International Meeting on Fundamental Physics

El Escorial, April 7, 2006

Search for single top quarks at DØ

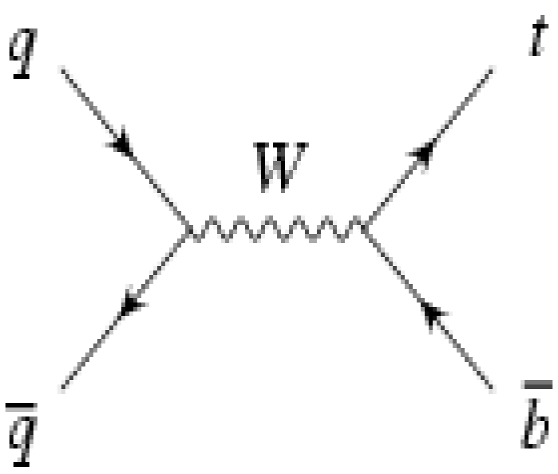
- Electroweak top quark production
- Signal and background modeling, b-tagging
- Analysis strategy:
  - Loose event selection
  - Multivariate analysis: Likelihood Discriminants
  - Binned likelihood for limit setting
- Results
- Conclusions and outlook

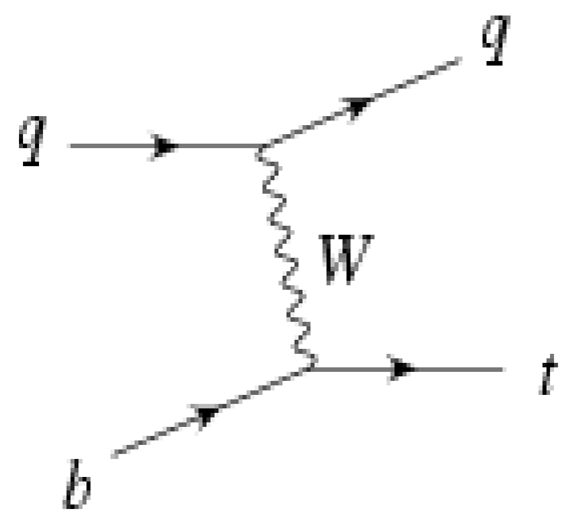
UNIVERSITY OF WASHINGTON

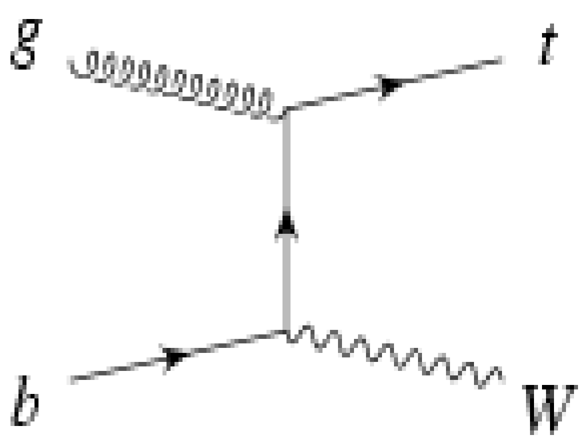
Arán García-Bellido

Top quark electroweak production

s-channel	**t-channel**	**associated tW**
TeV: 0.88±0.11 pb	**1.98±0.25 pb**	**<0.1 pb**
LHC: 10.6±1.1 pb	**246.6±0.25 pb**	**$62.0^{+16.6}_{-3.6}$ pb**





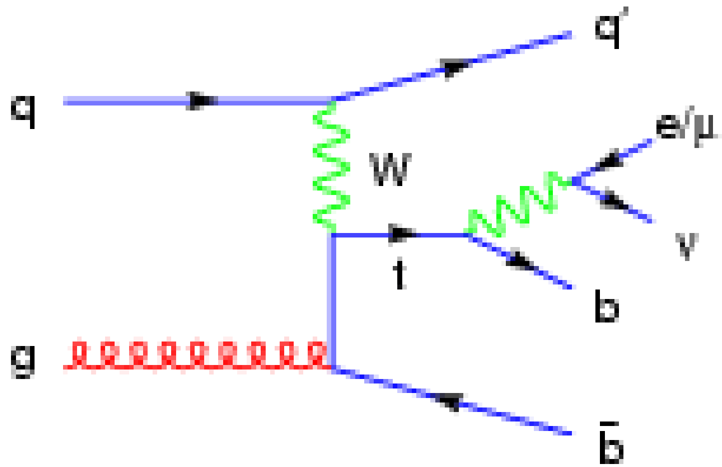


Harris, Laenen, Phaf, Sullivan, Weinzierl: hep-ph/0207055
Cao and Yuan: hep-ph/0408180

Tait: hep-ph/9909352
Belyaev, Boos: hep-ph/0003260

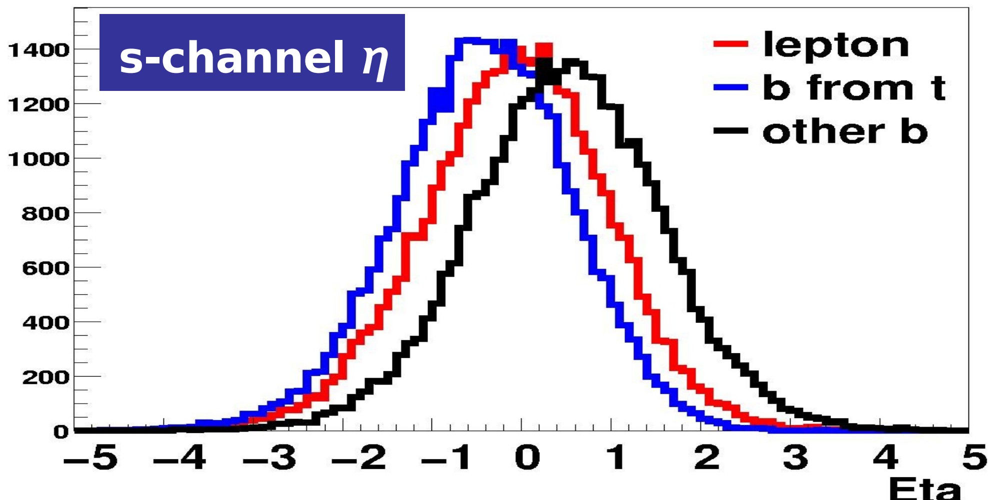
- Observe single top quark production in Run II ← Final Goal!
- Measure directly the V_{tb} CKM element
- Look for new physics: FCNC, 4th generation, W', SUSY,...
- Probe W+jets understanding and help SM higgs searches

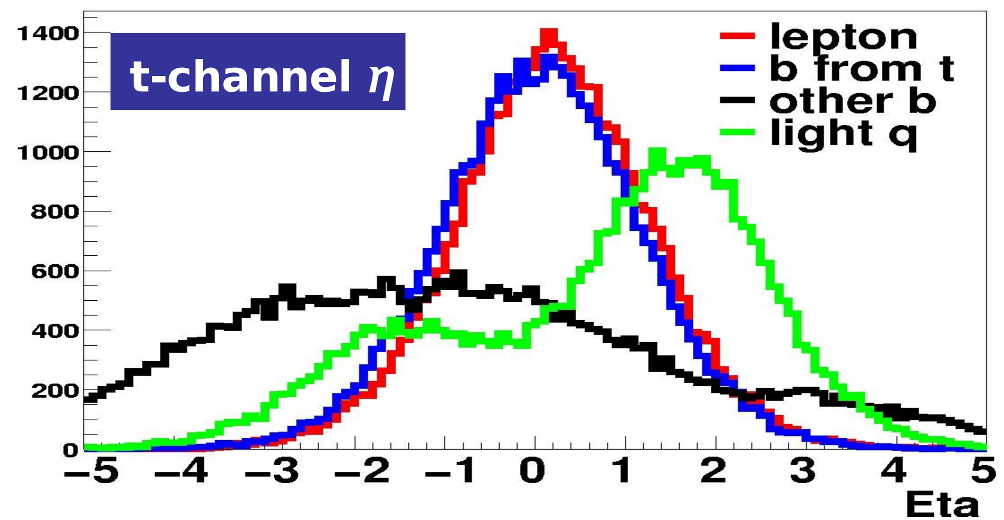
Signal topology

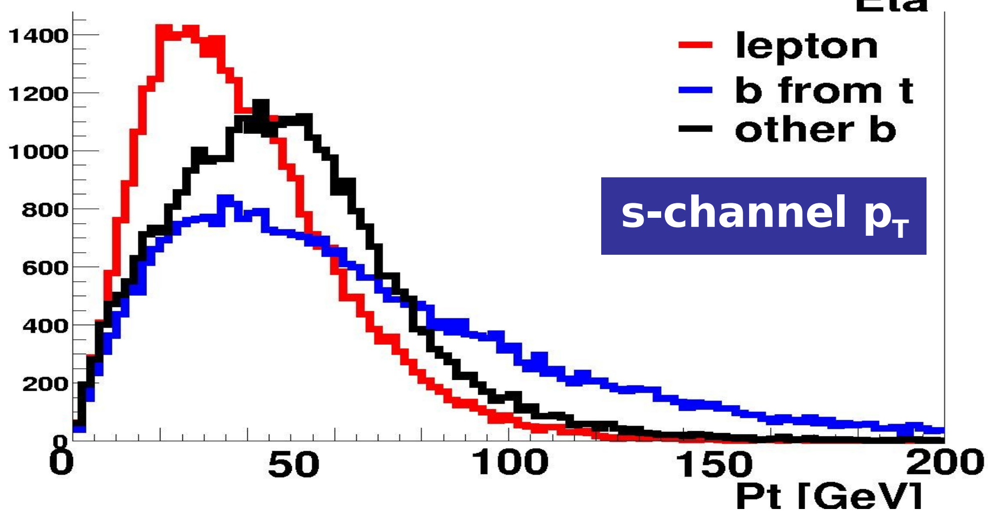


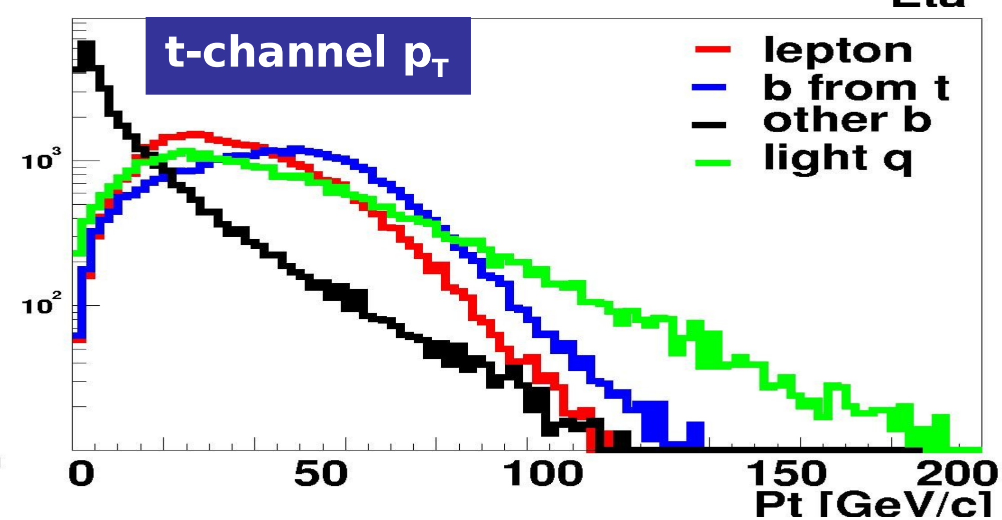
We are looking for:

- 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



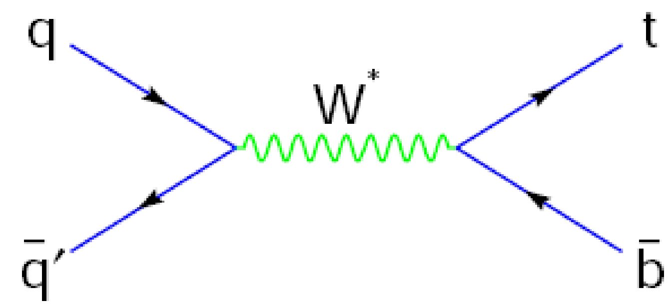


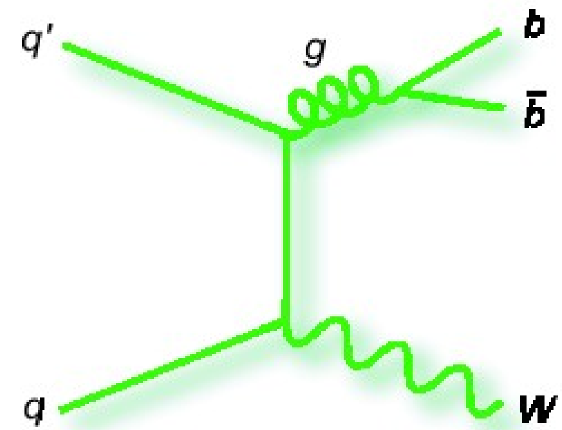


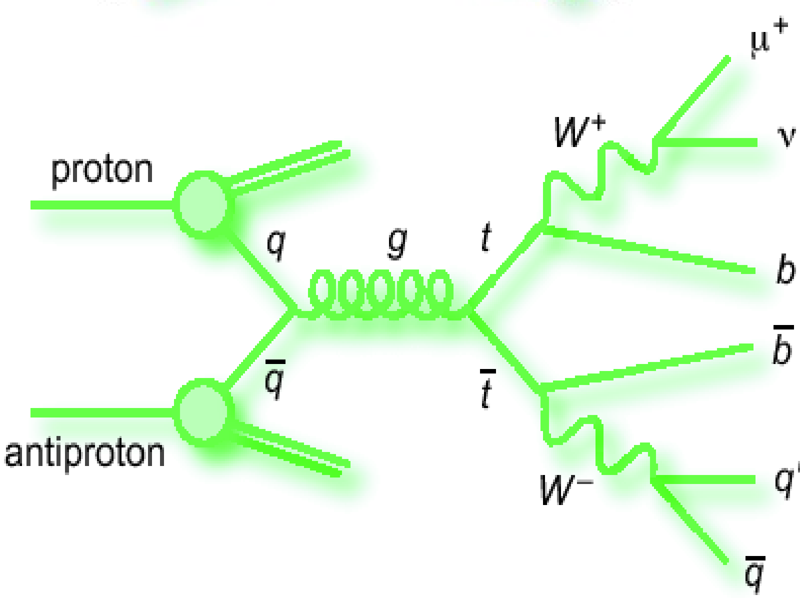


Signal and background modeling

- Signal: “Effective” NLO CompHEP
  - Match 2→2 and 2→3 processes using b p_T for cross over, normalize to NLO
  - Resulting distributions agree well with ZTOP & MCFM
- W+jets
  - Distributions from Alpgen
  - Wbb to Wjj fraction from NLO
  - Normalization from data
- Top pairs: Alpgen
- Multijet events: from data
  - Misidentified lepton
- Diboson (WZ, WW): Alpgen

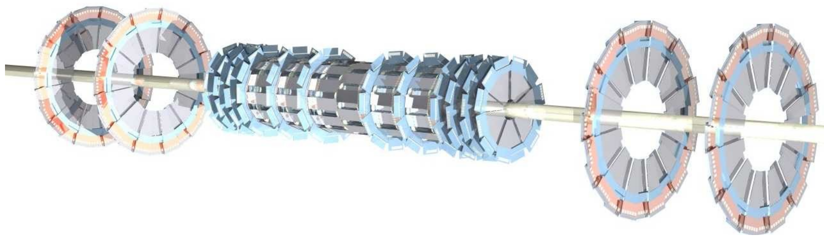


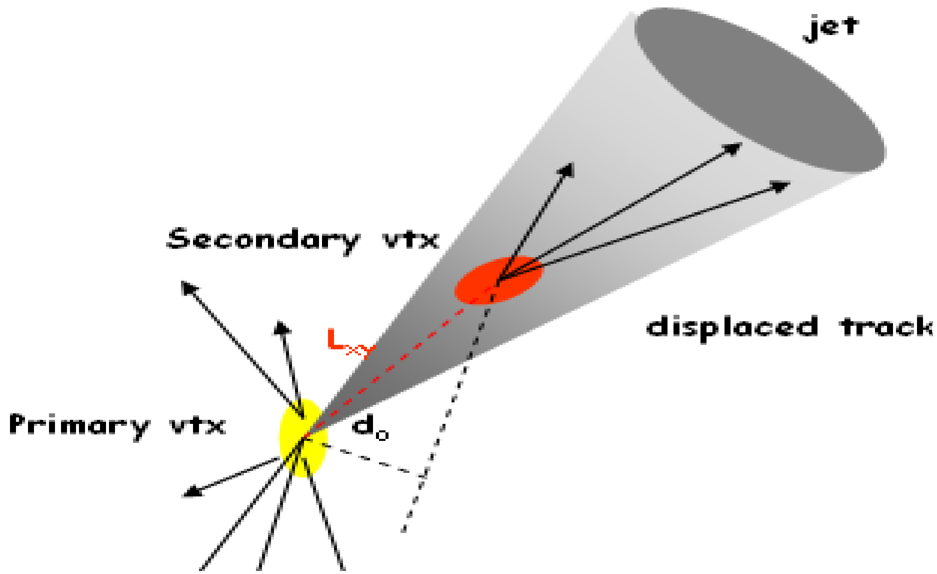




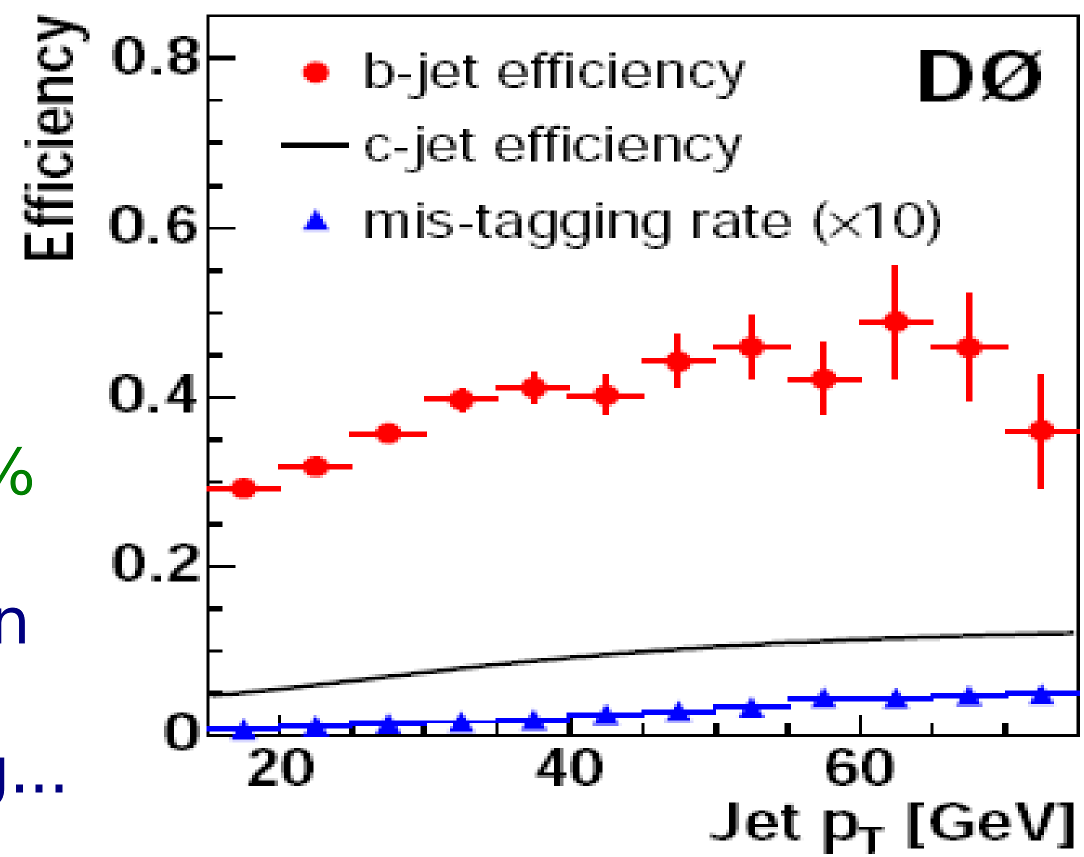
Tagging b-jets

- Use jet lifetime probability
- Good tracking and vertexing capabilities are crucial!





b-jet efficiency: ~40%
c-jet efficiency: ~10%
light jet efficiency: <0.2%



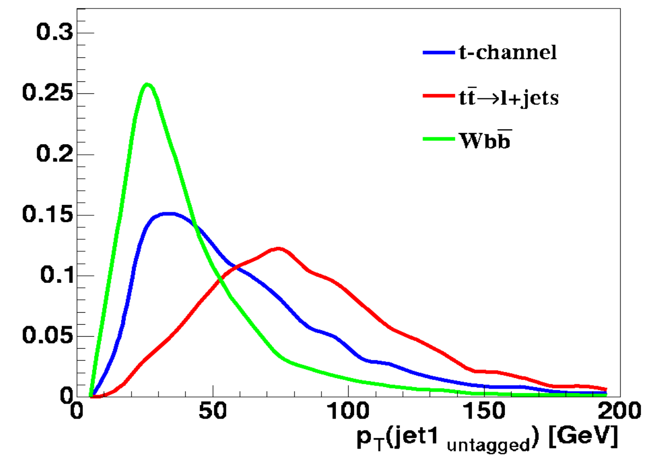
- Tagging improves S/B ratio
- But heavy flavor composition after tagging is complicated
- Admixture of Wbb, Wcj, Wcg...

New NN tagger now certified with much better efficiency!

Analysis strategy

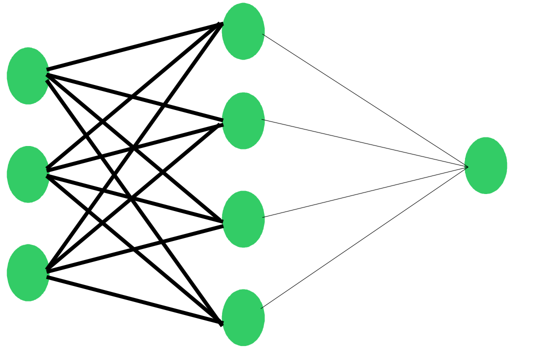
1) Event selection

- Select W like events
- Maximize acceptance
- Model backgrounds well



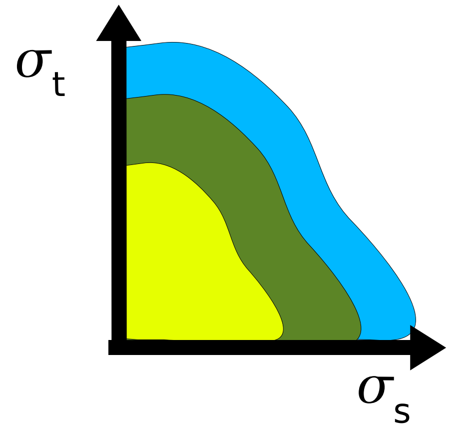
2) Separate signals from backgrounds

- Find discriminating variables
- Multivariate analysis:
  Cuts, Decision Trees, Neural Networks, Likelihoods



3) Determine cross section

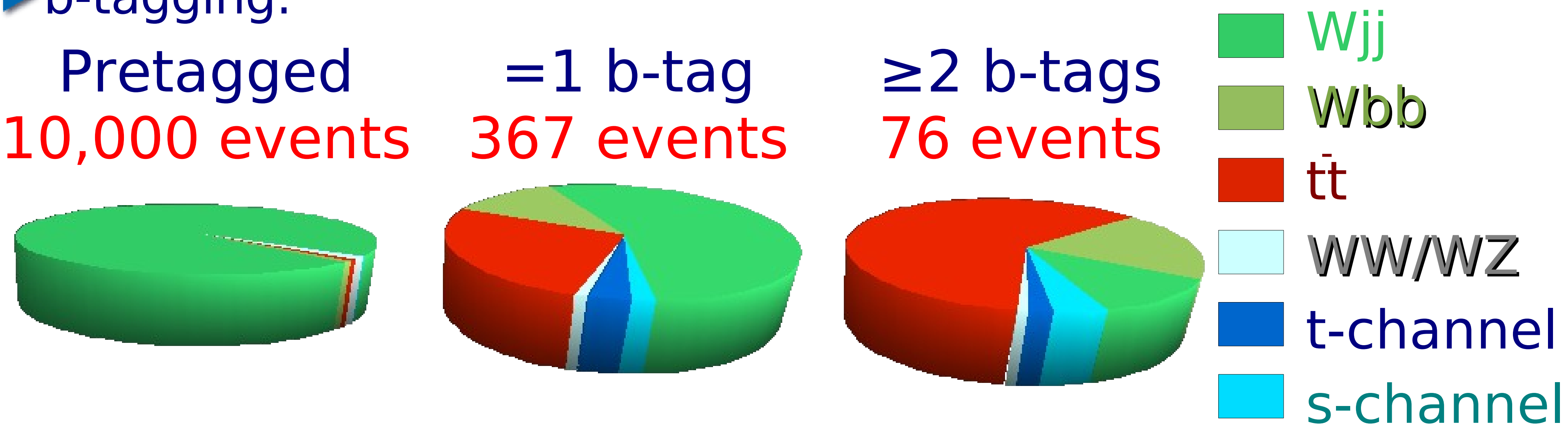
- Use shape information
- Bayesian statistical analysis



1) Event Selection and yields (370 pb^{-1})

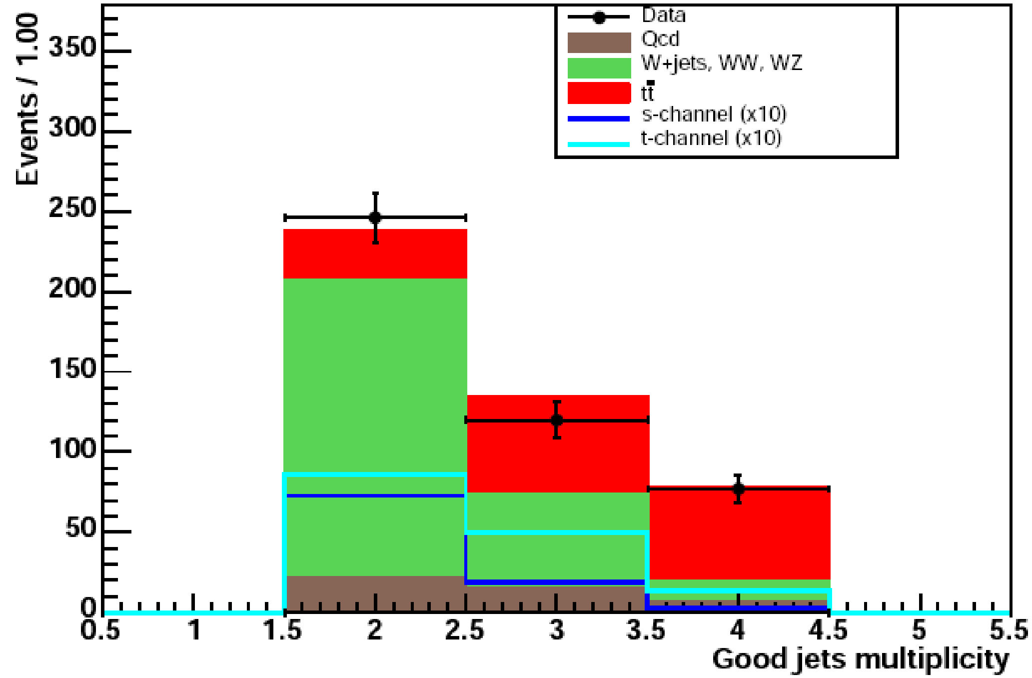
- 1 isolated lepton (e,μ)
  - p_T>15 GeV
  - $|\eta_{e(\mu)}|$<1.1 (2.0)
- MET: MET>15 GeV
- Jets: $2 \leq N_{jets} \leq 4$
  - E_T>15 GeV
  - $|\eta|$<3.4
  - Jet1: E_T>25 GeV
- Other clean-up cuts
- b-tagging:

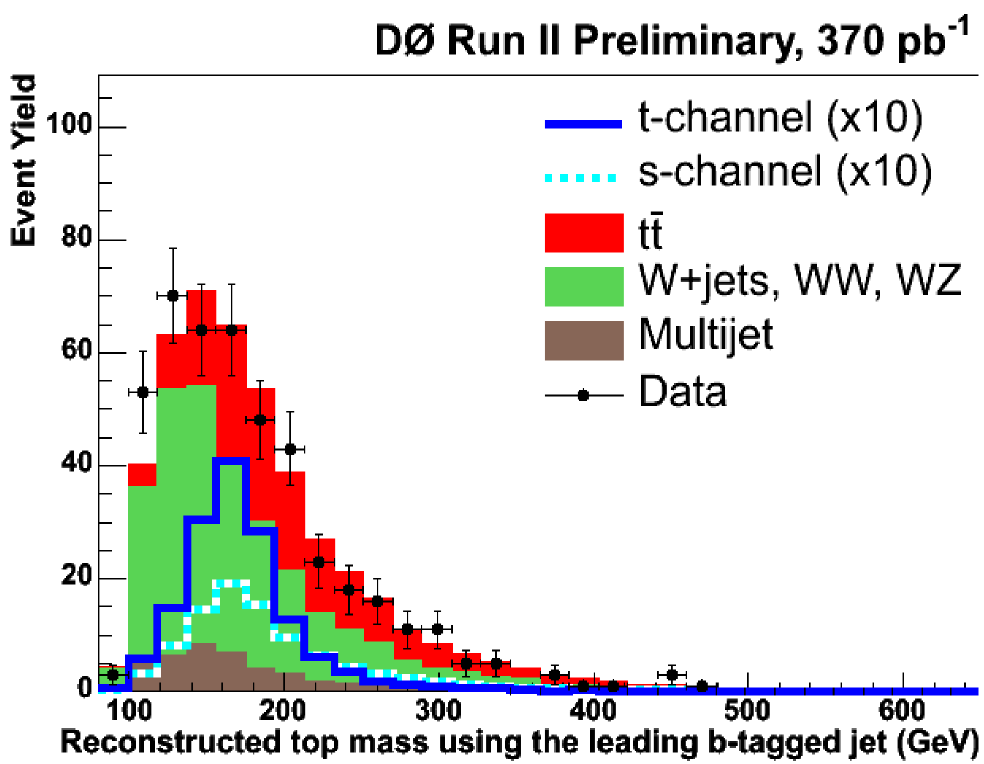
	s-channel	t-channel
Cuts acceptance	23%	22%
b-tag eff	54%	38%
Signal yield	9.5	15.0
Backgnd yield	452	
Data	443	
S/B	1/50	1/30

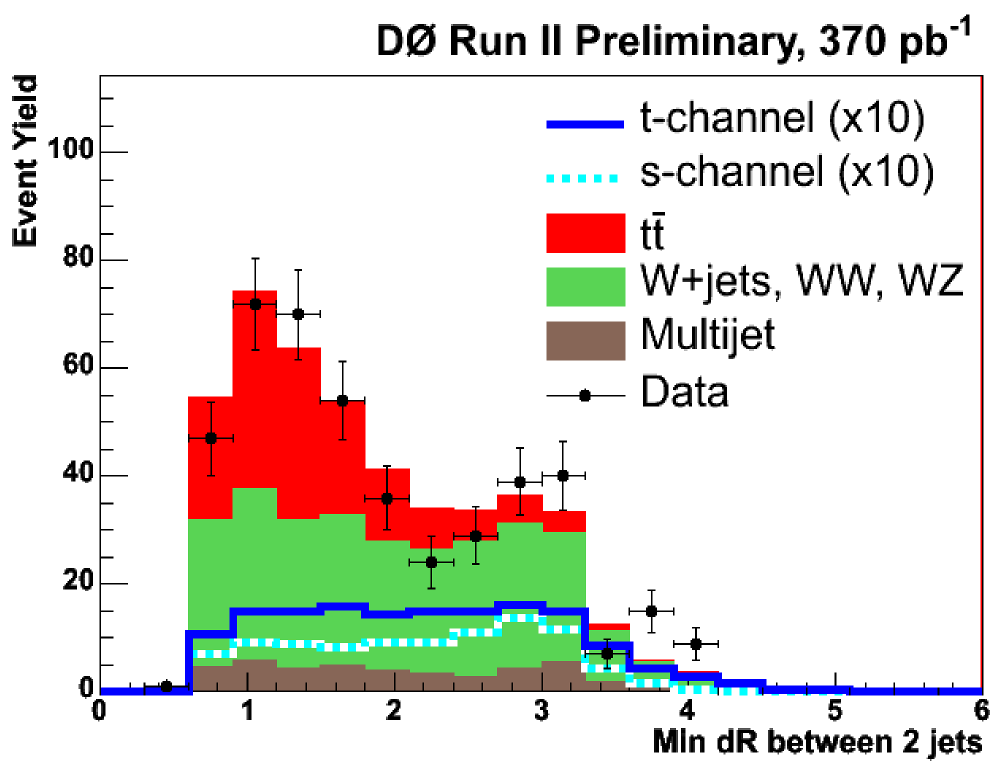


2) Separate signals from backgrounds

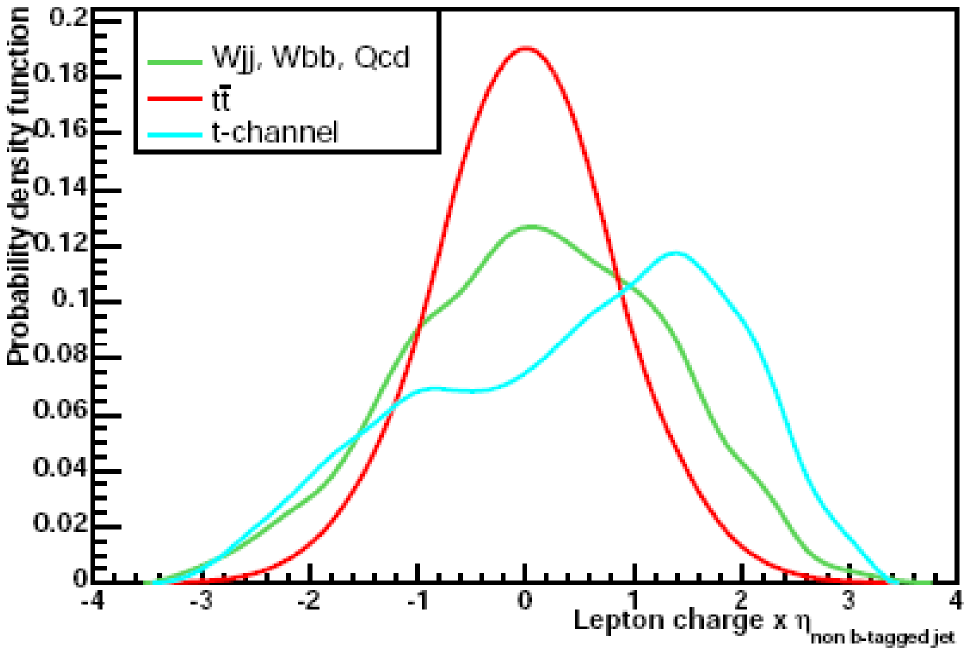
- Make sure your data agrees well with your prediction
- Choose variables that show good discrimination
- Use all handles: angular correlations, event kinematics, object kinematics

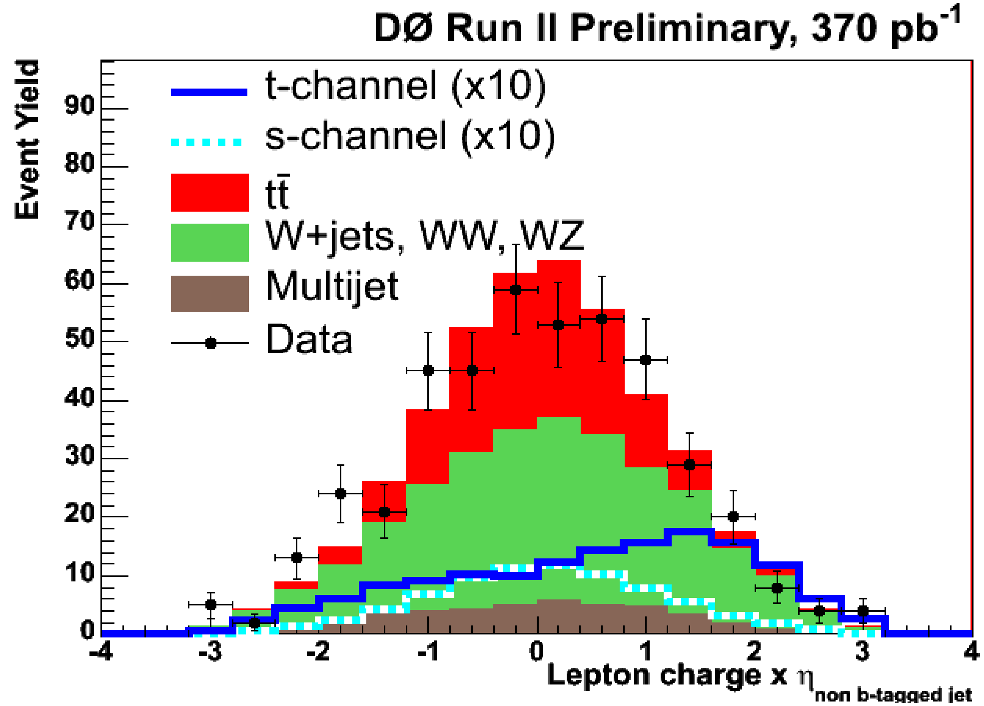


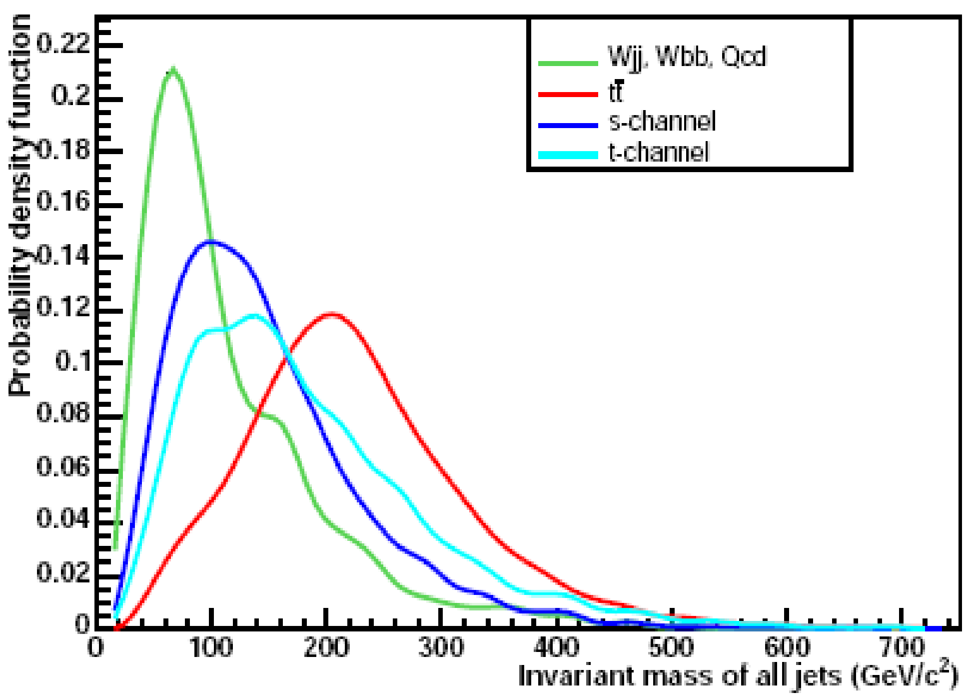


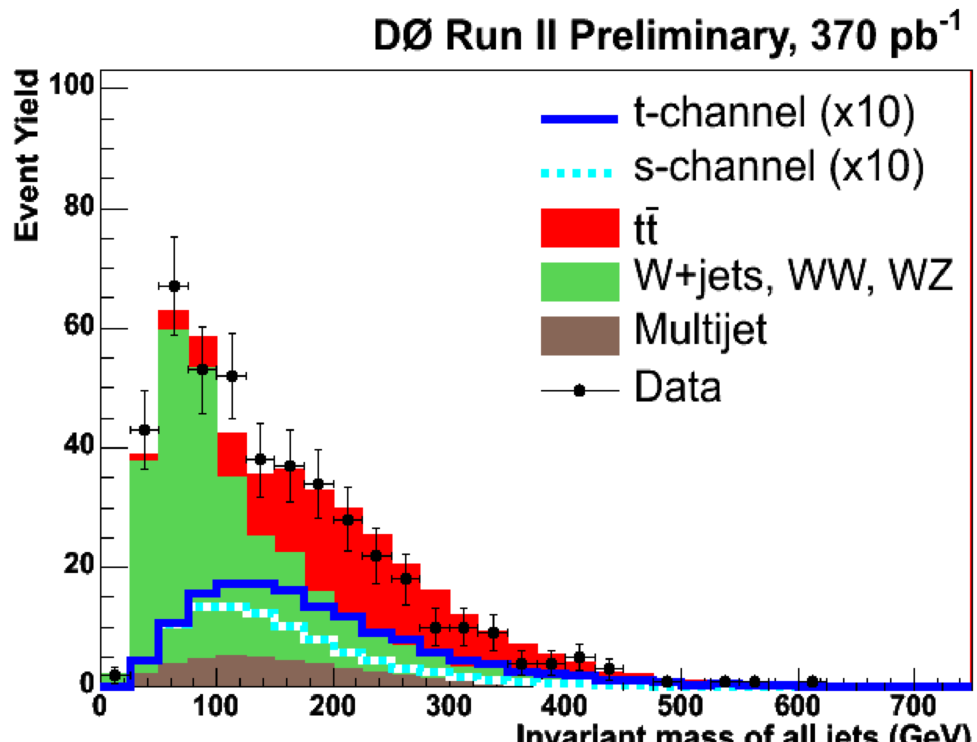


More distributions





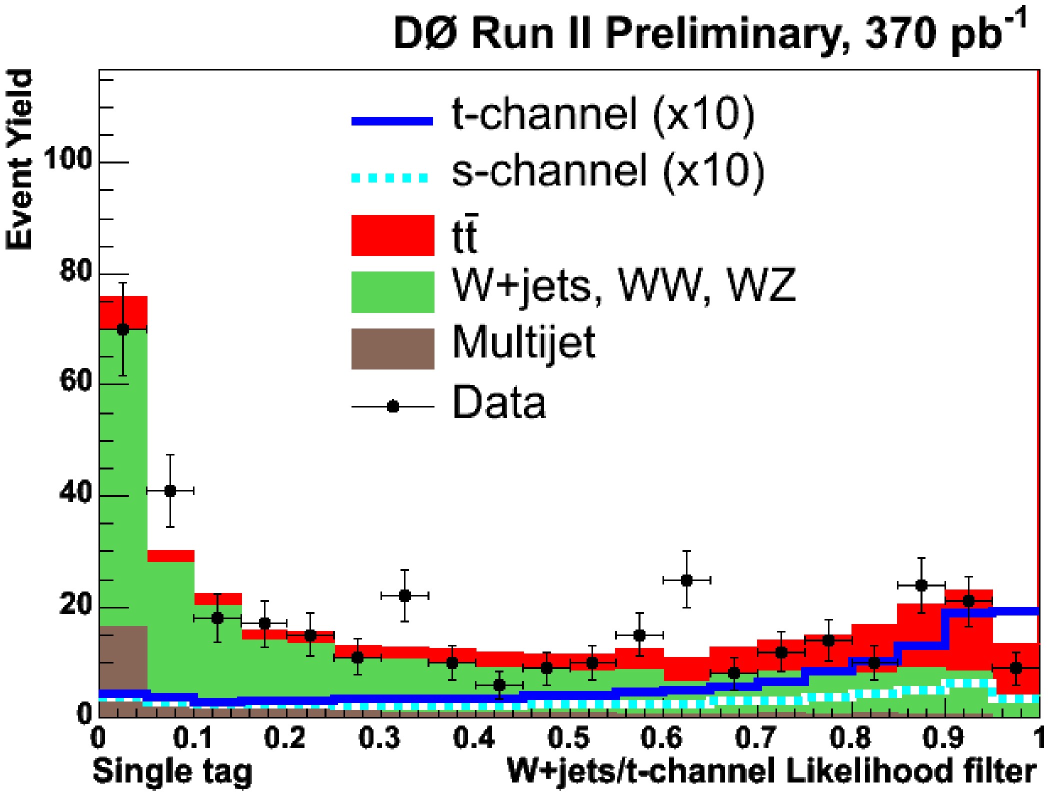


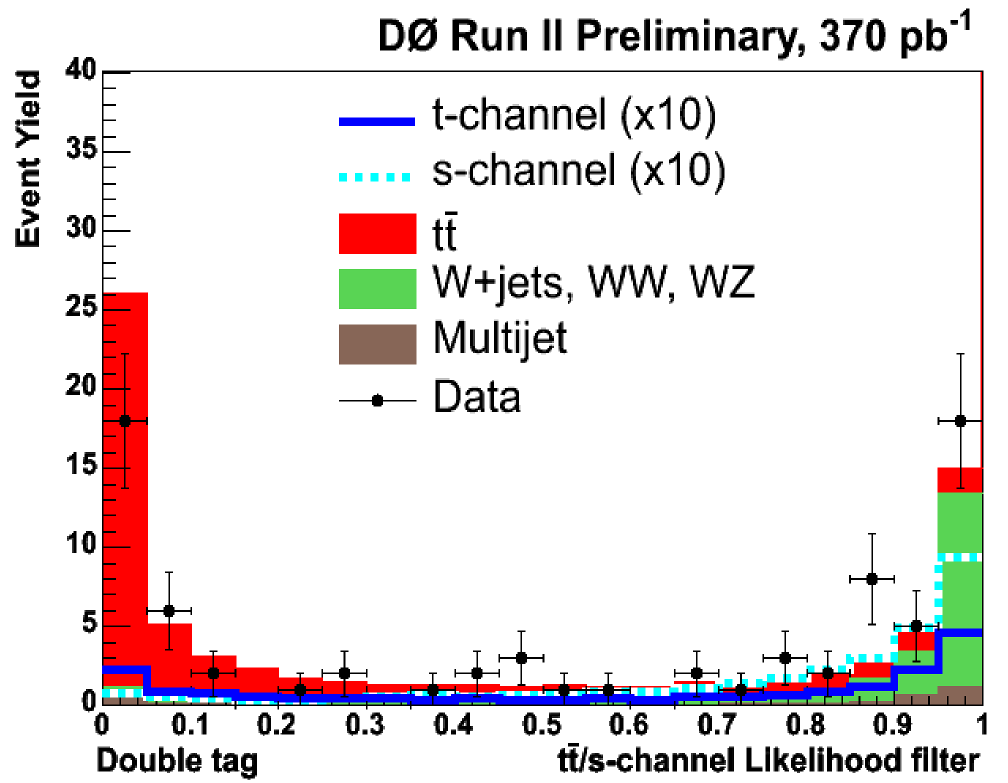


Likelihood Discriminants

$$L = \frac{P_{signal}(\vec{x})}{P_{signal}(\vec{x}) + P_{background}(\vec{x})} \quad ; \quad P(\vec{x}) = \prod_{vars} P(x_i)$$

- Build Wbb and tt discriminants for s- and t-channel, each for e, μ and single and double tags: a total of 16 likelihoods
- Use one tight and one loose tagged jets in the double tag sample
- Between 7-10 variables in the likelihoods
- Achieved similar sensitivity to Neural Networks

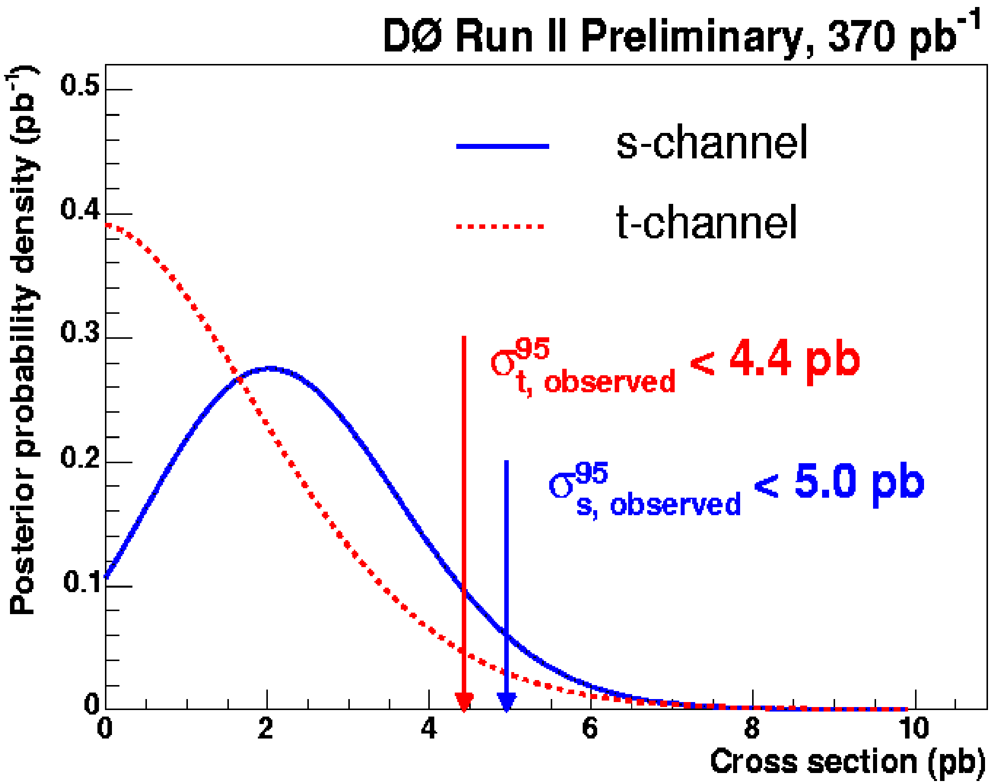




3) Results

- Use 2D histograms as input for binned likelihood
- Including bin-by-bin systematics and correlations
- Uncertainties: b-tag 6-17%, JES 5%, trigger 5%, theory σ18%
- Set 95% CL upper cross section limit with Bayesian approach to combine channels (e, μ and 1 tag, 2 tags)
- Multivariate analysis + shape information from output:

  → **factor 2 better than simple cuts**

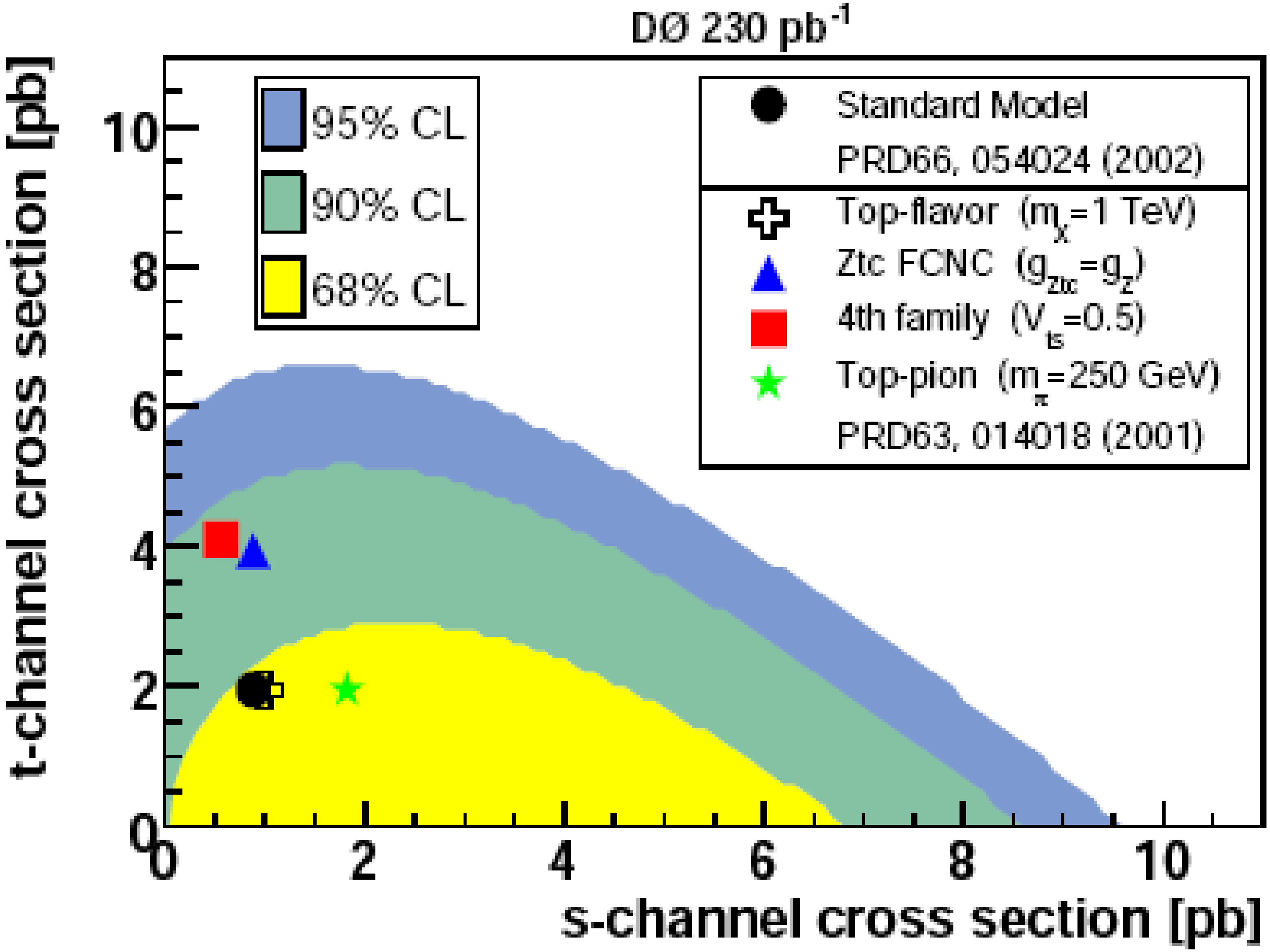


- Current upper limits:

		s-channel	t-channel
DØ Lhood 370/pb	Observed	**5.0**	**4.4**
	Expected	3.3	4.3
CDF NN 695/pb	Observed	**3.2**	**3.1**
	Expected	3.7	4.2

Model independent limits

- Relax assumption of SM cross sections for s- and t-channel
- Study sensitivity to models beyond SM in the (σ_s,σ_t) plane



Conclusions & outlook

- DØ embarked in new reconstruction with better CAL calibration, Jet Energy Scale, Data-MC agreement,...
- Will have single top results with 1 fb^{-1} soon!
- Improved b-tagging and new Matrix Element analysis
- Getting very close to observation!

New result March 06: m(W')>650 GeV @ 95% CL

Working on FCNC search

Latest DØ and CDF results:

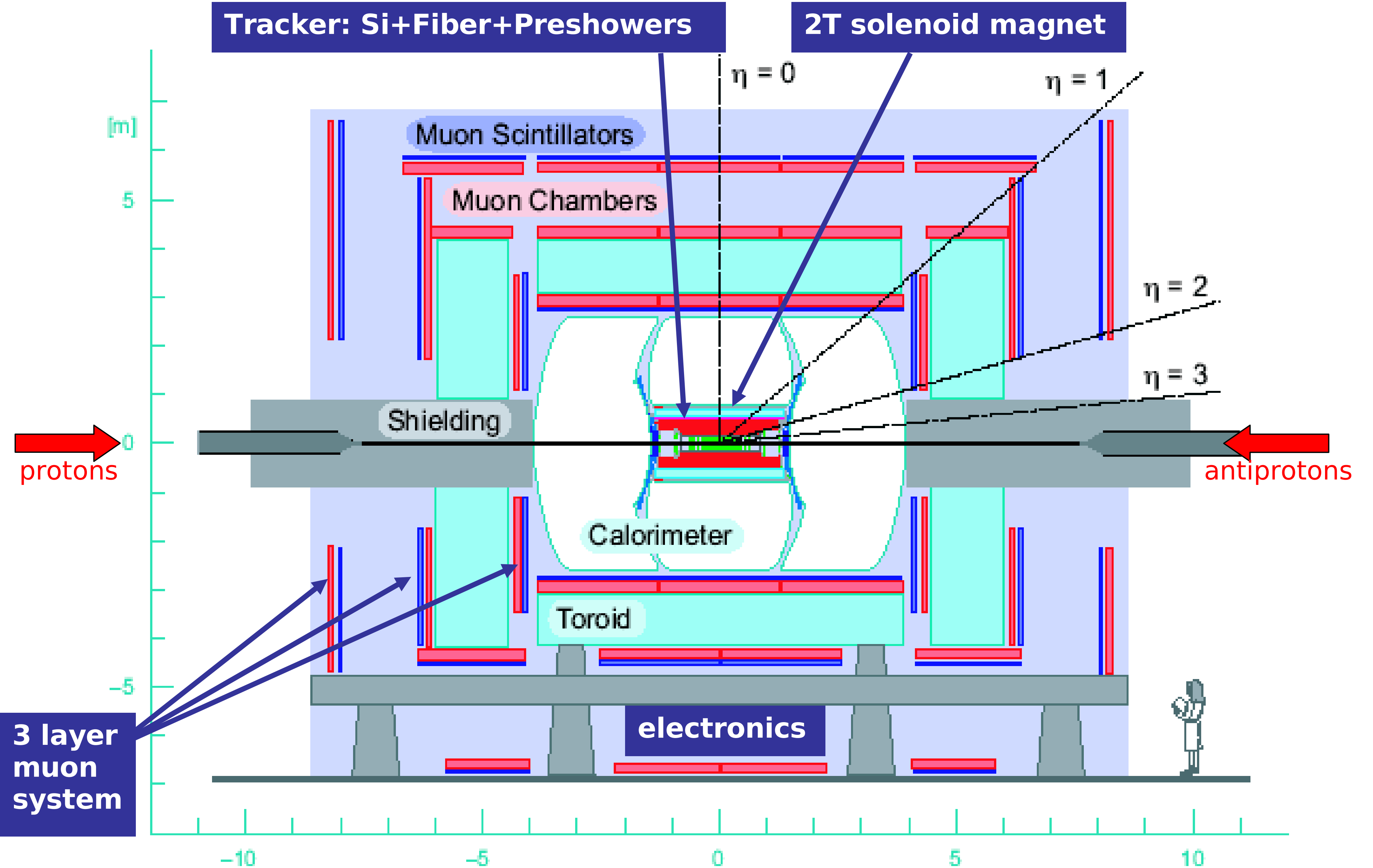
http://www-d0.fnal.gov/Run2Physics/WWW/results.htm

http://www-cdf.fnal.gov/physics/physics.html

TeV4LHC: http://conferences.fnal.gov/tev4lhc/

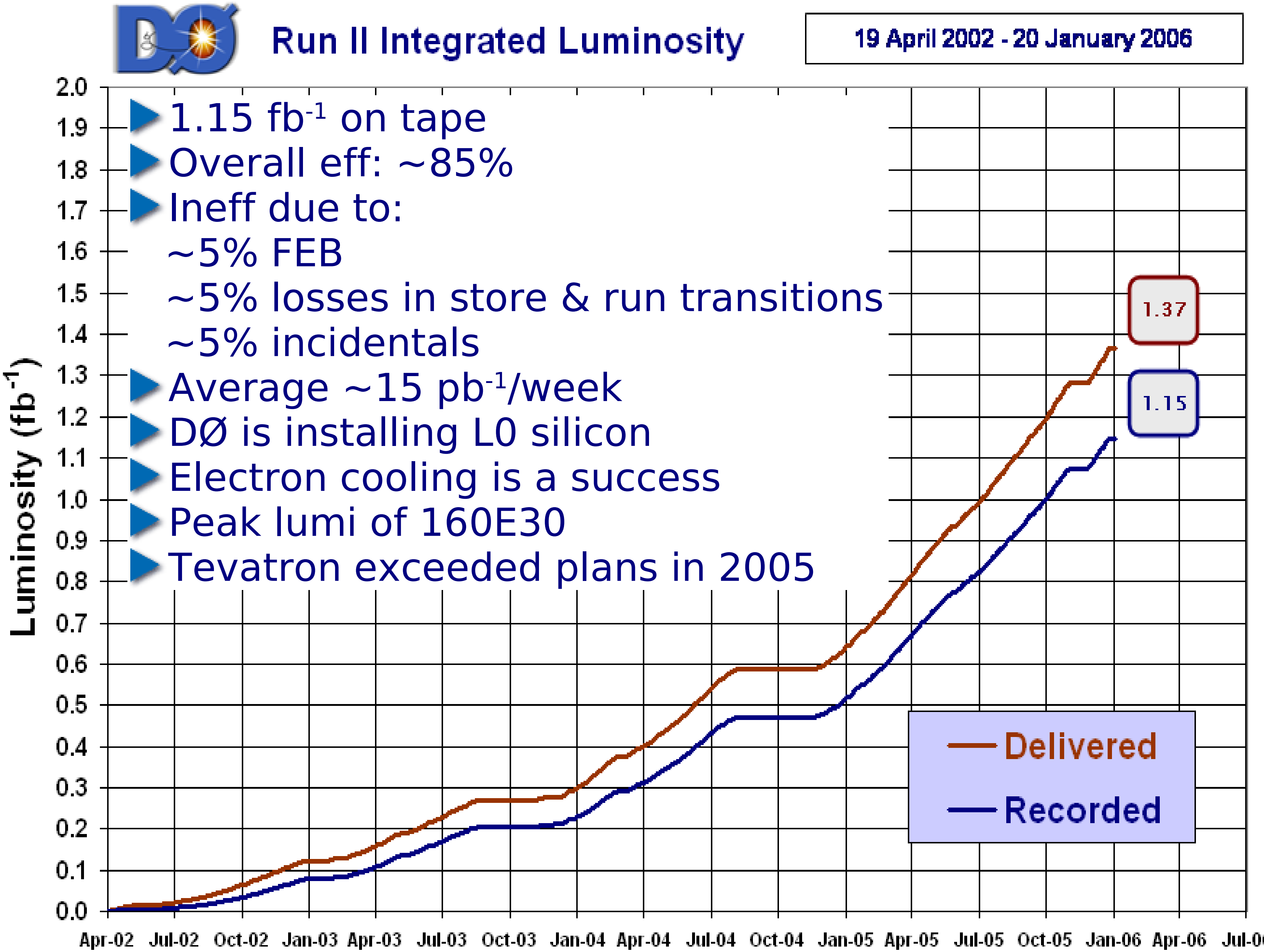
Extra Slides

DØ for Run II



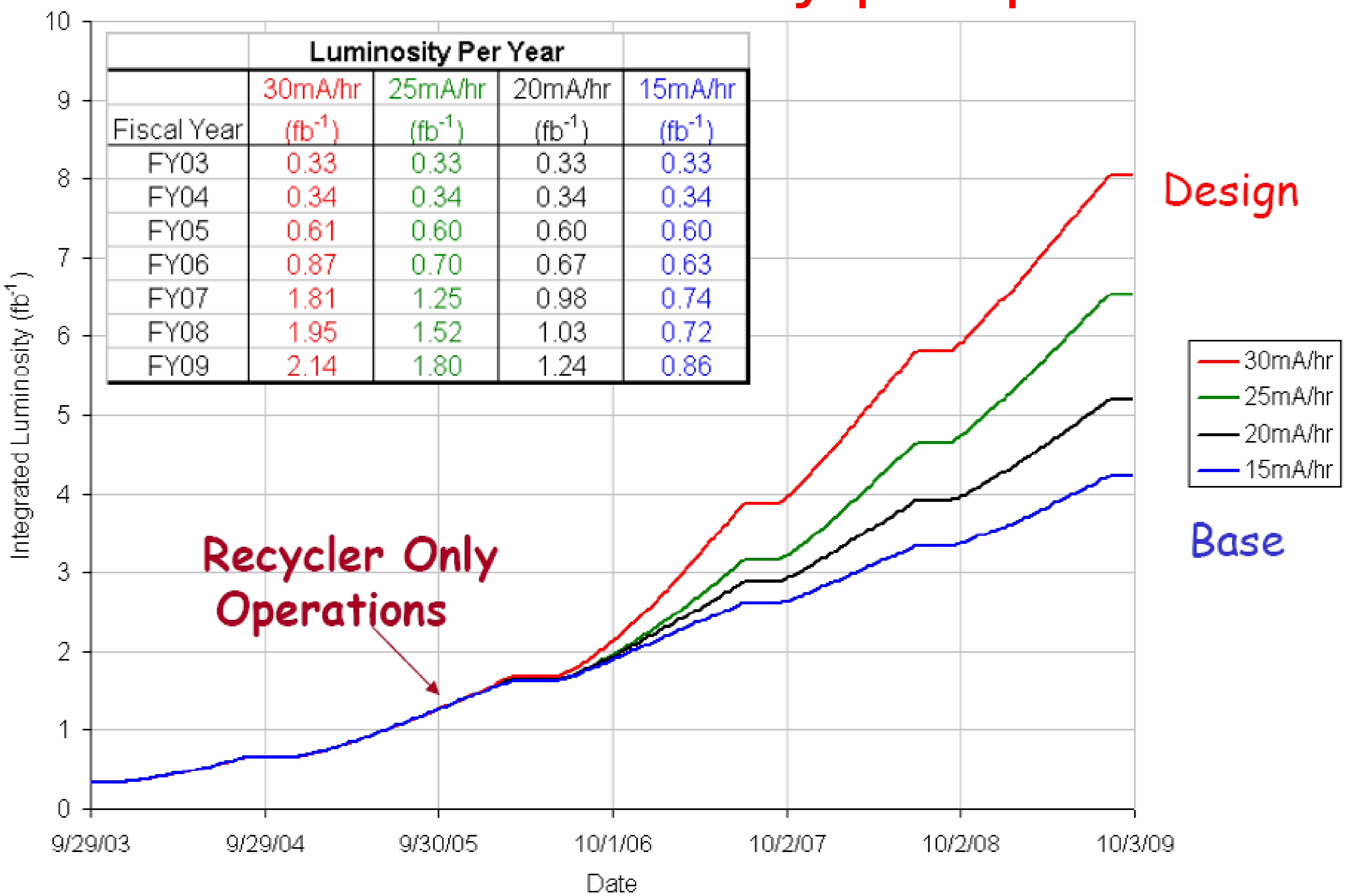
Tracker: Si+Fiber+Preshowers
2T solenoid magnet
η = 0
η = 1
η = 2
η = 3
[m]
5
0
−5
Muon Scintillators
Muon Chambers
Shielding
Calorimeter
Toroid
electronics
protons
antiprotons
3 layer muon system
−10
−5
0
5
10

Data taking



- 1.15 fb^{-1} on tape
- Overall eff: ~85%
- Ineff due to:
  ~5% FEB
  ~5% losses in store & run transitions
  ~5% incidentals
- Average ~15 pb^{-1}/week
- DØ is installing L0 silicon
- Electron cooling is a success
- Peak lumi of 160E30
- Tevatron exceeded plans in 2005

Tevatron luminosity prospects



Luminosity Per Year				
Fiscal Year	30mA/hr (fb^{-1})	25mA/hr (fb^{-1})	20mA/hr (fb^{-1})	15mA/hr (fb^{-1})
FY03	0.33	0.33	0.33	0.33
FY04	0.34	0.34	0.34	0.34
FY05	0.61	0.60	0.60	0.60
FY06	0.87	0.70	0.67	0.63
FY07	1.81	1.25	0.98	0.74
FY08	1.95	1.52	1.03	0.72
FY09	2.14	1.80	1.24	0.86

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}

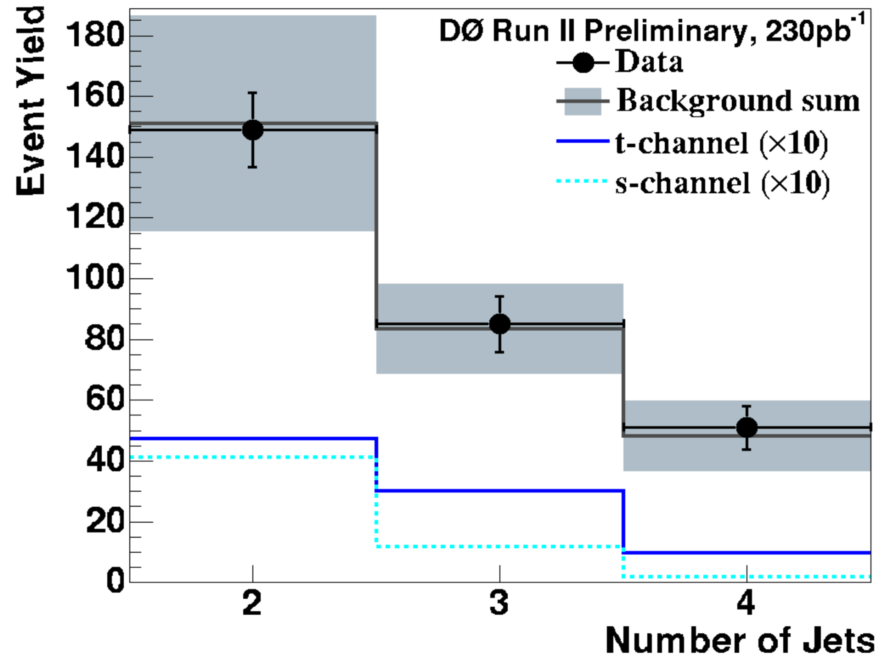
- Normalize the MC Wjj and Wbb samples to the real ℓ yield found in data, after correcting for the presence of tt and diboson events:

$$\varepsilon_{real} N^{loose}_{real} = SF[Y(Wjj) + Y(Wb\bar{b})] + Y(t\bar{t}) + Y(WW) + Y(WZ)$$ SF=1.05

- The sum Y(Wjj)+Y(Wbb) is done according to the NLO ratio of cross sections (MCFM) before b-tagging → 25% uncertainty
- Then apply b-tagging
  - Greatly reduce W+jets background (Wbb ~1% of Wjj)
  - Shift distributions, changes flavor composition

Systematic Uncertainites

Monte Carlo Systematic Uncertainties	
Theory cross sections	15 %
SVT modeling, single (double) tag	10 %(20 %)
Jet Energy Scale	10 %
Trigger Modeling	6 %
Jet Fragmentation	6 %
Jet ID	5 %
ℓ ID	5 %



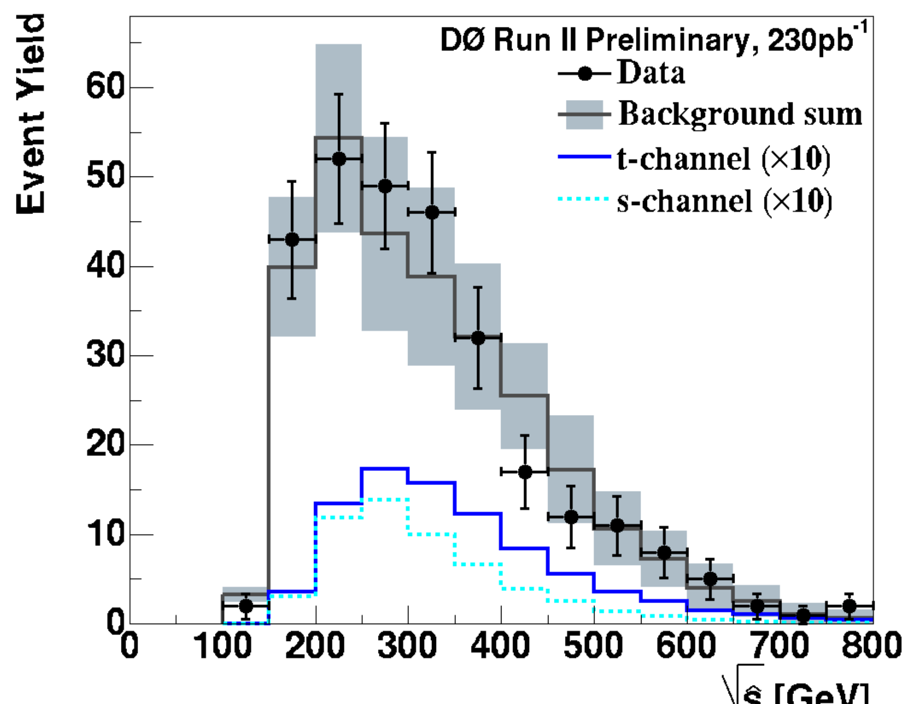
- Some systematic uncertainties also affect shape:

  JES, b-tag and trigger modeling

- Total uncertainty:

	1 tag	2 tags
Signal acceptance	15%	25%
Background sum	10%	26%

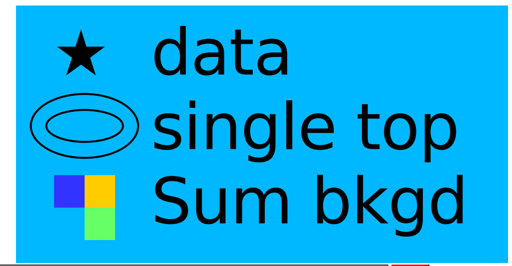
- Result is statistics limited

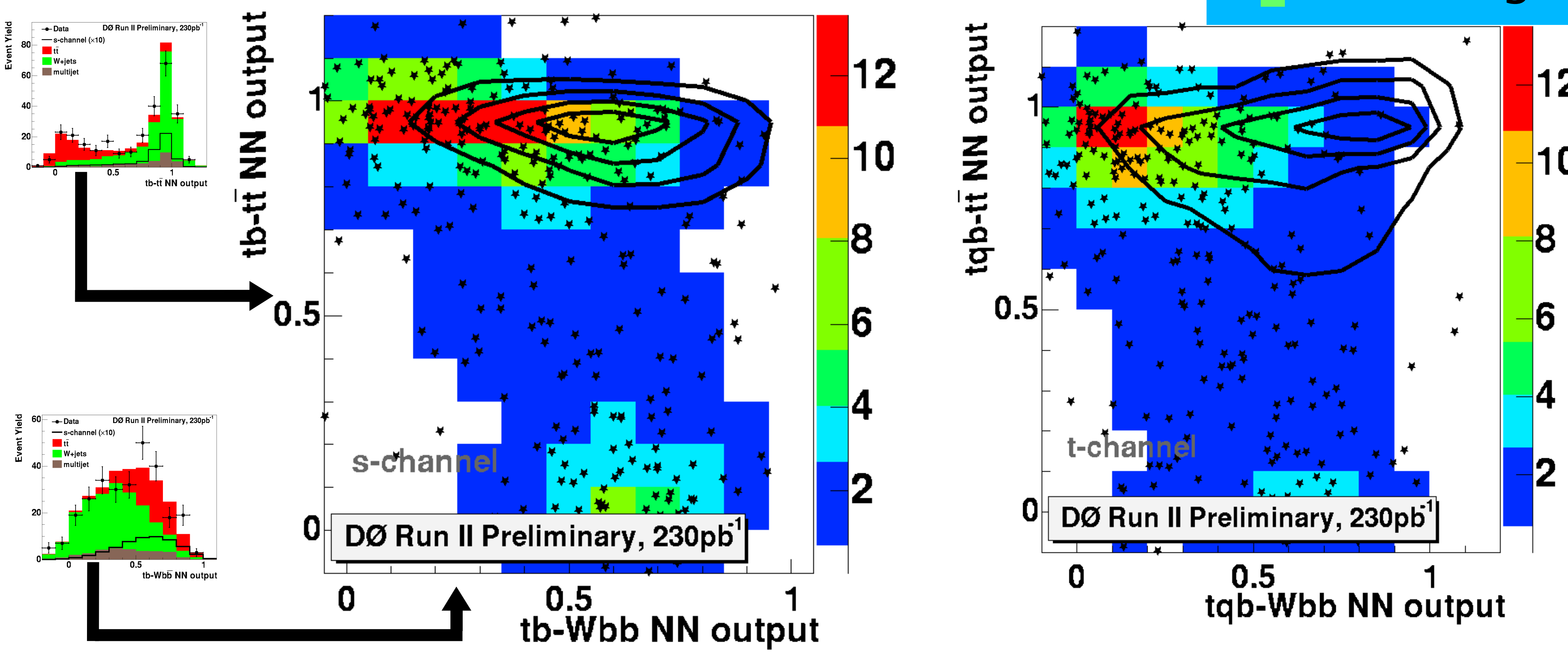


Limits from binned likelihood

- No evidence for single top signal
- Set 95% CL upper cross section limit with Bayesian approach
- Use 2D histograms as input for binned likelihood
- Including bin-by-bin systematics and correlations

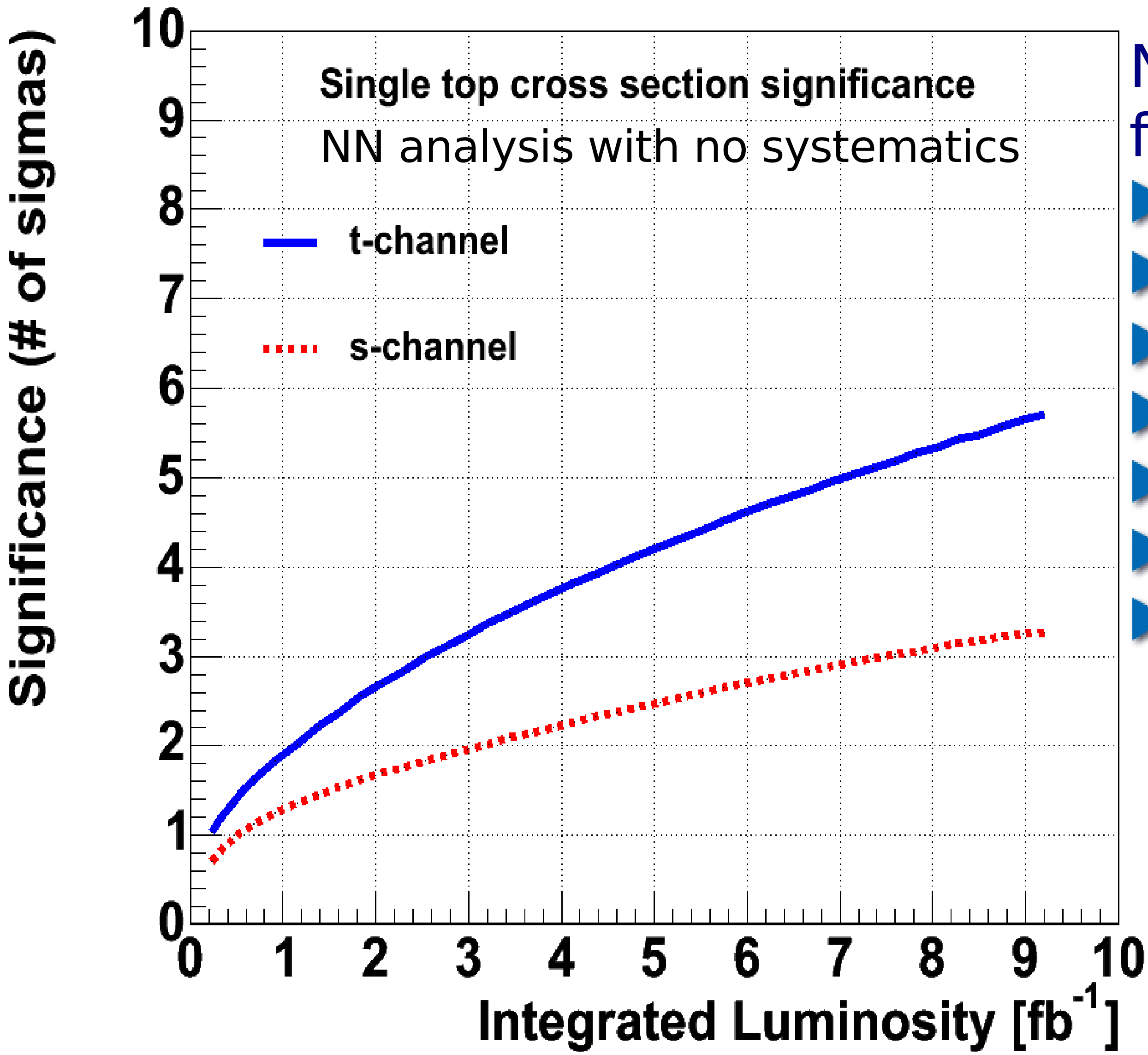
Used for DT and NN analyses
Cut-based analysis uses likelihood from event count





Sensitivity

With current analysis, we would need several fb^{-1} for an observation of SM single top



Need to work on many fronts to improve:

- Trigger efficiency
- Object ID: e, μ, jet, b
- Jets resolution
- Add more channels
- Background estimation
- Reduction of systematics
- Bkgnd-signal separation

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

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

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

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

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

P(A|B): "posterior" probability density $P(n_{predicted}|N_{observed})$

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

Preparing the way for the LHC

Studies at the Tevatron will help the LHC:

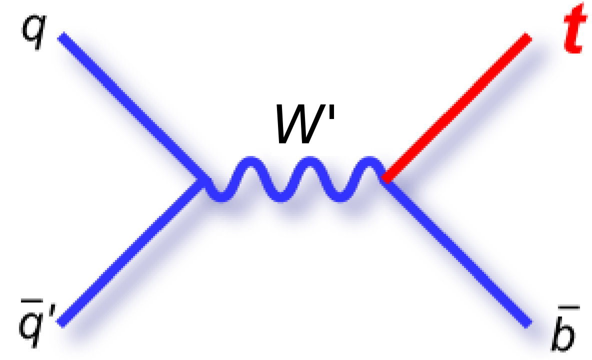
- Wbb measurement (will also help WH search) (DØ: hep-ex/0410062)
  - Current limit at 4.6 pb for $p_T(b)$>20GeV
- In general, W+jets background determination techniques
  - tt will be main background, but large uncertainties come from W+jets
  - Effect of jet vetoes (N_{jet}=2), check other methods planned in LHC analyses
- Study charge asymmetries (Bowen, Ellis, Strassler: hep-ph/0412223)
  - Signal shows asymmetry in ($Q_\ell \times \eta_j$, $Q_\ell \times \eta_\ell$) plane at TeV
- Study kinematics of forward jets in t-channel (WW→H at LHC)
- Even measure asymmetry in production rate (Yuan: hep-ph/9412214) (probe CP-violation in the top sector):

$$A_t = \frac{\sigma(p\bar{p} \to tX) - \sigma(p\bar{p} \to \bar{t}X)}{\sigma(p\bar{p} \to tX) + \sigma(p\bar{p} \to \bar{t}X)}$$

TeV4LHC workshop report to appear soon

W' search in single top events

- Heavy W' boson (SM-like coupling to fermions)
- Interference between SM s-channel and W'
- SM t-channel single top as background
- Discriminant distribution the invariant mass of all objects in the event
- Consider $400 < \sqrt{\hat{s}} < 1000$ GeV window
- Use only events with 2 or 3 jets



$\sigma(m_{W'}=600) < 1.7$ pb
$200 > m_{W'} > 650$ GeV