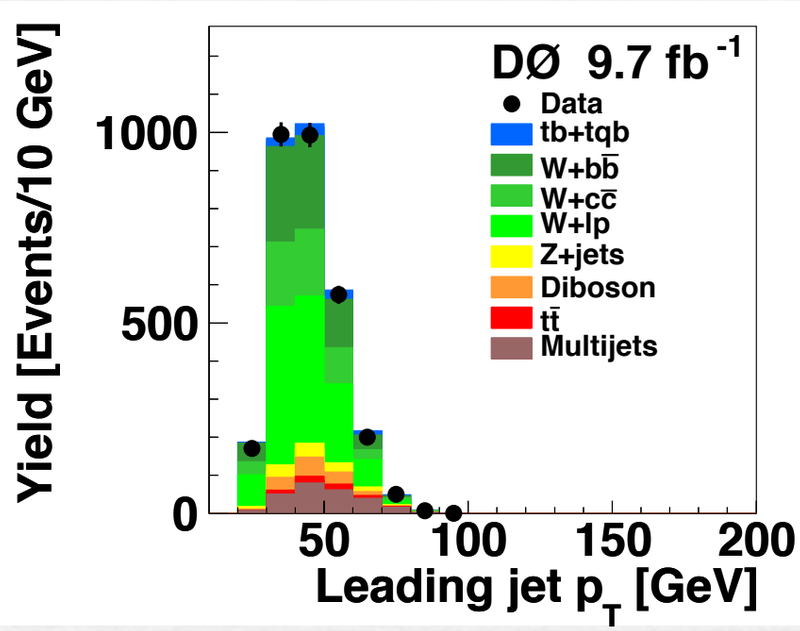
Data Background Comparison

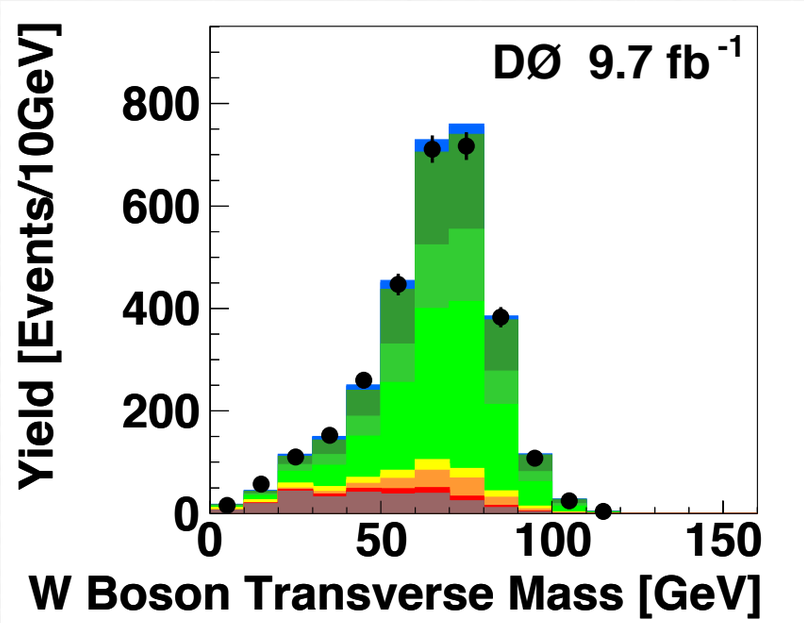


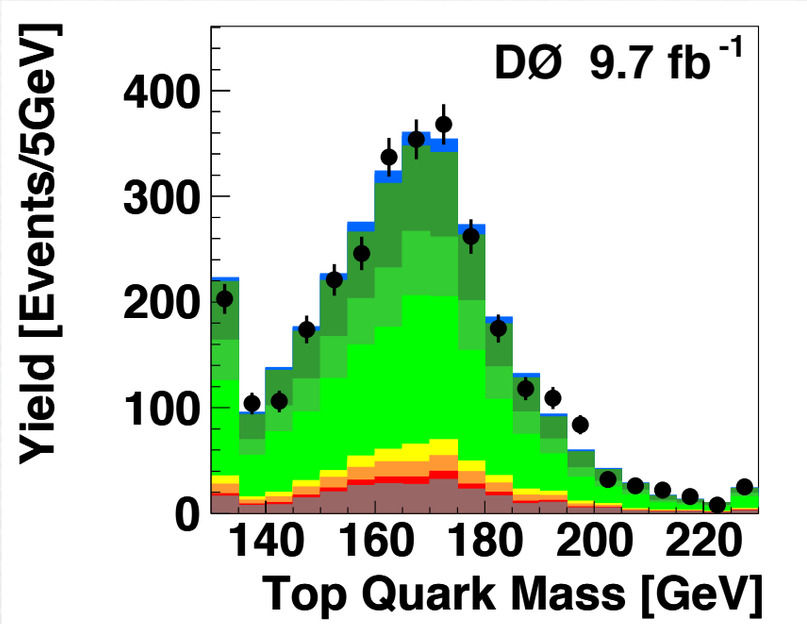
Good data-background agreement, but no clear signal excess...

Cross-Check Samples

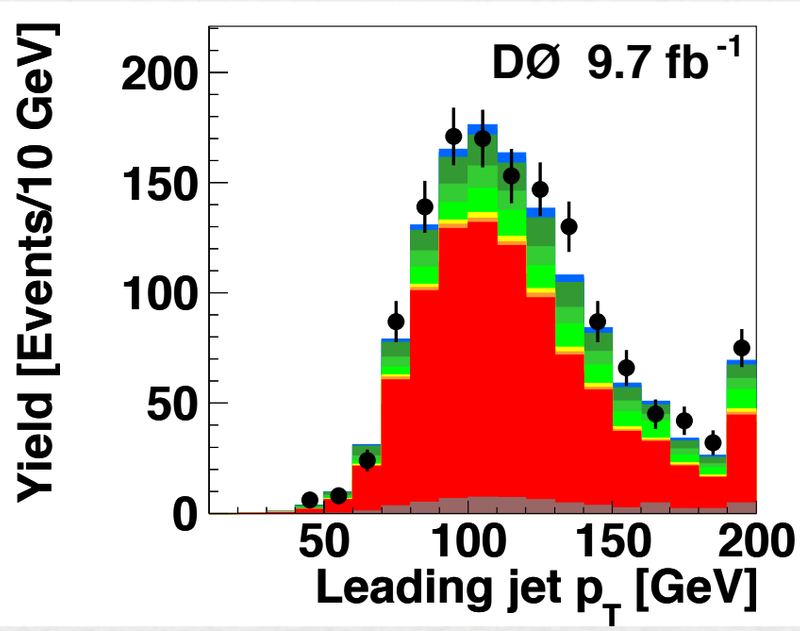
"W+jets" enriched : 2 jets, 1 btag, H_T < 175 GeV

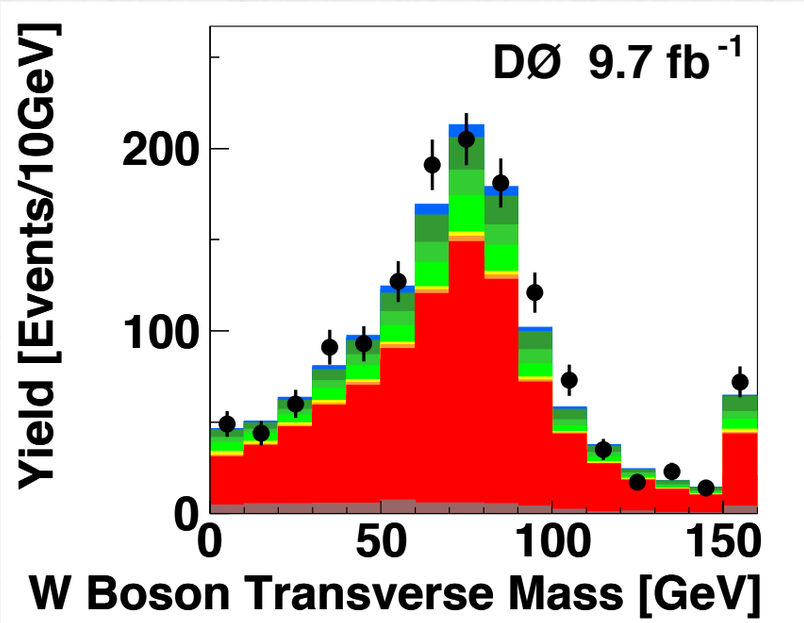


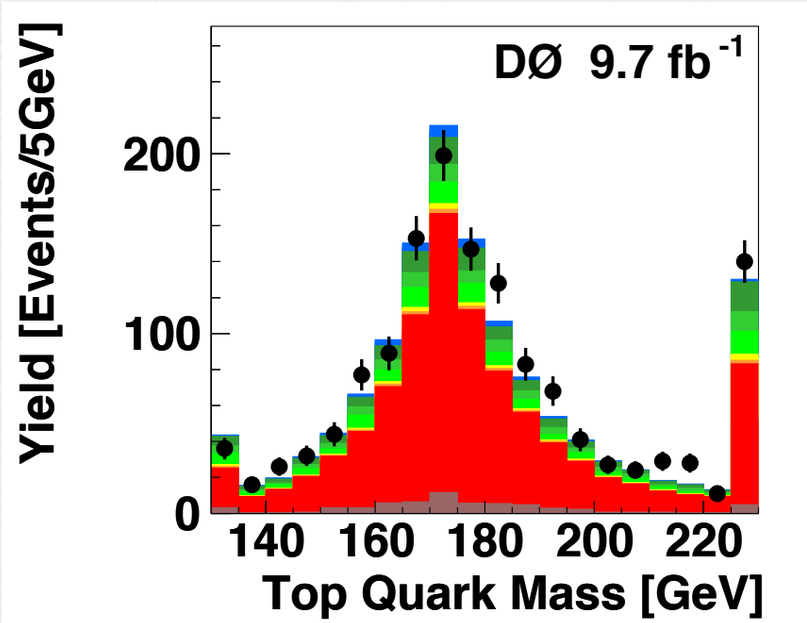




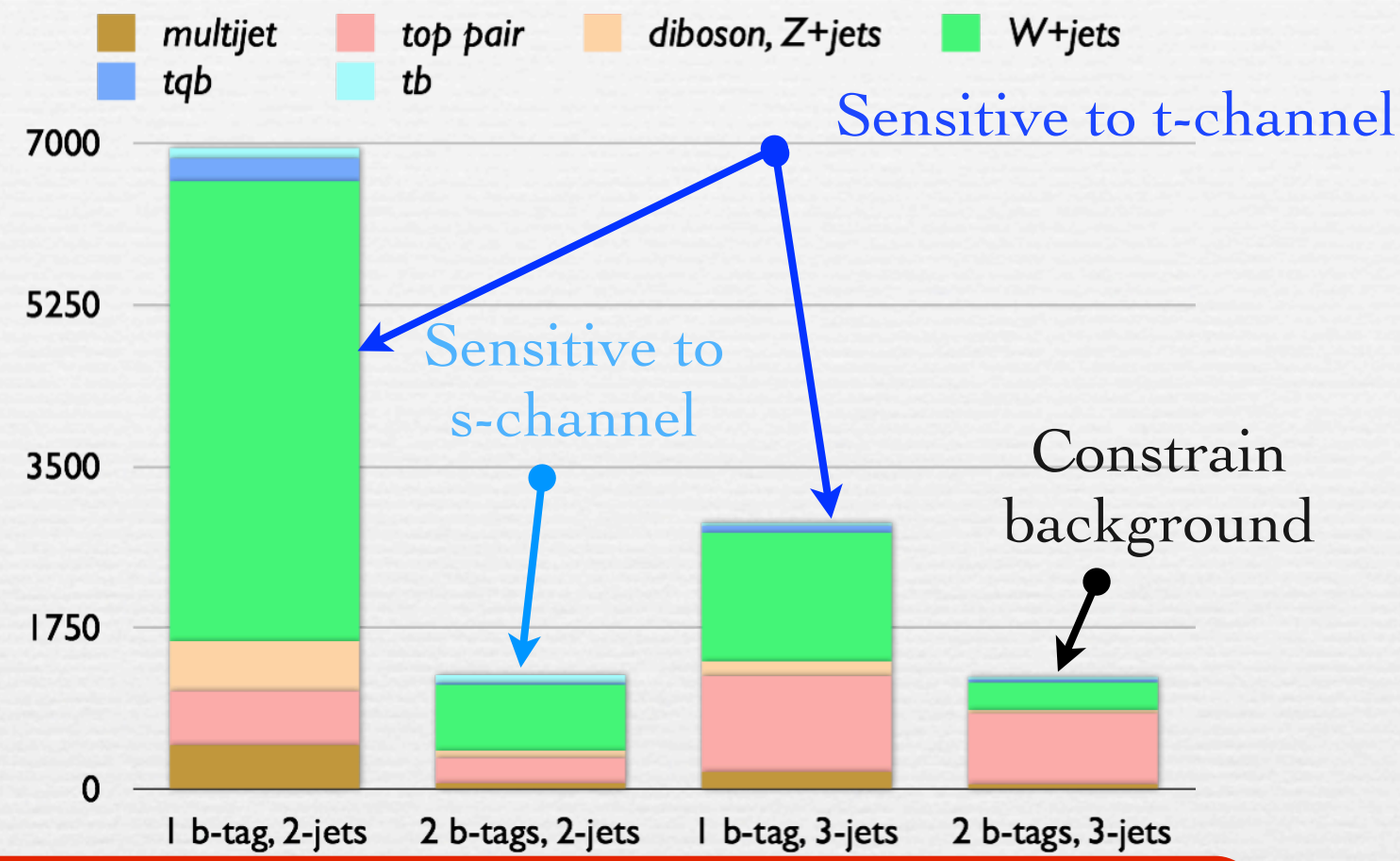
"ttbar" enriched : 3 jets, at-least 1 btag, H_T > 300 GeV







Not A Counting Experiment !

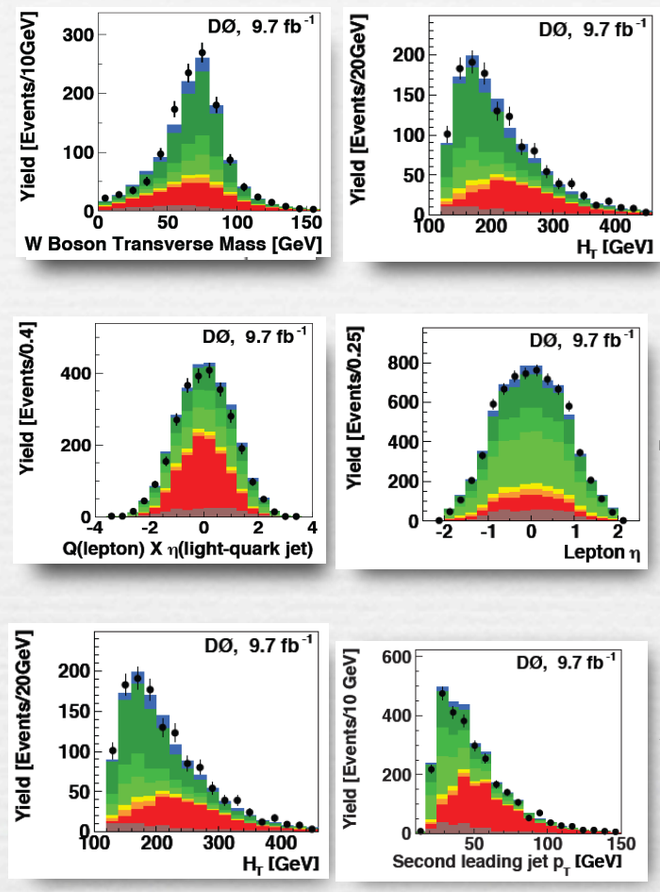


tb: tqb: B	1: 2: 59	1: 0.3: 14	1: 2: 85	1: 1: 40

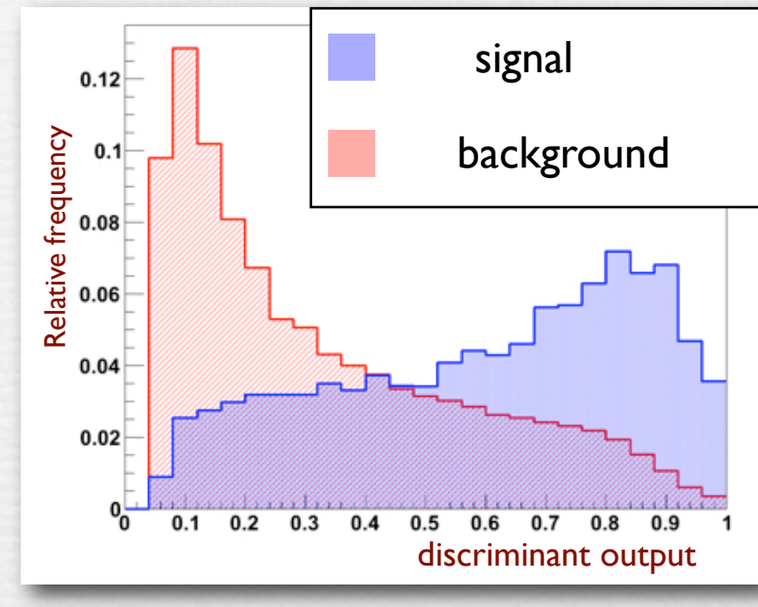
- The amount of signal is less than the uncertainty on the background.
- All the backgrounds ttbar, multijets, W+jets mimics signal signature very closely.
- Need a variable to separate the signals and backgrounds. Not Such single variable !
- Not feasible to perform a counting experiment. Need multivariate techniques.

Multivariate Analysis

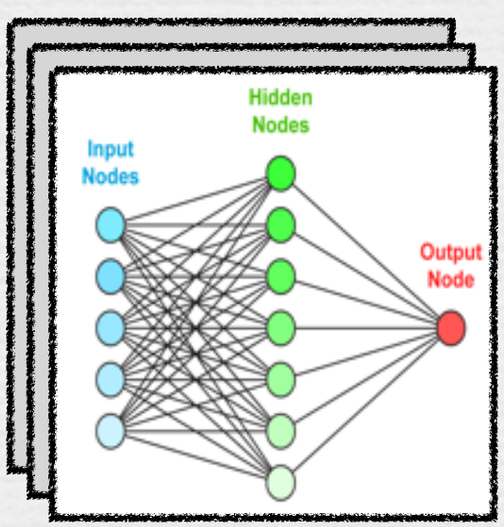
Combine different kinematic variables with some discrimination power into one variable with larger discrimination.



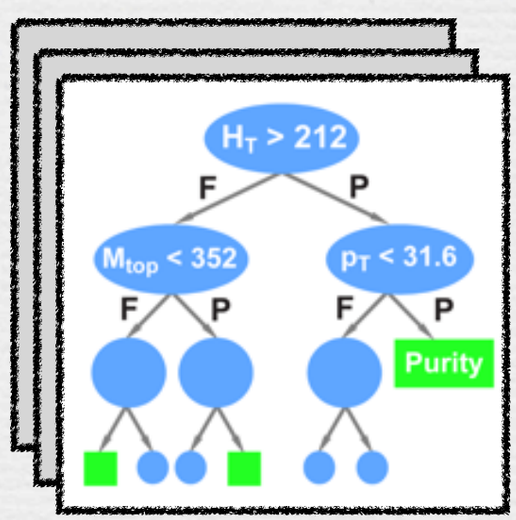
After training



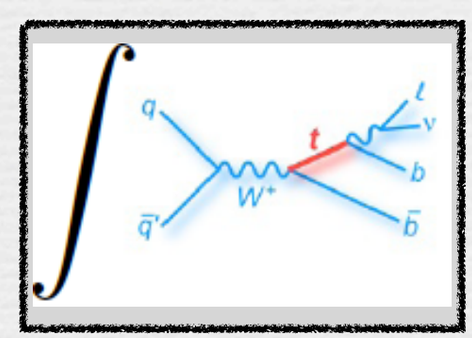
Bayesian Neural Networks (BNN)



Boosted Decision Tree (BDT)

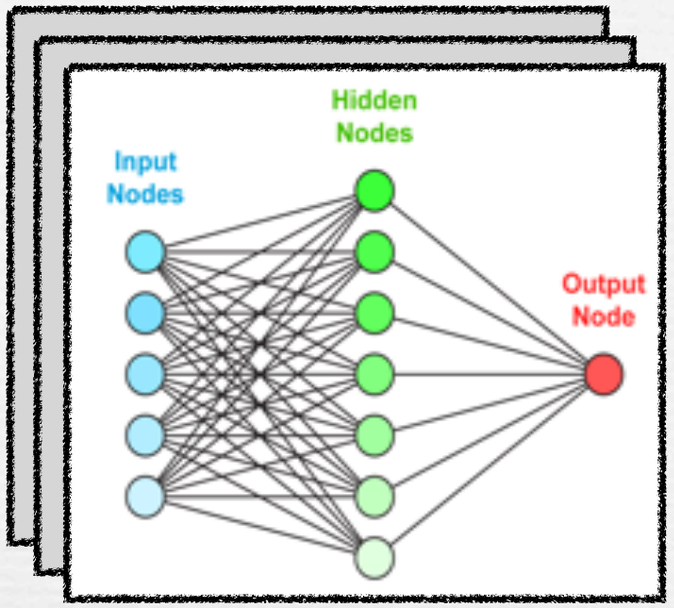


Matrix Element Method (ME)



Bayesian Neural Networks

From the structural point of view, **Neural Network (NN)** is an interconnected group of nodes. It can be used to find complex relationships between inputs and outputs, or to find patterns in data.



The NNs we used have:

N_{var} Input nodes
H Hidden nodes
1 Output node -> **gives the prob. for an event to be signal**

From the mathematical point of view, NN is a **non-linear function** n(x,w) which approximated the discriminant D(x).

- $D(x) = Prob(sig|x) = f(x|sig) / (f(x|sig) + f(x|bkg))$; x = vector of input variables
- $n(x,w) = 1 / (1 + exp.(-g(x,w)))$ **where**
  $g(x,w) = b + \sum_{j=1 \text{ to } H} v_j \tanh(a_j + \sum_{i=1 \text{ to } Nvar} u_{ji} x_i)$; $w=(u_{ji}, a_j, v_j, b)$ are NN weights

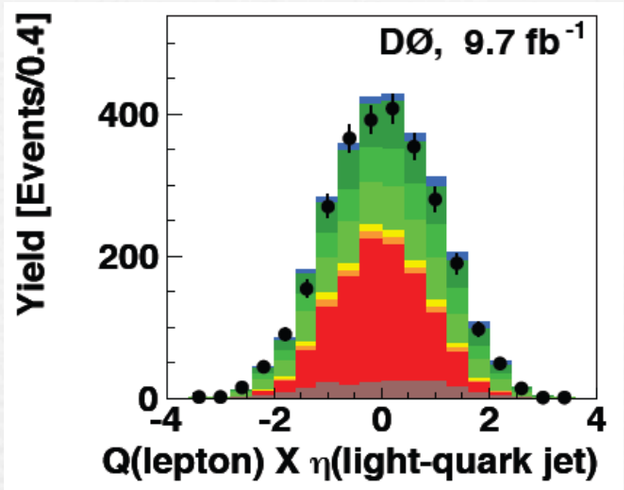
A Bayesian Neural Network is an average over the $(1 + H(2+N_{var}))$-dim parameter space w

- **Ideally,** $n(x) = \int_{\{w\}} n(x,w)\, p(w|T)\, dw$; T = set of training data
- **In practice,** $n(x) = 1/100 \sum_{k=201 \text{ to } 300} n(x,w_k(T))$; 300 NNs are iteratively trained, average over the last 100

BNN - Training and Validation

Variables Used : In order to decrease the complexity and make it less correlated with other methods, tried to optimize the lesser number of variables. We used :

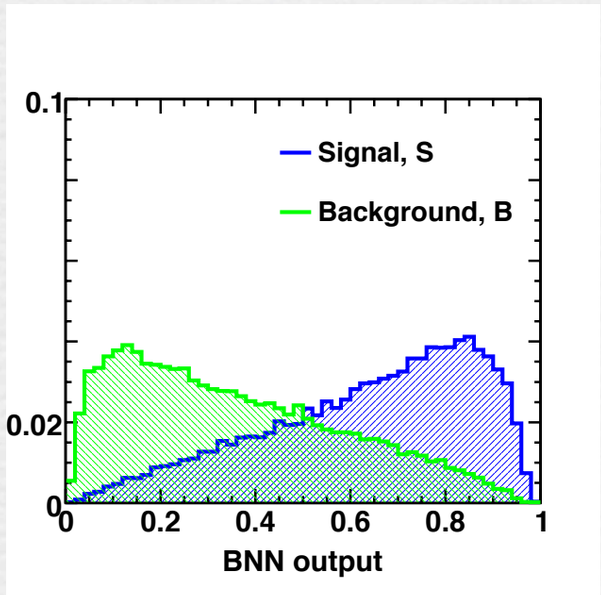
- lepton and jets 4-vectors
- missing transverse energy 2-vector
- charge information
- btagging information

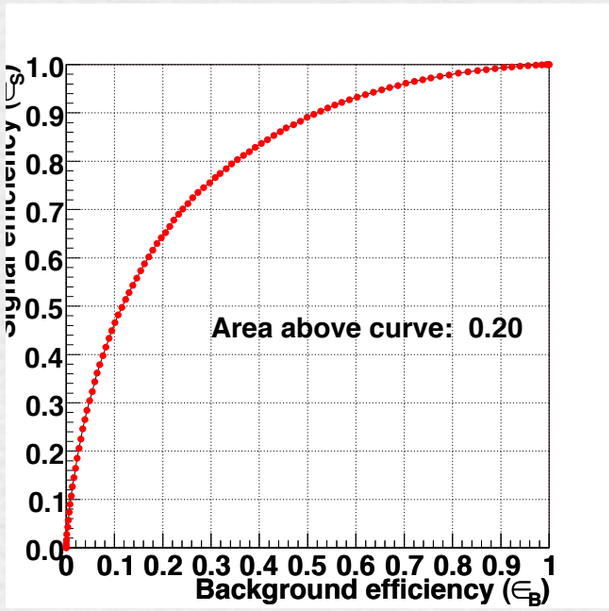


Sample is divided into three independent sub-samples - training, testing and yield

Training : For each analysis channel, a training set 'T' with 10,000 signal and 10,000 background random events is built from "training" sub-sample.

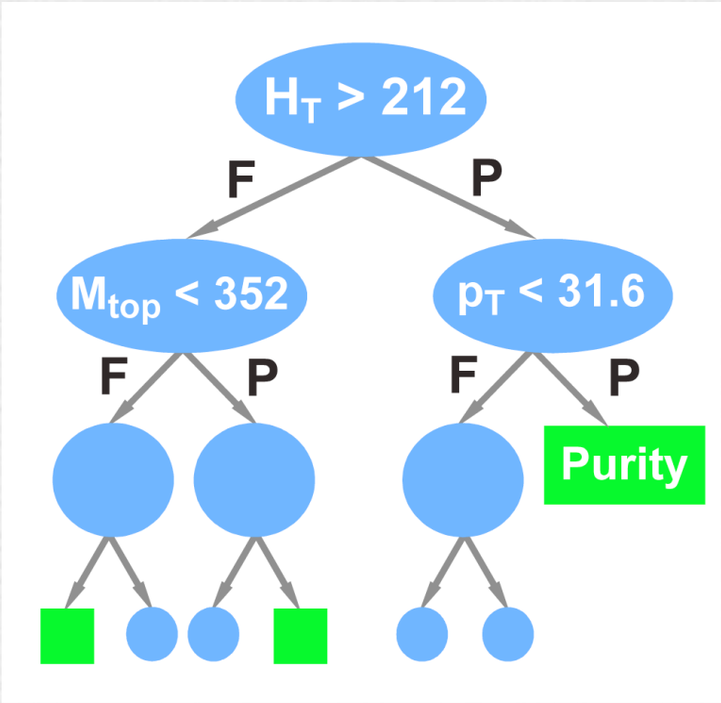
Trained separately for **s-channel** **and** **t-channel**

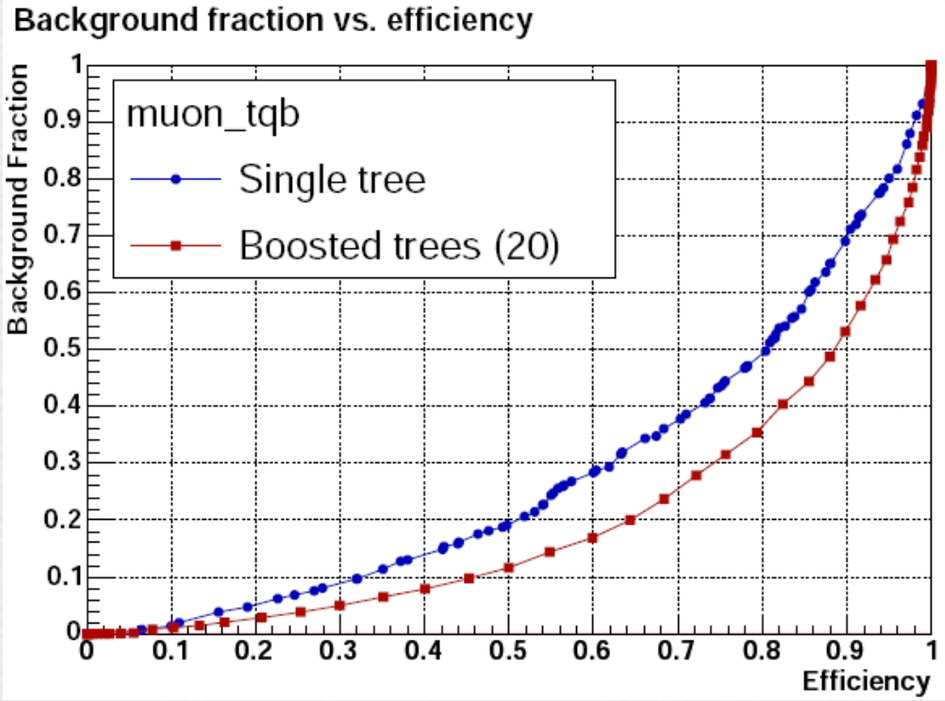




Boosted Decision Trees

- Apply sequential cuts but keep the failing the events
  - List of 30 variables optimized for s-channel (KS > 0.25)
  - Trained separately for **s-channel** **and** **t-channel**
- Train another tree produced by enhancing misclassified events
- Boosting improves the performance.





#	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	$\text{Centrality}(\text{alljets})$
27	$M\text{jet1}, \text{jet2})$
28	$p_T(\text{jet1}, \text{jet2})$
29	$M_T(W)$
30	$p_T(W)$

Matrix Element Method

- The idea is to use all available kinematic information from a fully differential cross section measurement.
- Calculate an event probability for signal and background hypothesis

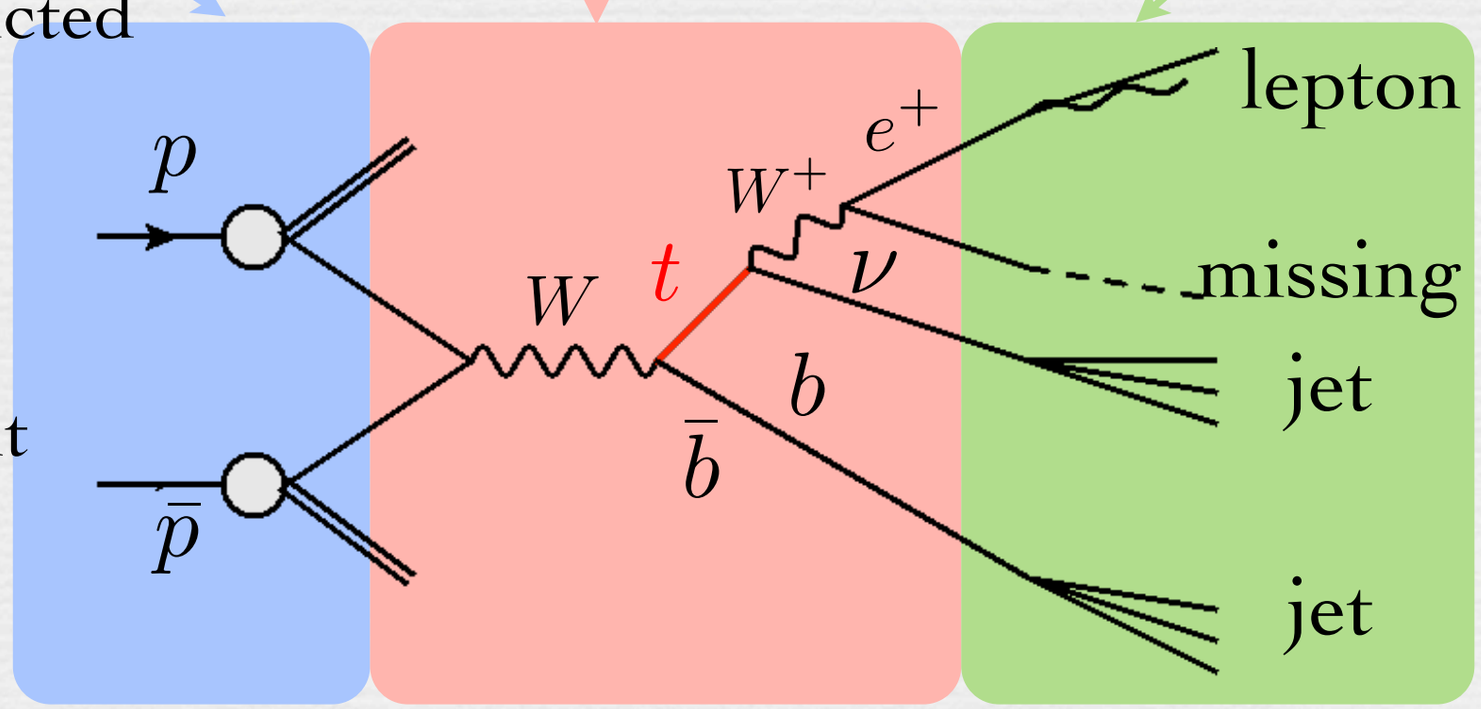
$$P(\vec{x}) = \frac{1}{\sigma^{obs}} \sum_{x,y} \sum_{i,j} \int f_i(q_1) dq_1 f_j(q_2) dq_2 \frac{\partial \sigma_{hs,ij}(\vec{y})}{\partial \vec{y}} W(\vec{x}, \vec{y}) d\vec{y}$$

CTEQ6L PDF

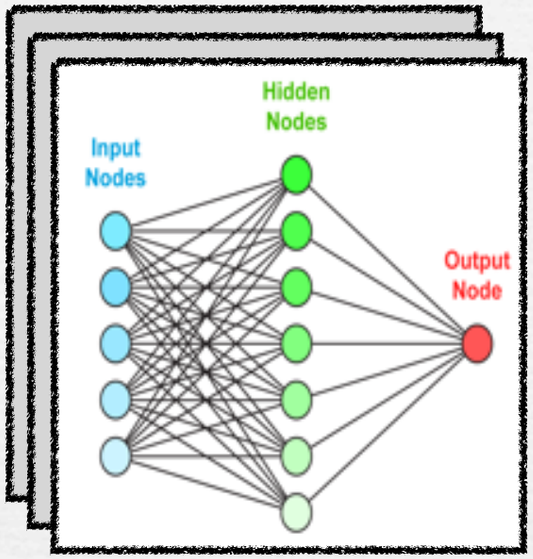
LO Matrix Element

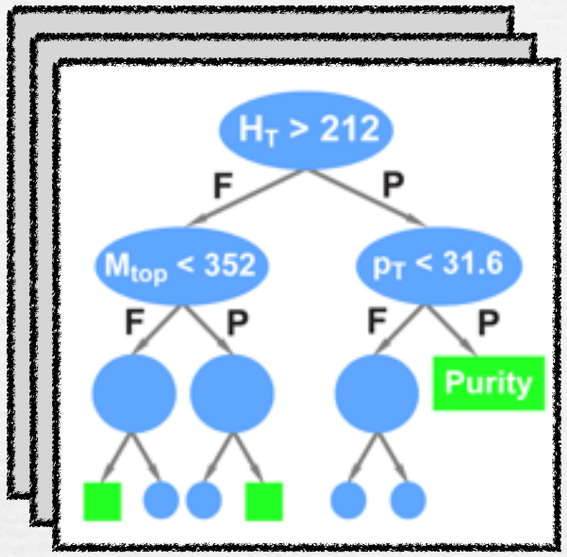
Transfer Function : maps parton level (y) to reconstructed variables (x)

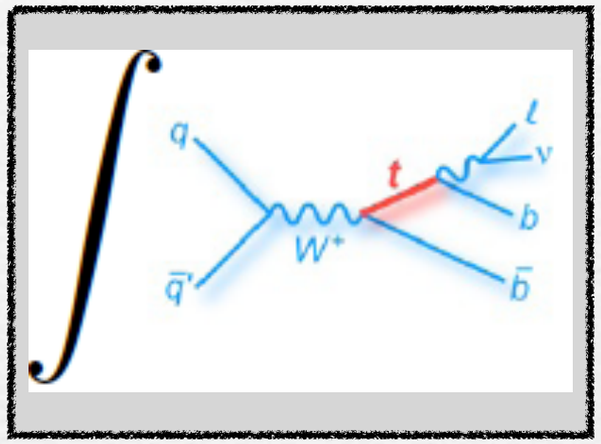
- Uses 4-vectors of reconstructed lepton and jets.
- jet-parton assignment use: btag information.
- Efficiently returns full event information

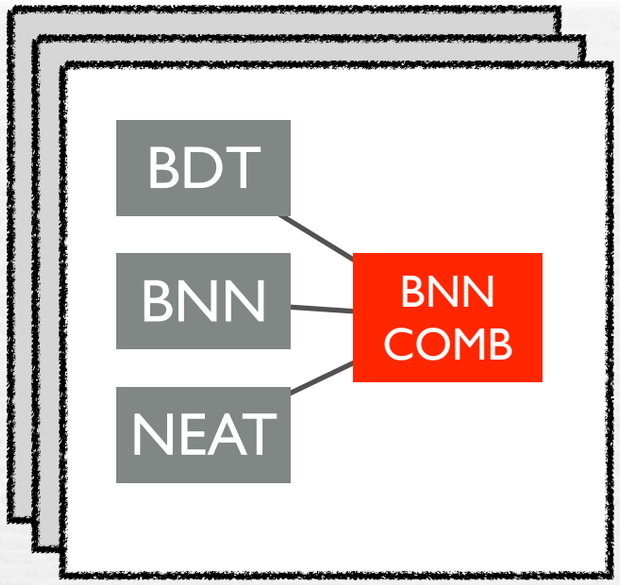


Combination Analysis

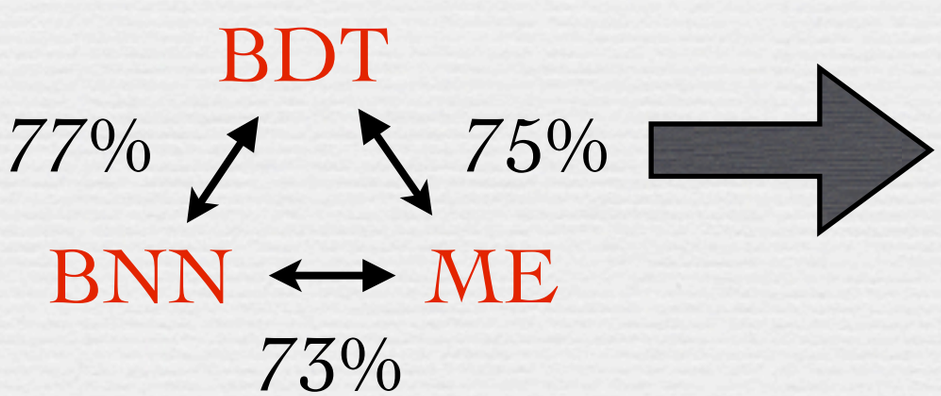








Correlations

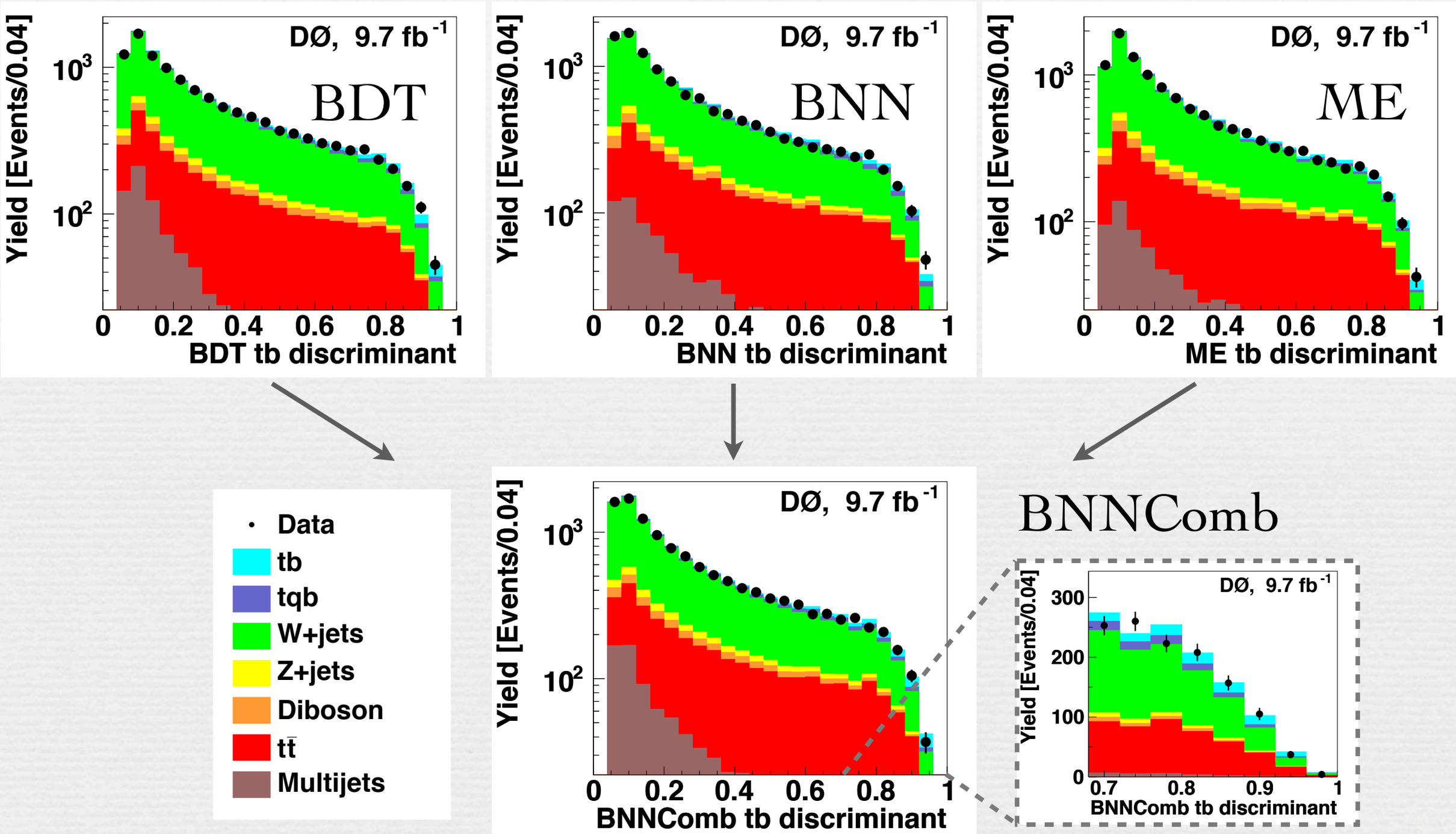


Increase sensitivity by combining

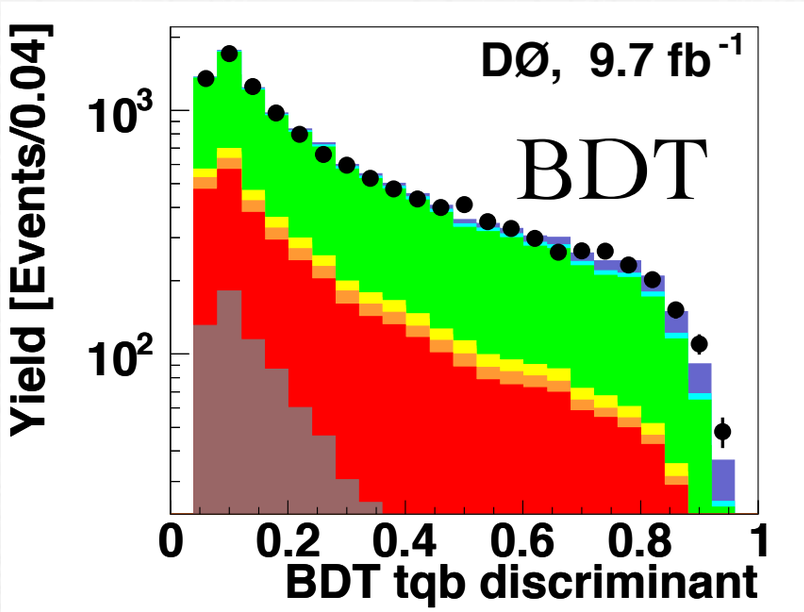
Training :

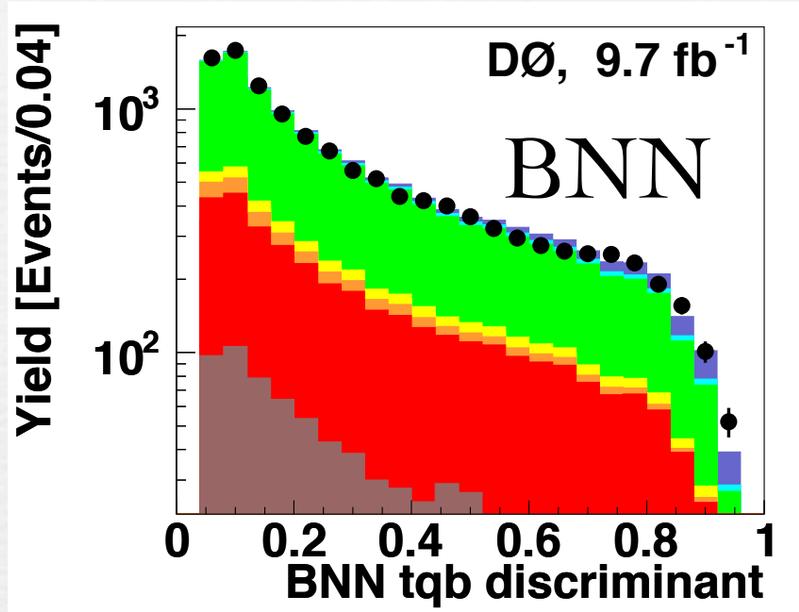
Training for BNN Combination is done on independent 'testing' sub-sample.

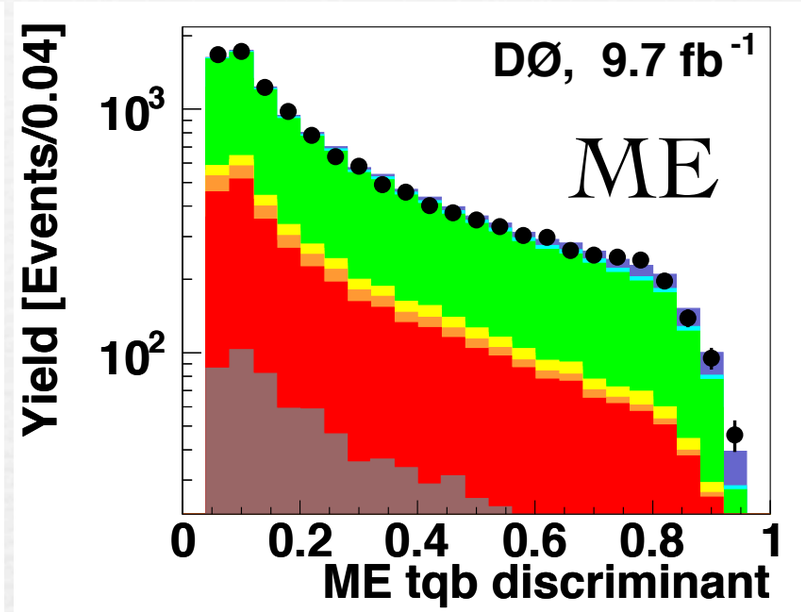
BNN tb Combination

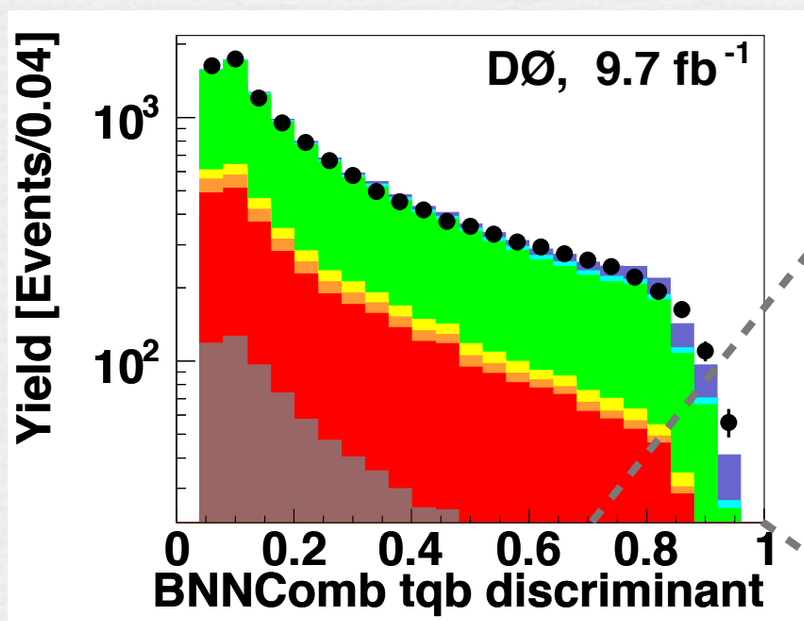


BNN tqb Combination

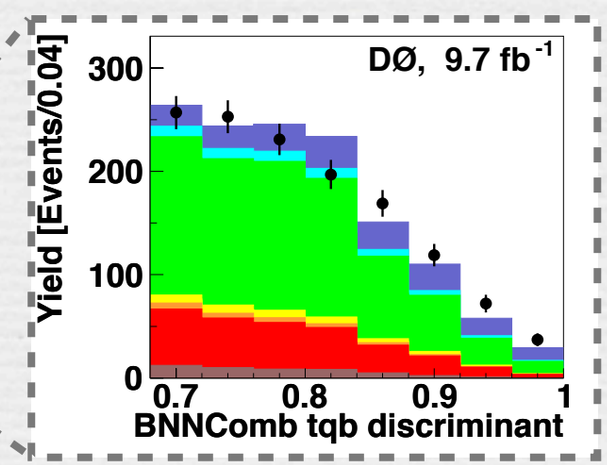






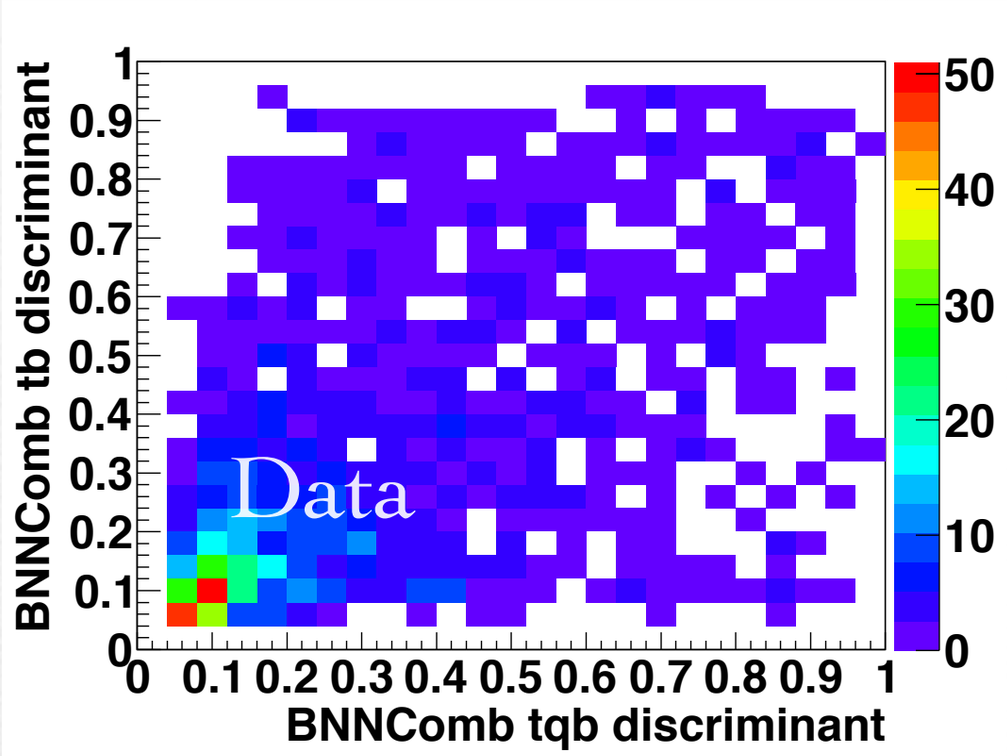


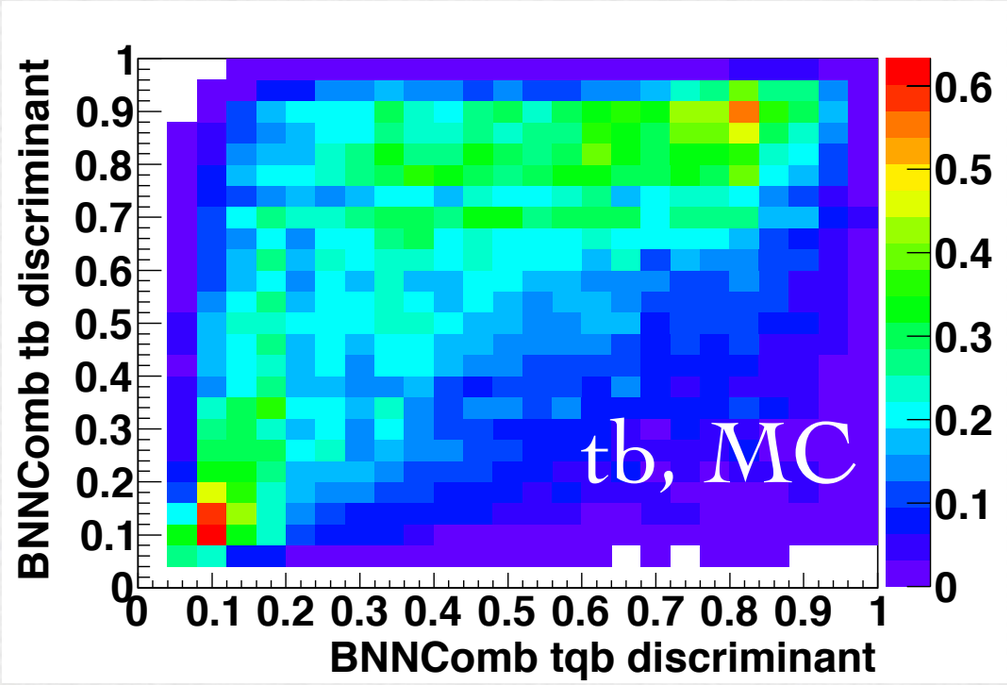
BNNComb



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 containing enough statistics to have a stable measurement
- Split every event based on whether $D_{tb} > D_{tqb}$ or $D_{tqb} > D_{tb}$

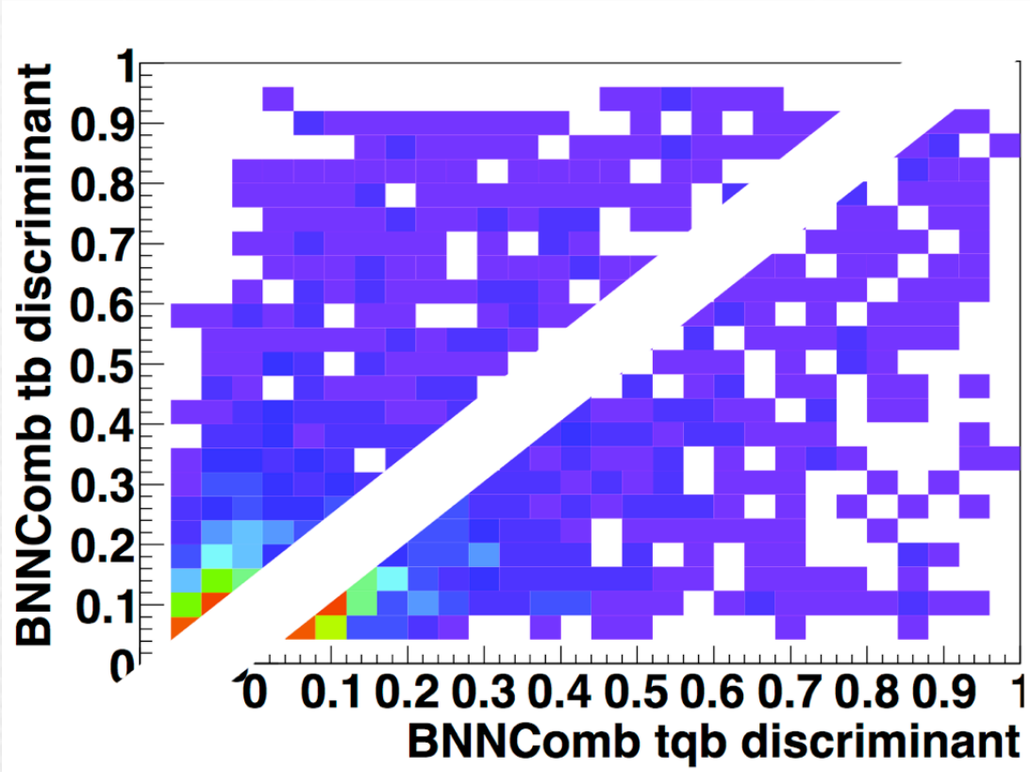




New Discriminant

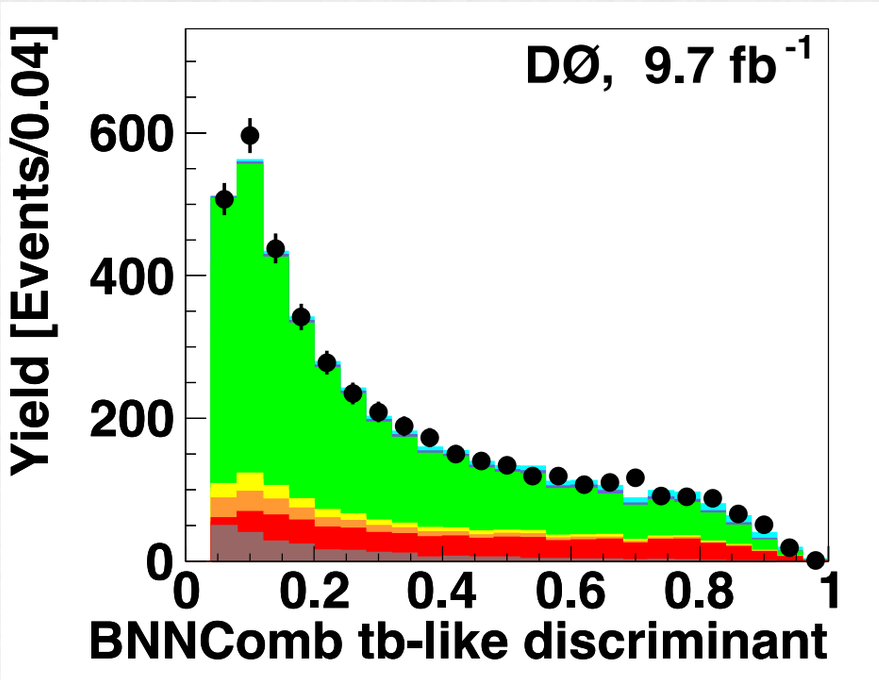
If $D_{tb} > D_{tqb}$:

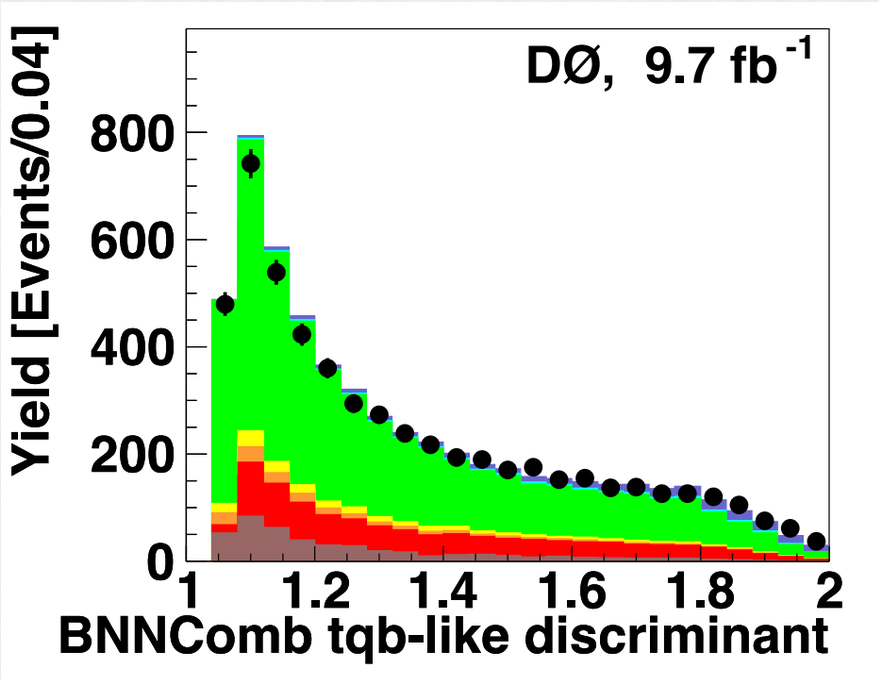
- tb category
- Use D_{tb}
- Plot in the range [0, 1]



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

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





Systematics 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 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)%		

Cross Section Measurement

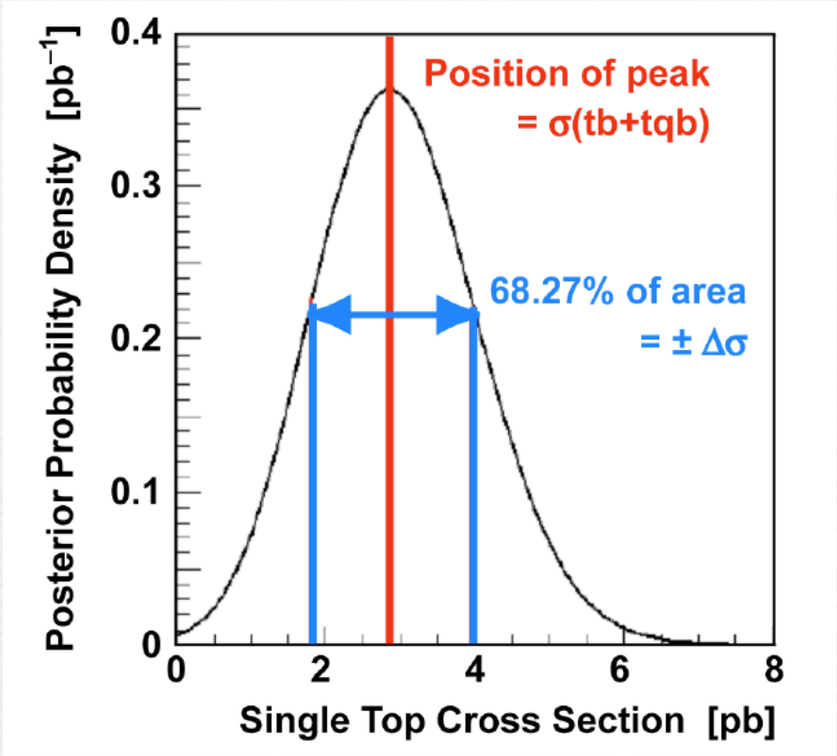
- Cross sections are measured by building a Bayesian posterior probability density.
  - Binned likelihood.
  - No cut on the discriminant.

- For each analysis, the single top cross section is given by the position of the posterior density peak, with 68% asymmetric interval as uncertainty.

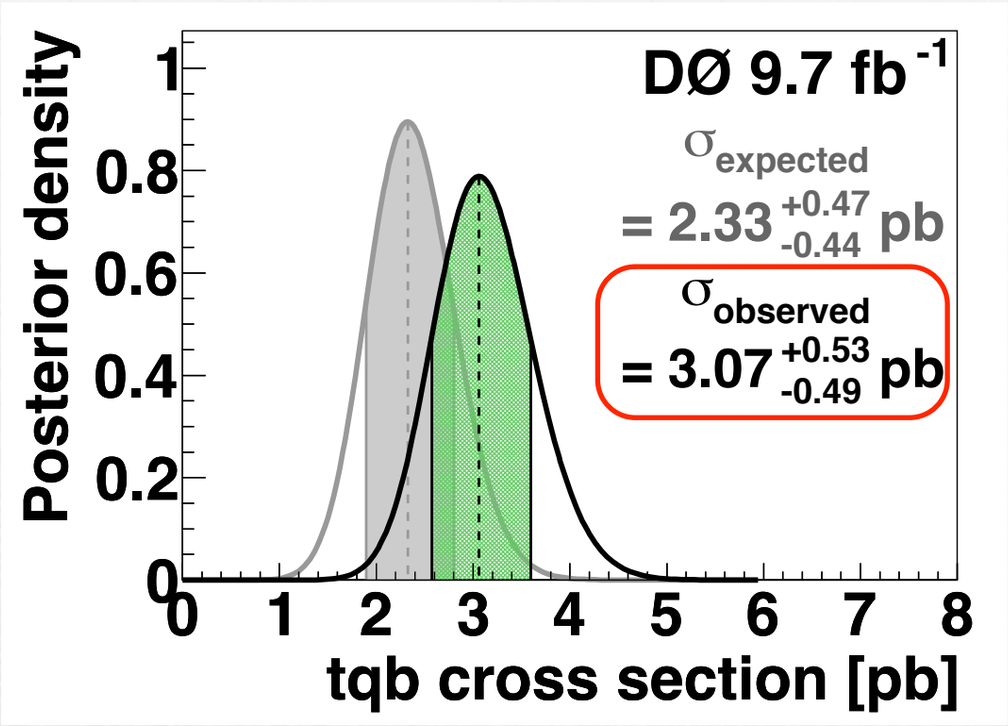
- Gaussian prior for systematic uncertainties.

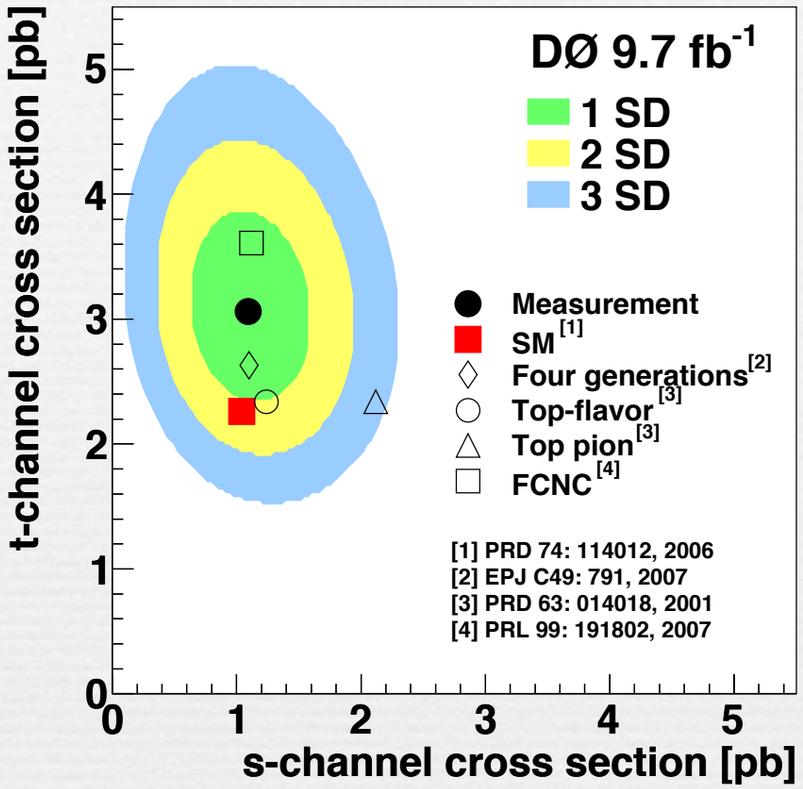
- Correlations of uncertainties are properly taken into account.

- Flat, non-negative prior in signal cross sections.

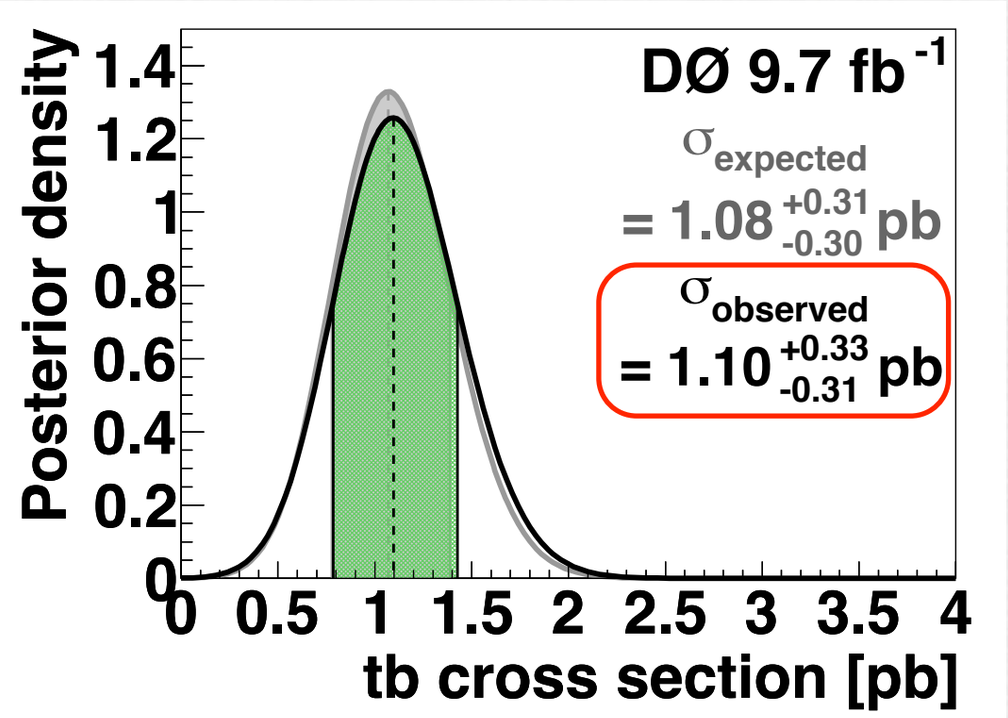


2D Posterior & Measured Cross Section



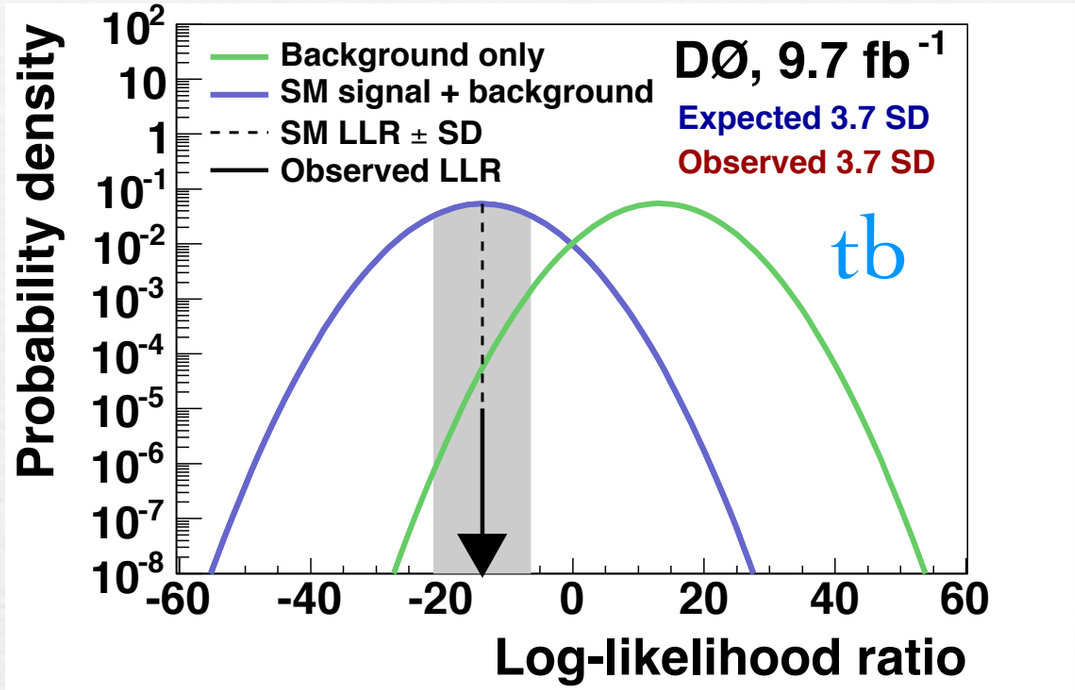


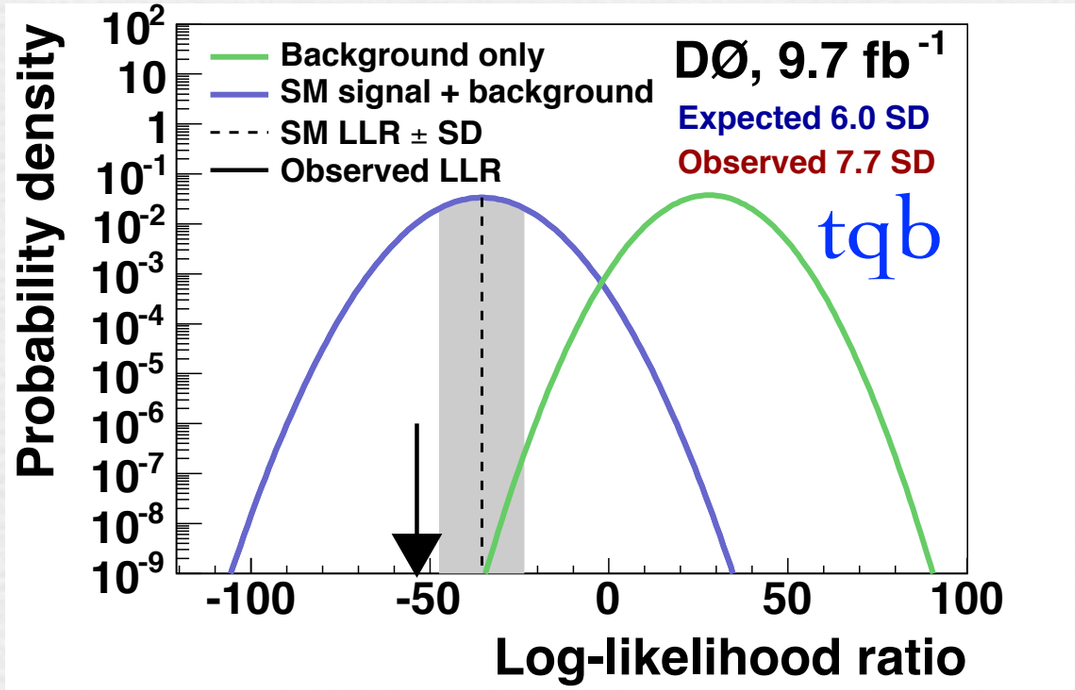
No assumption on SM σ_{tb}/σ_{tqb}!



Significance

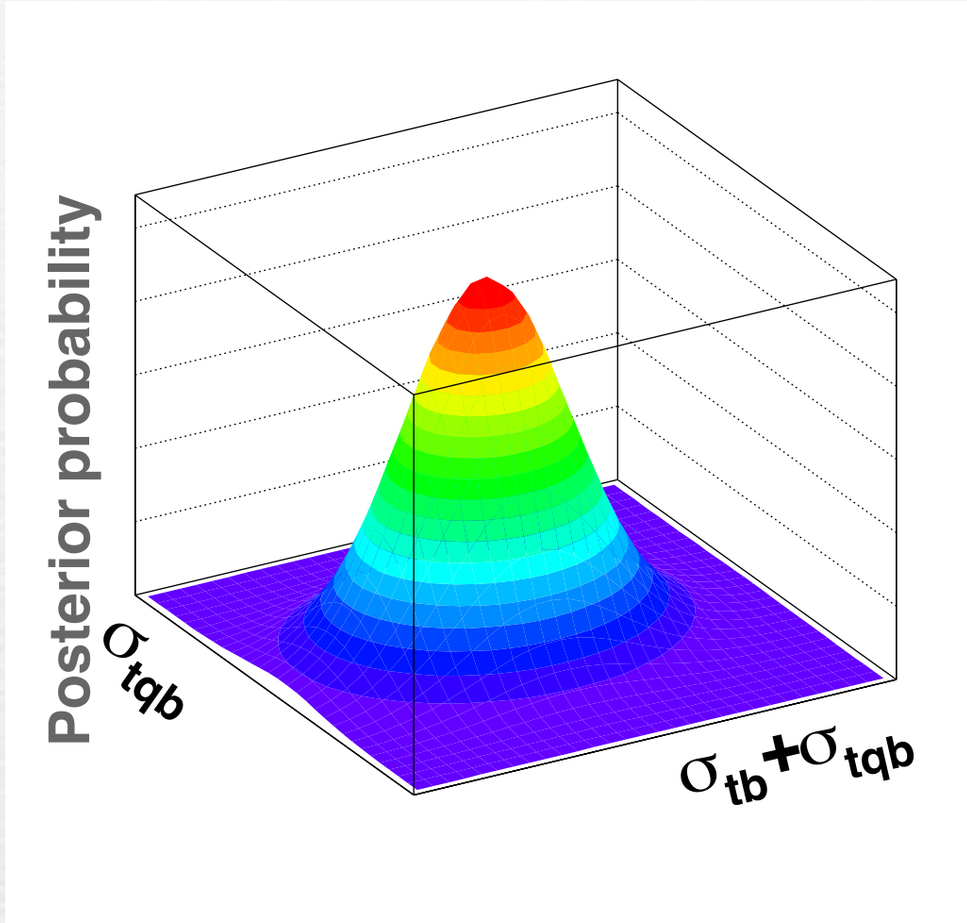
- Asymptotic approximation of the log-likelihood ratio
- Tests how likely the data is to fluctuate to the measured σ value, in the absence of the signals
- Expected p-values:
  - tb: 1.0×10^{-4} (3.7 SD)
  - tqb: 9.9×10^{-10} (6.0 SD)
- Observed p-values:
  - tb: 1.0×10^{-4} (3.7 SD)
  - tqb: 6.1×10^{-15}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

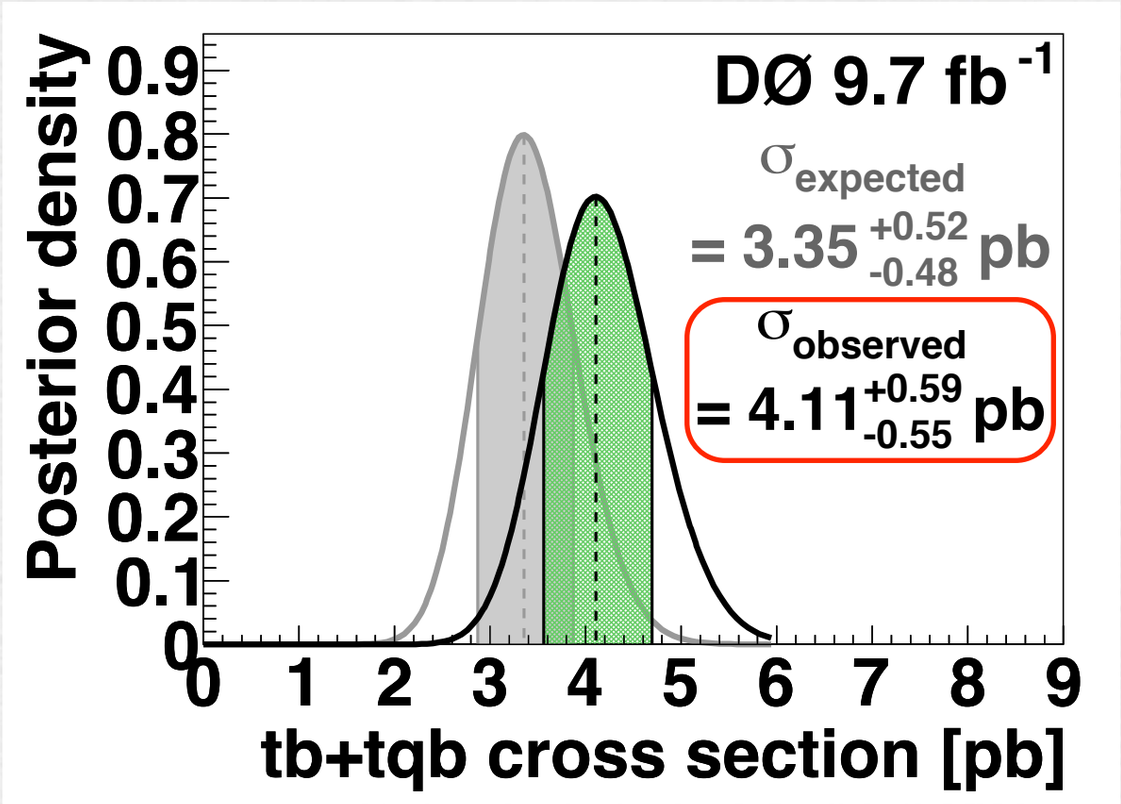




Single top total tb+tqb Cross Section

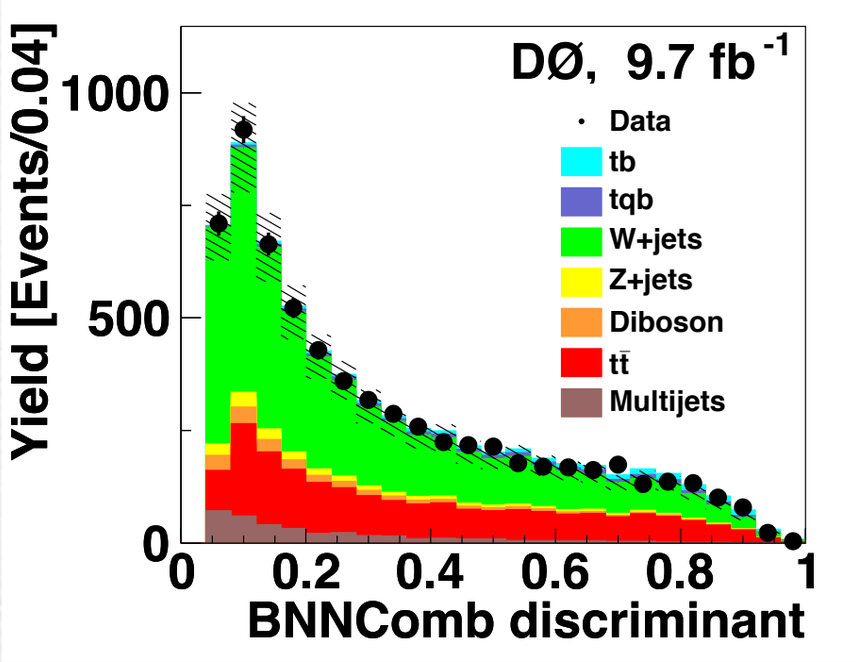
- Measure σ_{tb+tqb} without assuming the SM σ_{tb}/σ_{tqb}
- Use 2D posterior p.d.f.
- Integrate over σ_{tqb} and obtain a 1D p.d.f of σ_{tb+tqb}

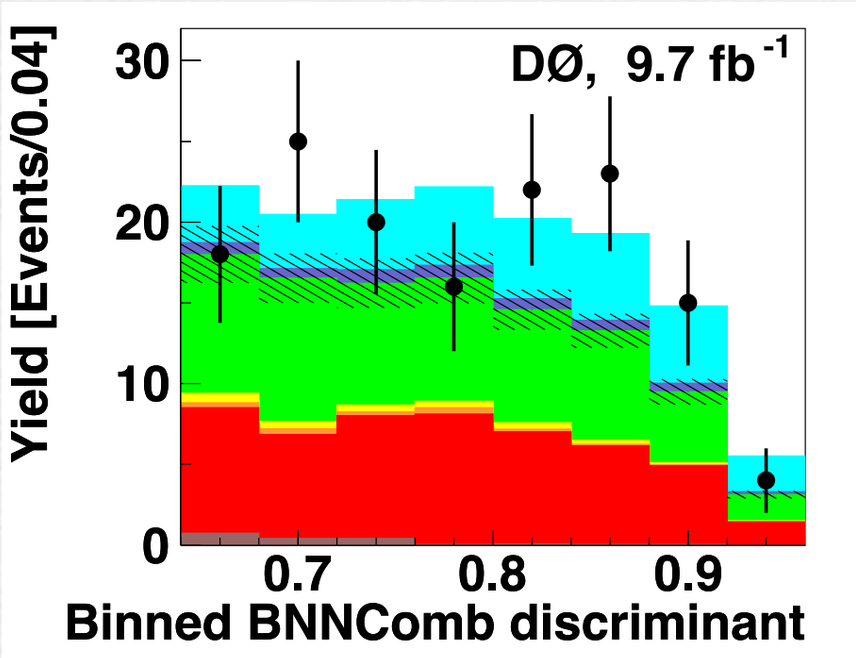


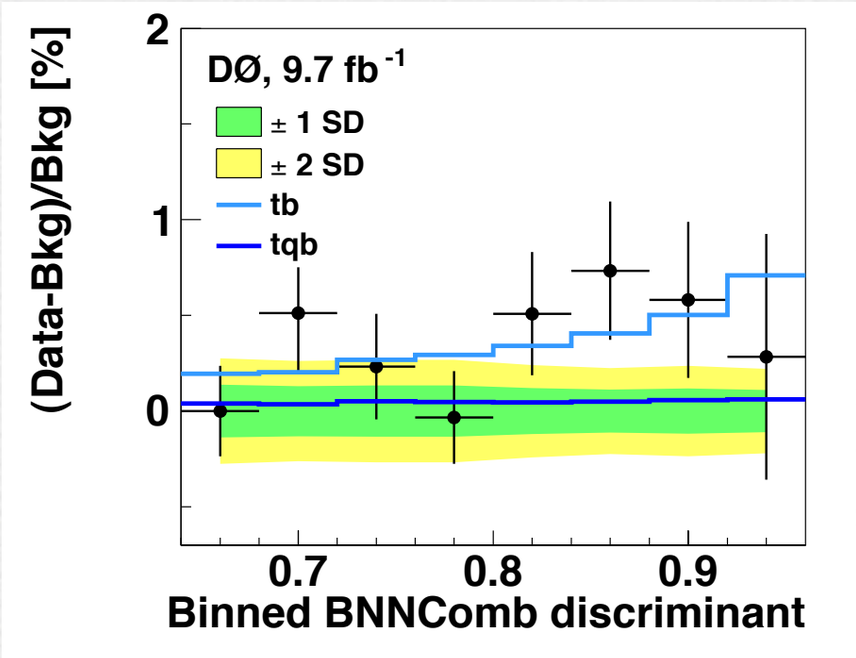


Data favors the presence of tb

- BNNComb discriminant with the post-fit uncertainty
- Examine the most sensitive bins (largest S/B)
- Data favors the presence of tb !

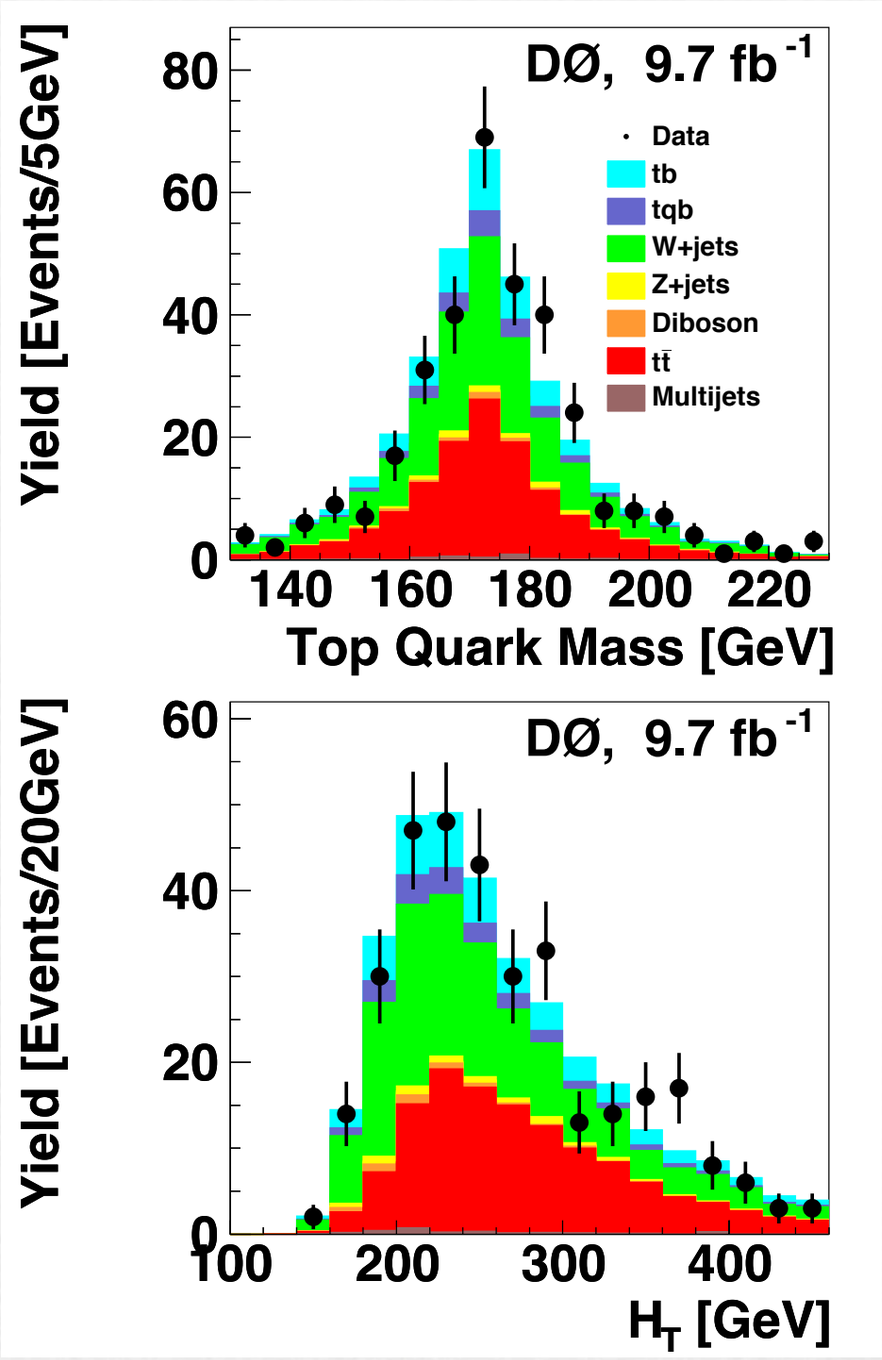




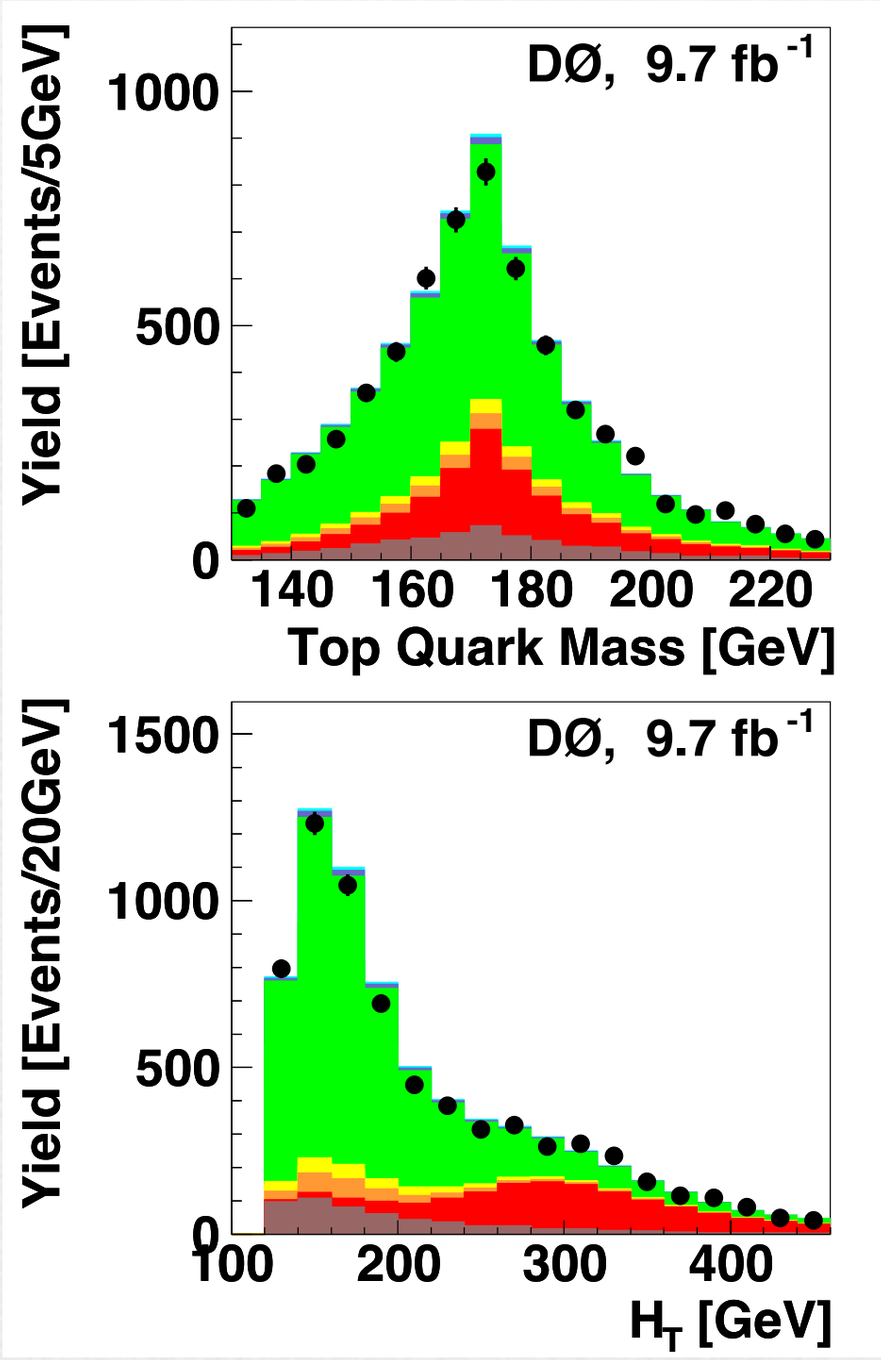


tb Event Characteristics

tb Category: $D_{tb} > 0.8$

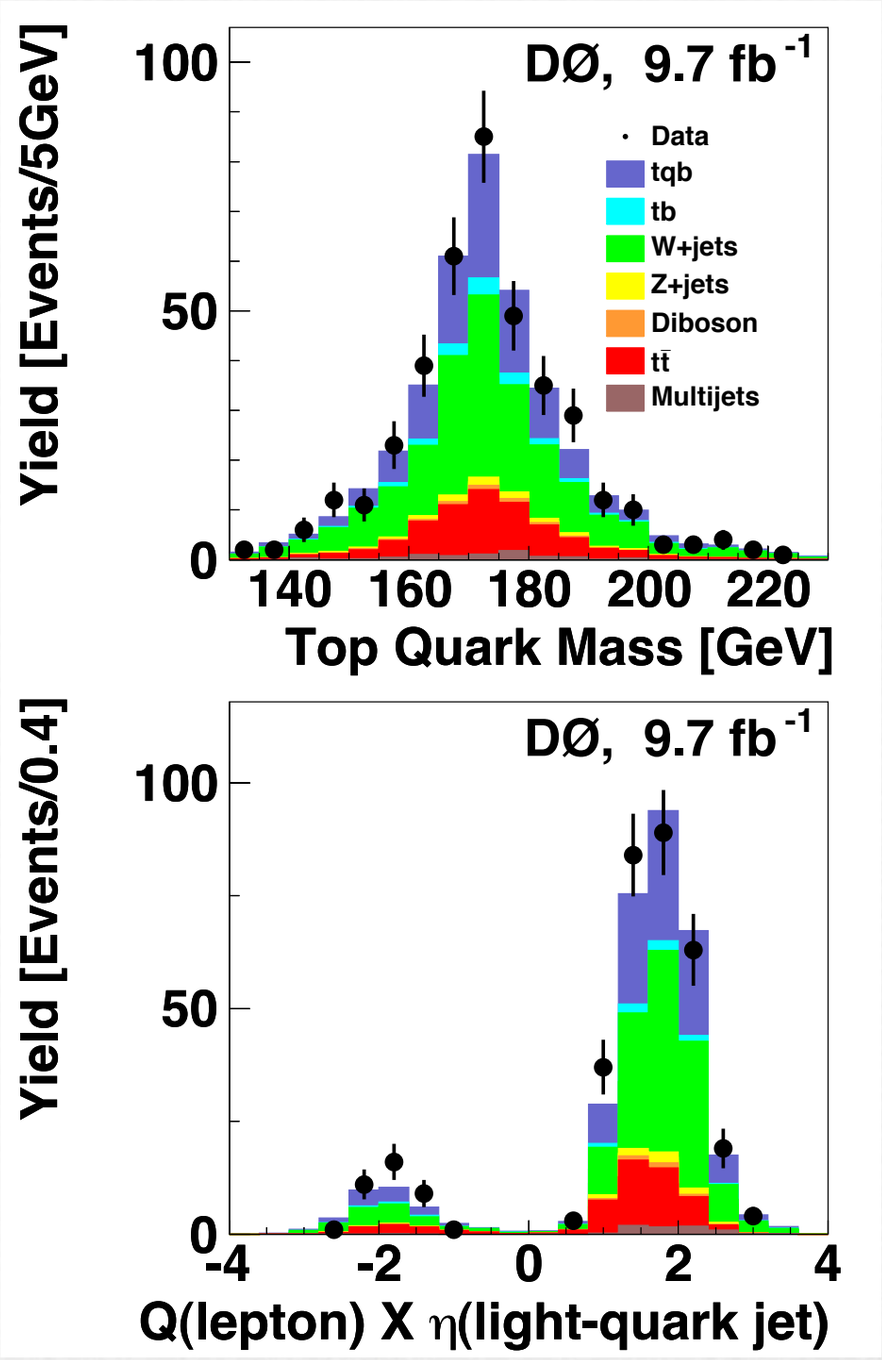


tb & tqb Depleted Region

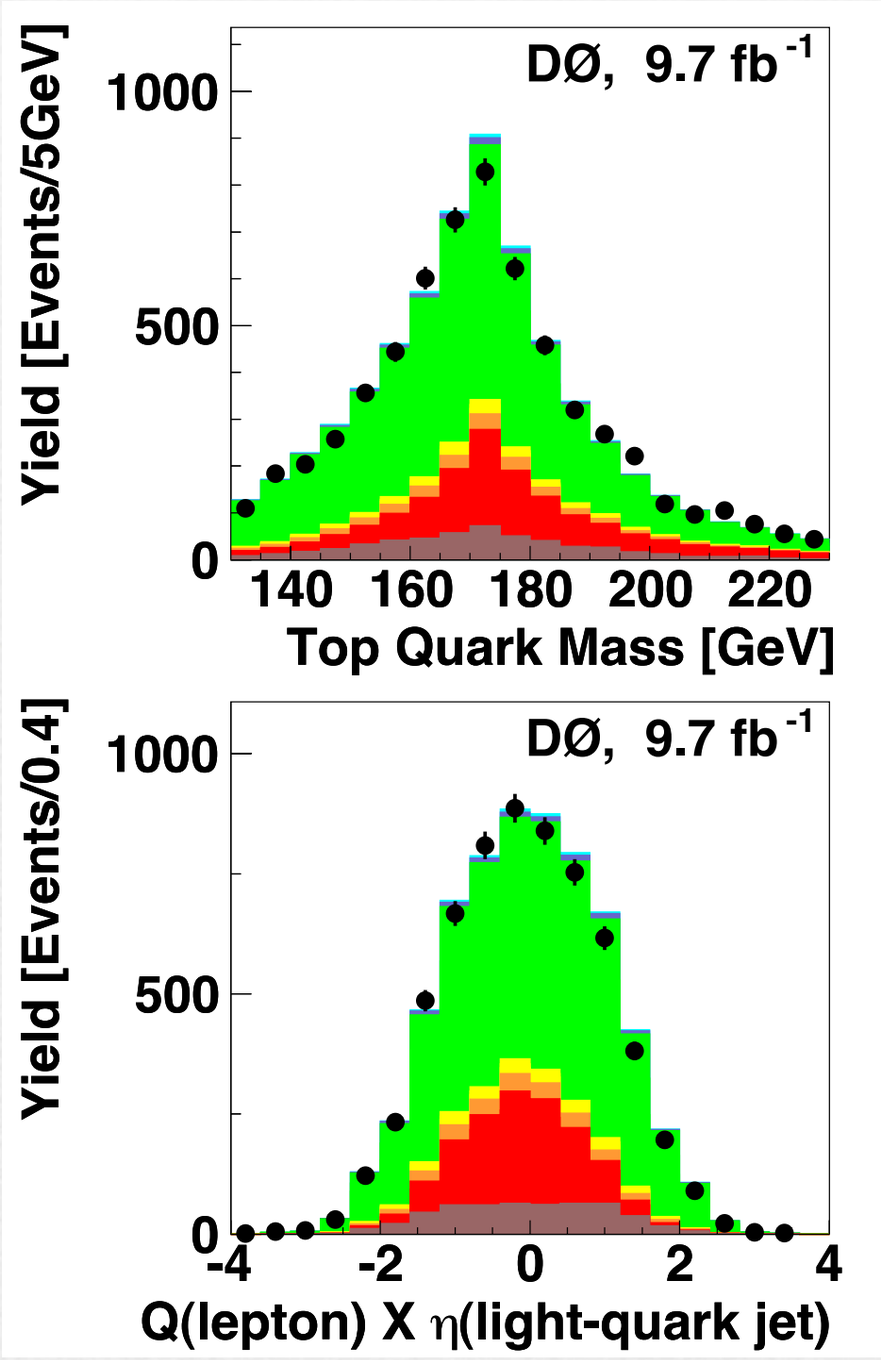


tqb Event Characteristics

tqb Category: $D_{tqb} > 0.8$



tb & tqb Depleted Region



Current Experimental Status

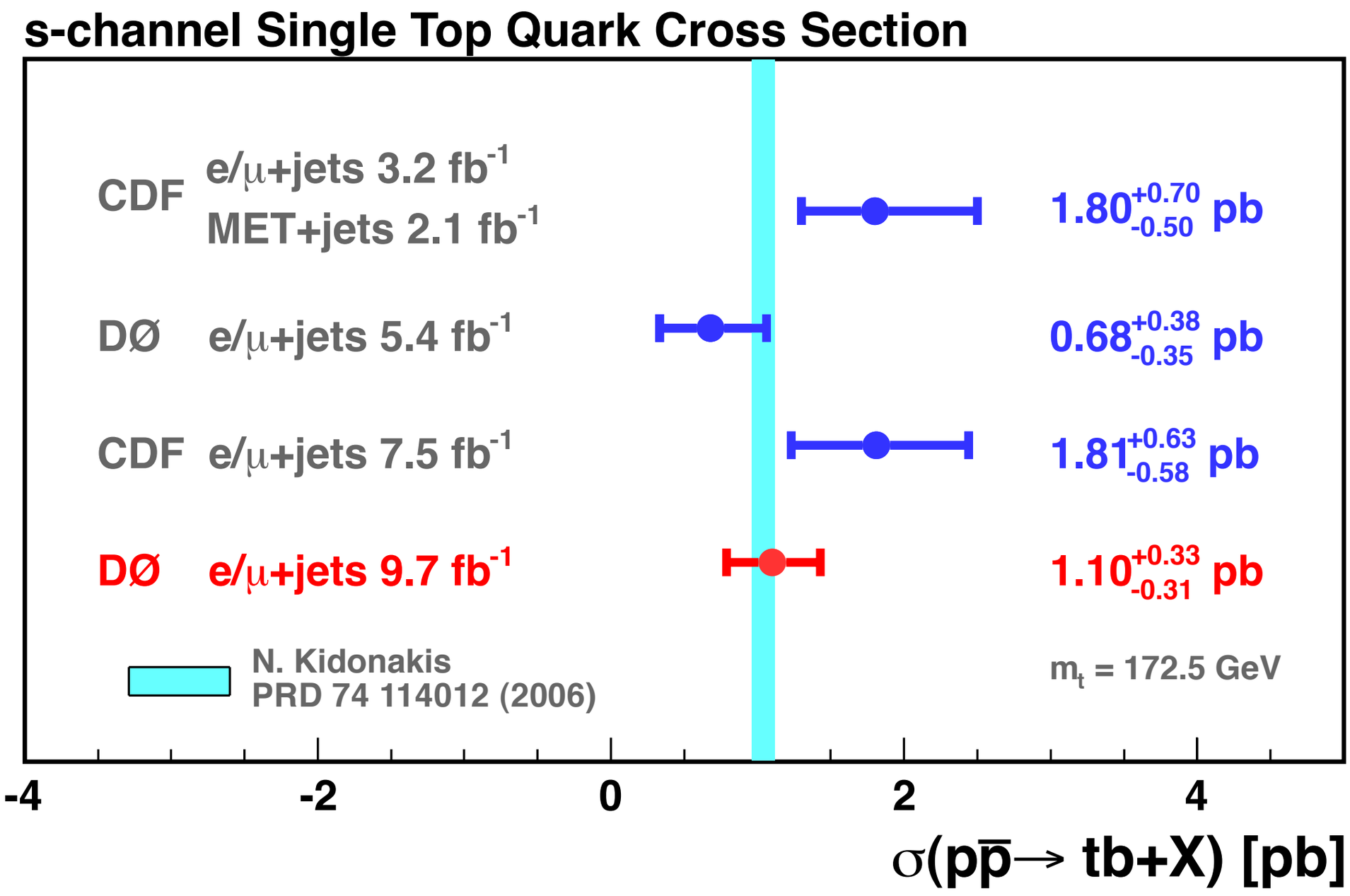
σ (pb) ~ NNLO	tb [pb]	tqb [pb]	tW [pb]
Tevatron Prediction[1] (1.96 TeV)	1.04	2.26	0.3
CDF	1.81 ± 0.6	1.49 ± 0.45	
DØ	1.10 ± 0.33	3.07 ± 0.53	
LHC Prediction[2] (7 TeV)	4.59	64.6	15.7
ATLAS	< 20.5 (95% C.L.)	83 ± 20	17 ± 6
CMS		67 ± 6	16 ± 5

Observation (5 SD)

Evidence (3 SD)

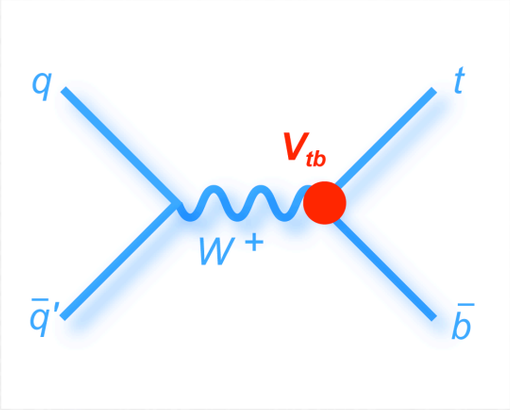
1: PRD 74, 114012 (2006)
2: PRD 81, 054028 (2010)
PRD 83, 091503 (2011)

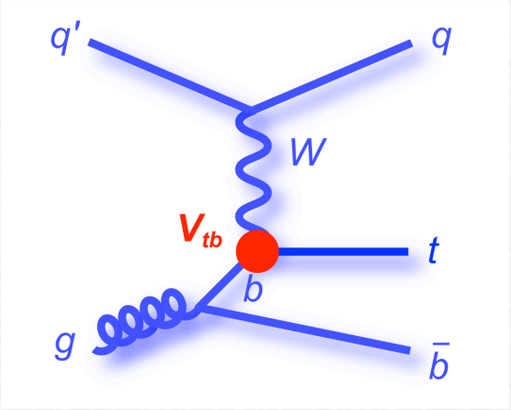
Tevatron latest s-channel Measurements



Measuring CKM Matrix Element $|V_{tb}|$

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}$$





$$|V_{tb} f_L^1|^2 \propto \sigma(s + t\text{-channel})$$

The most general, *CP*-conserving *Wtb* vertex can be parameterized with an effective Lagrangian by :

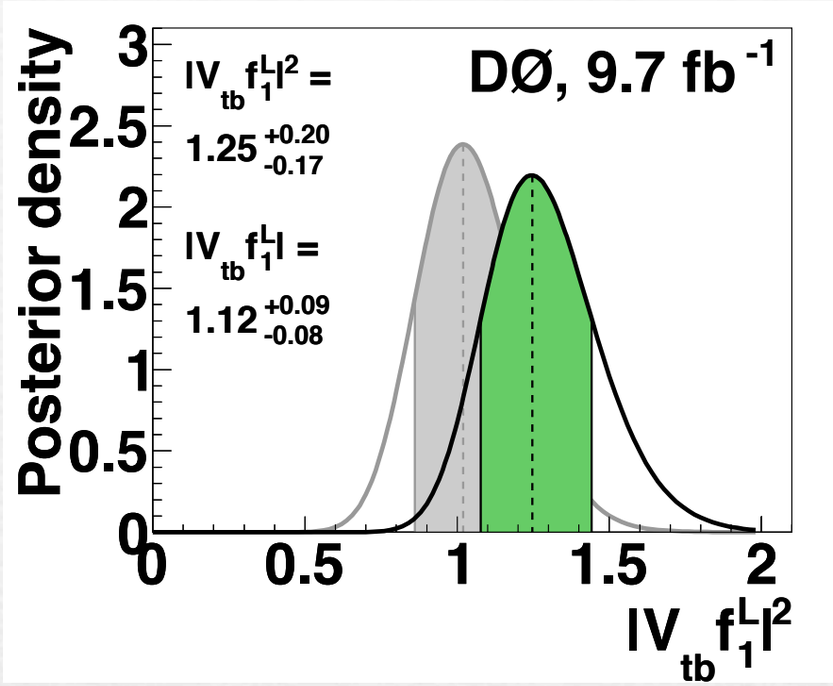
$$\mathcal{L} = -\frac{g}{\sqrt{2}} V_{tb}\, \bar{b}\gamma^\mu (f_1^L P_L) t\, W_\mu^-$$

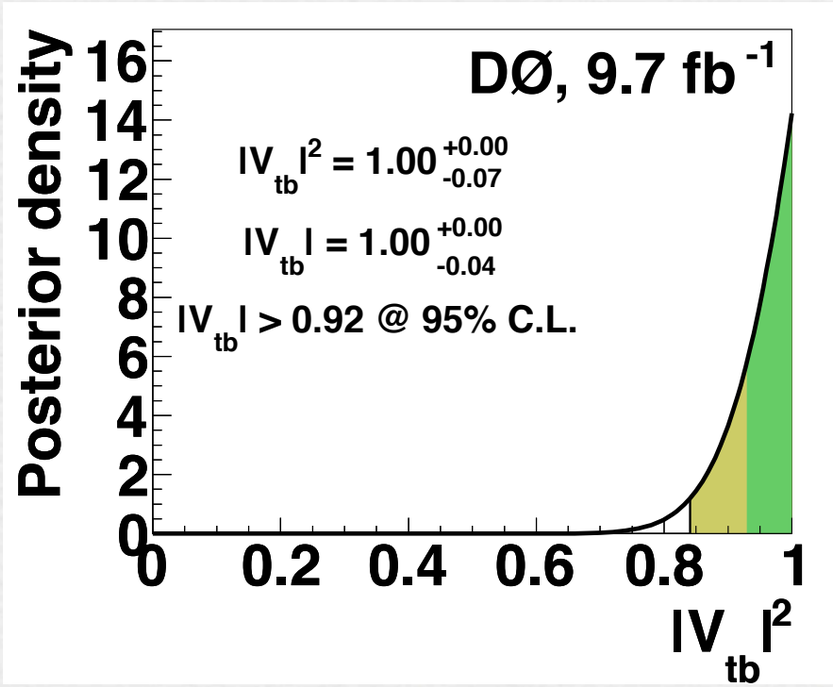
SM: $f_1^L = 1$

- Assume:
  - SM top decay: $|V_{td}|^2 + |V_{ts}|^2 << |V_{tb}|^2$
  - Pure V-A interaction
  - CP conservation
- DO NOT assume
  - 3 generations
  - unitarity of the CKM matrix; allow $|V_{tb} f_1^L|^2 > 1$
  - σ_{tb}/σ_{tqb} (NEW)

New $|V_{tb}|$ measurement

- 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
- Current $|V_{tb}|$ limits @ 95% C.L.
  - CDF (7.5 fb^{-1}) : $0.78 < |V_{tb}| \leq 1$
  - CMS (6 fb^{-1} 8TeV) : $0.81 < |V_{tb}| \leq 1$
  - ATLAS (5 fb^{-1} 8TeV): $0.80 < |V_{tb}| \leq 1$





Conclusions

- First evidence of s-channel single top quark production.

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

  Measurement submitted to Phys. Lett. B, arxiv: 1307:0731
- Simultaneously measure σ_{tb} and σ_{tqb}, without assuming the SM predictions for either.
- Also measure σ_{tb+tqb} and $|V_{tb}|$ without assuming the SM ratio of σ_{tb}/σ_{tqb}

  $|V_{tb}| > 0.92$ @ 95% C.L.
- The results shown are consistent with the SM predictions.
- A legacy measurement at the Tevatron.
- Combination with CDF is in the pipeline.

Thanks !

Back-up

W+jets & Multijets Normalization

- Normalize W+jets and multijets to data simultaneously before b-tagging using Matrix Method.
  - Split data sample in events with real and fake isolated lepton
  - Measure the probability to have an isolated lepton in each sample.

Known from the selected data samples

$$N_{\text{loose}} = N_{\text{loose}}^{\text{fake-}\ell} + N_{\text{loose}}^{\text{real-}\ell}$$

$$N_{\text{tight}} = \epsilon_{\text{fake-}\ell} N_{\text{loose}}^{\text{fake-}\ell} + \epsilon_{\text{real-}\ell} N_{\text{loose}}^{\text{real-}\ell}$$

Determined from MC and data samples

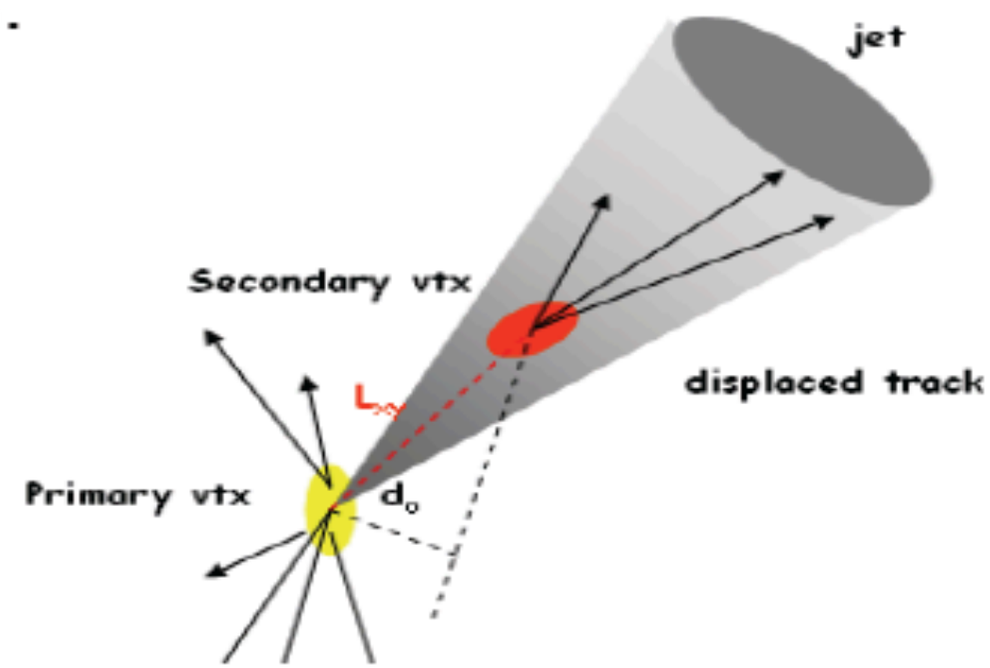
where,

Efficiency(real-l) -> for real isolated lepton to pass the tight selection in MC sample.
Efficiency(fake-l)-> for fake isolated lepton to pass the tight lepton criteria in a multijet enriched sample.
N_{loose} and N_{tight} are the number of events in the loose and tight samples, respectively.

By solving above equations for $N_{\text{loose}}^{\text{fake}-\ell}$ and $N_{\text{loose}}^{\text{real}-\ell}$, the W+jets and multijets contributions in the tight sample can be determined

b-jet Identification

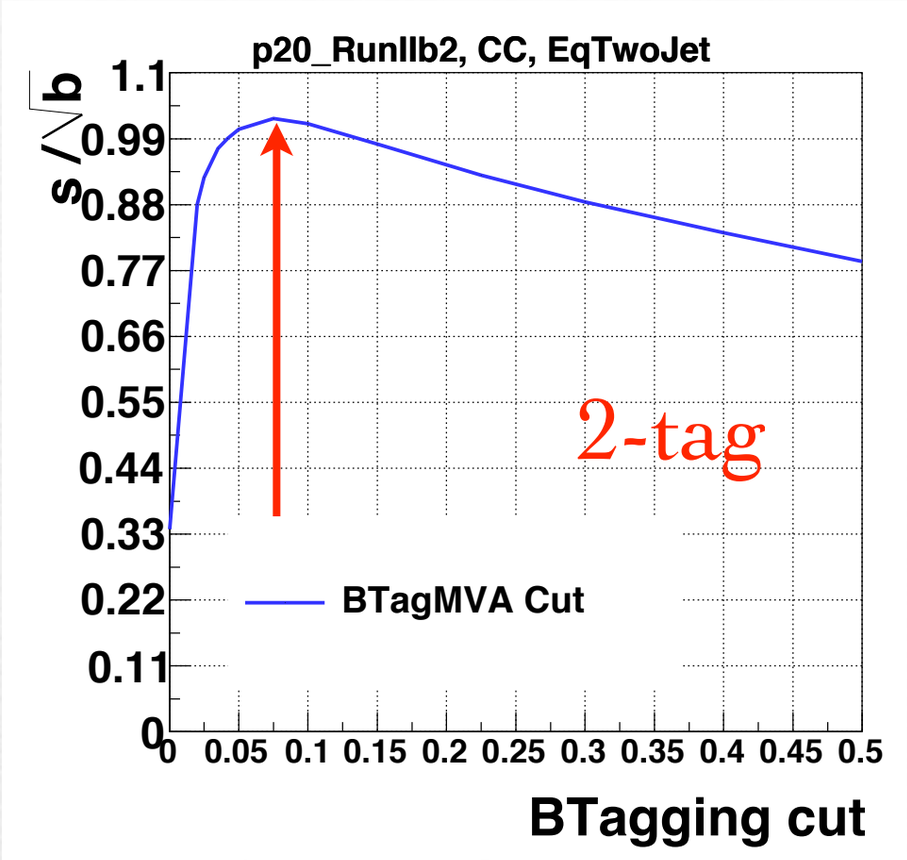
- Separate b-jets from light-quark and gluon jets to reject most W +jets background
- DØ uses a neural network algorithm with seven input variables based on impact parameter and reconstructed vertex
- Two operating points used in analysis:
  - TIGHT ($\varepsilon_b = 40\%$, $\varepsilon_c = 9\%$, $\varepsilon_l = 0.4\%$)
  - LOOSE ($\varepsilon_b = 50\%$, $\varepsilon_c = 14\%$, $\varepsilon_l = 1.5\%$)
- Leading b-jet $p_T > 20$ GeV
- Define two exclusive samples
  - EqOneTag: 1T, no L
  - EqTwoTag: 2L

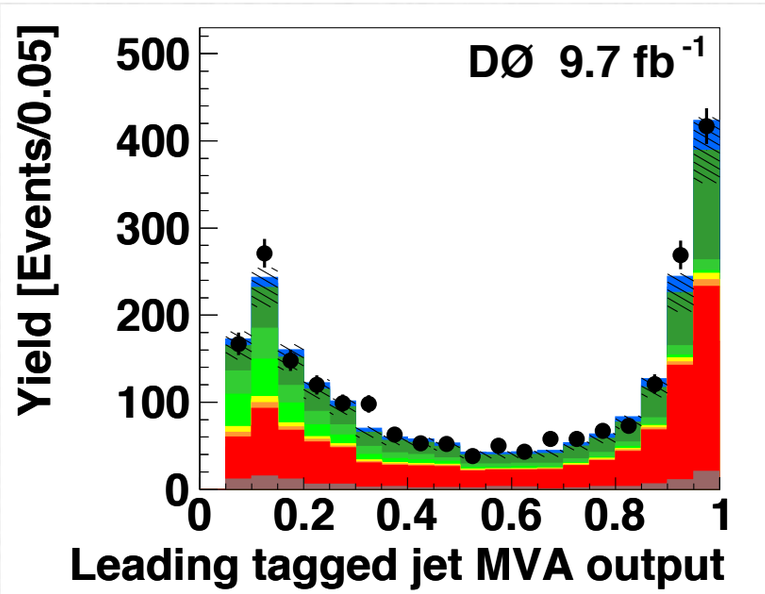


- Uncertainties dominated by variation in data samples used to measure the efficiencies.
- Smaller contribution from MC sample dependence

b-ID Cut Optimization

- Calculate s/√b from the predicted signal and background events with each *b*-ID cut
- Maximize the tb signal rate
  - The tb events have two central *b*-jets
  - 2-tag channel: 2 jets with Loose b-ID
  - 1-tag channel: 1 jet with Tight b-ID, veto the 2nd jet with Loose b-ID
  - Non-overlapping categories



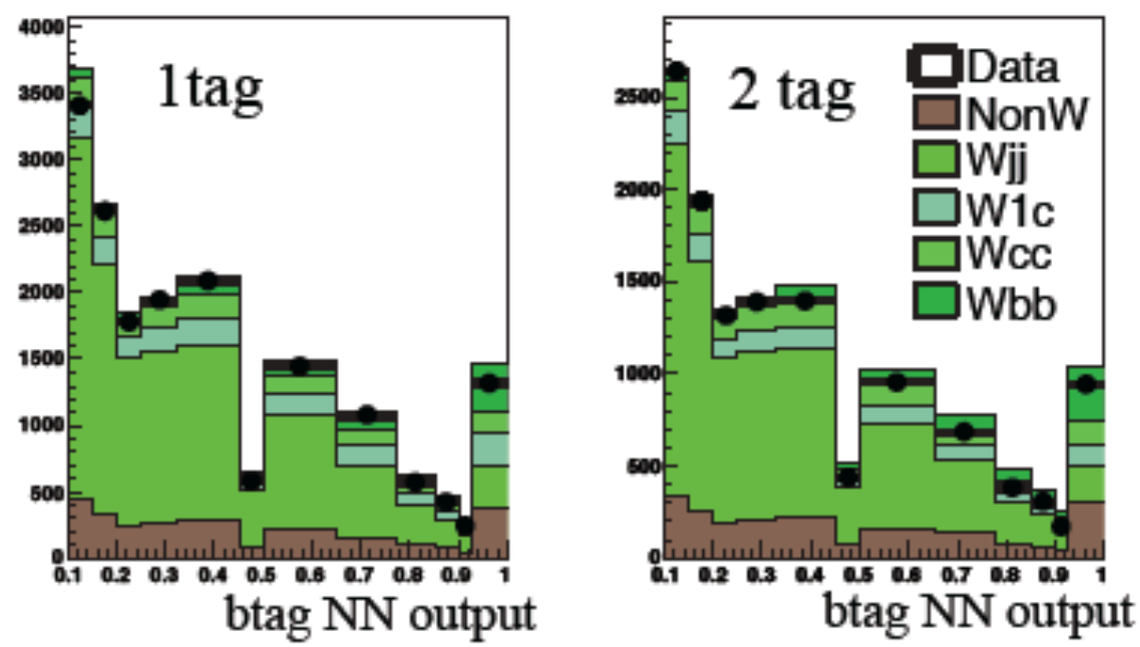


W+ Heavy Flavor Normalization

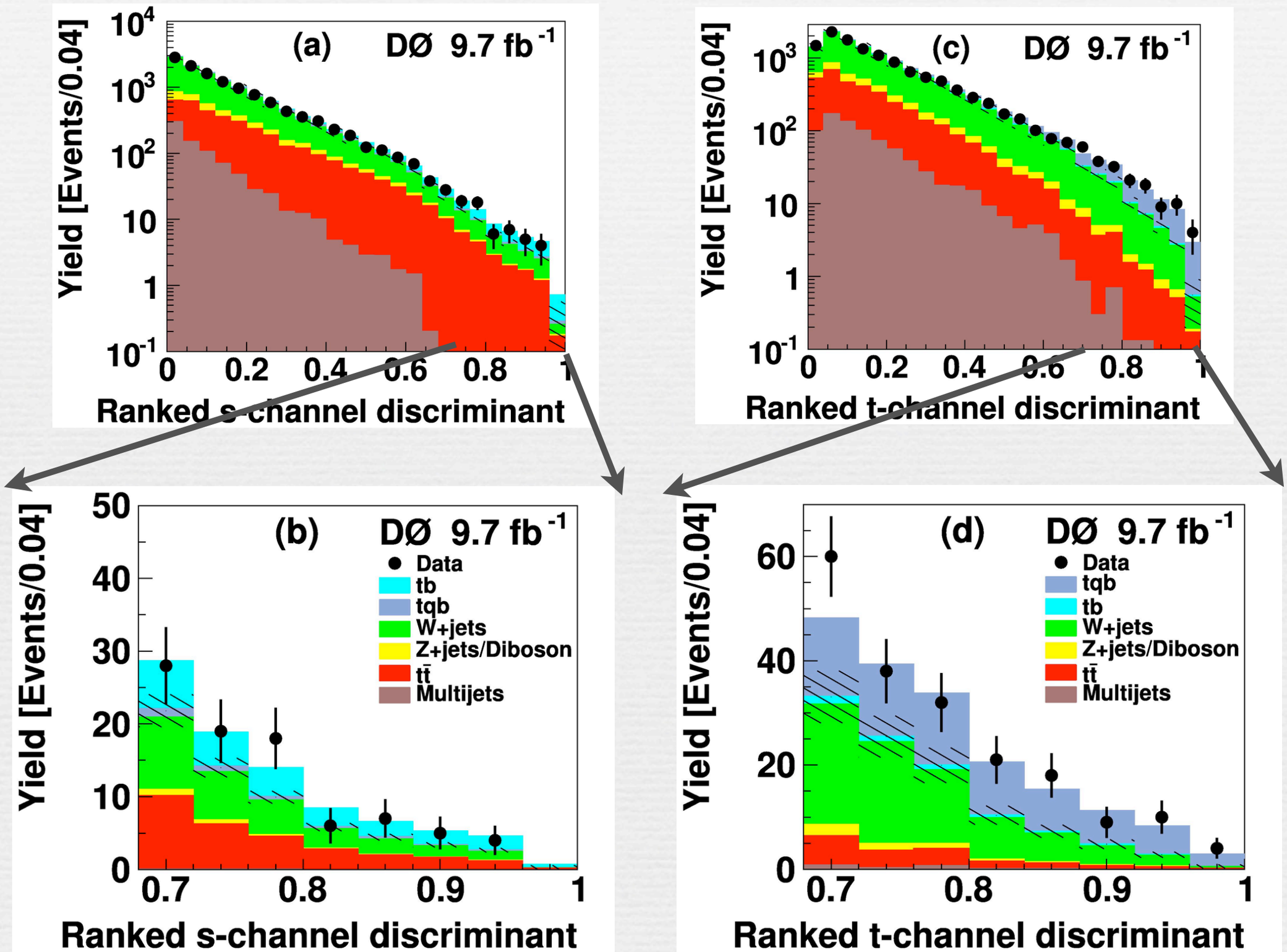
- W/Z + heavy flavor normalized to theory (MCFM-NLO)
  - 1.47 (Wbb,Wcc), 1.32 (Wcj), 1.52 (Zbb), 1.67 (Zcc)
- Normalization checked in 2 jets/0 tag sample

$$N^{(0)} = N^{(0)}_{\mathrm{Wlp}} + \lambda_{\mathrm{HF}} N^{(0)}_{\mathrm{Whp}}$$

  - λ_{HF} found to be consistent with 1
- Uncertainties considered
  - ± 40% single top cross section
    - ± 1% in λ_{HF}
  - ± 10% on the Wcj theory SF
    - ± 7% in λ_{HF}
  - Additional ± 10% Wbb/Wcc
    - ± 8% in λ_{HF}
  - For a total uncertainty of 12%



Combination Discriminants



Cross - Section Measurement

- **Bayesian approach is used**

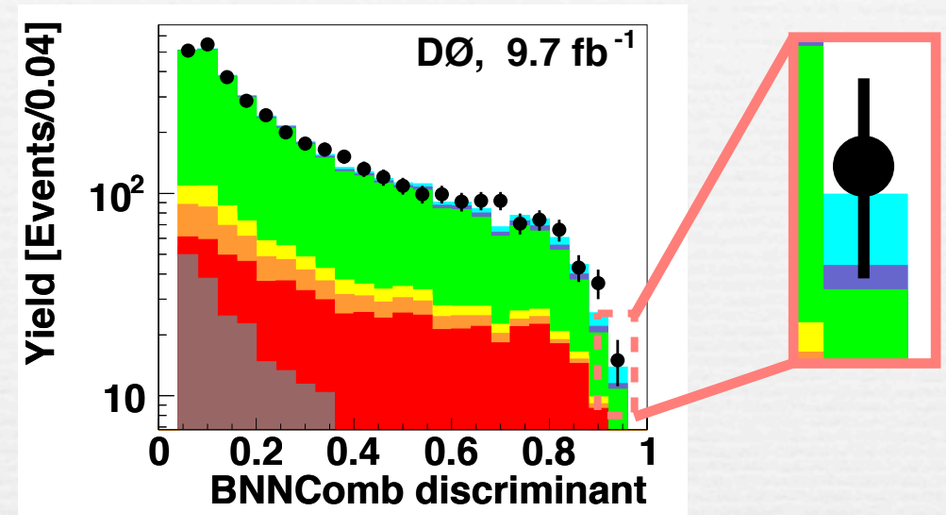
$$d = S + B = \sigma \mathcal{A} \mathcal{L} + B = \sigma a + \sum_{i=1}^{\text{Nbkgds}} b_i$$

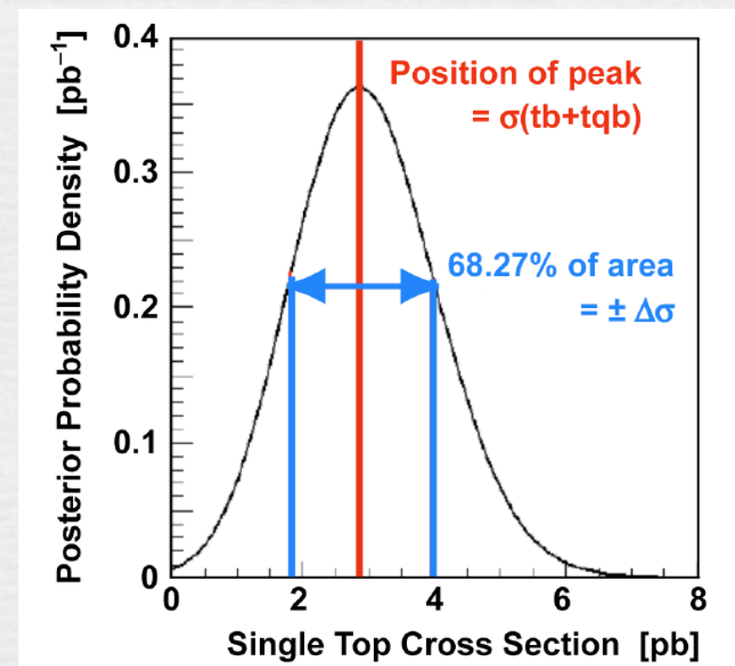
d = Predicted number of data events
S = Predicted number of signal events
B = Predicted number of background events
σ = Cross section
$\mathcal{A}$ = Signal acceptance
$\mathcal{L}$ = Integrated luminosity
a = Effective luminosity
b_i = No. of events in each background component

$$\text{Prob}(D|d) \equiv \text{Prob}(D|\sigma, a, \mathbf{b}) = \prod_{i=1}^{\text{Nbins}} \text{Prob}(D_i|d_i)$$

D = Observed number of data events
$\mathbf{b}$ = Vector of background components

$$\text{Posterior Probability Density}(\sigma|D) \propto \int_a \int_{\mathbf{b}} \text{Prob}(D|\sigma, a, \mathbf{b})\ \text{Prior}(a, \mathbf{b})\ \text{Prior}(\sigma)\ da\ d\mathbf{b}$$





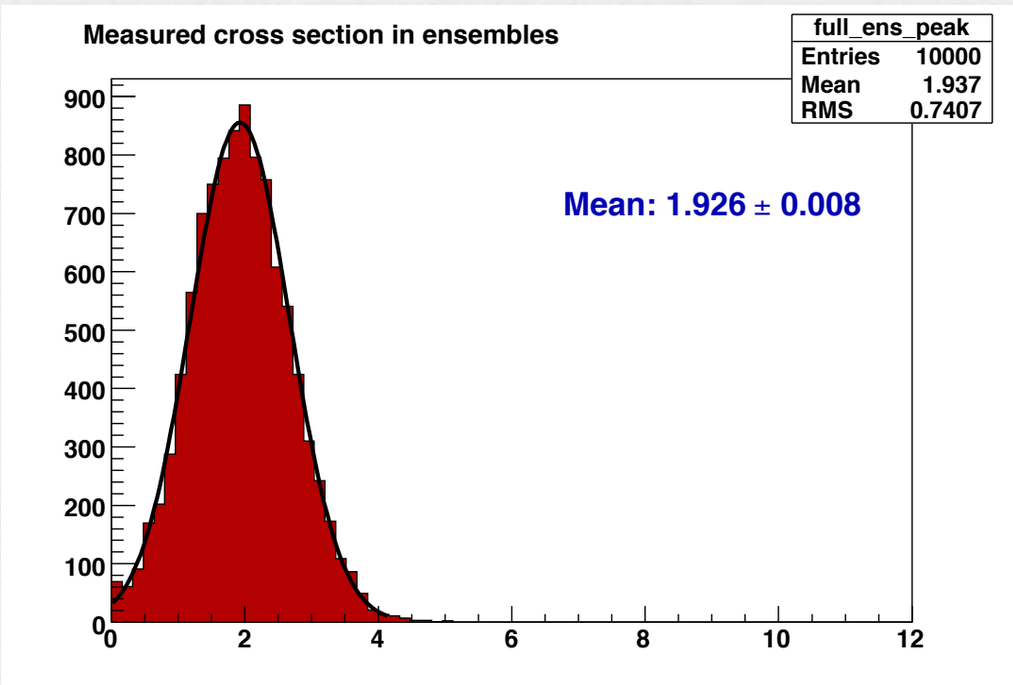
- A flat prior for σ (Prior(σ)) is used.
- Prior(a,**b**) include the shape and normalization systematic uncertainties.

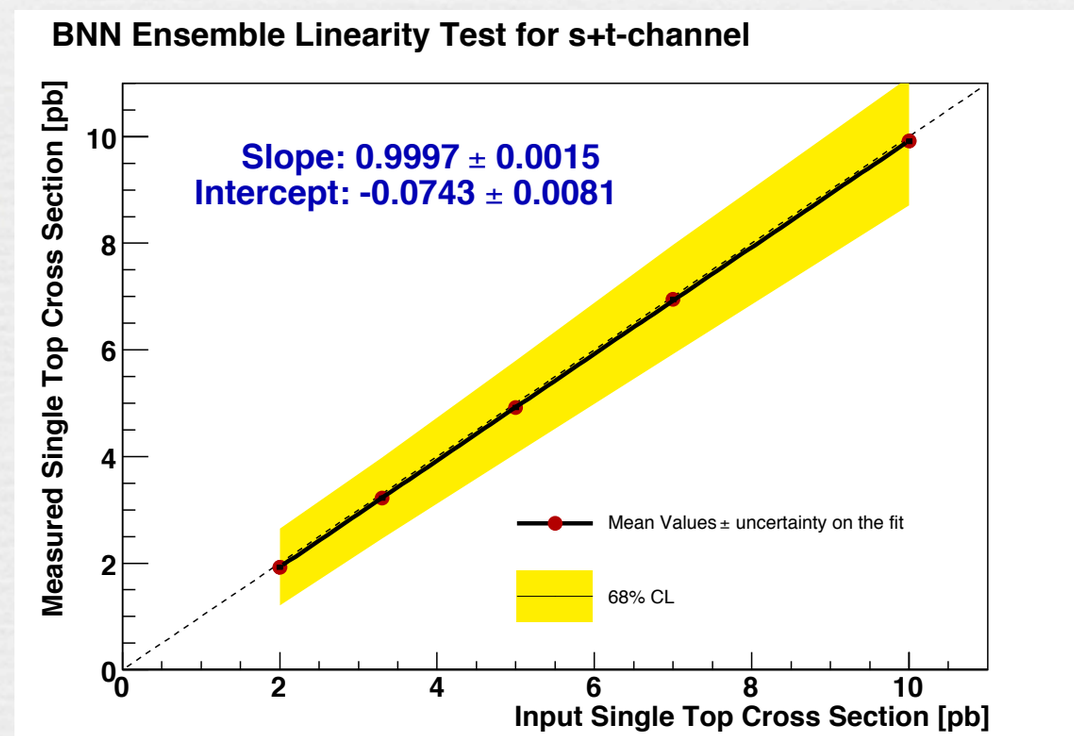
Cross-Section is obtained from the peak position of the posterior prob. density.

Ensembles - Linearity test

Test of machinery with many sets of pseudo-data. To ensure no bias in the cross section measurement.

- Subset of our total pool of background events.
- Systematic uncertainties are fully taken into consideration
- Generated several ensembles with different single top content
- The bias is determined from straight-line fit to the mean of distributions vs. input cross-sections.





- We verify that on average the method measures the same value that was used for top quark cross section in the pseudo-experiments.

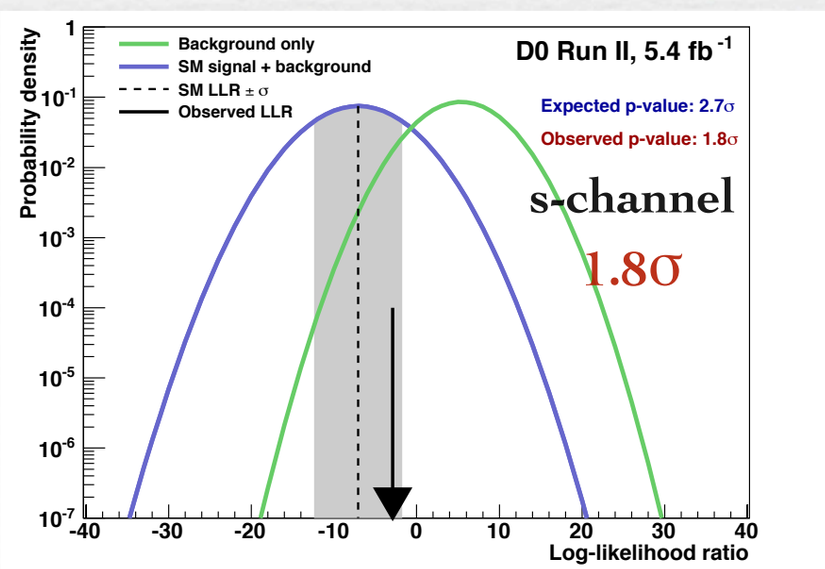
Significance

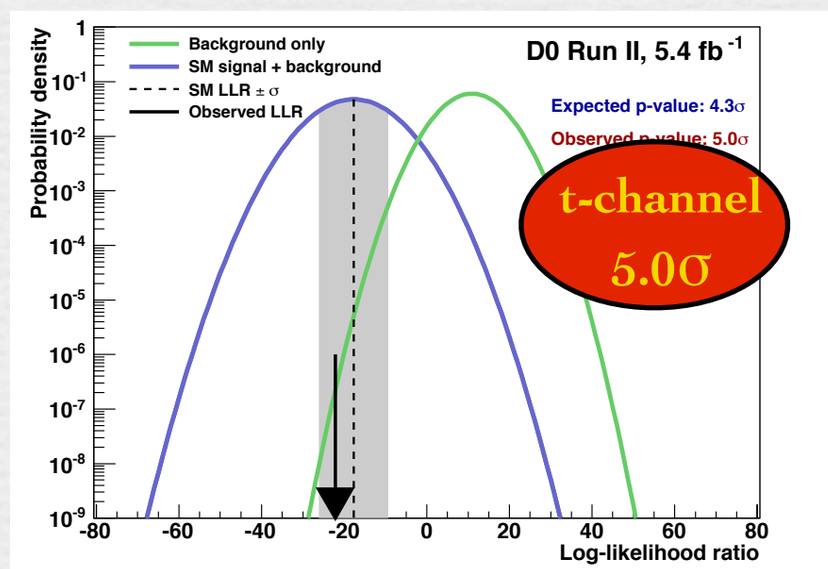
Significance is computed as :

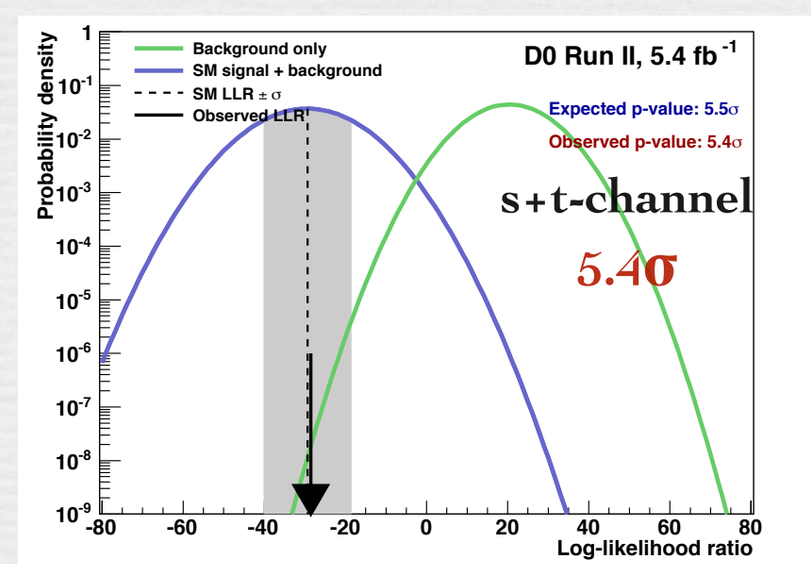
$$Q = -2 \ln \frac{L(\mu = 1)}{L(\mu = 0)}$$

where $L(\mu)$ is the likelihood after integrating over all the systematics for a cross section $\sigma = \mu\sigma_0$, with μ -> strength parameter and σ_0 is the theoretical cross section for signal.
The ratio tests the data compatibility with two null hypothesis: ($\mu = 0$) only background & ($\mu = 1$) background+signal.

The p-value is the probability that the log-likelihood ratio is smaller or equal to the observed result in the case when there is no signal.







Measuring Cross Section

Single-bin likelihood

$$L(D|d) = \frac{e^{-d} d^D}{\Gamma(D+1)}$$

$$d = \alpha \mathcal{L} \sigma + \sum_{i=1}^{N} b_i = a\sigma + \sum_{i=1}^{N} b_i$$

Binned likelihood

$$L(\mathbf{D}|\mathbf{d}) \equiv L(\mathbf{D}|\sigma, \mathbf{a}, \mathbf{b}) = \prod_{i=1}^{M} L(D_i|d_i)$$

Bayes' Theorem

$$p(\sigma|\mathbf{D}) = \frac{1}{\mathcal{N}} \int L(\mathbf{D}|\sigma, \mathbf{a}, \mathbf{b}) \pi(\sigma, \mathbf{a}, \mathbf{b}) d\mathbf{a} d\mathbf{b}$$

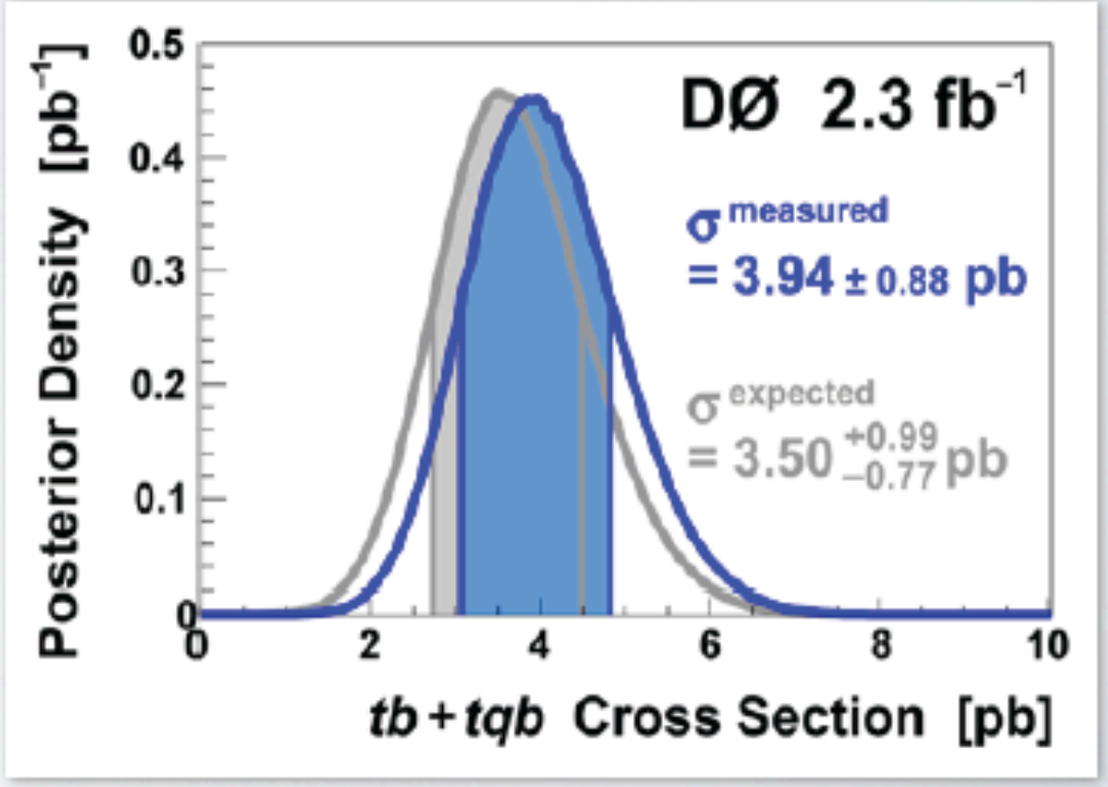
Assuming nonnegative flat prior

$$p(\sigma|\mathbf{D}) = \frac{1}{\mathcal{N}\sigma_{\max}} \int L(\mathbf{D}|\sigma, \mathbf{a}, \mathbf{b}) \pi(\mathbf{a}, \mathbf{b}) d\mathbf{a} d\mathbf{b}$$

Numerical integration

$$p(\sigma|\mathbf{D}) \propto \int L(\mathbf{D}|\sigma, \mathbf{a}, \mathbf{b}) \pi(\mathbf{a}, \mathbf{b}) d\mathbf{a} d\mathbf{b}$$

$$\approx \frac{1}{K} \sum_{k=1}^{K} L(\mathbf{D}|\sigma, \mathbf{a}_k, \mathbf{b}_k)$$



Bayesian Statistical analysis in 2D

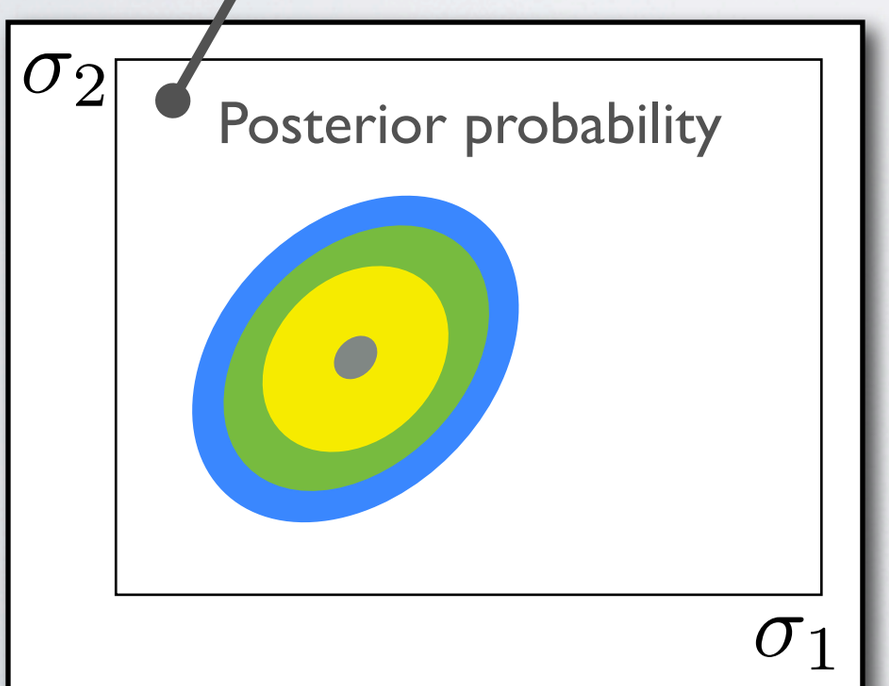
In this case both the signal prior and posterior distributions have 2 dimensions

Mean event count

$$\boldsymbol{d} = \sigma_1 \overbrace{\boldsymbol{a}_1}^{\text{signal acceptances}} + \sigma_2 \overbrace{\boldsymbol{a}_2} + \underbrace{\boldsymbol{b}}_{\text{background event yields}}$$

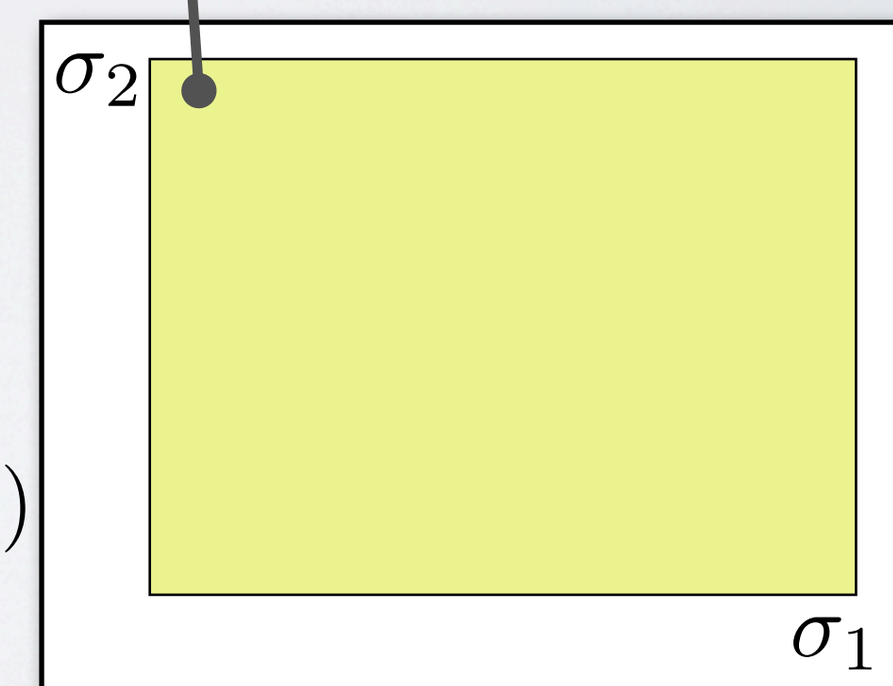
signal production rates

$$p(\sigma_1, \sigma_2) = \frac{1}{\mathcal{N}} \int L(\boldsymbol{D}|\sigma_1, \sigma_2, \boldsymbol{a}, \boldsymbol{b}) \pi(\sigma_1, \sigma_2) \pi(\boldsymbol{a}, \boldsymbol{b}) d\boldsymbol{a} d\boldsymbol{b}$$

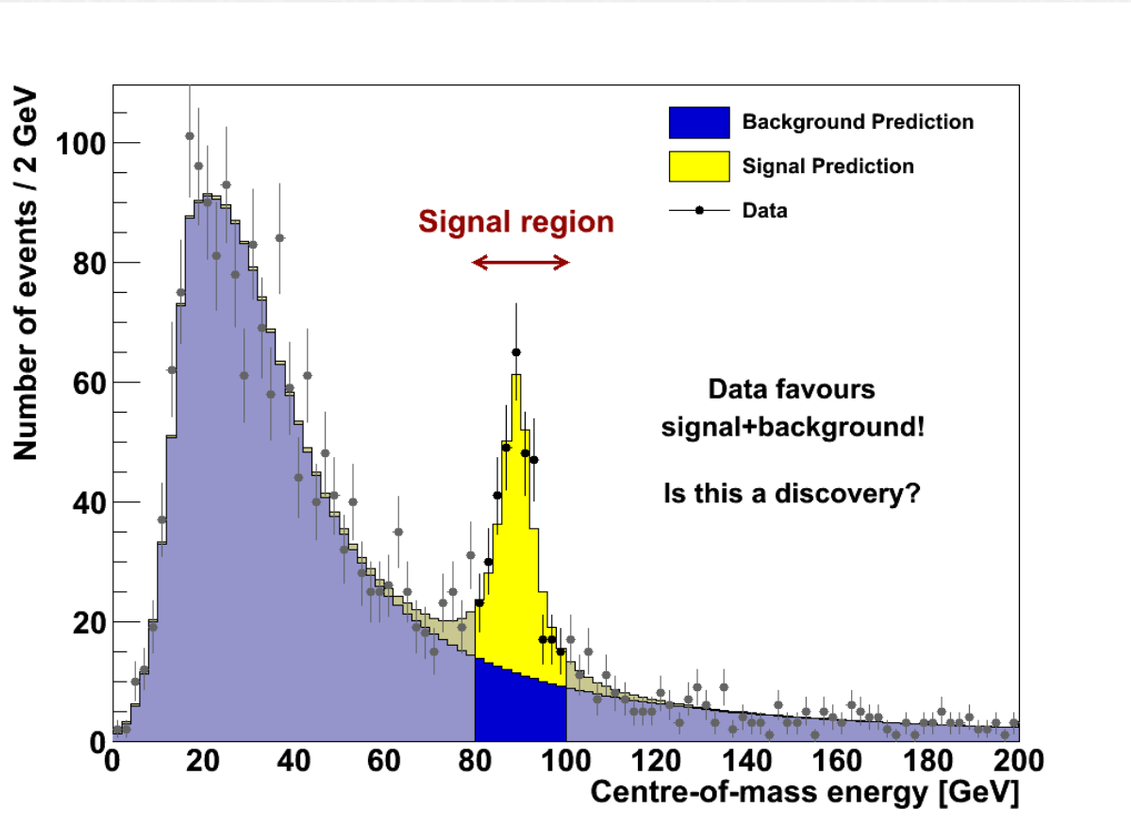


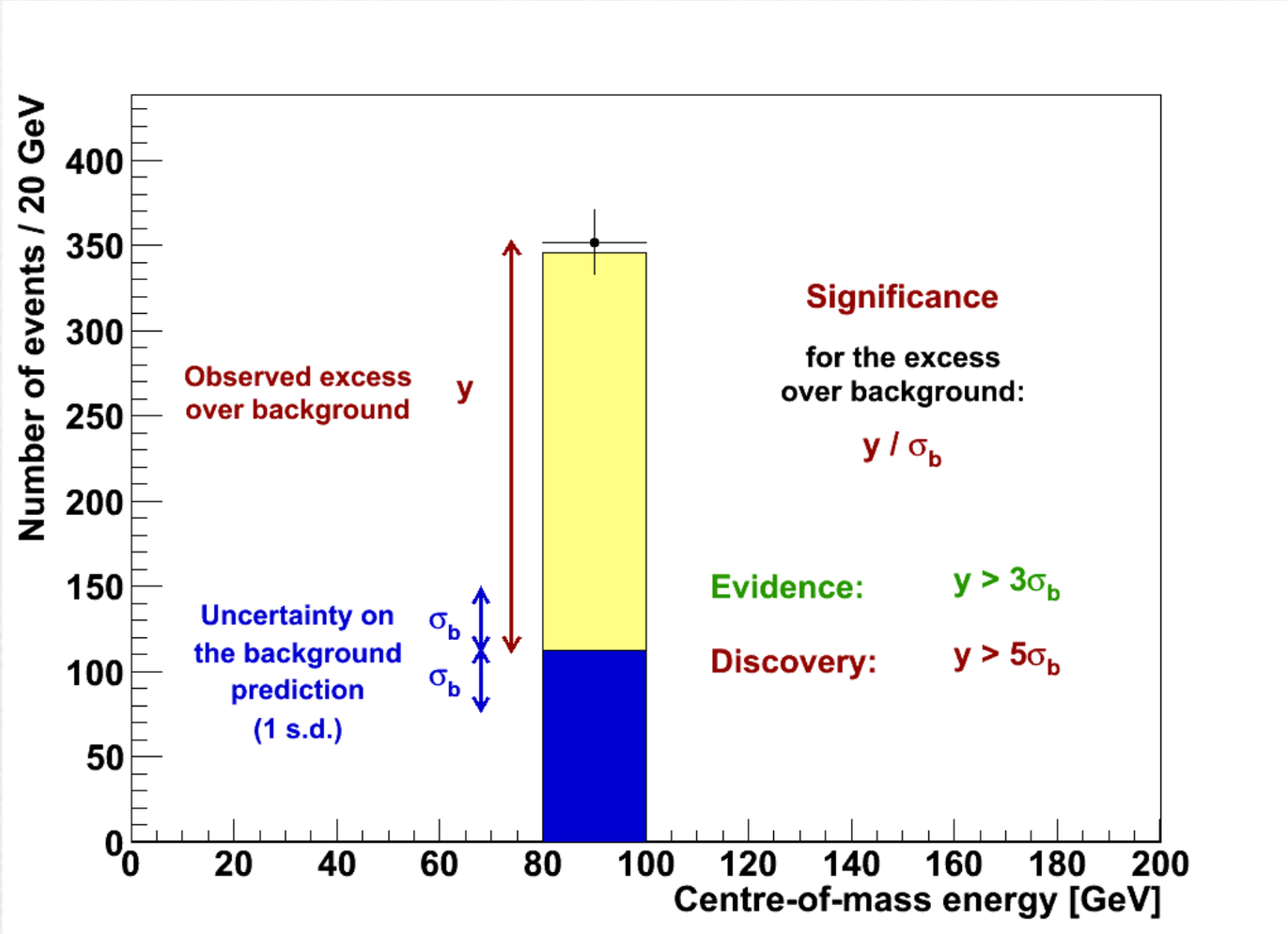
Update

$$L(\boldsymbol{D}|\sigma_1, \sigma_2, \boldsymbol{a}, \boldsymbol{b})$$



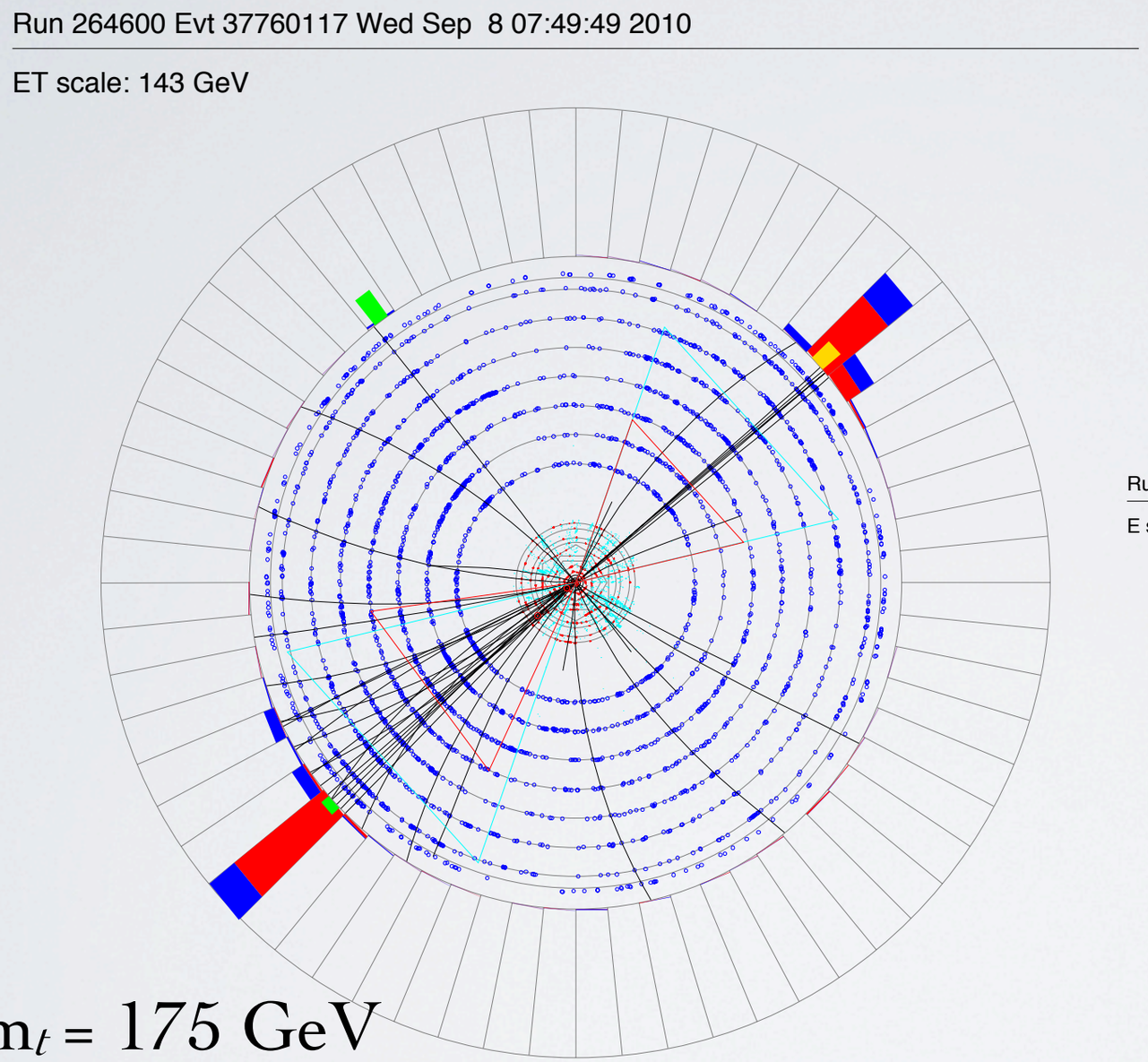
Significance ?



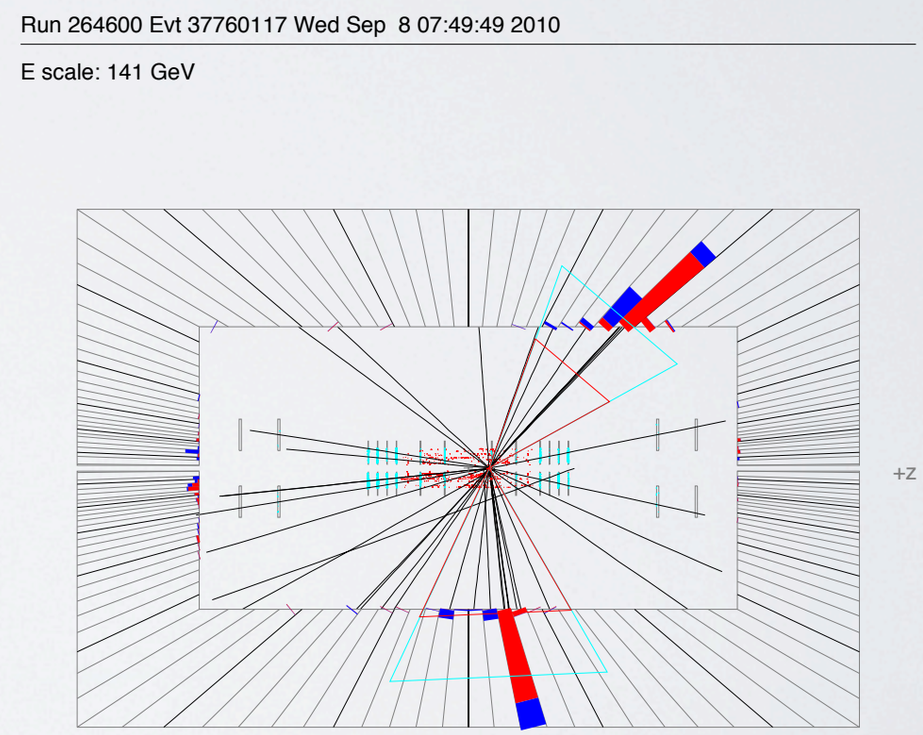


Probability for 3σ upward fluctuation: 0.135%, 5σ: 0.000029%

tb 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