W + *b*-jet Production at CDF: A Prerequisite to Higgs

Dr. Christopher Neu

Penn
University of Pennsylvania

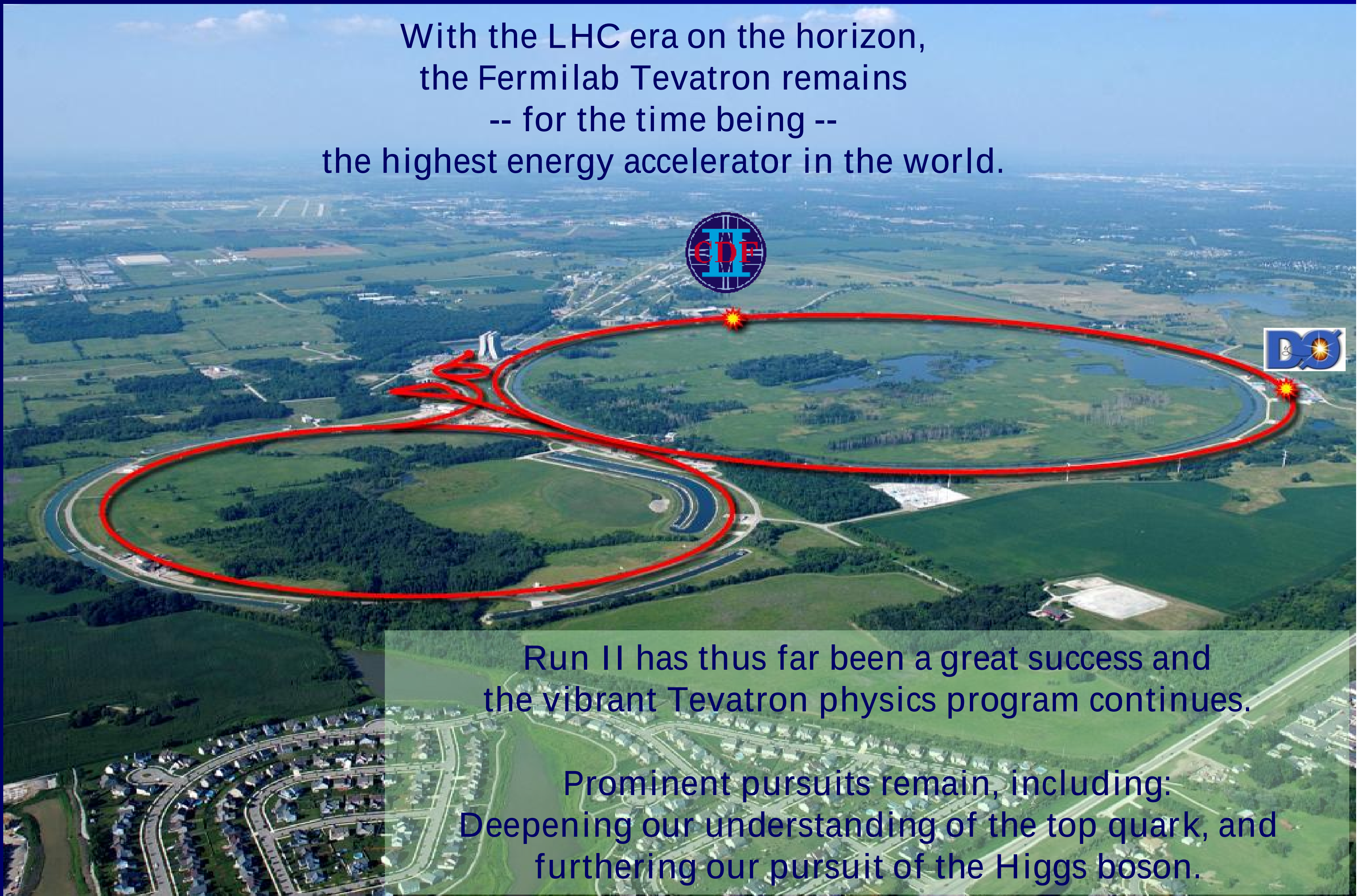
Outline

- Motivation
- Measurement Definition
- Experimental apparatus
- Measurement strategy
- Results

Outline

- *Motivation*
- Measurement Definition
- Experimental apparatus
- Measurement strategy
- Results

The Current Energy Frontier



With the LHC era on the horizon,
the Fermilab Tevatron remains
-- for the time being --
the highest energy accelerator in the world.

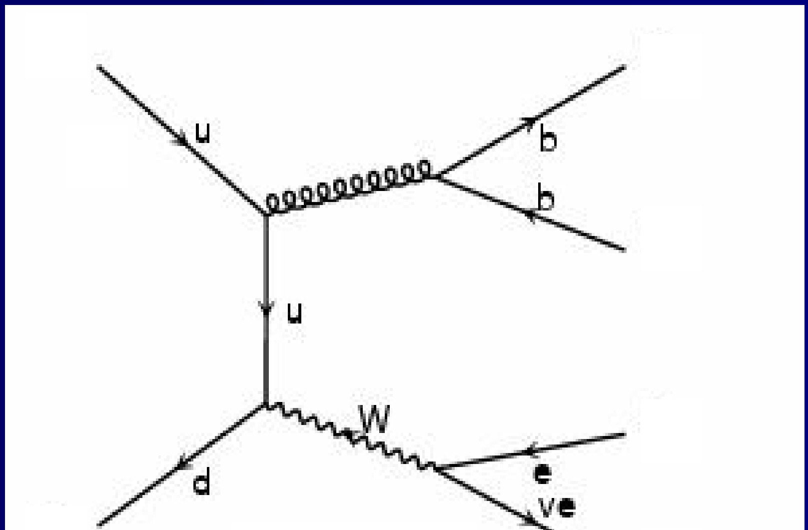
Run II has thus far been a great success and
the vibrant Tevatron physics program continues.

Prominent pursuits remain, including:
Deepening our understanding of the top quark, and
furthering our pursuit of the Higgs boson.

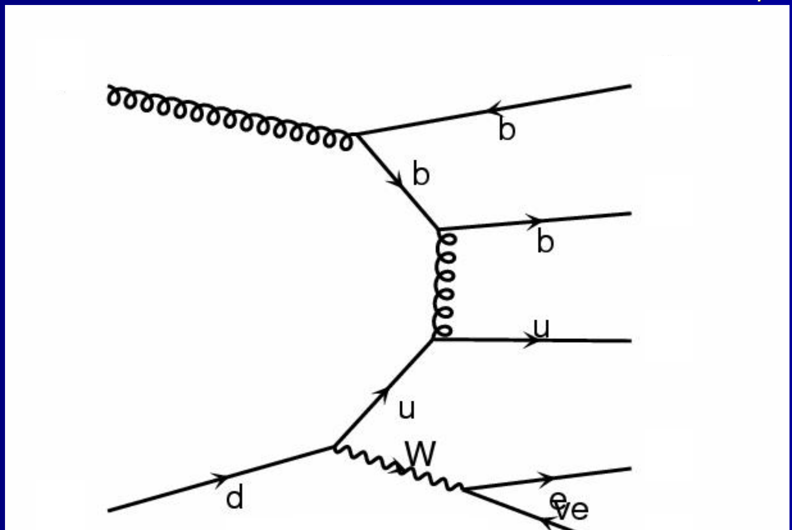
Why Study *W*+*b*-jet Production?

- First, a definition:
  - *W*+*b*-jets refers to QCD production of *b*-jets in events with a *W* boson

Made with MadGraph



$O(\alpha_s^2)$



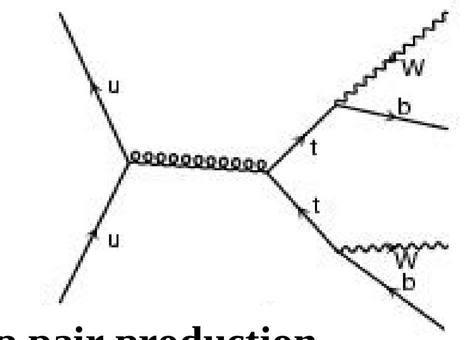
$O(\alpha_s^3)$

Examples of *W*+*b*-jets production at tree level

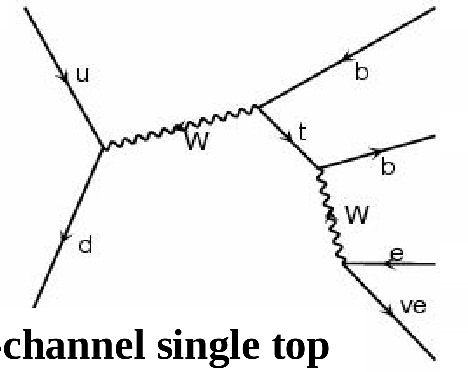
- Why is *W*+*b*-jets interesting?
  - Consider some primary Run II targets…

Signatures with *W*'s and *b*'s

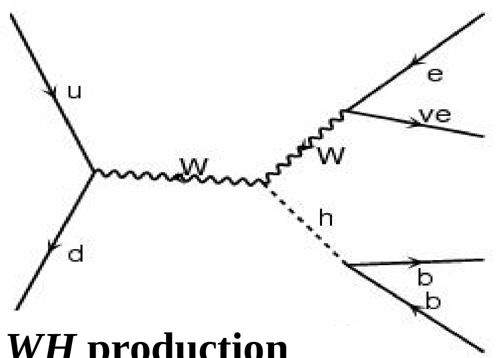
- Rich top physics program at the Tevatron
  - BR ($t \to Wb$) ~ 100%
  - top pair production
    - $p\,\bar{p} \to t\,\bar{t} \to W^+\,b\,W^-\,\bar{b}$
    - Production cross section = ~7 pb
- Current hot topic: single top production
  - $p\,\bar{p} \to W^* \to t\,\bar{b} \to W^+\,b\,\bar{b}$: ~0.3 pb
  - $p\,\bar{p} \to t\,b\,q \to W^+\,b\,\bar{b}\,q$: ~0.6 pb
  - Insight on $|V_{tb}|$
- The Search for the Higgs
  - Promising Tevatron production mode: $p\,\bar{p} \to W^* \to W^\pm\,H$: ~0.1-0.2 pb
  - Higgs decays to *b* quarks if its mass is low: BR($H \to b\,\bar{b}$) = ~70% for M_H = 120 GeV/c^2



top pair production



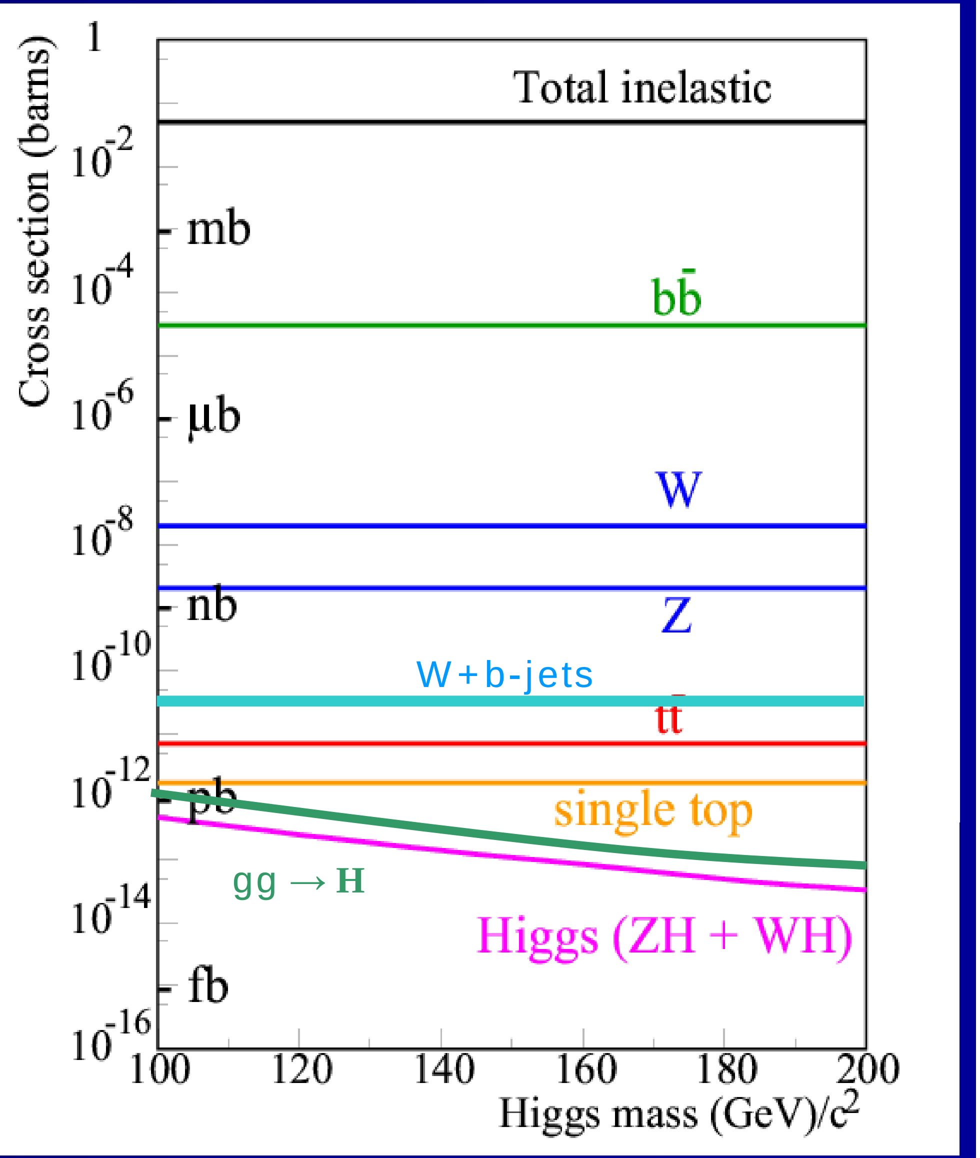
s-channel single top



***WH* production**

Importance of *W*+*b*-jet Production

- Common trait of those prominent signatures:
  - *W*'s and *b*'s
- *W*+*b*-jet production casts a long shadow:
  - Largest background source
  - Rate for *W*+*b*-jets exceeds these others significantly
  - Theory prediction: 10 - 15 pb
- Good understanding of the *W*+*b*-jets process is essential for success

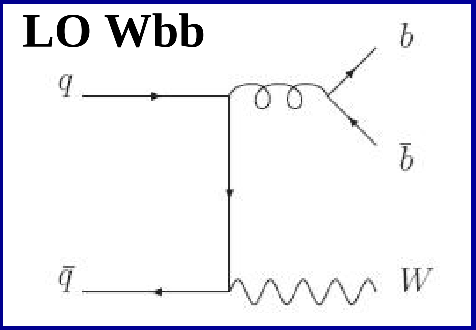


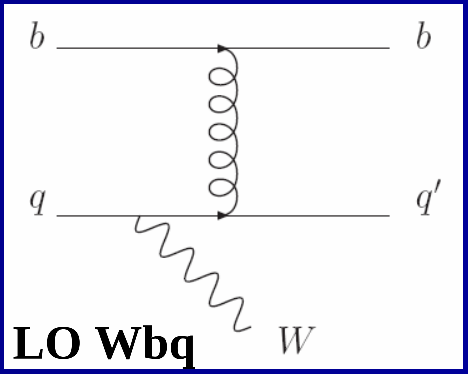
W+*b*-jets: Theory

- Several theory groups have tackled the *W*+*b*-jets calculation:
- Campbell, Ellis, Maltoni, Willenbrock at LO & NLO using MCFM (hep-ph/0611348)
  - Several processes make up *W*+*b*-jets
  - Categorized according to outgoing partons
  - Wbb and Wbq categories ~80% of total NLO

	Inclusive Cross Section (pb)	
	LO	NLO
Wbb	4.96	6.28
Wbq	2.12	5.08

p_T>15 GeV/c^2, $|\eta|$ < 2.0
for all outgoing b, q
in this calculation





- Mangano, et al., at tree level – original motivation for ALPGEN (hep-ph/0108069)
  - Qualitative agreement w/ LO MCFM results
  - Wide use of ALPGEN at CDF for *W*+jets shapes (*W*+*b*-jets, *W*+*c*-jets, *W*+LF-jets)

Example: *W*+*b*-jets Prediction in *WH* Search

- WH→ªνbb analysis needs prediction for *W*+*b*-jet yield
- Predicted rates from MC distrusted for *W*+*b*-jets
- Predictions for these events use data to set the overall scale of *W*+jets production
- Ultimate prediction is frought with systematic error
- Small *WH* signal obscured by error on the background
- We must be able to do better.

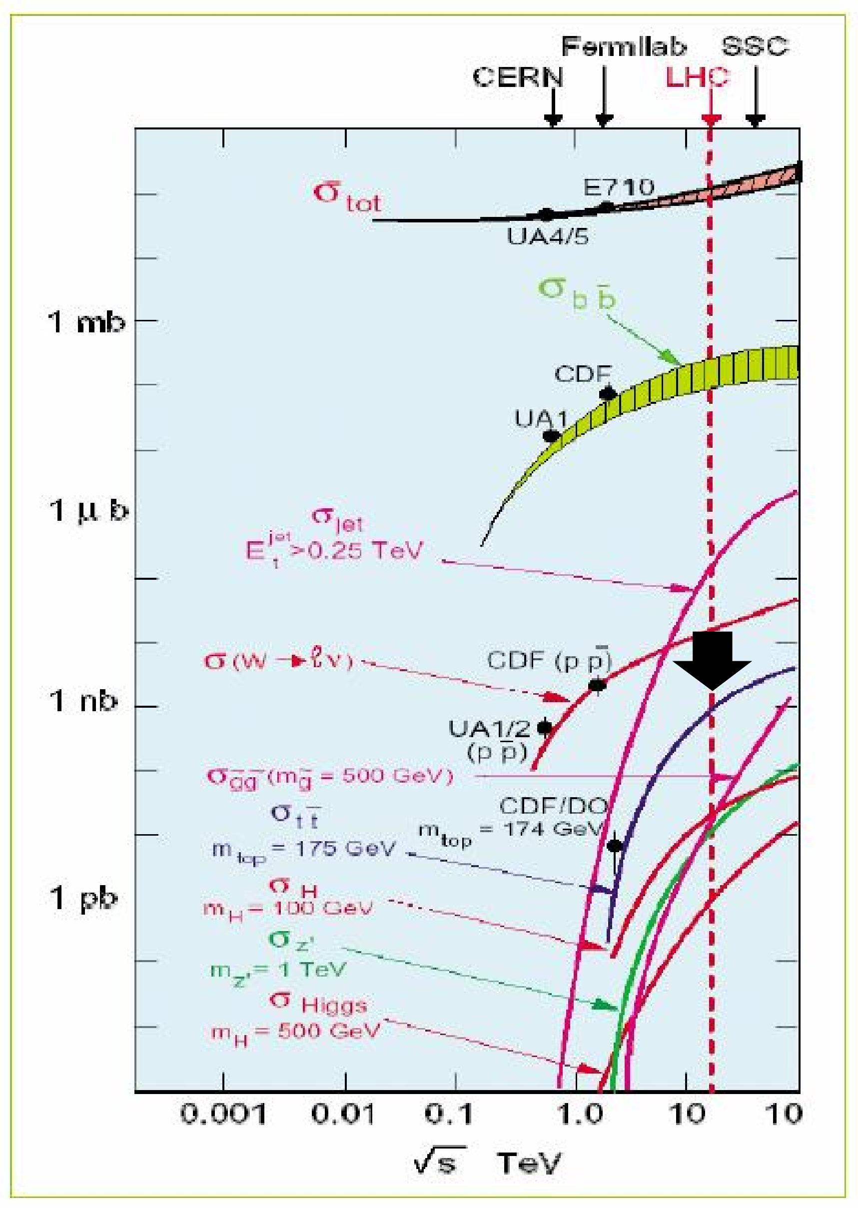
Predicted Event Yields

Jet Multiplicity	1 jet	2 jets	3 jets	≥ 4 jets
WLF	139.7 ± 27.3	53.9 ± 10.7	15.7 ± 3.1	4.2 ± 0.8
$Wb\bar{b}$	306.9 ± 106.9	144.7 ± 49.4	29.9 ± 9.7	6.4 ± 2.5
$Wc\bar{c}$	63.1 ± 22.0	43.0 ± 14.7	8.7 ± 2.8	1.9 ± 0.8
Wc	185.7 ± 47.2	34.4 ± 9.0	3.4 ± 0.9	0.6 ± 0.2
$t\bar{t}$(6.7pb)	6.9 ± 1.2	42.0 ± 6.6	84.9 ± 12.8	98.6 ± 14.3
Single Top	16.7 ± 1.8	23.5 ± 2.4	4.8 ± 0.5	0.8 ± 0.1
Diboson/$Z^0 \rightarrow \tau\tau$	11.7 ± 2.2	14.2 ± 2.3	3.9 ± 0.9	1.0 ± 0.3
non-W QCD	84.2 ± 14.1	38.9 ± 6.7	12.1 ± 2.3	5.5 ± 1.2
Total Background	814.9 ± 140.7	394.4 ± 66.6	163.4 ± 18.7	118.9 ± 14.9
Observed Events	856	421	177	139
Expected Signal Events Higgs Mass 120		1.26 ± 0.12		

Question:
Can we ***measure*** *W+b-jets and*
improve these predictions?
Ultimately ***improve the models****?*

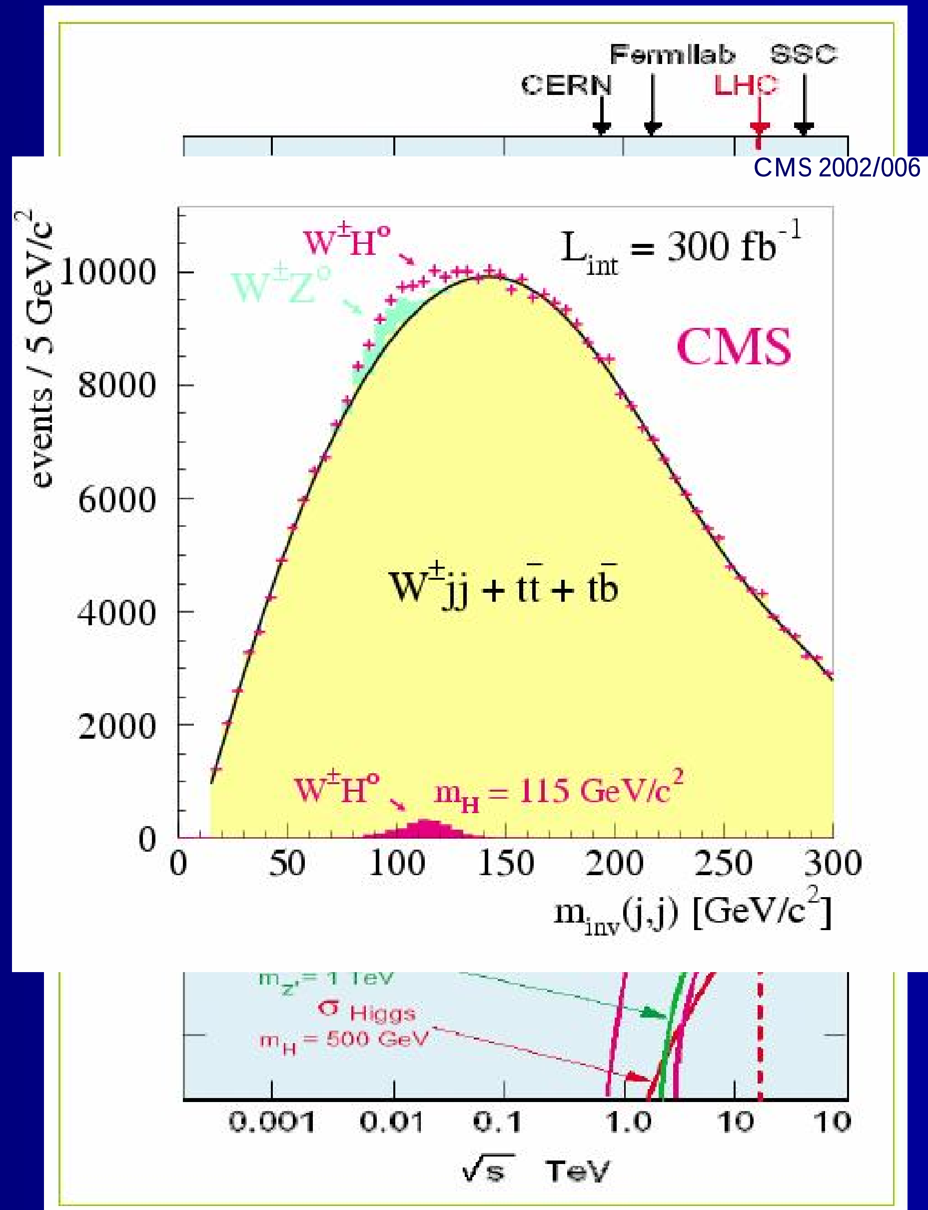
W+b-jets: Relevance for LHC

- Understanding *W+b*-jets at the Tevatron is important also for LHC



W+b-jets: Relevance for LHC

- Understanding *W+b*-jets at the Tevatron is important also for LHC
  - *WH* observation in 300/fb only possible w/ precise background modeling – mostly *W+b*-jets
  - Not a discovery mode!
  - But this channel plays a vital role in understanding a Higgs discovered through other avenues
  - Lessons learned at the Tevatron can help build better models for ATLAS and CMS



Outline

- Motivation
- *Measurement Definition*
- Experimental apparatus
- Measurement strategy
- Results

W+b-jets Cross Section Definition

- Seek to improve our understanding of W+b-jet production
- Design the analysis to focus on the sample that is most relevant for Higgs and single top searches:
  - Leptonically decaying W
  - Exactly 1 or 2 total jets

- Seek a result that is insulated from theory dependence
  - MC events are used for shape and acceptance studies
    - Restrict phase space of considered events. Require:
      - MC e or μ w/ $p_T > 20$, $|\eta| < 1.1$
      - MC ν w/ $p_T > 25$
      - Exactly 1 or 2 $E_T > 20$, $|\eta| < 2.0$ MC jets
  - Measure b jet cross section rather than inclusive event cross section
    - Models have difficulty matching the definition of "event" when requiring precisely 1 or 2 jets

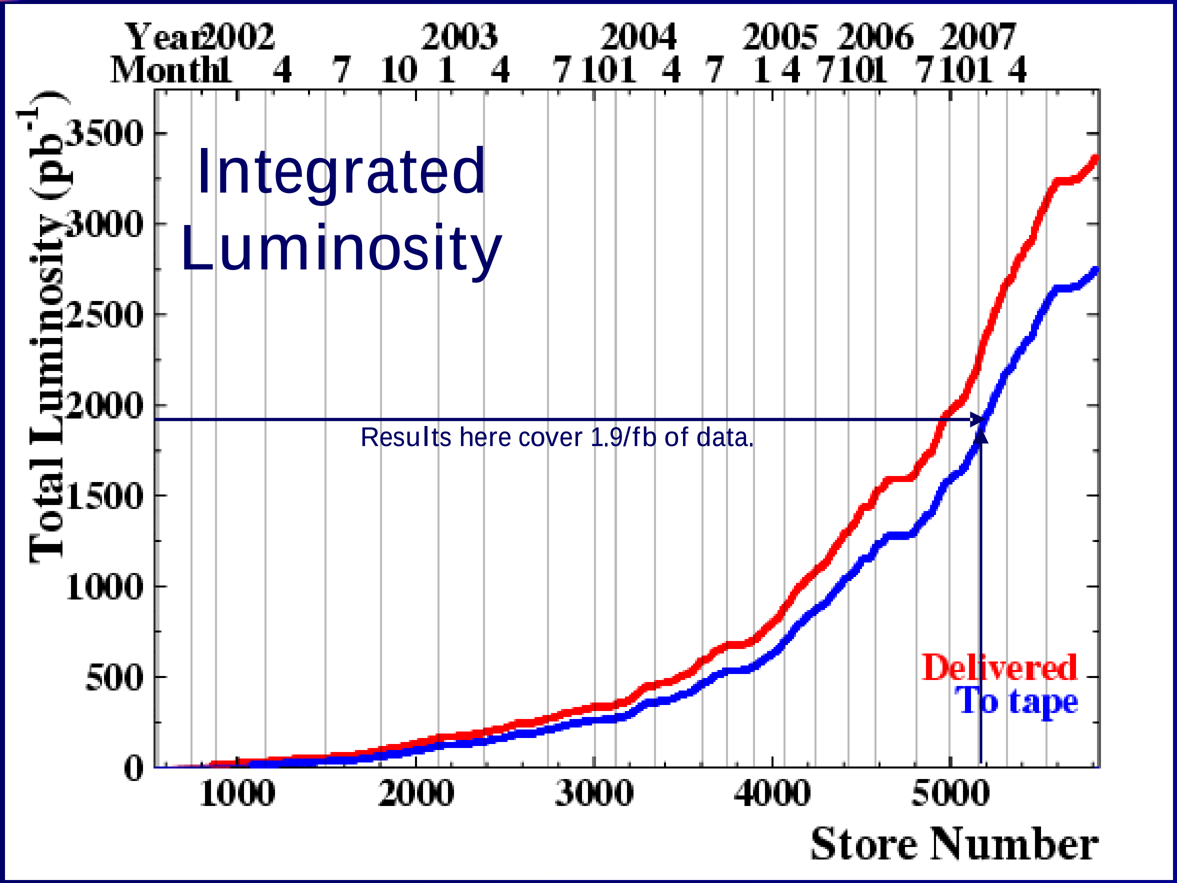
- Can calculate the b jet cross section prediction under such conditions for one model:

ALPGEN: $\boxed{\sigma_{\text{b-jets}}(W + b - \text{jets}) \times \text{BR}(W \to \ell\nu) = 0.78\ \text{pb}}$

Outline

- Motivation
- Measurement Definition
- *Experimental apparatus*
- Measurement strategy
- Results

Tevatron Performance

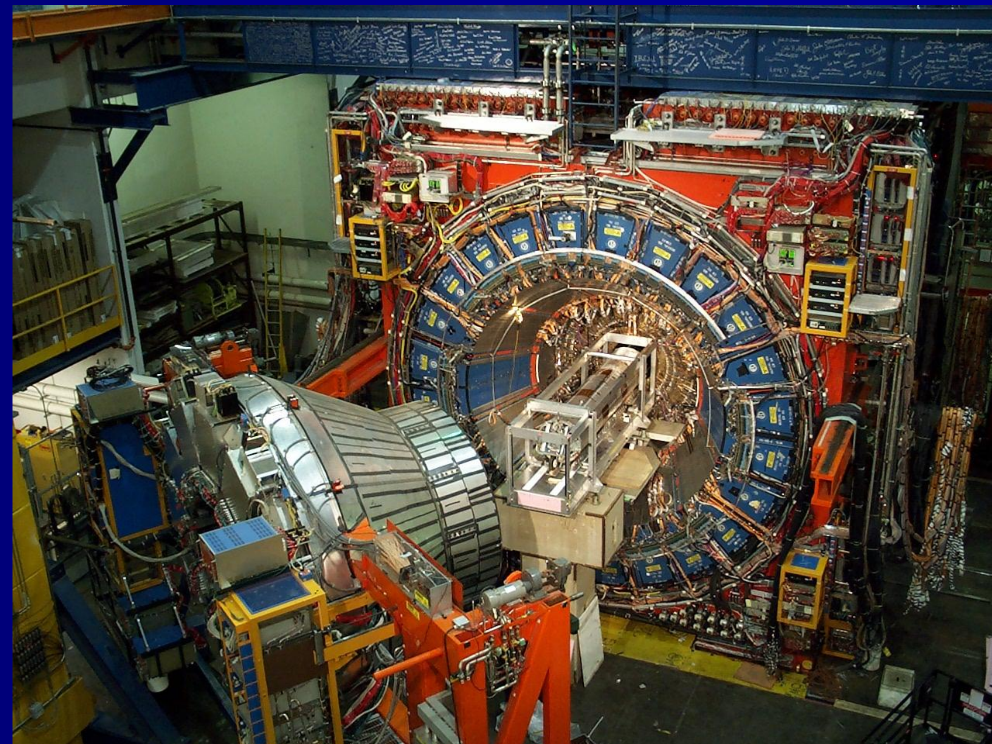


$$N_{evts} = \sigma_{evt} \cdot L_{int} \cdot A \cdot \varepsilon$$

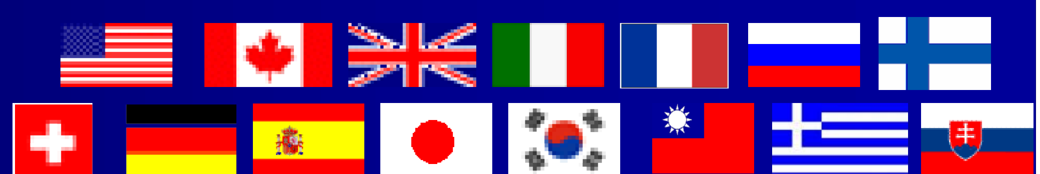
- Tevatron integrated luminosity climbing higher and higher
- Integrated lum goal is to collect 5.5-6.5/fb through 2009
- Discussions underway about running through 2010 – quite valuable
- Stable, reliable beams provided by FNAL's Accelerator Division allow us to get the most out of our experiments

The CDF Experiment

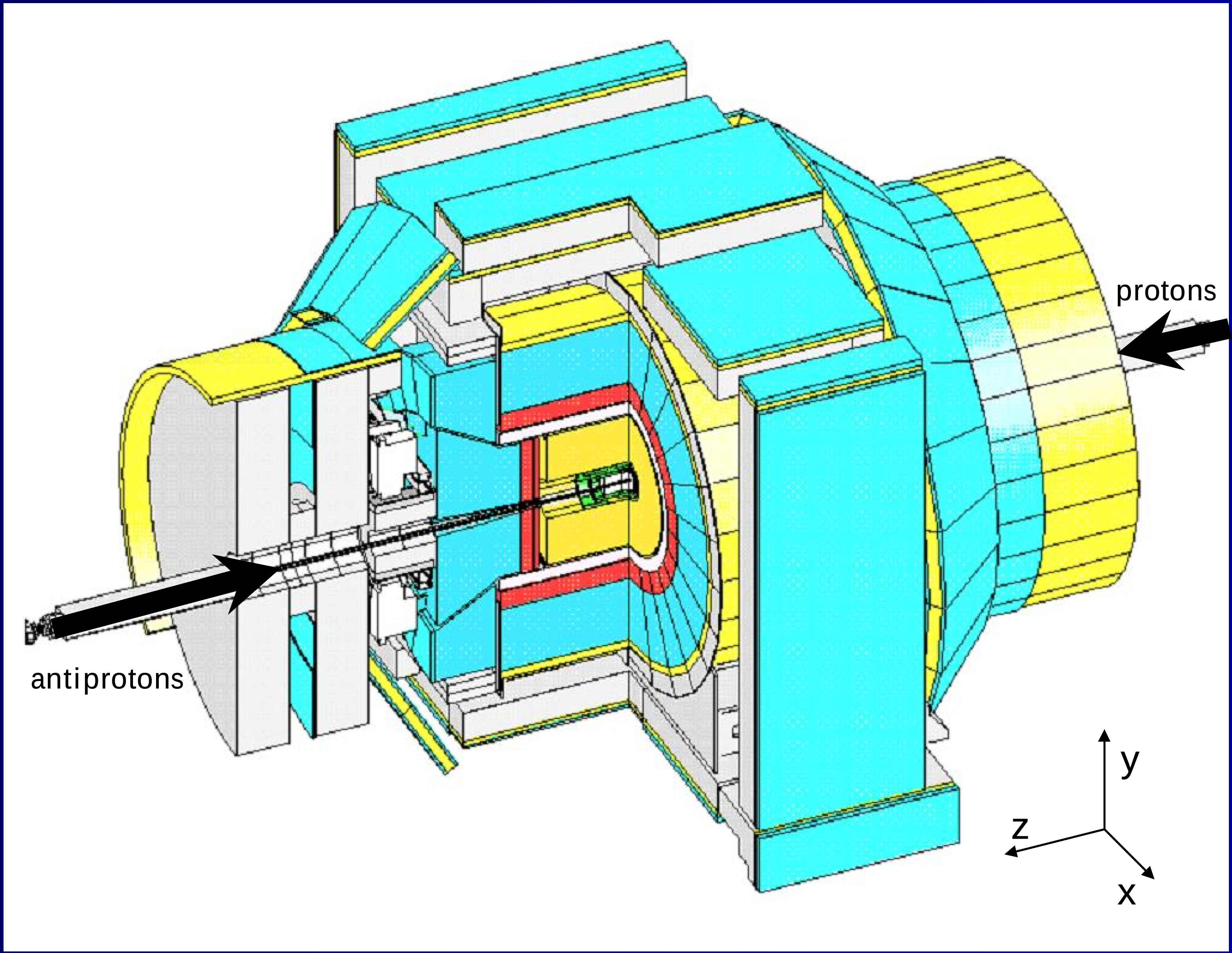
- Collider Detector at Fermilab Experiment
  - A collaborative effort
  - One of two collider physics experiments at the Tevatron
- CDF detector:
  - General-purpose
    - Can detect various decay products
    - Allows us to look for all sorts of phenomena
  - Handmade
    - Cannot buy these things at Radio Shack!



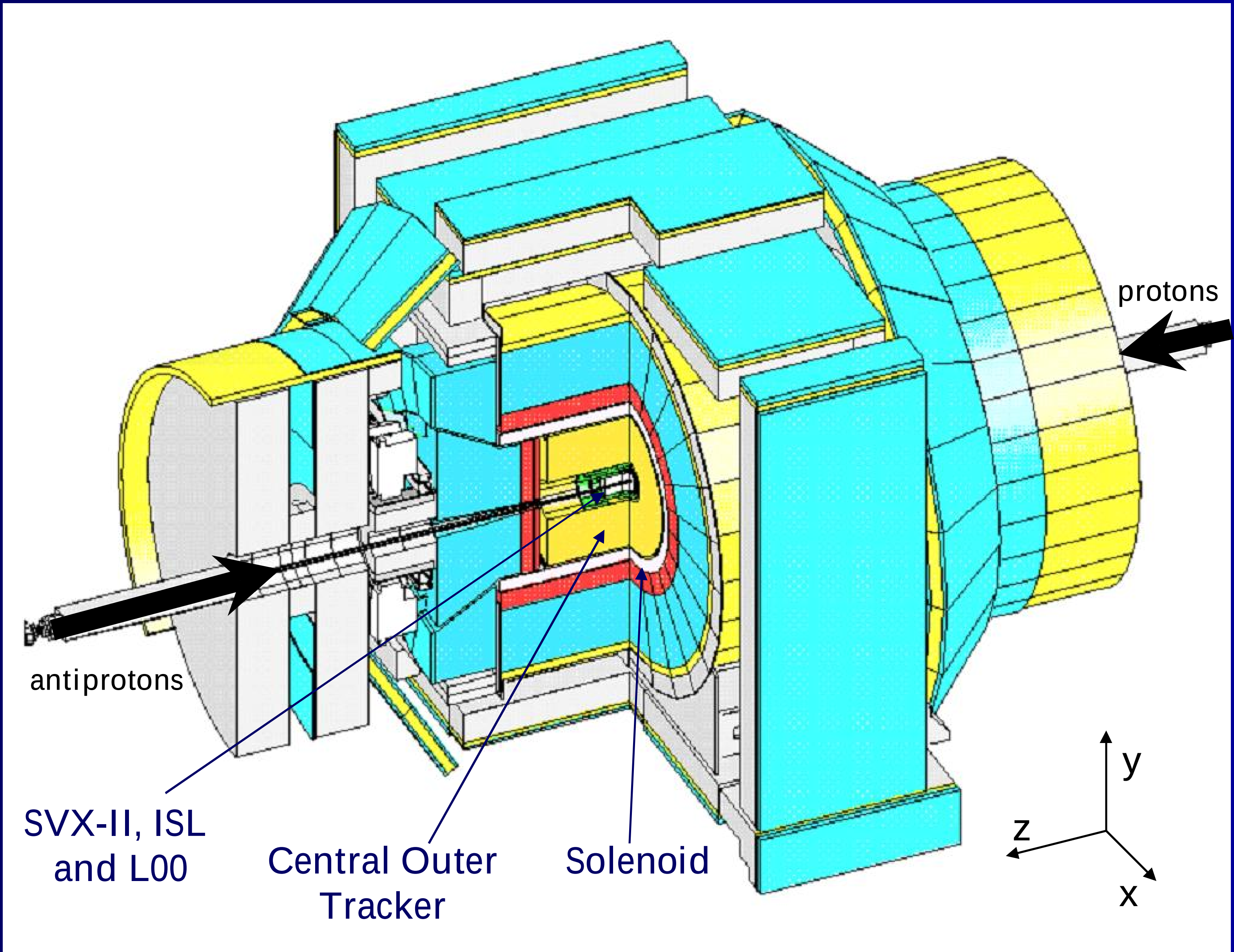
CDF Collaboration:
635 physicists
63 institutions
15 countries



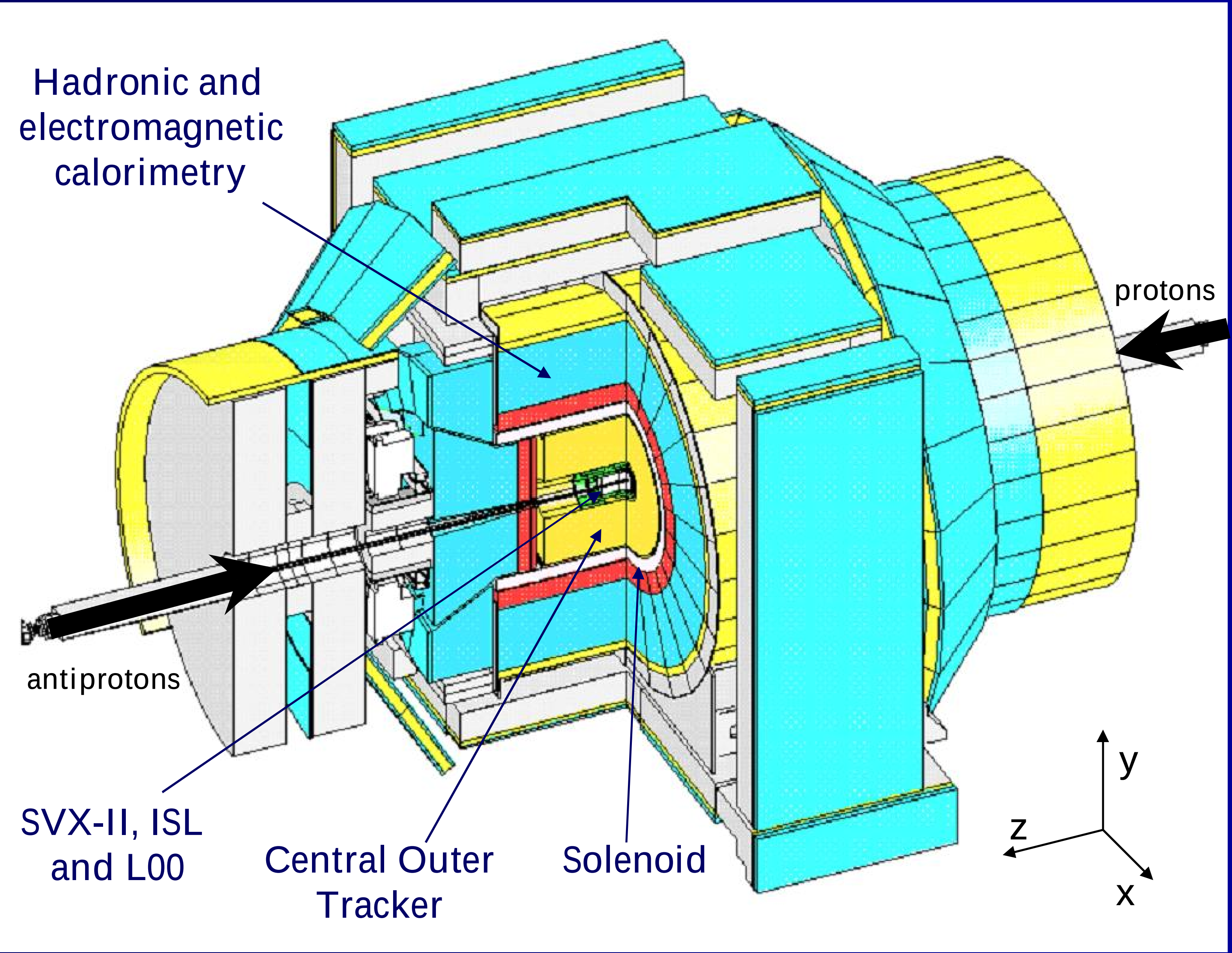
The CDF Detector



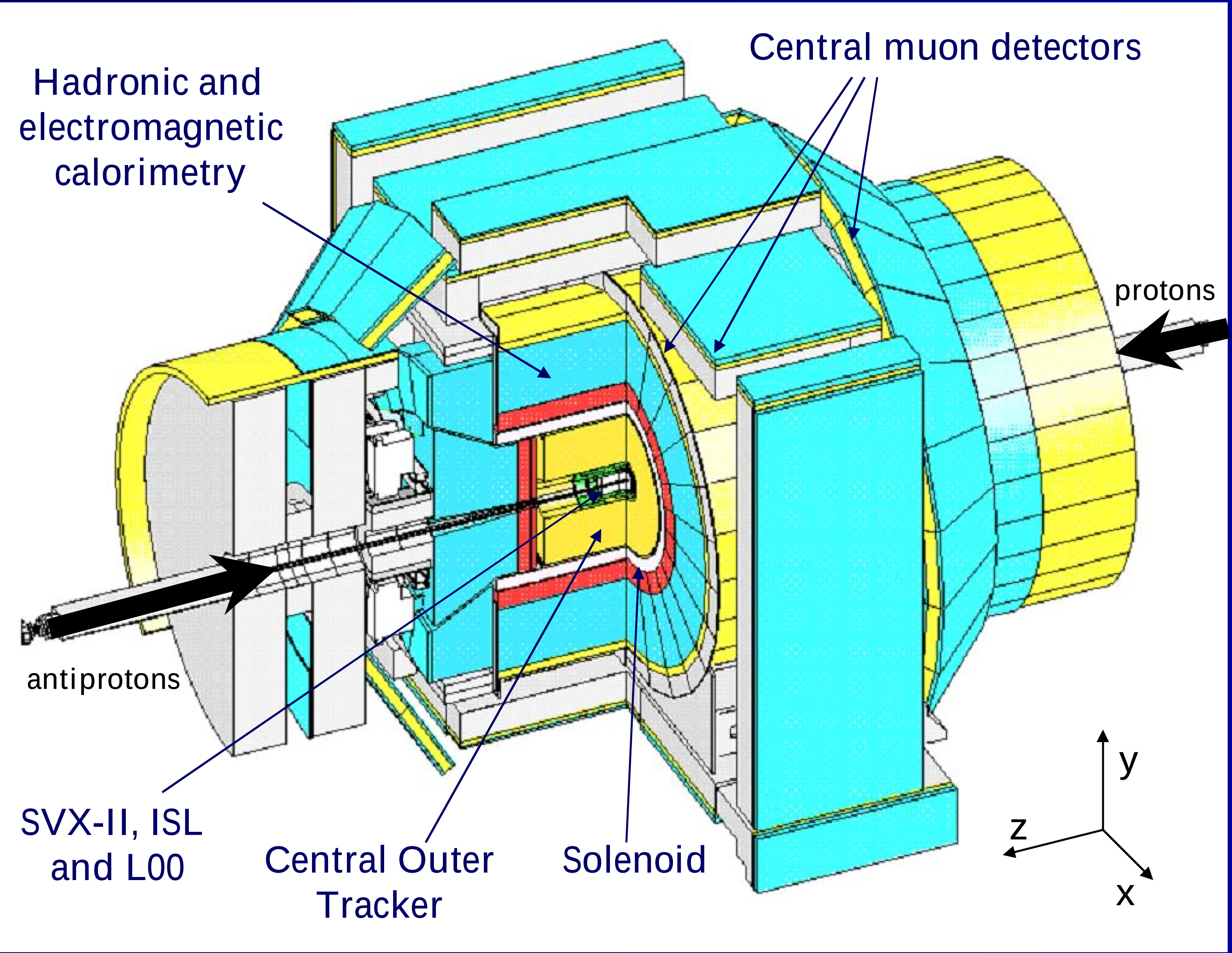
The CDF Detector



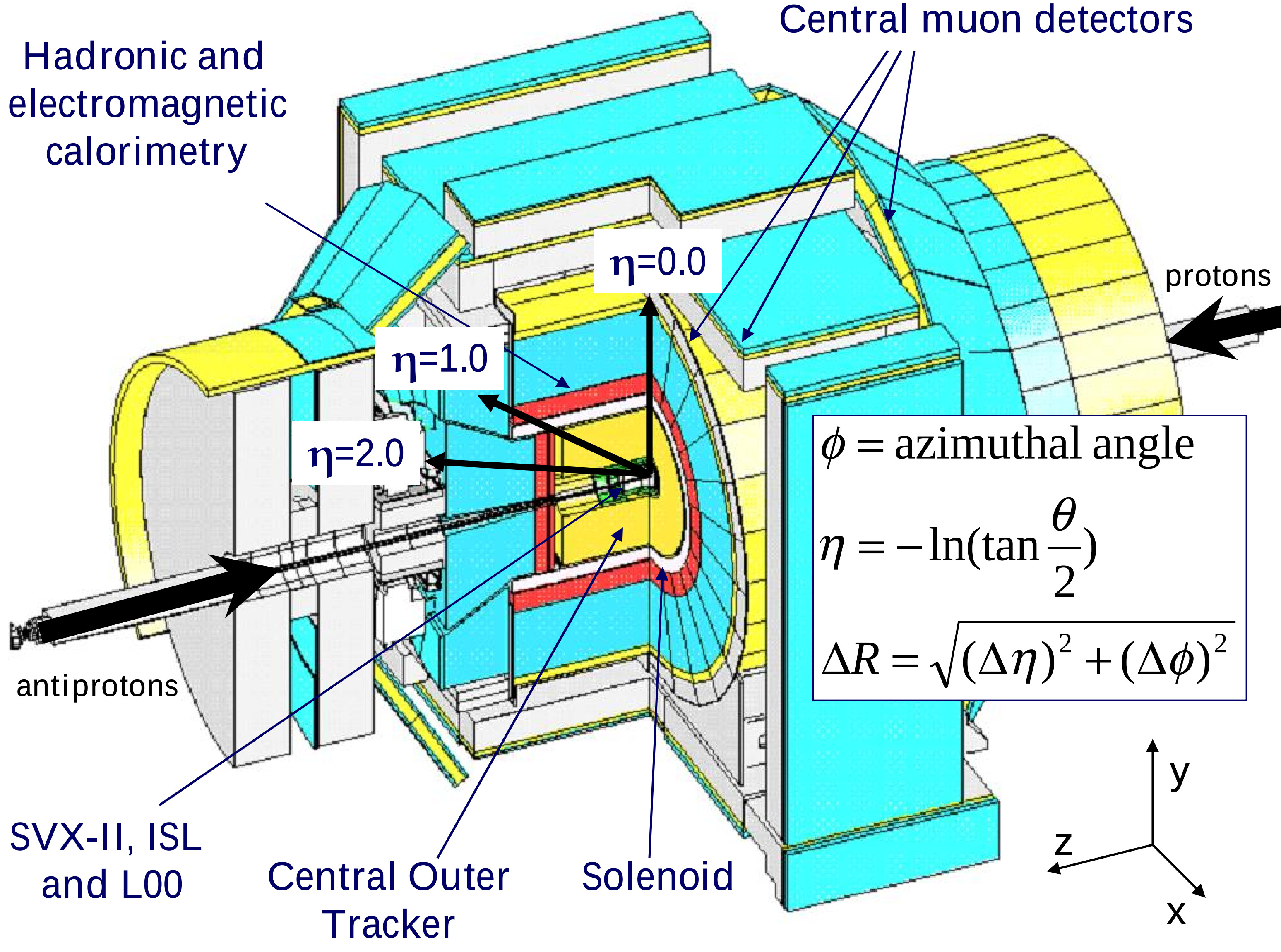
The CDF Detector



The CDF Detector



The CDF Detector



$$\phi = \text{azimuthal angle}$$

$$\eta = -\ln(\tan\frac{\theta}{2})$$

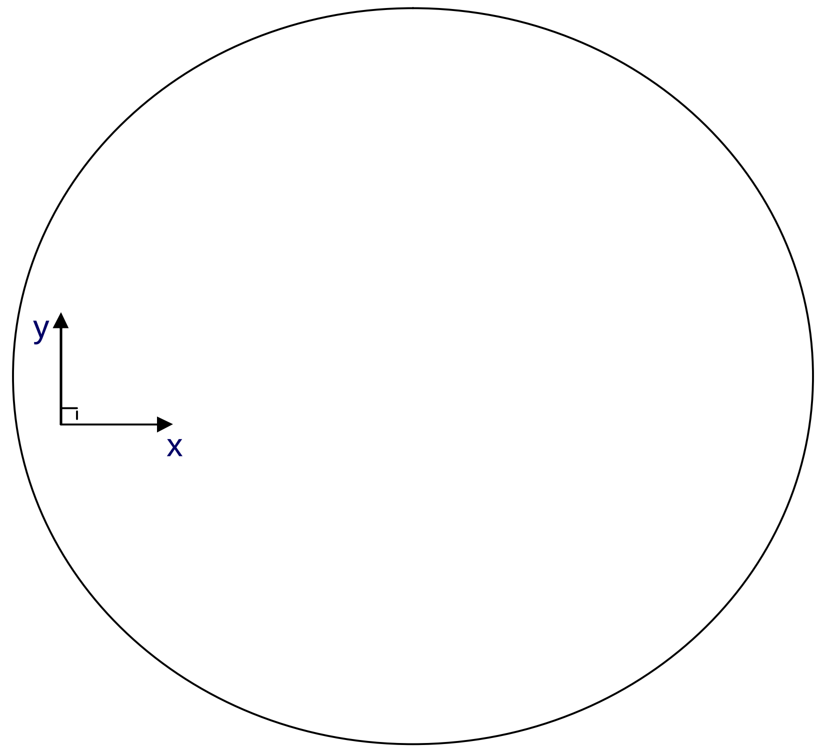
$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$$

Outline

- Motivation
- Measurement Definition
- Experimental apparatus
- *Measurement strategy*
- Results

Selecting *W*+*b*-jets Events

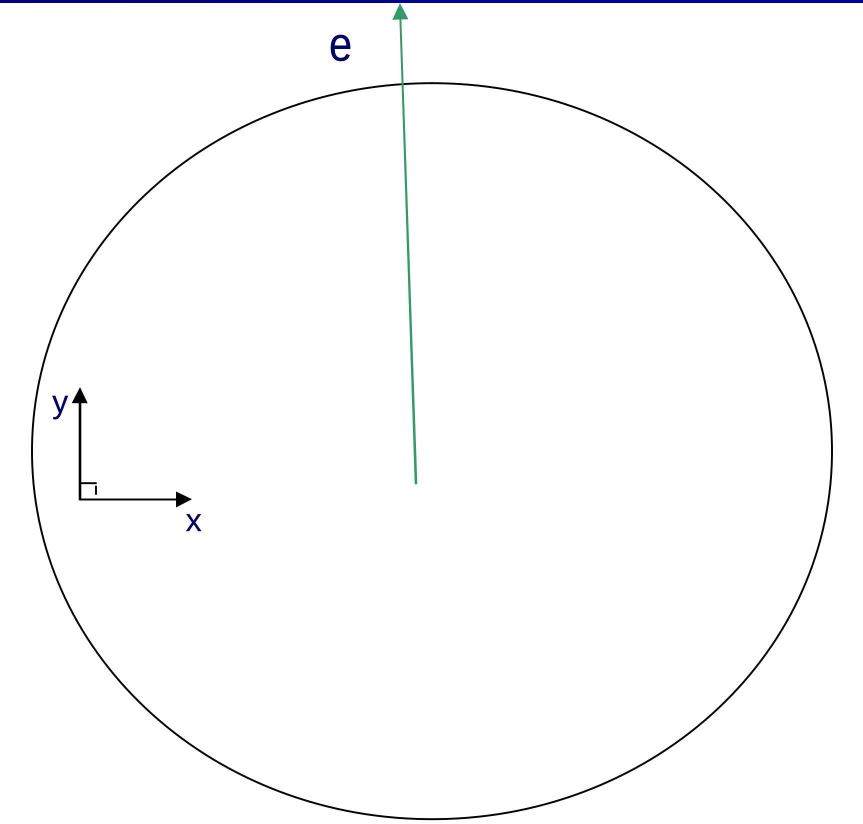
- Focus on leptonic *W* decays, $W \rightarrow \ell\nu$, $\ell=e,\mu$



CDF end view
Transverse plane

Selecting *W*+*b*-jets Events

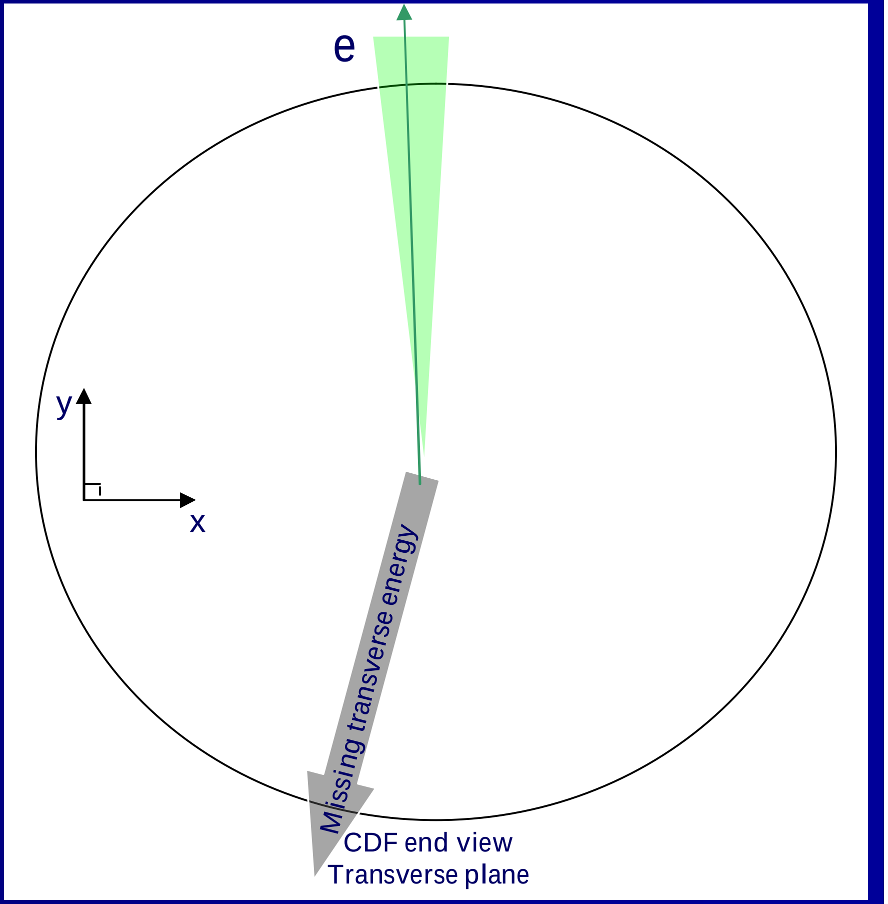
- Focus on leptonic *W* decays, $W \to \ell\nu$, $\ell = e,\mu$
- Online event trigger:
  - 18 GeV $|\eta| < 1.1$ electron OR
  - 18 GeV $|\eta| < 0.6$ muon OR
  - 18 GeV $0.6 < |\eta| < 1.0$ muon



CDF end view
Transverse plane

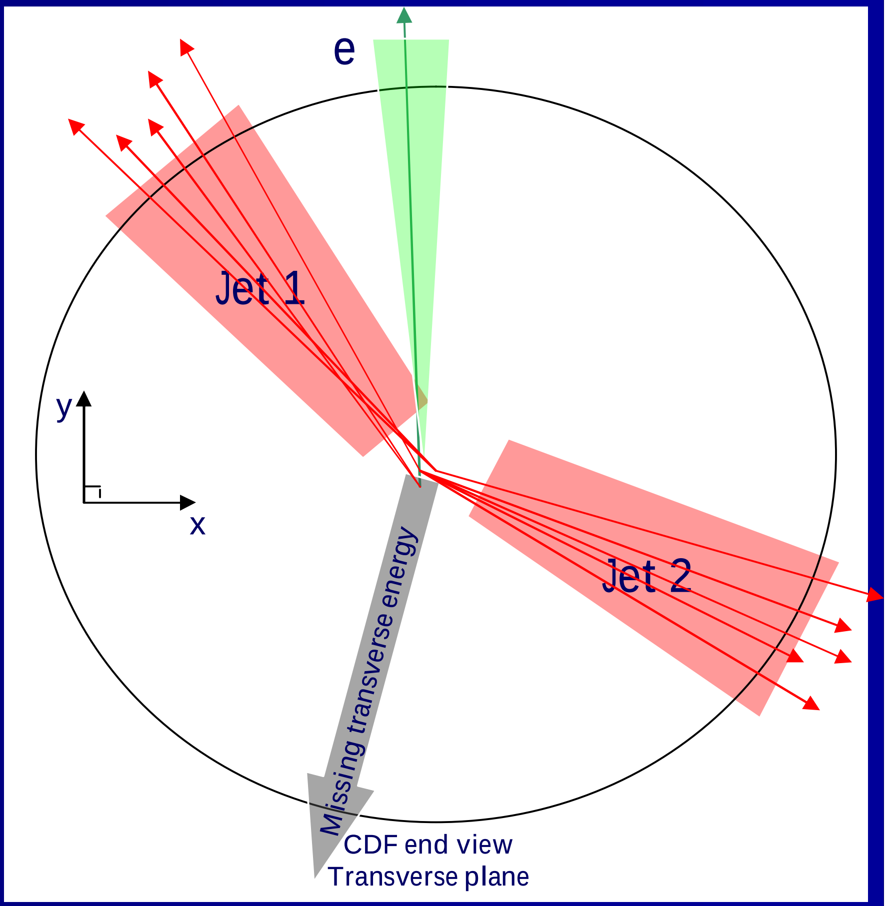
Selecting *W*+*b*-jets Events

- Focus on leptonic *W* decays, $W \rightarrow \ell\nu$, $\ell = e,\mu$
- Online event trigger:
  - 18 GeV $|\eta| < 1.1$ electron OR
  - 18 GeV $|\eta| < 0.6$ muon OR
  - 18 GeV $0.6 < |\eta| < 1.0$ muon
- *W* selection:
  - $p_T > 20$ GeV/c isolated central lepton
  - Large missing energy: MET > 25 GeV



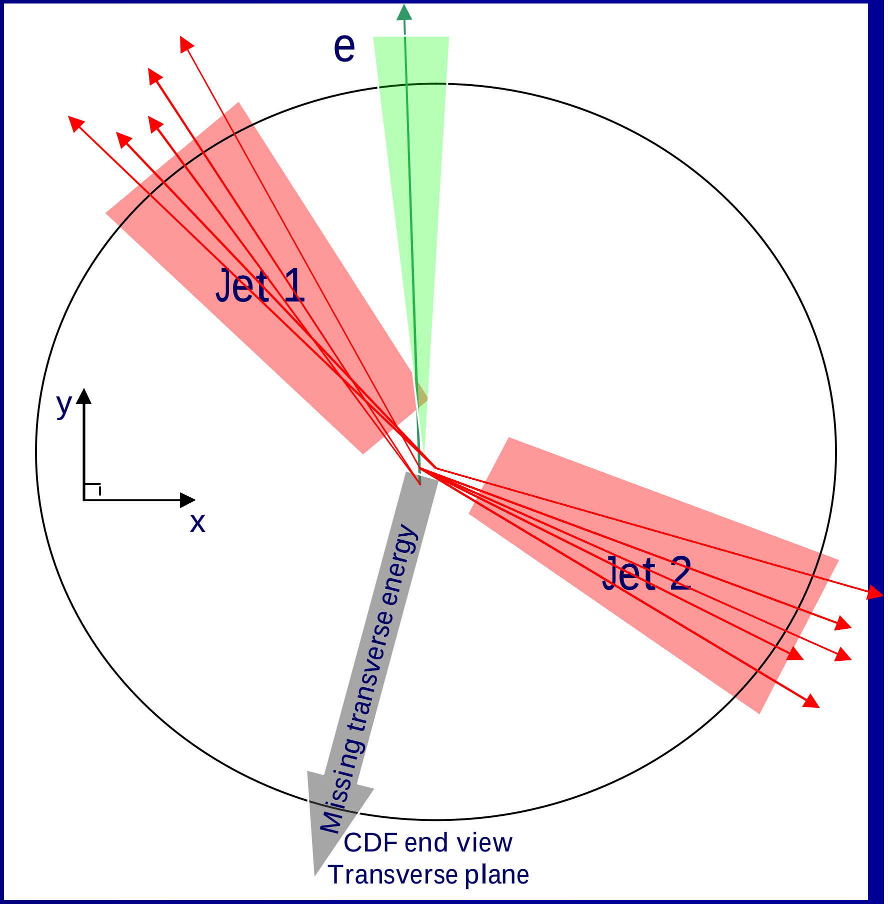
Selecting *W*+*b*-jets Events

- Focus on leptonic *W* decays, $W \to \ell\nu$, $\ell = e,\mu$
- Online event trigger:
  - 18 GeV $|\eta| < 1.1$ electron OR
  - 18 GeV $|\eta| < 0.6$ muon OR
  - 18 GeV $0.6 < |\eta| < 1.0$ muon
- *W* selection:
  - $p_T > 20$ GeV/c isolated central lepton
  - Large missing energy: MET > 25 GeV
- Jet selection:
  - Exactly 1 or 2 $E_T > 20$ GeV, $|\eta| < 2.0$ jets
  - JetClu clustering with R=0.4 cone



Selecting *W*+*b*-jets Events

- Focus on leptonic *W* decays, $W \rightarrow$ ªν, ª$=e,\mu$
- Online event trigger:
  - 18 GeV $|\eta|$ < 1.1 electron OR
  - 18 GeV $|\eta|$ < 0.6 muon OR
  - 18 GeV 0.6 < $|\eta|$ < 1.0 muon
- *W* selection:
  - p_T > 20 GeV/c isolated central lepton
  - Large missing energy: MET > 25 GeV
- Jet selection:
  - Exactly 1 or 2 E_T > 20 GeV, $|\eta|$ < 2.0 jets
  - JetClu clustering with R=0.4

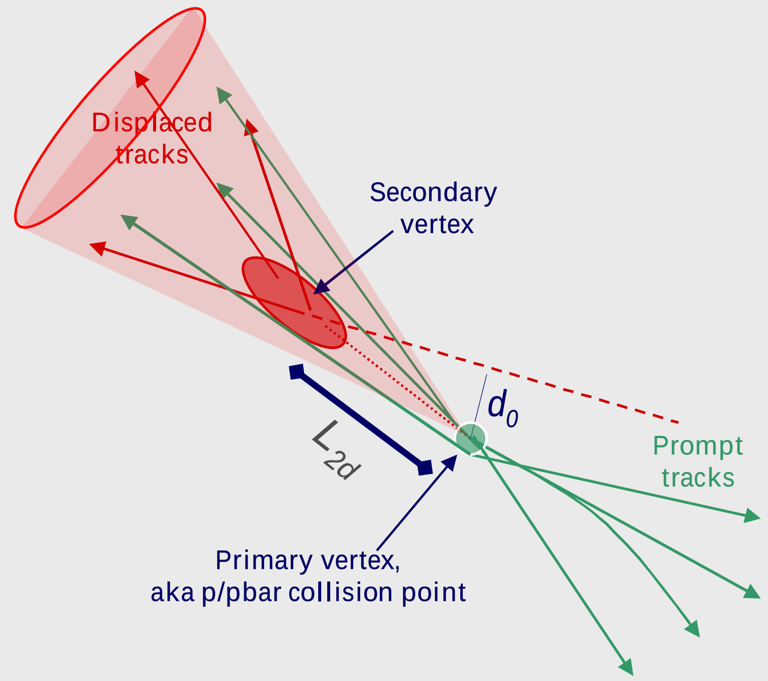


- Do not consider events from other processes:
  - Veto events w/ 2 high p_T leptons to avoid ttbar
  - Guard against $Z \rightarrow$ ªª production where one lepton is not fully reconstructed
  - Remove cosmic ray events, events with objects from different interactions
  - Veto fake *W* events

Identification of *b* Jets

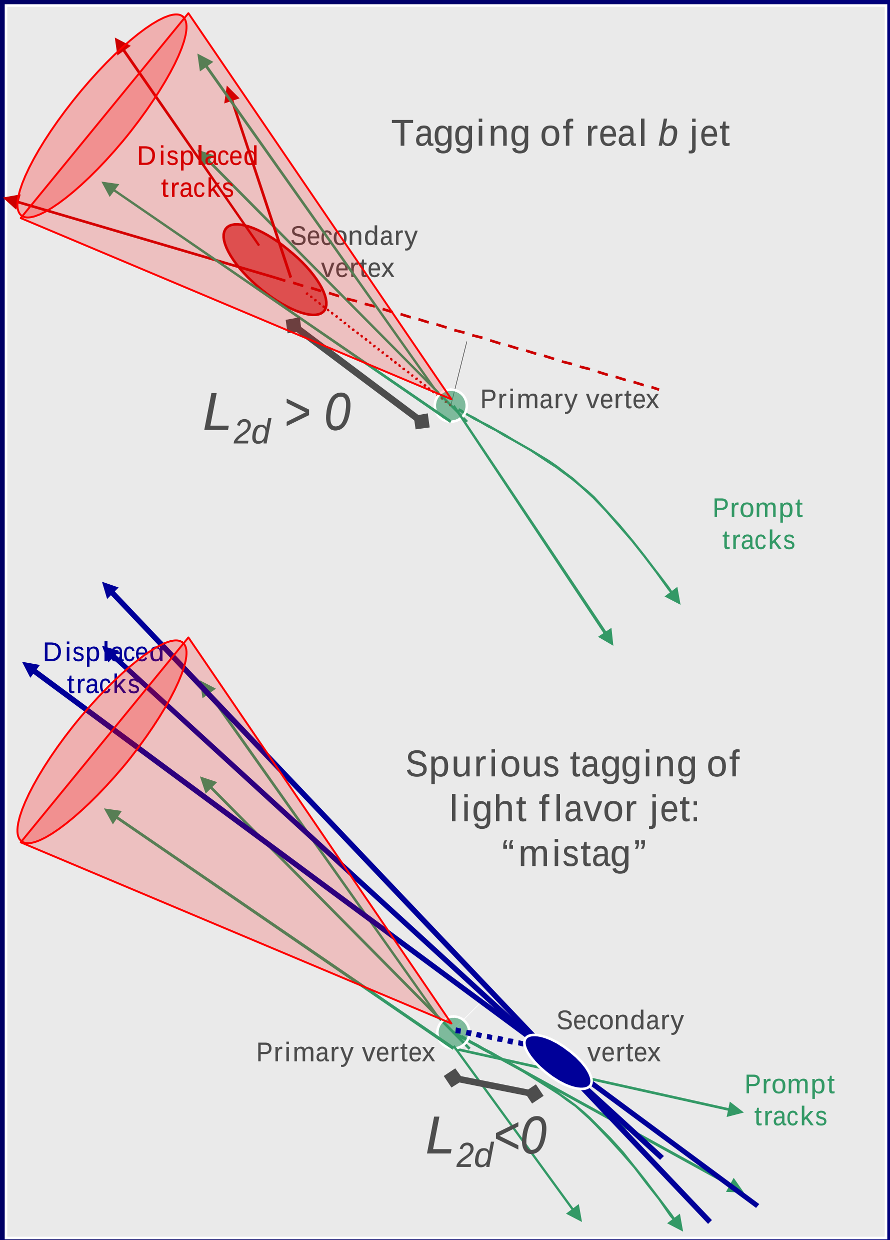
- What makes *b*-jets so special?
  - Long lifetime of the *b*
  - Large mass of *B* hadrons
  - High momentum decay products of *B* hadrons
- Some special relativity:
  - *b* quark lifetime: ~1.5 ps
  - Typical speed of *B* hadron is close to the speed of light
  - Moving clocks run slower…
  - Distance traveled in lab frame before decaying: <u>~2-3 mm</u>
- Exploit this feature:
  - Look within jets for displaced tracks
  - See if they intersect at a common point
  - Require the common point be significantly displaced from the primary interaction point

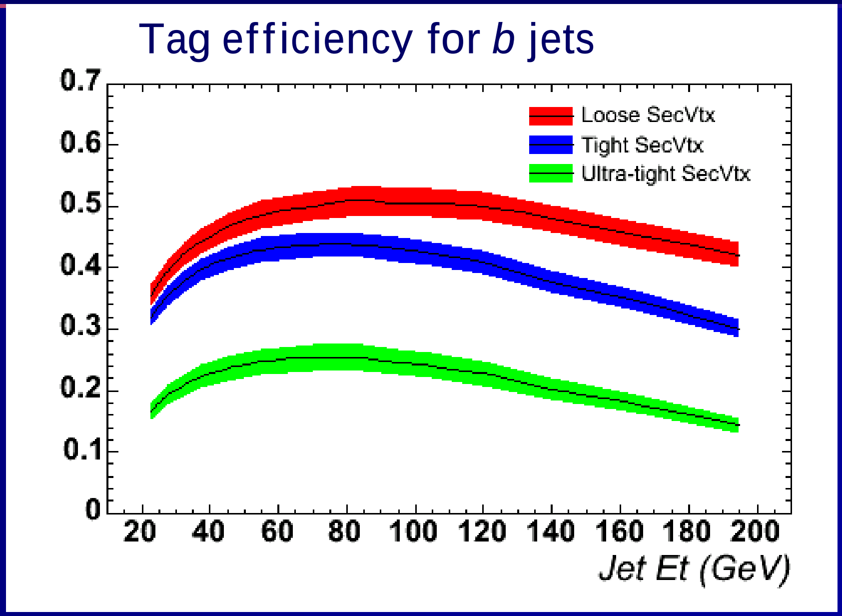
Long-lifetime yields secondary decay vertex "*b*-tagging"

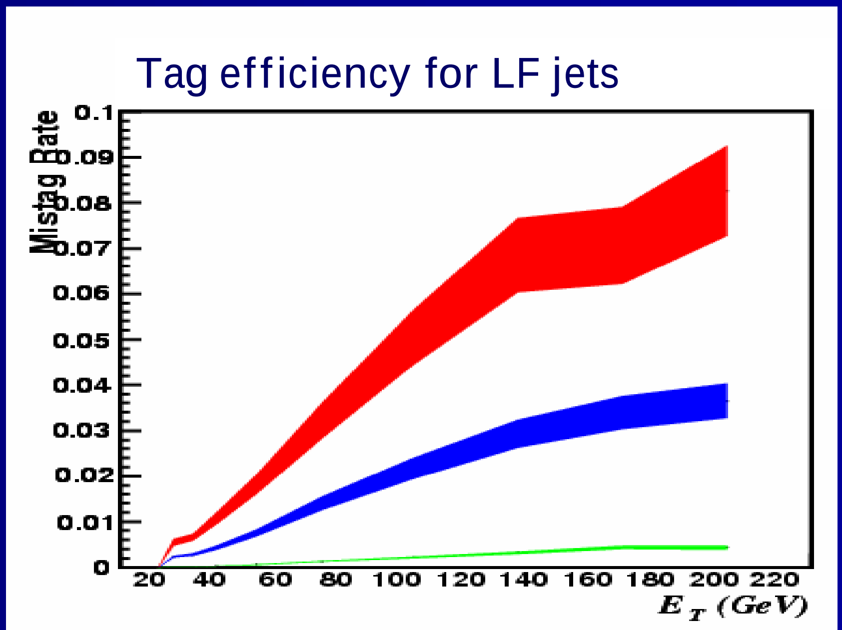


	Meaning	Typical	Resolution
d_0	Track impact parameter	150um	40um
L_{2d}	Vertex displacement	2-3mm	100um

b-tagging: *b*'s and Non-*b*'s

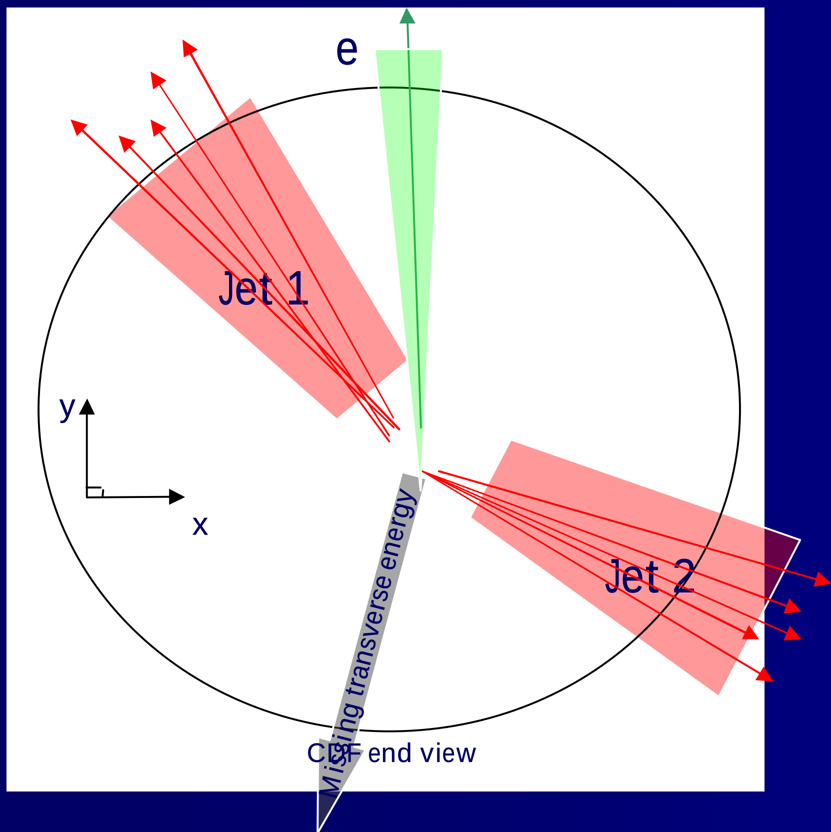




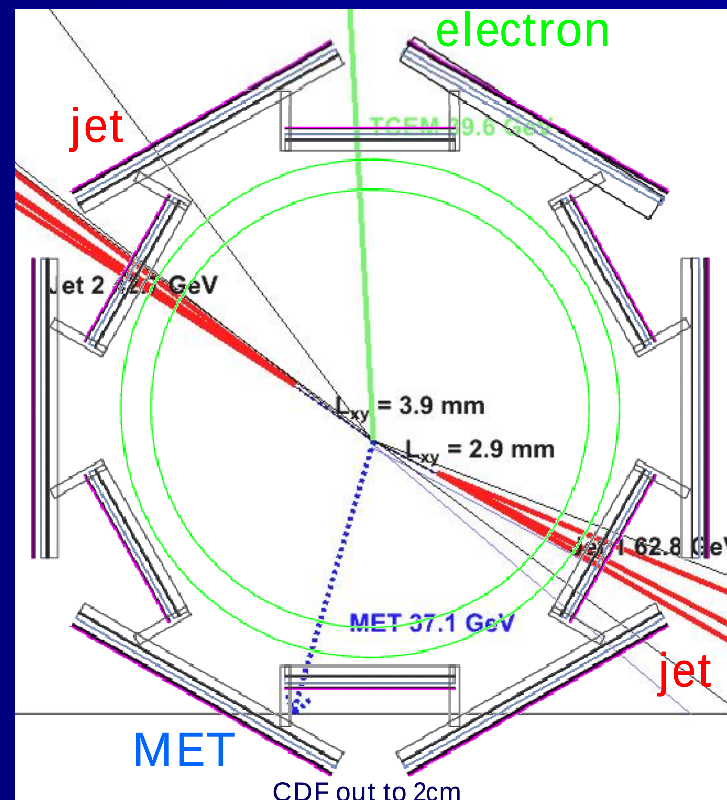


Yield of Tagged Jets

My cartoon:



The real thing: event recorded 10/2005



Event has 2 tagged jets!

Yield in 1.9/fb of data:

Selected Events (before tagging)	175712
Total Jets	199670
Tagged Jets	943

Ultratight SECVTX (UT) has low yield but increased purity.

W+*b*-jets: Measurement Strategy

$$\sigma_{b\ jets}(W+b-jets)\times BR(W\rightarrow \ell\nu)=\frac{n^{fit}_{bjets}-n^{\text{not}W+b}_{bjets}}{L\cdot A_{W+b\ bjets}\cdot\varepsilon}$$

Where do various pieces come from?

- Discriminate *b*/*c*/LF in tagged sample using **vertex mass** — n^{fit}_{bjets}

- Determine contribution from **background** tagged *b* jets and subtract from overall yield — n^{notW+b}_{bjets}

- Calculate **acceptance** for *W*+*b*-jet events — $A_{W+b\ bjets}$

- Measure **tag efficiency** for *b* jets in *W*+*b*-jet production in MC and correct to match that of data — ε

Extracting Species Content of Tagged Sample

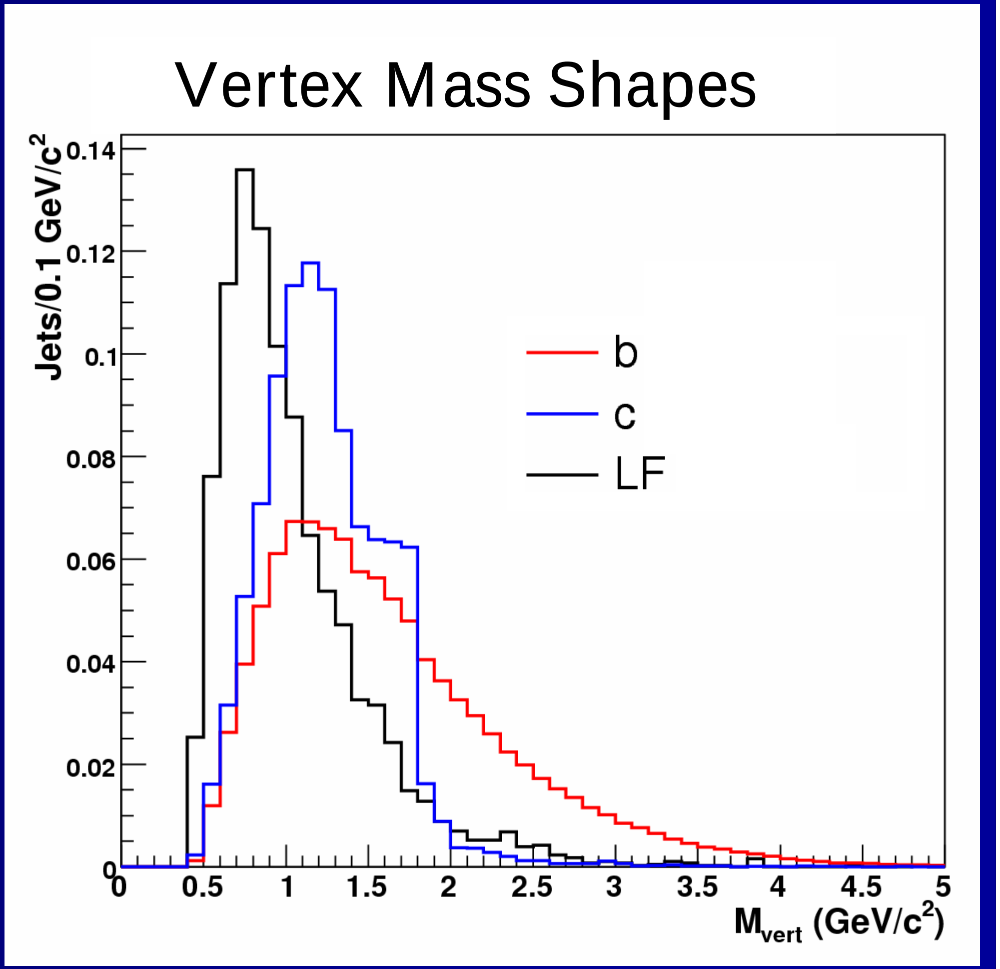
- Tagged jets are not guaranteed to be just from *b*'s
  - *b*, *c*, or LF (u/d/s/g)

- Discriminate the species of tagged jets via vertex mass, M_{vert}:
  - Invariant mass of tracks participating in found Ultratight secondary vertex
  - Correlated to mass of decaying hadron: Qualitatively,

$$M_{B\text{-hadrons}} > M_{C\text{-hadrons}} > M_{LF\text{-hadrons}}$$

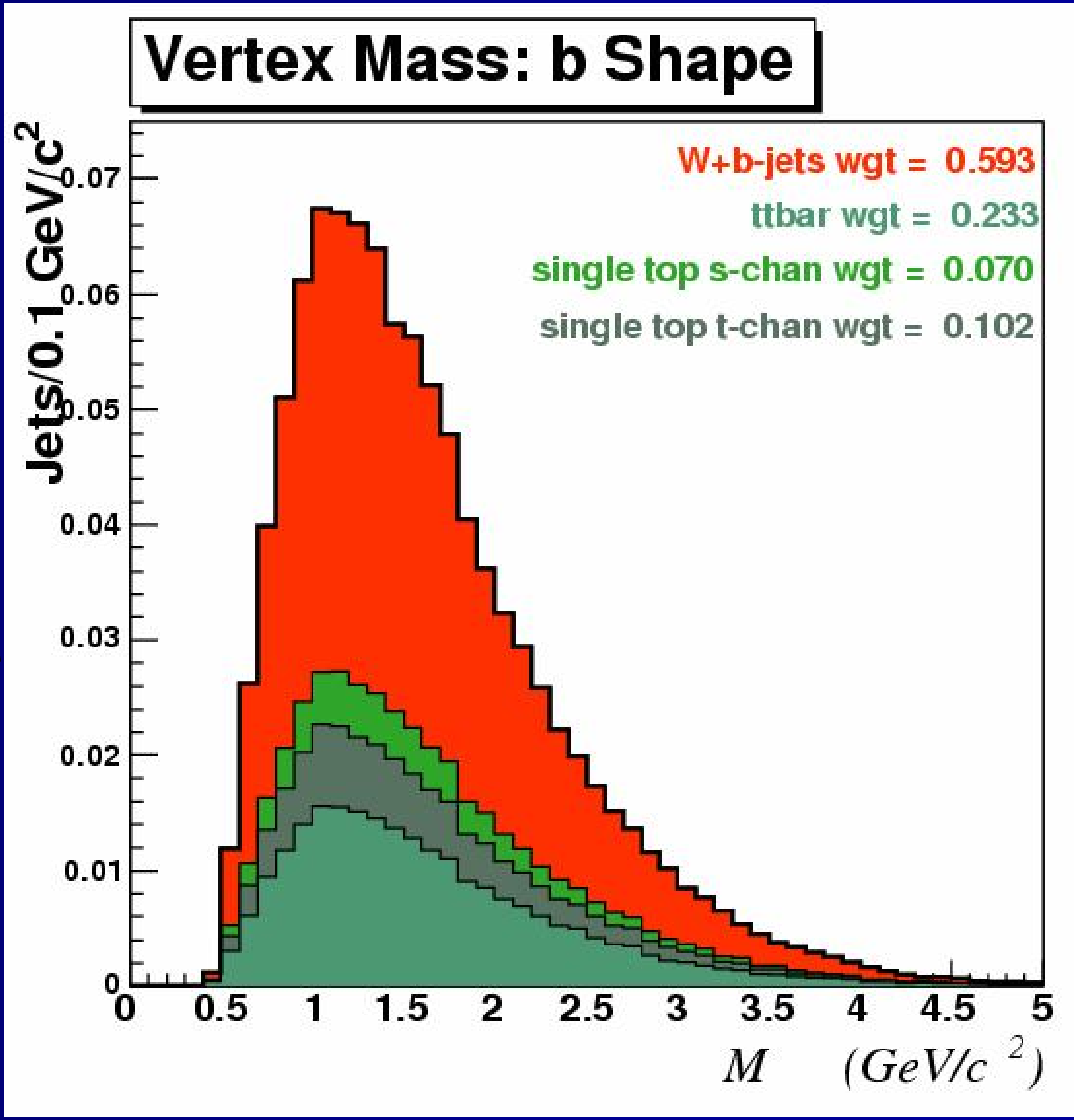
so

$$M^{b}_{vert} > M^{c}_{vert} > M^{LF}_{vert}$$

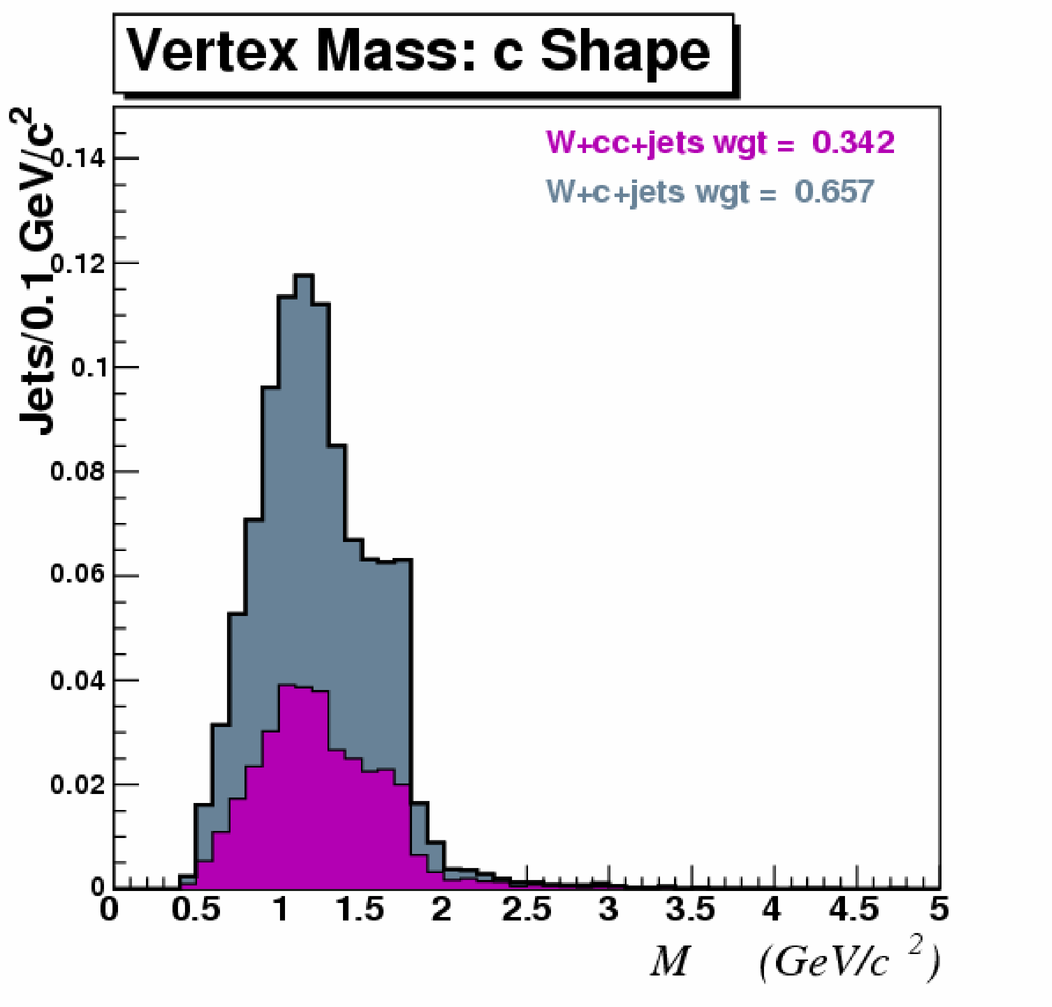


Species Content of Tagged Sample: *b* Shape

- Use MC events to build the shape for *b*:
  - Weighted contributions from main *b* sources to selected sample
    - *W+b*-jets
    - ttbar
    - Single top
  - Shapes for each process are similar: not sensitive to assumed weight of each
  - Insensitive to even large changes in top, single top cross sections

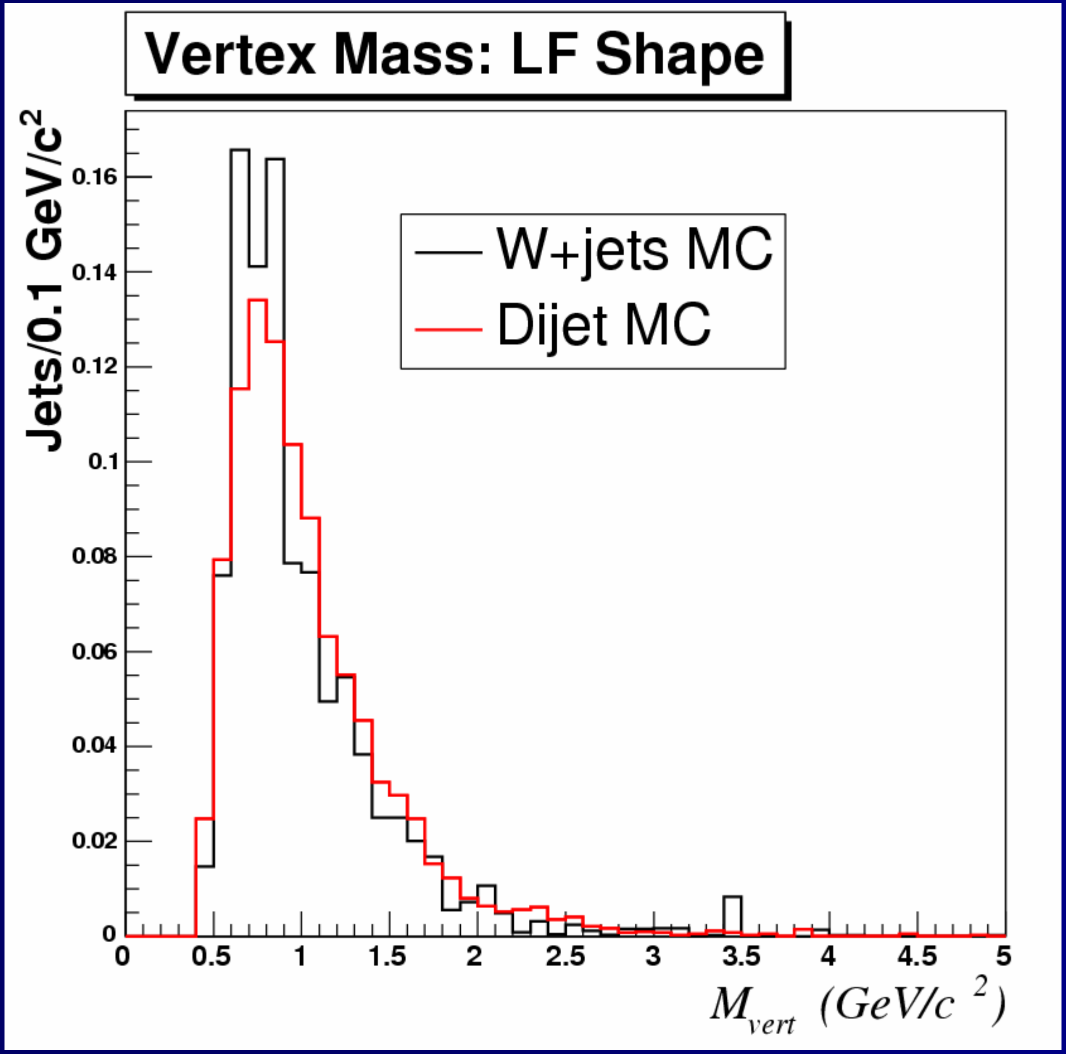


Species Content of Tagged Sample: *c* Shape



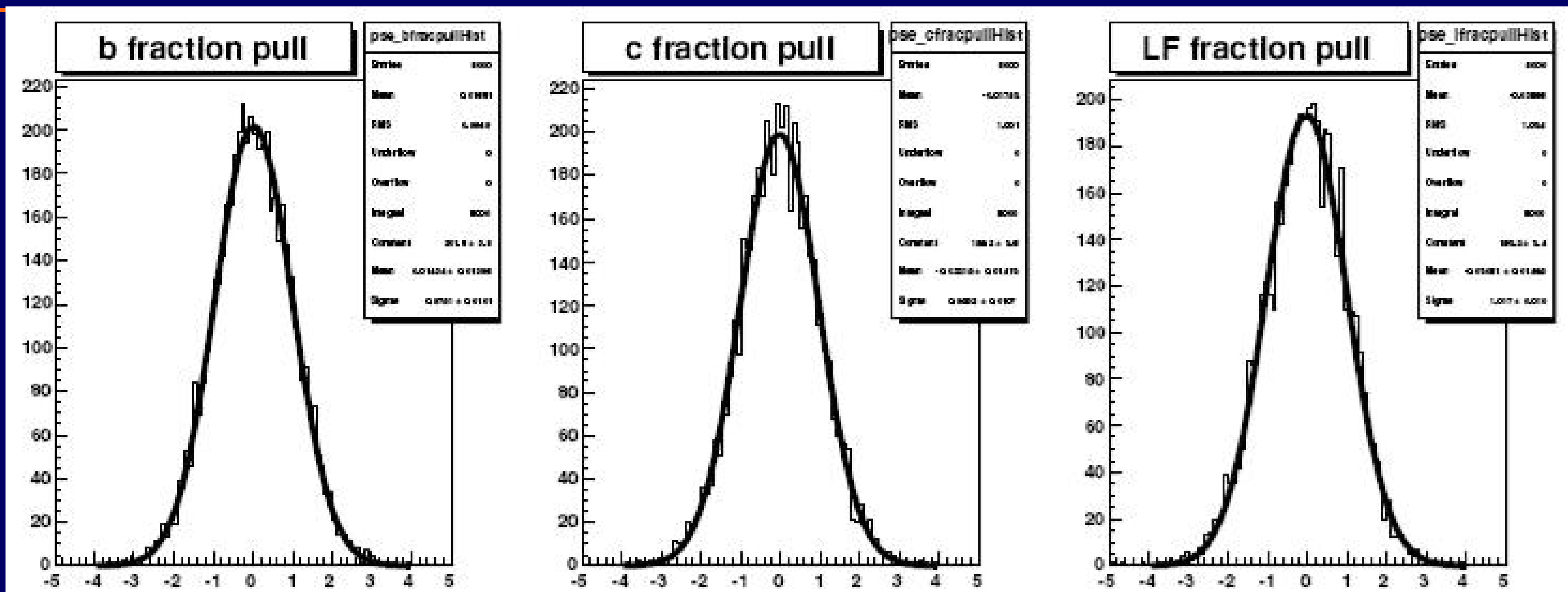
- Shape for charm: comes from significant *c* sources

Species Content of Tagged Sample: LF Shape

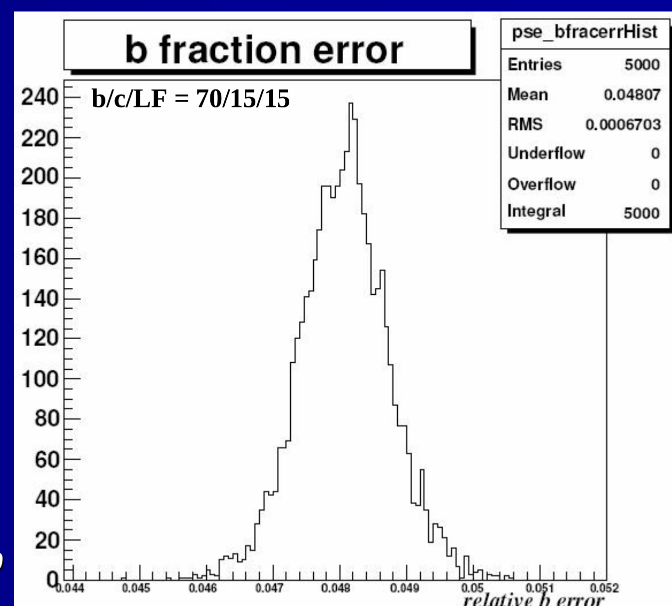


- Shape for LF comes from tags of LF-matched jets in MC
- Several MC samples were studied, including:
  - W+jets MC
  - Dijet MC w/ at least one pT>50 jet
- All shapes are reasonably consistent
- Chose to use the dijet MC shape for fitting and use high statistics alternative for setting a systematic

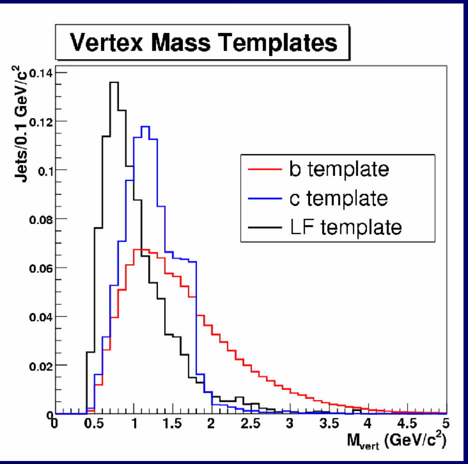
Likelihood Fit



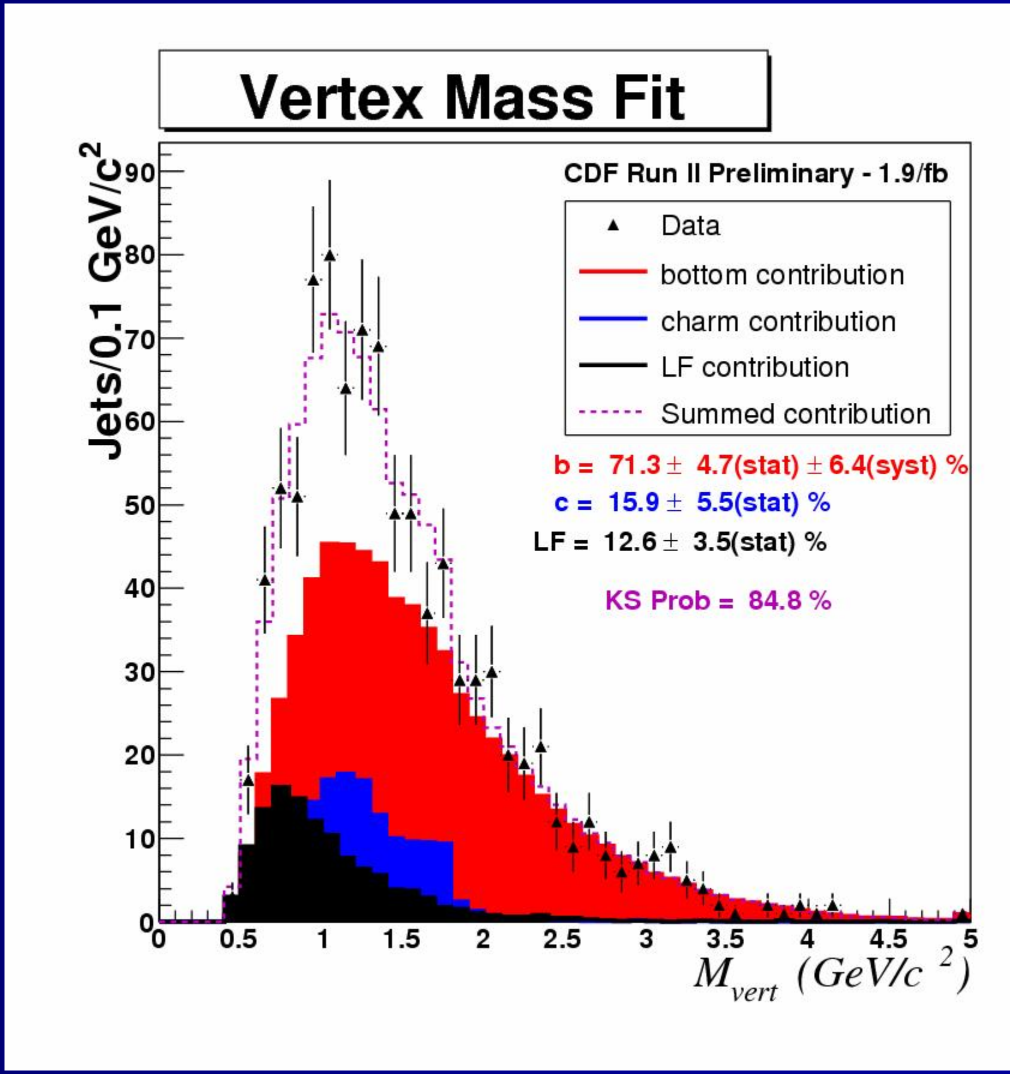
- Data is comprised of three species
- Use binned Poisson maximum likelihood fit to extract contribution from each source
- Pseudoexperiment studies showed fit results were accurate and had relative fit error of ~5% on f_b



Species Content of Tagged Sample: Fit Results



- Fit results in the CDF data!
- Fit claims ~71% of tagged jets are from *b*.
- Given the yield of 943 tagged jets that corresponds to

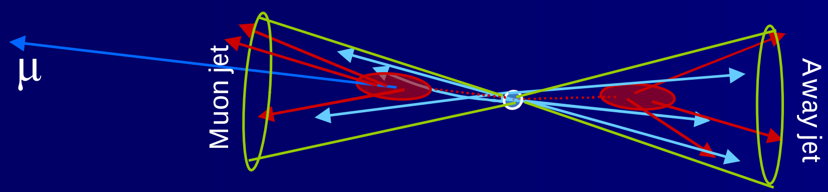


$$n_{bjets}^{fit} = 672.3 \pm 44.3(\text{stat}) \pm 60.4(\text{syst})$$

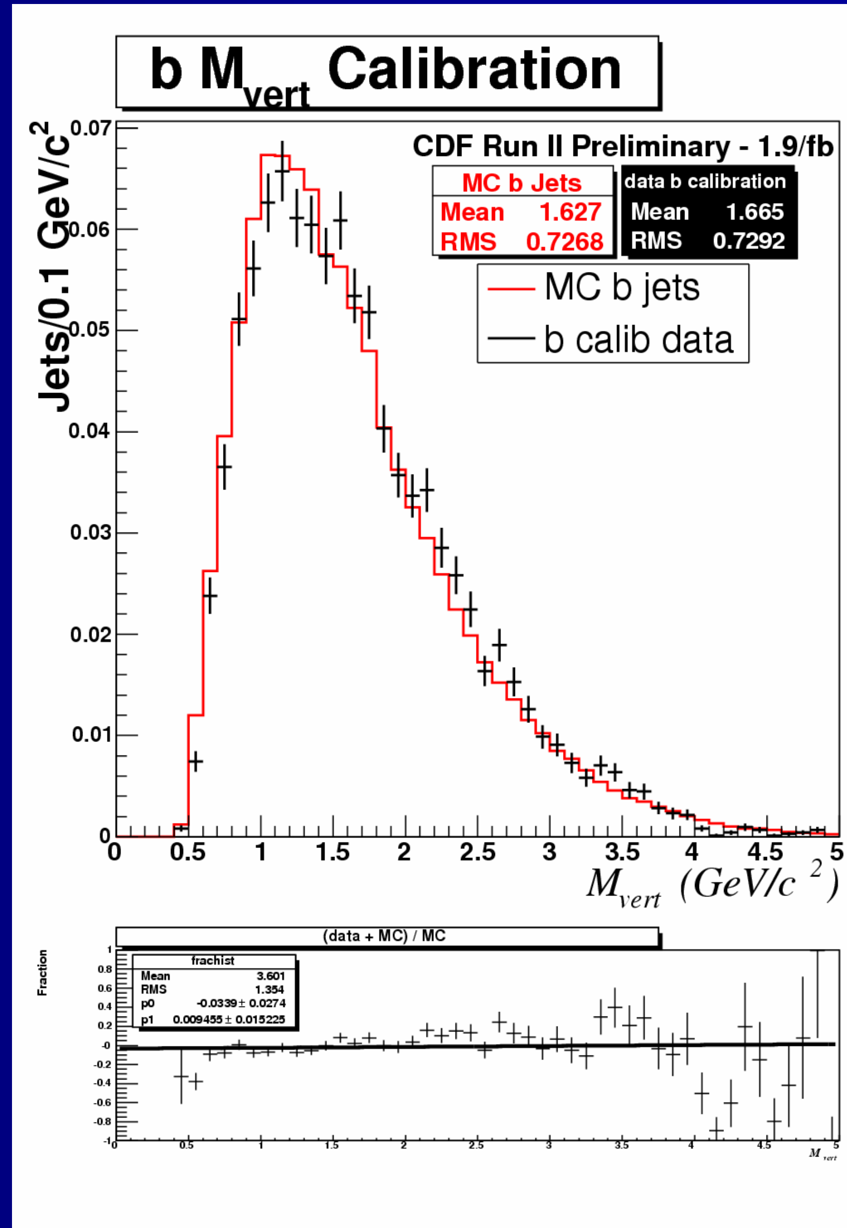
Where does this systematic error come from?

Species Content of Tagged Sample: Systematics

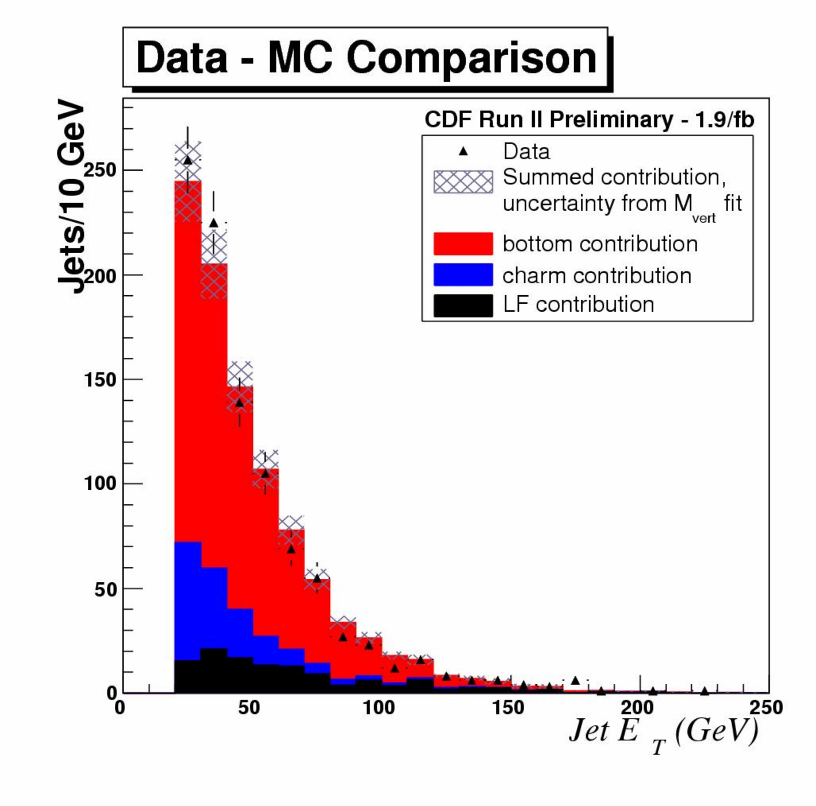
- Seek calibration of shape for *b*
- Can construct a pure sample of Ultratight tagged *b* jets in data:

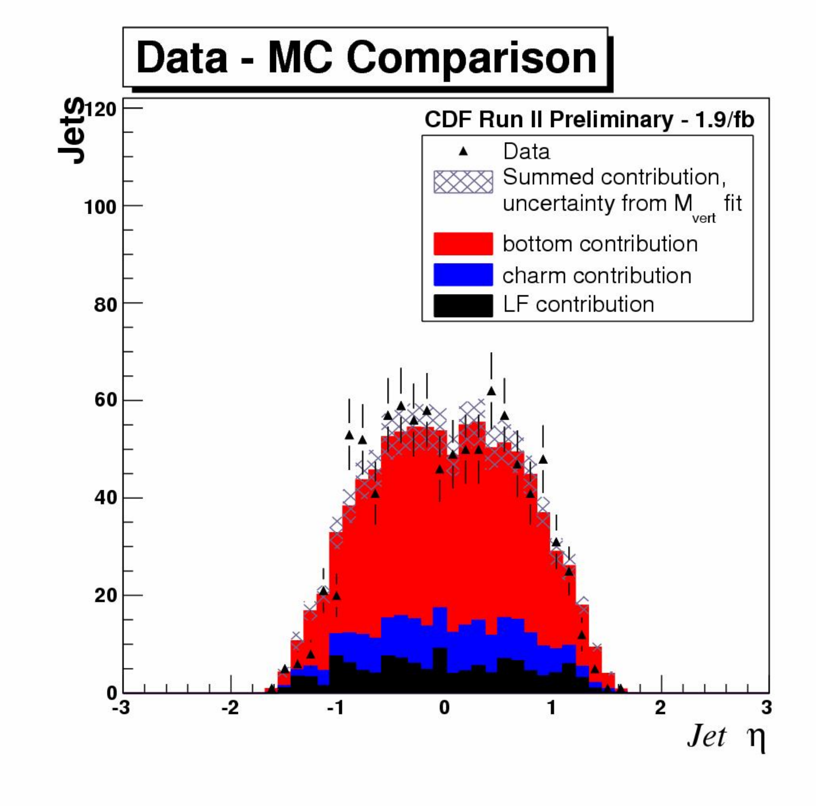


  - Trigger: 8 GeV muon
  - Construct back-to-back dijet system:
    - Muon jet: UT-tagged, M_{vert}>1.7GeV
    - Away jet: UT-tagged
  - Away jet *b* purity > 99% in Pythia
- Shape difference: a $\delta f_b/f_b$ = 8% effect
- *c*, LF shape systs have smaller effect on f_b



Vertex Mass Fit Consistency Check





- Check species fractions from fit in other variables
- Things look reasonable in these and other distributions

W+b-jets: Measurement Strategy

$$\sigma_{b\,jets}(W+b-jets)\times BR(W\to \ell\nu)=\frac{n^{fit}_{bjets}-n^{\mathrm{not}W+b}_{bjets}}{L\cdot A_{W+b\,bjets}\cdot\varepsilon}$$

Where do various pieces come from?

- Discriminate *b*/*c*/LF in tagged sample using **vertex mass**

$$n^{fit}_{bjets}=672.3\pm 44.3(\mathrm{stat})\pm 60.4(\mathrm{syst})$$

- Determine contribution from **background** tagged *b* jets and subtract from overall yield

$$n^{notW+b}_{bjets}$$

- Calculate **acceptance** for *W*+*b*-jet events

$$A_{W+b\,bjets}$$

- Measure **tag efficiency** for *b* jets in *W*+*b*-jet production in MC and correct to match that of data

$$\varepsilon$$

Background Sources of *b* Jets

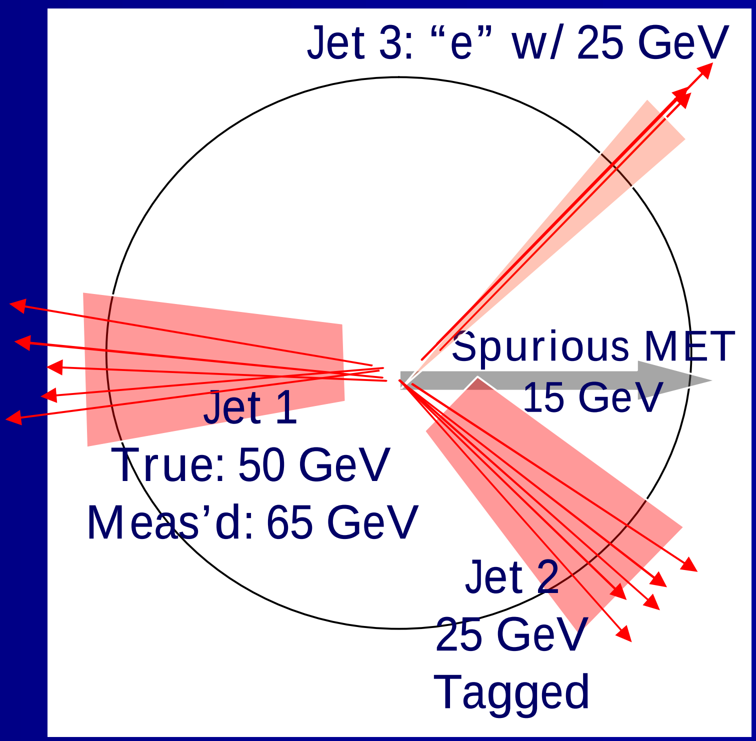
- Various processes contribute to *b*-tags in *W*+1,2 jet sample
- Two categories treated here:
  - MC-driven (ttbar, single top, dibosons, others)
  - Data-driven (Fake *W*)
- MC-based backgrounds:
  - Use Pythia, MadEvent, ALPGEN
  - Apply event selection, get efficiency,
  - Use production cross section to estimate yield
- Top-based processes largest contributors

Process	n^b_{W+12j}
$t\bar{t}$	73.1 ± 10.1 *
s-channel	22.2 ± 9.6 *
t-channel	33.4 ± 15.0 *
WZ	9.1 ± 0.9
ZZ	0.28 ± 0.03
WW	0.83 ± 0.12
$W + bb$+Np, $W \to \tau\nu$	7.3 ± 0.8
$Z + bb$+Np, $Z \to e^+e^-$	0.67 ± 0.08
$Z + bb$+Np, $Z \to \mu^+\mu^-$	4.1 ± 0.4
$Z + bb+ \geq$Np, $Z \to \tau^+\tau^-$	1.48 ± 0.20

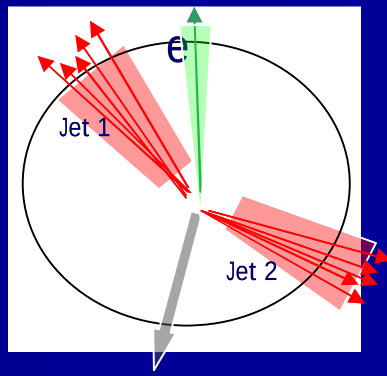
Background Sources of *b* Jets: *b* Jets in Fake *W* Events

- What are fake *W* events?
  - Mostly QCD multijet production mimicking isolated lepton w/ spurious missing energy from mismeasured jets
  - Tagged jets found elsewhere in the event
  - Characterized by:
    - small MET
    - large MET error
    - small W transverse mass
- Strategy here:
  - Remove as much as possible from the start
  - Model what remains using data
- Model for fake *W*: "antielectrons"
  - Most fake electrons *just barely* satisfy electron identification
  - Construct a sample of objects that nearly satisfy electron ID - marginal failures

Fake W Event:

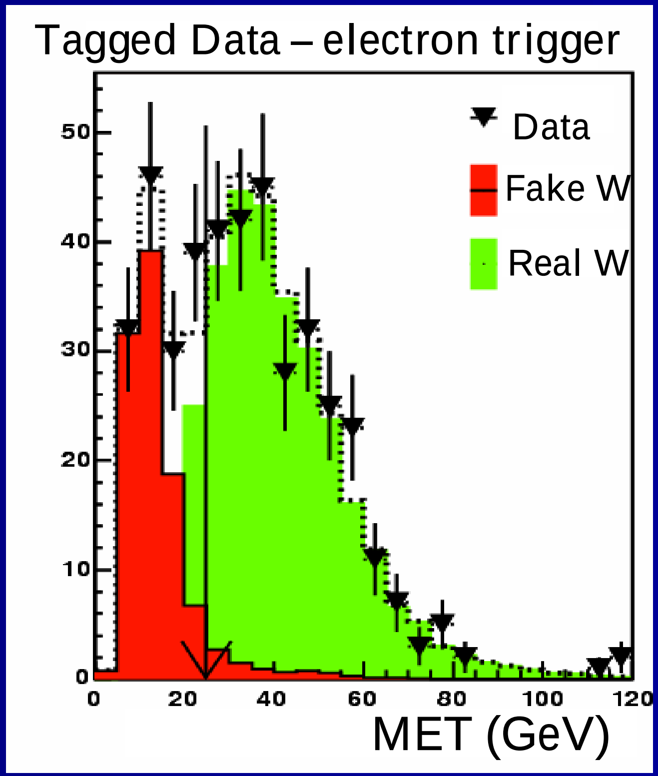


Recall, picture of real *W* event:



Background Sources of *b* Jets: *b* Jets in Fake *W* Events

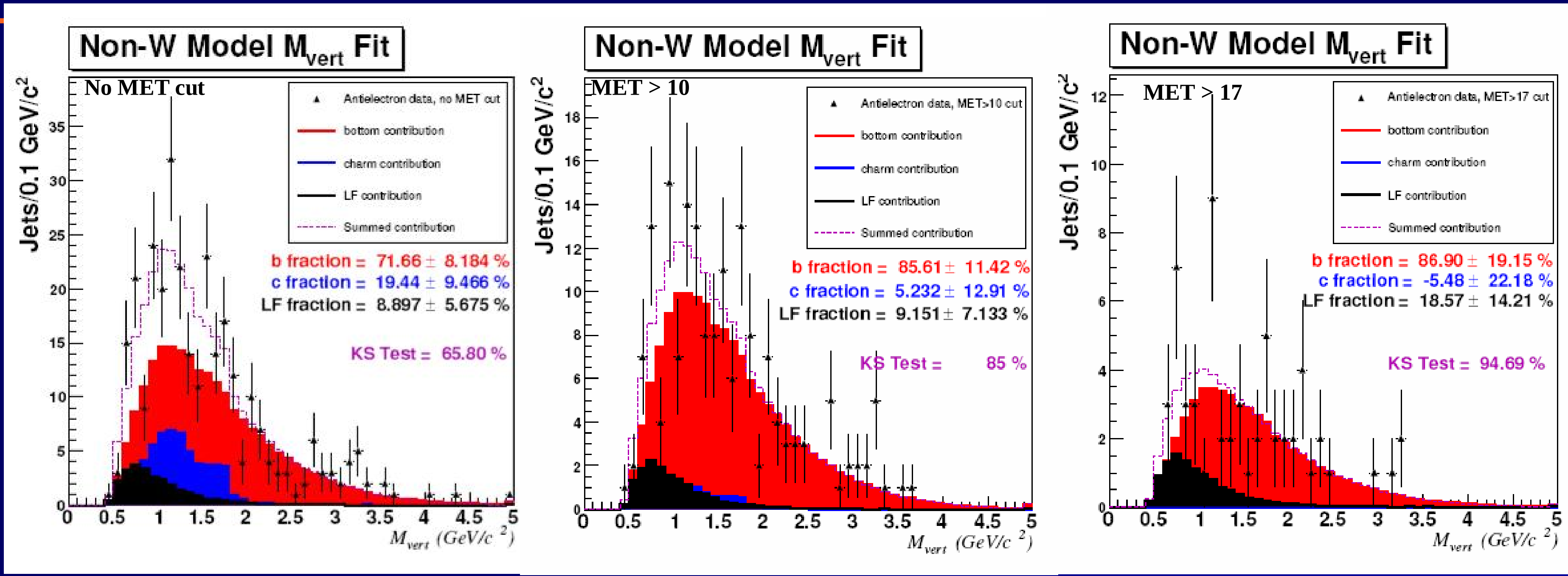
- With model in place can now determine how many tagged jets come from fake *W*
- Procedure:
  1. Use MET – discriminates between real and fake W events
  2. Relax MET cut (for lever arm)
  3. Fit entire data MET dist to shapes from top, single top, *W*+jets, Fake-*W*
  4. Return to MET>25 cut after fit and obtain Fake-*W* fraction
  5. Fit vertex mass of tagged jets to get *b* fraction



From this fit, fake *W* is responsible for 2.9% of tagged jets in electron trigger data.

NB: Here antielectron shape used to model fake W's in the muon trigger sample as well. Antimuons will be adopted in the future.

Background Sources of *b* Jets: *b* Jets in Fake *W* Events



- Step 5 fails – insufficient stats in antielectron sample with MET>25
- Step through different MET cuts, examine behavior
- As one tightens the MET cut f_b^{QCD} increases

	W + 1 jet	W + 2 jet	W + 1,2 jet
Fake W tags	11.8 +- 3.6	18.8 +- 6.3	**30.6 +- 7.4**
Fake W tagged b	9.4 +- 3.7	15.1 +- 6.3	24.5 +- 8.4

- Reasonable choice: $f_b^{FakeW} = 0.8 \pm 0.2$

Fit errors only

Summary: Background Sources of *b* Jets

Process	n^b_{W+12j}
$t\bar{t}$	73.1 ± 10.1
s-channel	22.2 ± 9.6
t-channel	33.4 ± 15.0
WZ	9.1 ± 0.9
ZZ	0.28 ± 0.03
WW	0.83 ± 0.12
$W + bb$+Np, $W \to \tau\nu$	7.3 ± 0.8
$Z + bb$+Np, $Z \to e^+e^-$	0.67 ± 0.08
$Z + bb$+Np, $Z \to \mu^+\mu^-$	4.1 ± 0.4
$Z + bb+ \geq$Np, $Z \to \tau^+\tau^-$	1.48 ± 0.20
Non-W	24.5 ± 8.4
Total	176.8 ± 22.3

Predicted yields from all *b*-jet backgrounds in 1.9/fb.

W+b-jets: Measurement Strategy

$$\sigma_{b\,jets}(W+b-jets)\times BR(W\rightarrow \ell\nu)=\frac{n_{bjets}^{fit}-n_{bjets}^{notW+b}}{L\cdot A_{W+b\;bjets}\cdot\varepsilon}$$

Where do various pieces come from?

- Discriminate b/c/LF in tagged sample using **vertex mass**

$$n_{bjets}^{fit}=672.3\pm 44.3(\mathrm{stat})\pm 60.4(\mathrm{syst})$$

- Determine contribution from **background** tagged b jets and subtract from overall yield

$$n_{bjets}^{notW+b}=176.8\pm 22.3(\mathrm{syst})$$

- Calculate **acceptance** for W+b-jet events

$$A_{W+b\;bjets}$$

- Measure **tag efficiency** for b jets in W+b-jet production in MC and correct to match that of data

$$\varepsilon$$

Acceptance Definition

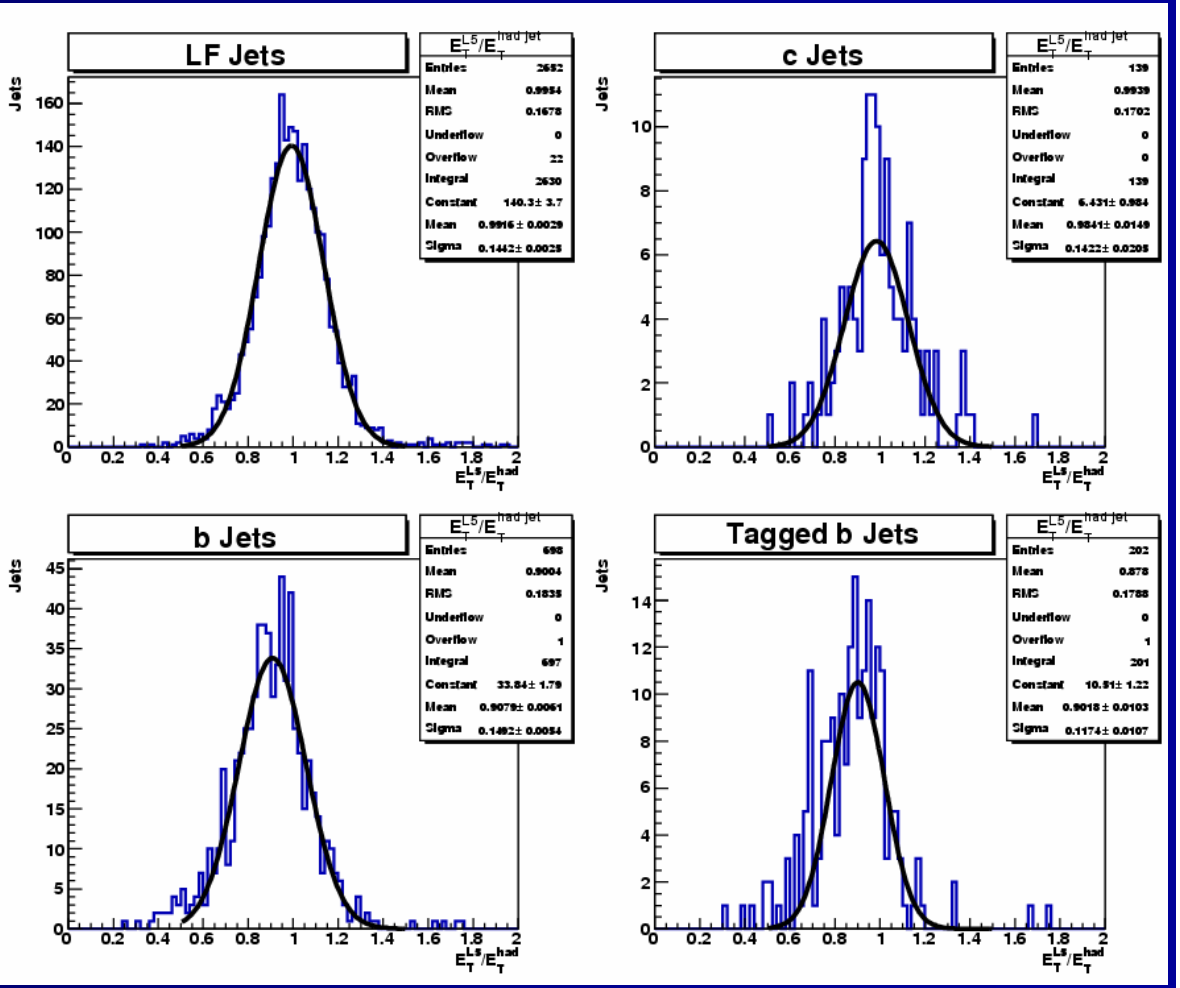
$$\mathcal{A}_{W+bjets} = \frac{\#\ \text{reconstructed}\ b\ \text{jets in events passing all selection}}{\#\ \text{MC}\ b\ \text{jets in event sample}}$$

- There are two effects the acceptance encodes:
  - Smearing provided by fragmentation effects, CDF detector, etc
    MC events migrate in and out:
    - a true 25 GeV jet can be reconstructed at 18 GeV – fail jet requirement!
    - a true 15 GeV jet can be reconstructed at 22 GeV – pass!
  - Reduction in sample through selection cuts designed to isolate signal
    MC events only migrate out:
    - eg, event vetos

- Denominator of *A:*
  - Number of *b*-jets in MC before detector simulation in events satisfying the phase space restrictions
  - Jets without the detector?

MC Jets

- Jets without the calorimeter!
  - SpartyJet: Software provides jet clustering on raw particles
  - Some knowledge of CDF geometry
  - Glimpse of “truth” jets
- Convention: exclude *W* daughters but make jets out of everything else
- Natural mismatch wrt measured jet E_Ts
  - effect largest for *b* jets



Measured b-jet energies are ~10% low on average wrt “truth”.
Agreement better for LF, *c* jets.

Acceptance Results

- Denominator of the acceptance is defined wrt these MC jets
- Use ALPGEN MC to evaluate the acceptance
  - Recall the phase space restrictions are attempt to insulate result from theory dependence

- Folded into this acceptance term
  - Efficiency for being in the luminous region of CDF $|z_0| < 60$ cm
  - Trigger efficiency
  - Lepton identification efficiency

- Sources of systematic error on the acceptance
  - Jet Energy Corrections (3% effect on *A*)
  - Renormalization/factorization scale choice (3%)
  - Parton distribution functions (2%)

$$A_{W+b\,bjets} = 0.593 \pm 0.017(\mathrm{syst})$$ integrated over all three trigger paths

W+*b*-jets: Measurement Strategy

$$\sigma_{b\ jets}(W+b-jets)\times BR(W\to \text{ª}\nu)=\frac{n^{fit}_{bjets}-n^{\mathrm{not}W+b}_{bjets}}{L\cdot A_{W+b\ bjets}\cdot\varepsilon}$$

Where do various pieces come from?

- Discriminate *b/c*/LF in tagged sample using **vertex mass**

$$n^{fit}_{bjets}=672.3\pm 44.3(\mathrm{stat})\pm 60.4(\mathrm{syst})$$

- Determine contribution from **background** tagged *b* jets and subtract from overall yield

$$n^{notW+b}_{bjets}=176.8\pm 22.3(\mathrm{syst})$$

- Calculate **acceptance** for *W*+*b*-jet events

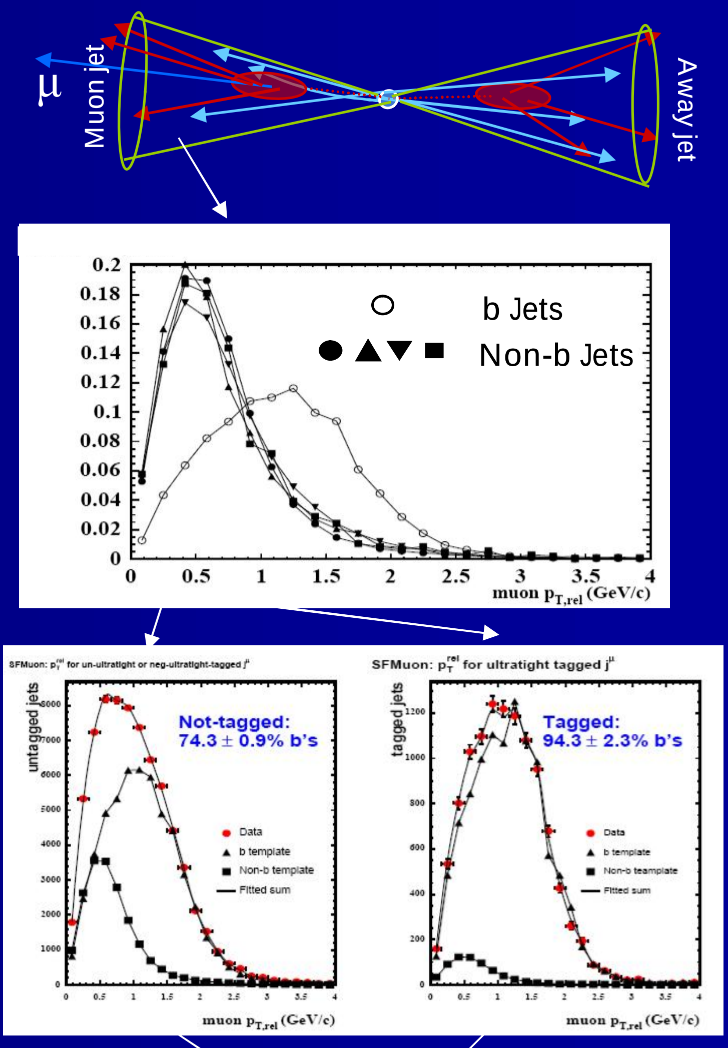
$$A_{W+b\ bjets}=0.593\pm 0.017(\mathrm{syst})$$

- Measure **tag efficiency** for *b* jets in *W*+*b*-jet production in MC and correct to match that of data

$$\varepsilon$$

Tag Efficiency

- Measure tag efficiency for signal *b*-jets in *W*+*b*-jet simulated samples
- Correct via known MC-to-data scale factor
  - Tracking simulation is generally more optimistic than reality
- Measuring the tag efficiency in the data
  - Dijet events, enhance HF content
  - Require probe jet to have a semileptonic hadron decay (muon inside the cone)
  - Muon's relative momentum discriminates *b*/non-*b*
  - Fits for *b* fraction in tagged, untagged samples allow one to extract efficiency



Efficiency for tagging *b* jets in data

ALPGEN W+b-jet MC, after scaling: $\varepsilon = 0.16 \pm 0.01\,(\mathrm{syst})$

Systematic driven by E_T dependence

W+b-jets: Measurement Strategy

$$\sigma_{b\ jets}(W+b-jets)\times BR(W\rightarrow \ell\nu)=\frac{n^{fit}_{bjets}-n^{\text{not}W+b}_{bjets}}{L\cdot A_{W+b\ bjets}\cdot\varepsilon}$$

Where do various pieces come from?

- Discriminate *b*/*c*/LF in tagged sample using **vertex mass**

$$n^{fit}_{bjets}=672.3\pm 44.3(\text{stat})\pm 60.4(\text{syst})$$

- Determine contribution from **background** tagged *b* jets and subtract from overall yield

$$n^{notW+b}_{bjets}=176.8\pm 22.3(\text{syst})$$

- Calculate **acceptance** for $W+b$-jet events

$$A_{W+b\ bjets}=0.593\pm 0.017(\text{syst})$$

- Measure **tag efficiency** for *b* jets in $W+b$-jet production in MC and correct to match that of data

$$\varepsilon=0.16\pm 0.01(\text{syst})$$

Outline

- Motivation
- Measurement Definition
- Experimental apparatus
- Measurement strategy
- *Results*

W+b-jets Cross Section Result

Pieces:

$$n_{bjets}^{fit} = 672.3 \pm 44.3(\text{stat}) \pm 60.4(\text{syst})$$

$$n_{bjets}^{notW+b} = 176.8 \pm 22.3(\text{syst})$$

$$A_{W+b\ bjets} = 0.593 \pm 0.017(\text{syst})$$

$$\varepsilon = 0.16 \pm 0.01(\text{syst})$$

What about the luminosity?

$$L = 1905 pb^{-1}$$

averaged over three trigger paths

Insert pieces here:

$$\sigma_{b\ jets}(W+b-jets) \times BR(W \to \ell\nu) = \frac{n_{bjets}^{fit} - n_{bjets}^{\text{not}W+b}}{L \cdot A_{W+b\ bjets} \cdot \varepsilon}$$

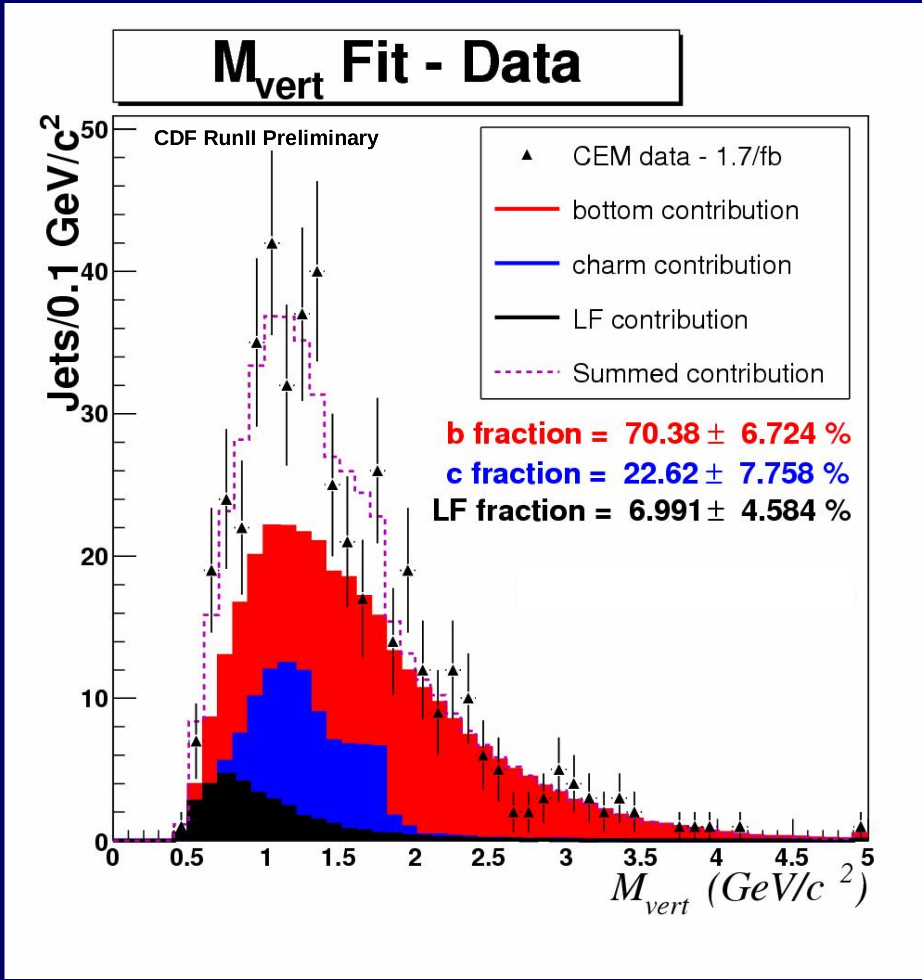
And finally:

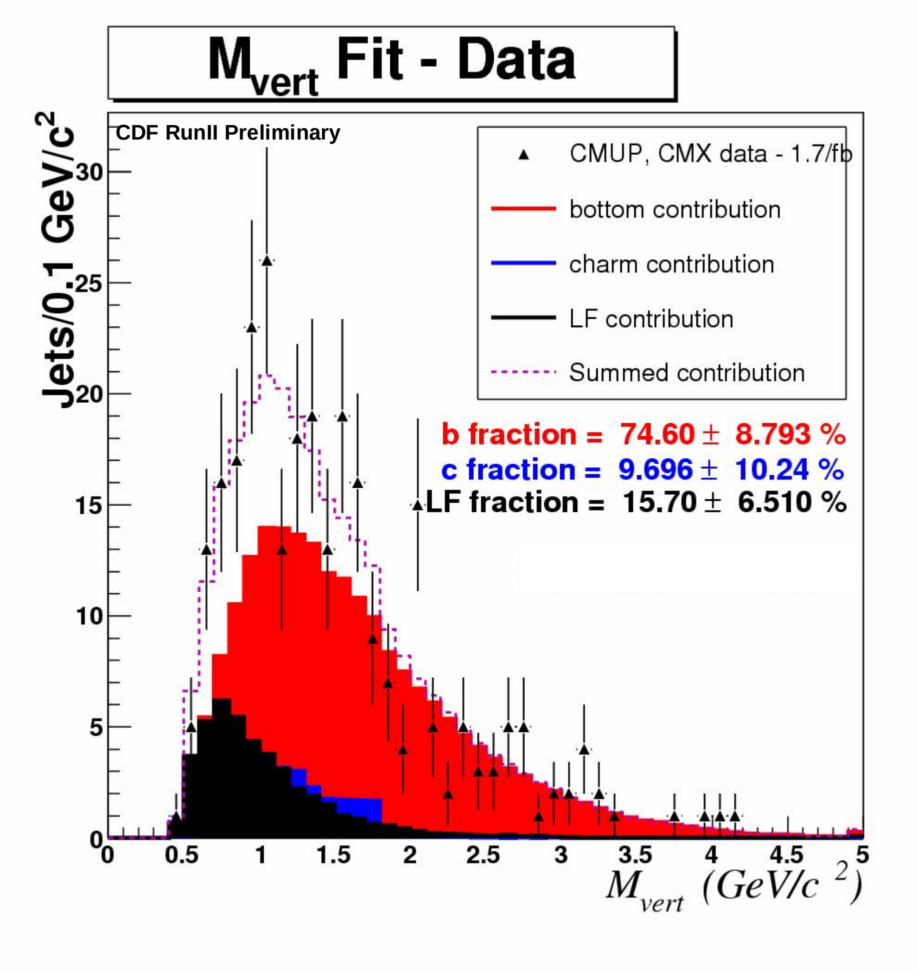
$$\sigma_{\text{b-jets}}(W+b-\text{jets}) \times \text{BR}(W \to \ell\nu) = 2.74 \pm 0.25(\text{stat}) \pm 0.44(\text{syst})\ \text{pb}$$

CDF RunII Preliminary – 1.9/fb

This cross section is for b jets from W+b-jet production in events with a high p_T central lepton, high p_T neutrino and 1 or 2 total jets.

Cross Check: $W \to e\nu$ and $W \to \mu\nu$ Exclusive Results, 1700/pb





$$\sigma_{W+b\ jets} \cdot BR(W \to e\nu) = 2.5 \pm 0.5\,\text{pb}$$

$$\sigma_{W+b\ jets} \cdot BR(W \to \mu\nu) = 3.0 \pm 0.6\,\text{pb}$$

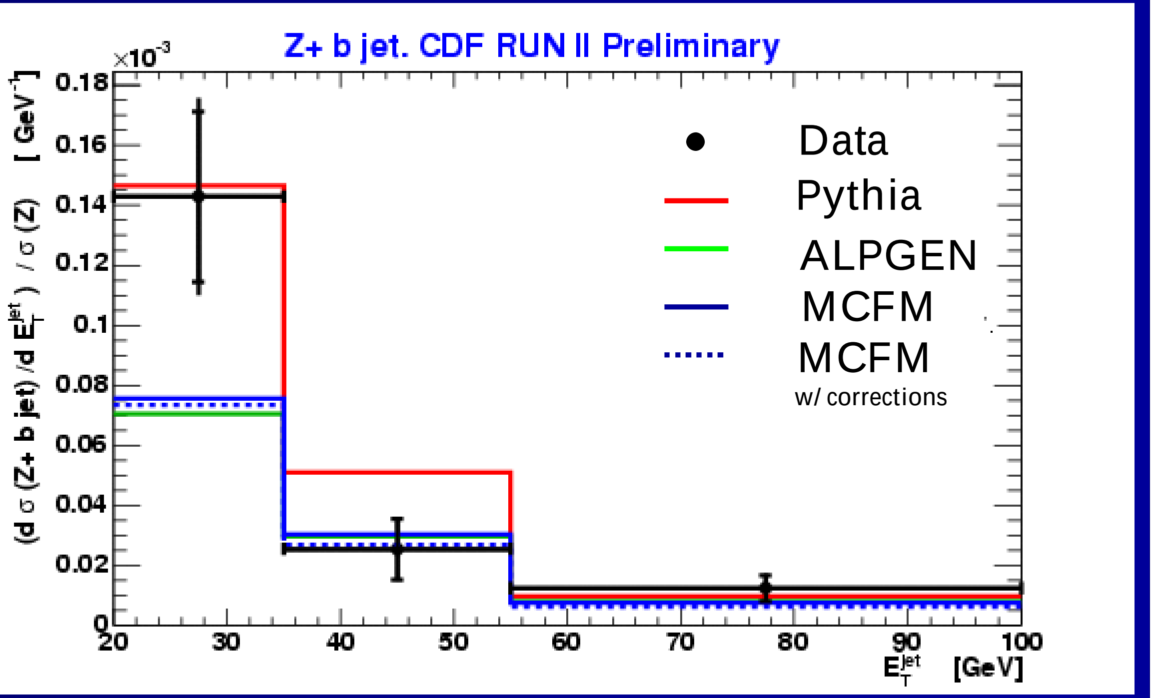
Results for electron and muon triggers are consistent.

Discussion

- Measured *b*-jet cross section in *W*+*b*-jets: 2.74 ± 0.25(stat) ± 0.44(syst) pb
  ALPGEN *b*-jet cross section in *W*+*b*-jets: 0.78 pb
  - Mismatch not unexpected
  - ALPGEN (tree level only) should underestimate cross section, potentially by a large amount if NLO contributions are large
    - Comparison with MCFM at NLO is in the queue

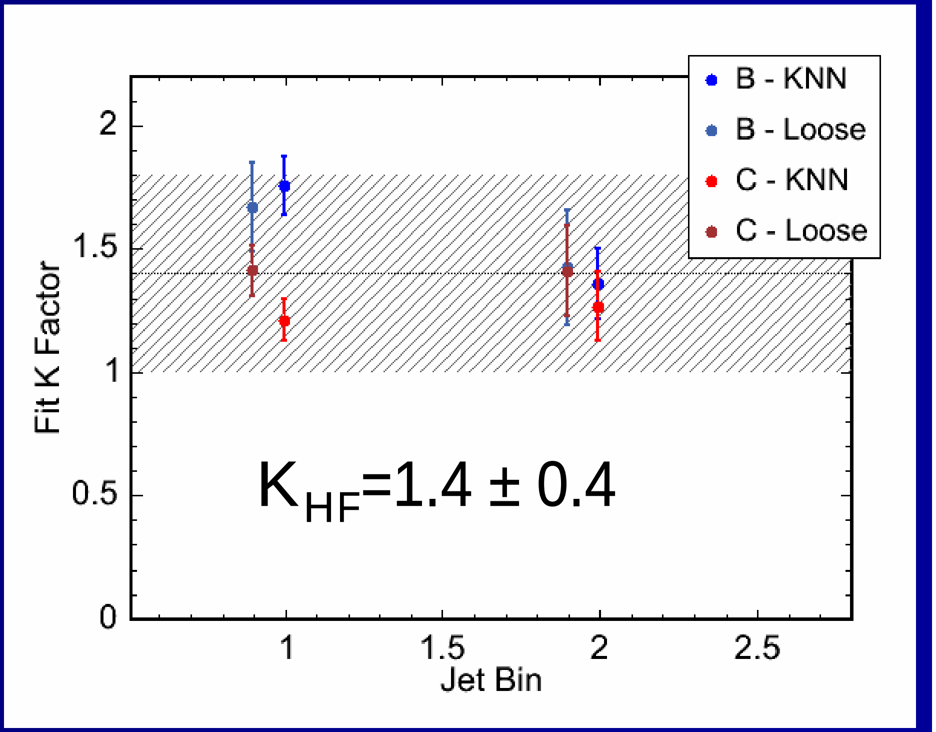
x3.5 deficit in prediction

- CDF sees a mismatch in *Z*+*b*-jets as well
  - Pythia seems to do a reasonable job at low ET
    - Pythia comparison in the queue
  - MCFM prediction similar to ALPGEN – surprising



What Is Next

- $\delta_{syst}(\sigma xBR)/(\sigma xBR) = 16\%$
  - Recall predictions for *W*+*b*-jets had ~40% uncertainty
  - Prediction relies on a fudge factor, K_{HF}, that reconciles the HF content of W+jets in MC to that of a data control sample (eg., W+1 jet)
  - K_{HF} is then used in the prediction for the sample used for the search (eg., W+2 jets)



This result probes the exact same issues.
Minimally this result could contribute to a more precise value for K_{HF}.

More preferable would be to feed this result back to the MC developers for the purpose of improving the models.

Summary

- *W*+*b*-jet production is a formidable obstacle to measuring signatures containing *W*'s and *b*'s

- In an effort to understand the process at a deeper level, we at CDF have measured the *b* jet cross section in events with a *W* boson in 1.9/fb

- Find measured cross section to be <u>2.74 ± 0.25(stat) ± 0.44(syst) pb</u>

- ALPGEN prediction for this process is x3.5 lower than measurement.

- Current work is focused on:
  - Getting more information out of this measurement
  - Using this result to improve the precision on the *W*+*b*-jet predictions that are necessary for single top and Higgs searches.

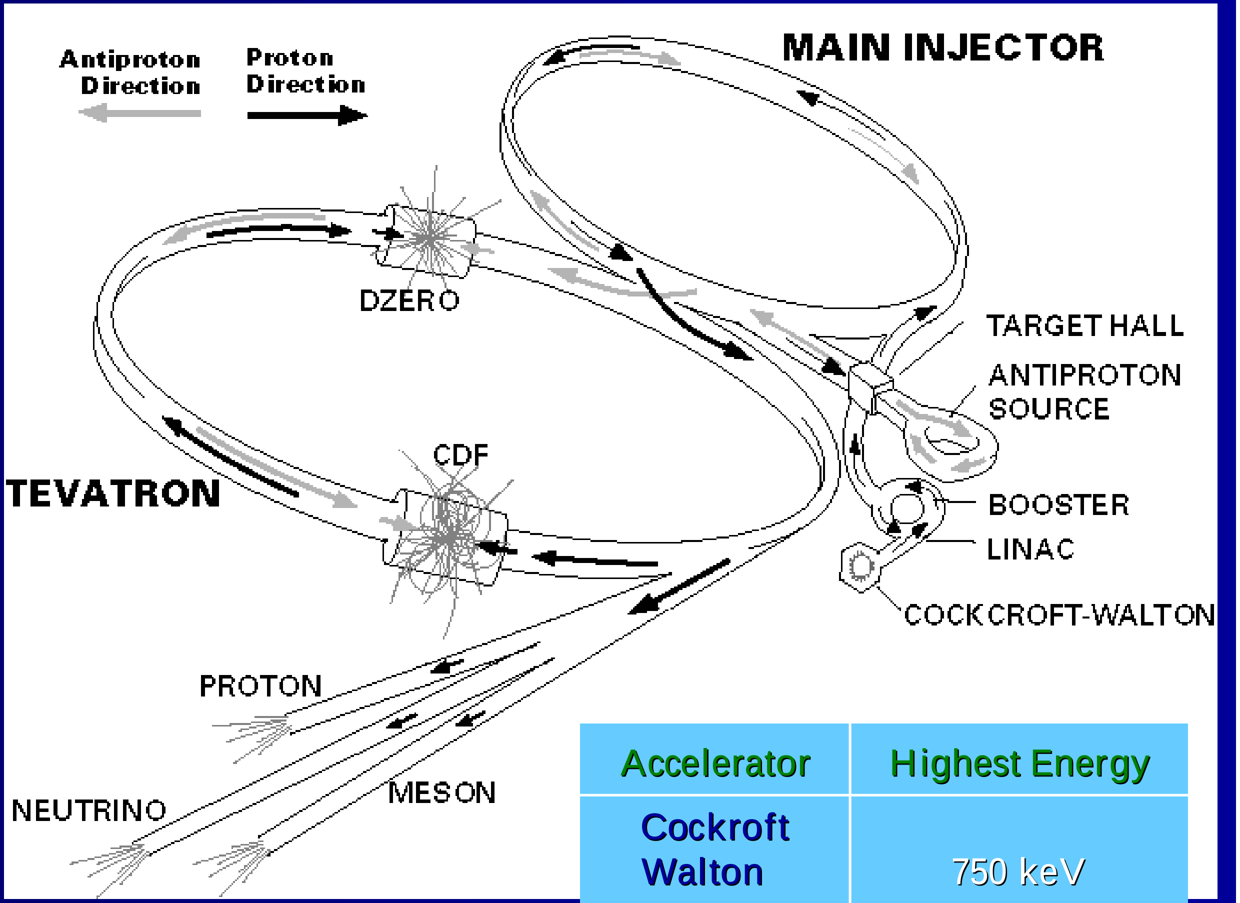
- Goal is to understand this process as best we can both from theory and experimental perspectives – we are on our way

Backup Slides

Fermilab Accelerator Complex

Source: FNAL Accelerator Division

- Series of staged accelerators, culminates in main workhorse, the Tevatron
- Proton and antiproton beams
- Beam is actually discrete packets
  - 396ns between collisions
  - 7 MHz crossing rate
  - ~5E12 particles per packet
- Beams collide with energy 1.96 TeV
- LHC: protons only, 25ns, 40MHz, 1E11 protons per packet, 14TeV



Accelerator	Highest Energy
Cockroft Walton	750 keV
Linac	400 MeV
Booster	8 GeV
Main injector	150 GeV
TEVATRON	980 GeV

RunI: 1992-1995
RunIIa: 2001-2005
RunIIb: 2005-pres

Jet Cross Section Definition

Sample	DSID	$\sigma_{\text{evt}} \times BR$ (pb)	N_{evt}	$n_{b-\text{jets}}^{1\text{or}2}$	$\sigma_{b-\text{jets}} \times BR$ (pb)	w
Wevbb0p	btop0w	2.98	1542539	2.915e+05	0.5631	0.722
Wevbb1p	btop1w	0.89	1545970	2.76e+05	0.1589	0.204
Wevbb2p	btop2w	0.29	1498550	1.196e+05	0.02314	0.030
Wevcc0p	ctop0w	5.00	2005399	49	0.0001222	0.000
Wevcc1p	ctop1w	1.79	1968365	68	6.184e-05	0.000
Wevcc2p	ctop2w	0.628	1885915	55	1.831e-05	0.000
Wevc0p	stopw0	17.1	1943317	44	0.0003872	0.000
Wevc1p	stopw1	3.39	1896728	72	0.0001287	0.000
Wevc2p	stopw2	0.507	1837070	60	1.656e-05	0.000
Wevc3p	stopw3	0.083	1745440	28	1.331e-06	0.000
Wev0p	ptopw0	1800	4868357	65	0.02403	0.031
Wev1p	ptopw1	225	4563248	168	0.008284	0.011
Wev2p	ptop2w	35.3	872814	43	0.001739	0.002
Wev3p	ptop3w	5.59	831222	33	0.0002219	0.000
Wev4p	ptop4w	1.03	775589	8	1.062e-05	0.000
Total					0.780	

Wmvbb0p	btop5w	2.98	1524880	2.897e+05	0.5661	0.721
Wmvbb1p	btop6w	0.89	1508029	2.716e+05	0.1603	0.204
Wmvbb2p	btop7w	0.29	1506613	1.209e+05	0.02328	0.030
Wmvcc0p	ctop5w	5.00	1982424	49	0.0001236	0.000
Wmvcc1p	ctop6w	1.79	1961120	77	7.028e-05	0.000
Wmvcc2p	ctop7w	0.628	1949189	72	2.32e-05	0.000
Wmvc0p	stopw5	17.1	1975397	56	0.0004848	0.001
Wmvc1p	stopw6	3.39	1911713	78	0.0001383	0.000
Wmvc2p	stopw7	0.507	1840847	73	2.011e-05	0.000
Wmvc3p	stopw8	0.507	1754673	36	1.04e-05	0.000
Wmv0p	ptopw5	1800	4955756	72	0.02615	0.033
Wmv1p	ptopw6	225	4648605	135	0.006534	0.008
Wmv2p	ptop7w	35.3	872511	46	0.001861	0.002
Wmv3p	ptop8w	5.59	839645	26	0.0001731	0.000
Wmv4p	ptop9w	5.59	774744	12	7.989e-05	0.000
Total					0.785	

Veto of Events with Fake *W* (aka NonW, aka QCD)

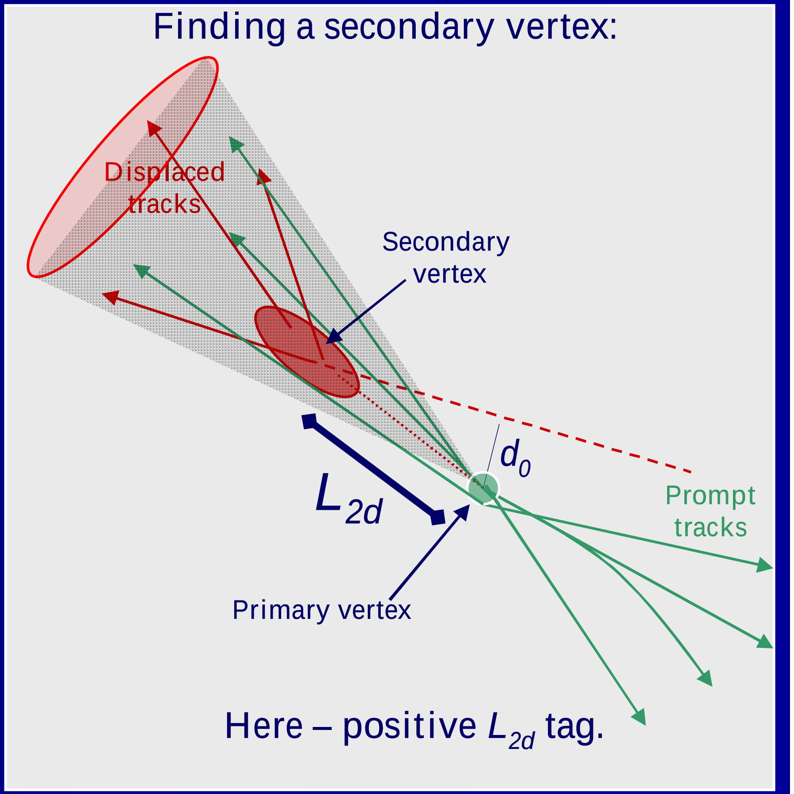
Keep event if…

	1 Jet	2 Jet
CEM	• $M_T(W) > 20$ GeV • $S_{MET} \geq -0.05^{*} M_T(W) + 3.5$ • $S_{MET} \geq -7.6 + 3.2^{*}$ $\Delta\varphi$(ª,j1)	• $M_T(W) > 20$ GeV • $S_{MET} \geq -0.05^{*} M_T(W) + 3.5$ • $S_{MET} \geq 2.5\text{-}3.125^{*}$ $\Delta\varphi$(MET,j2)
CMUP	• $M_T(W) > 10$ GeV • MET $\geq -145 + 60^{*}$ $\Delta\varphi$(ª,j1)	• $M_T(W) > 10$ GeV
CMX	• $M_T(W) > 10$ GeV	• $M_T(W) > 10$ GeV

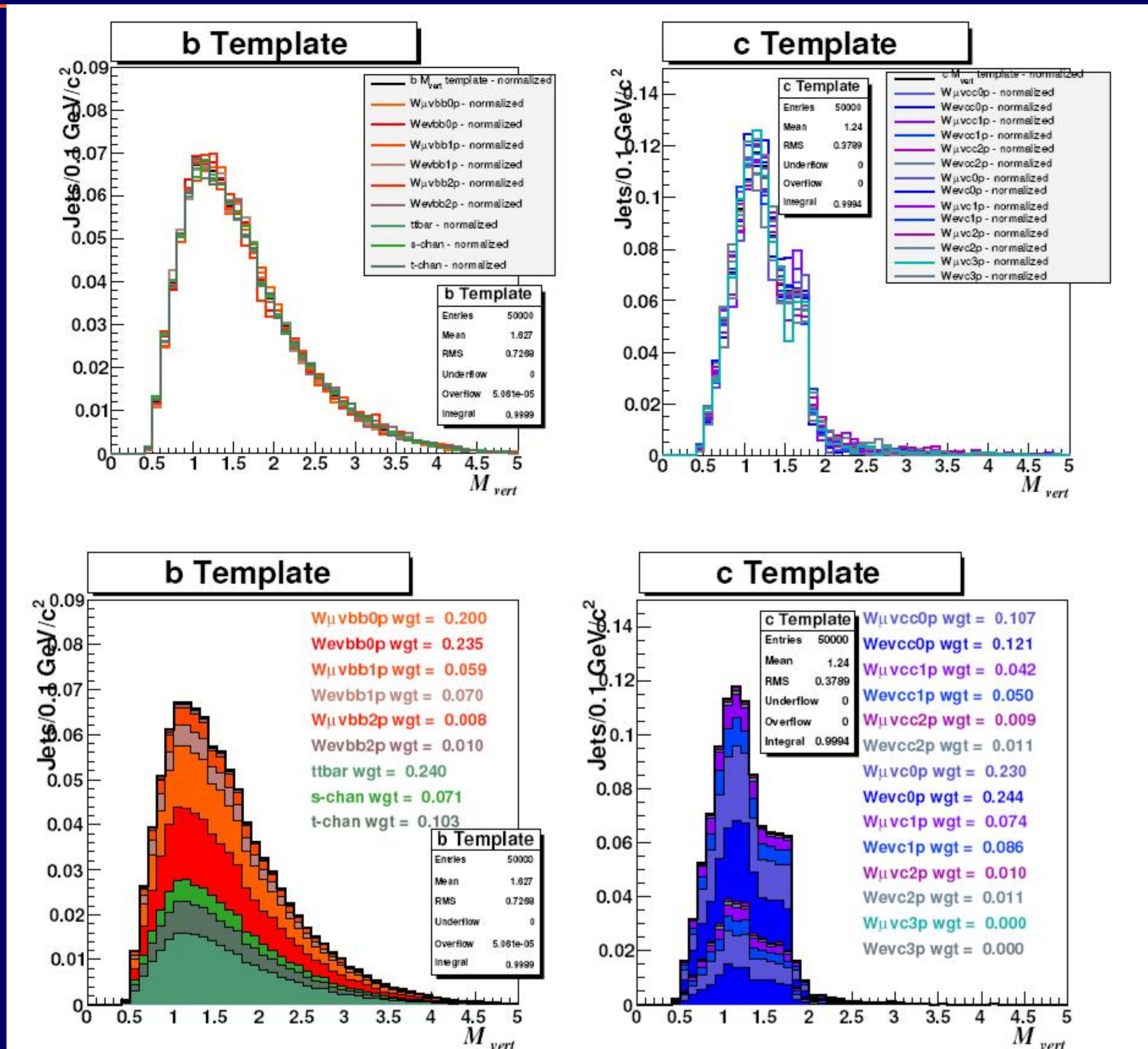
- Seek to eliminate fake *W* events – mostly QCD multijets – hard to model
- Effective non-*W* removal developed by Karlsruhe group for 1.5+/fb single top analyses
- Exploits features of fake *W* events:
  - Low transverse mass of spurious *W*
  - MET from spurious *W* is less significant
  - Correlations between jets and leptons and MET

Secondary vertex *b*-tagging at CDF

- SECVTX algorithm: attempt to construct a secondary vertex among large impact parameter (d_0) tracks using a two-pass scheme
  - Pass1:
    - Starts with construction of 2-track "seed" vertex
    - Attach all remaining tracks that are consistent with seed.
    - Construct the multitrack vertex, iteratively pruning away the attached tracks if they spoil vertex fit.
    - Resulting candidate vertex required to have 3 or more tracks
  - Pass2: tighter track d_0 significance requirement
    - Attempt to vertex all these tracks to a common point.
    - Remove any track that spoils the vertex fit, re-vertexing after each removal.
    - Resulting candidate vertex required to have 2 or more tracks
  - Apply vertex quality cuts
    - removal of K_s,Λ vertices
    - Removal of vertices in the material portion of CDF (beampipe, silicon ladders)
  - If the vertex survives, the jet is "tagged" if S_{L2D} >7.5
    - sign of transverse displacement of secondary vertex wrt interaction point, L_{xy}, determines positive tag or negative tag.

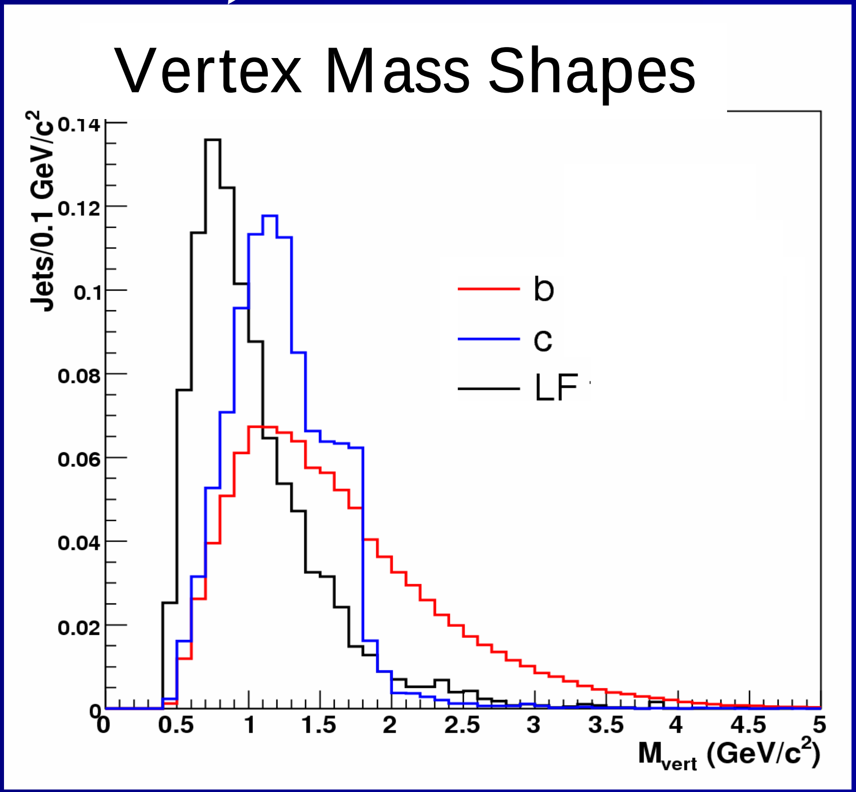


More on *b* and *c* Templates



Likelihood Maximization

$$\mu_i = N^{total}_{jets} [f_b^{fit} \cdot N_b^i + f_c^{fit} \cdot N_c^i + (1.0 - f_b^{fit} - f_b^{fit}) \cdot N^i_{LF}]$$



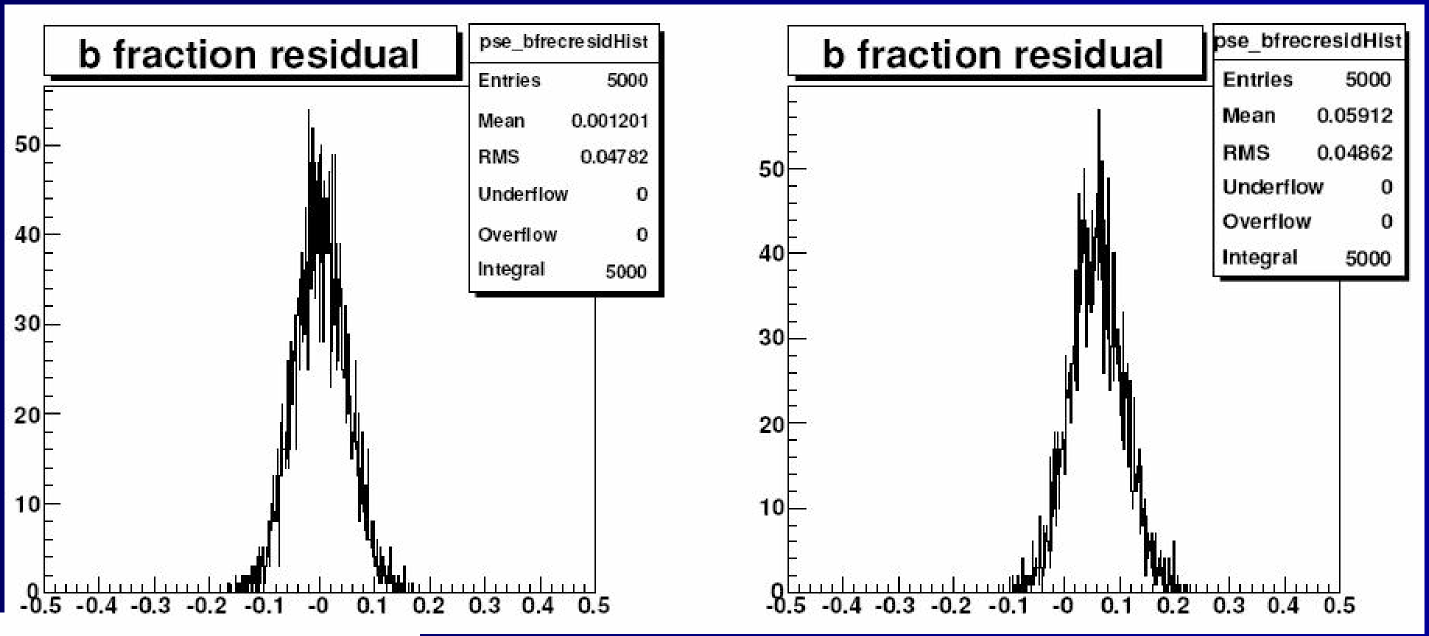
$$P(n_i \mid \mu_i) = \frac{e^{-\mu_i} \mu_i^{n_i}}{n_i!}$$

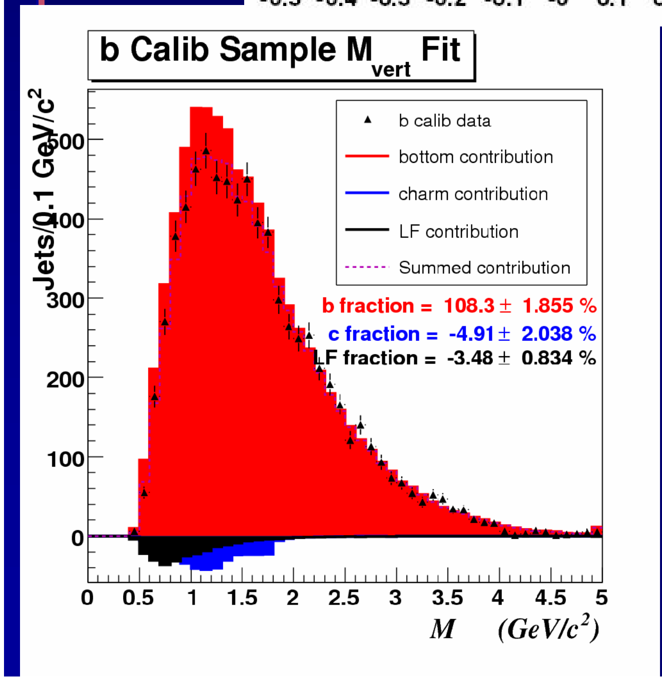
$$L = \prod_{i=1}^{N_{bins}} P(n_i \mid \mu_i)$$

$$\ln L = \ln \left[\prod_{i=1}^{N_{bins}} P(n_i \mid \mu_i) \right]$$

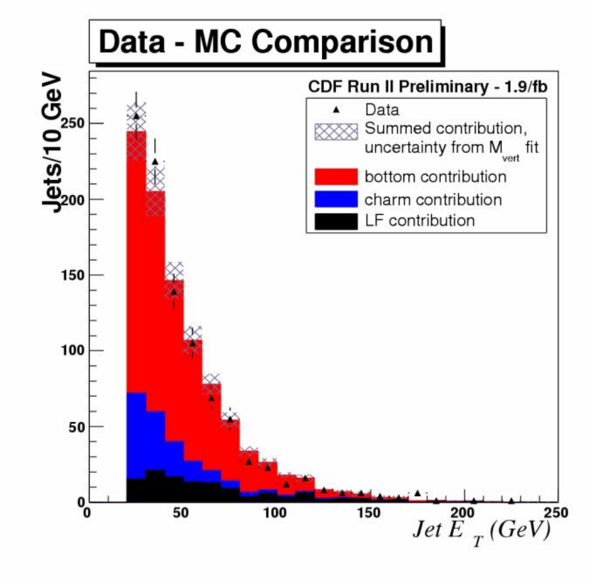
$$\ln L = \sum_{i=1}^{N_{bins}} [-\mu_i + n_i \ln \mu_i + const]$$

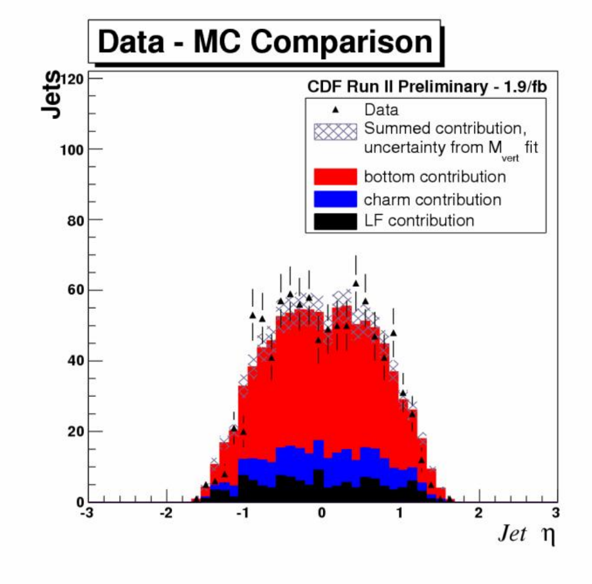
More on *b* Calibration

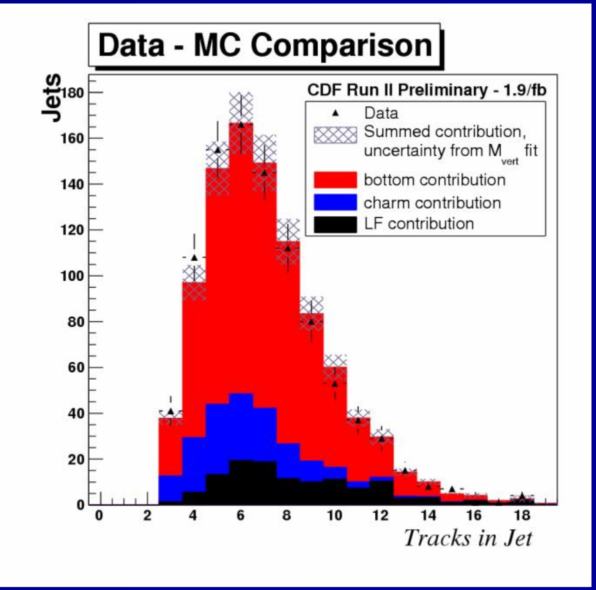


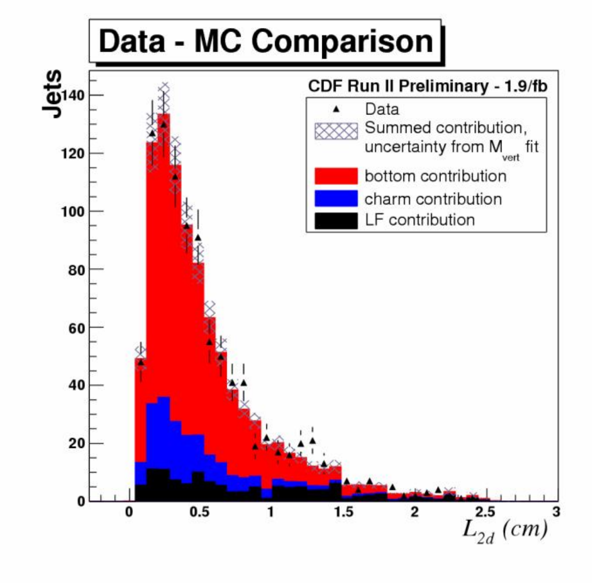


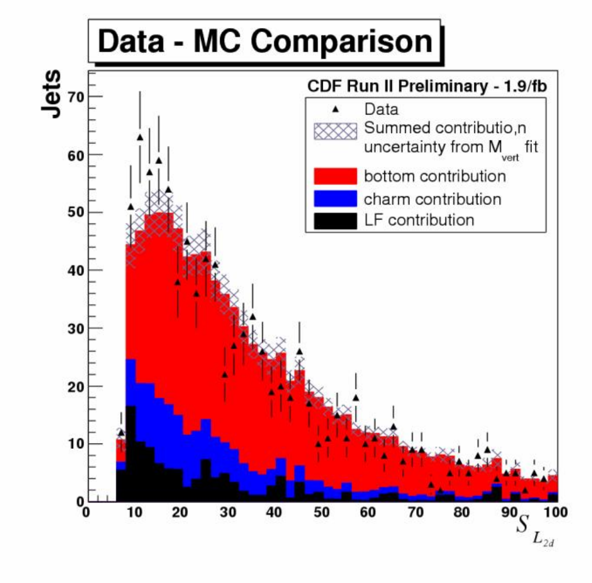
Vertex Mass Fit Consistency Check

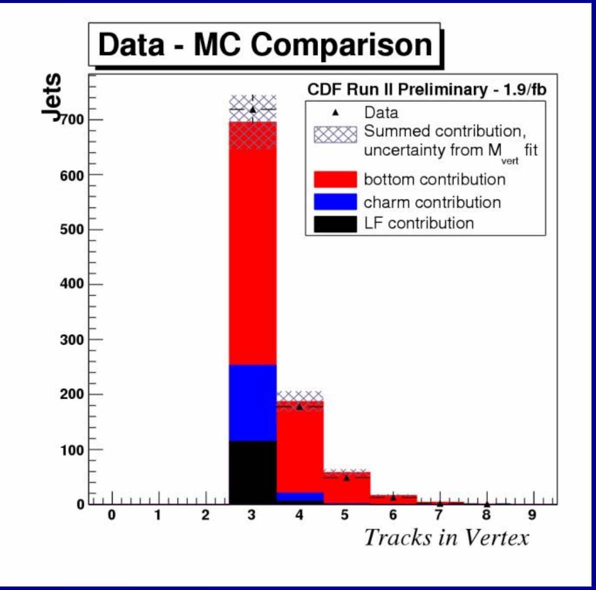




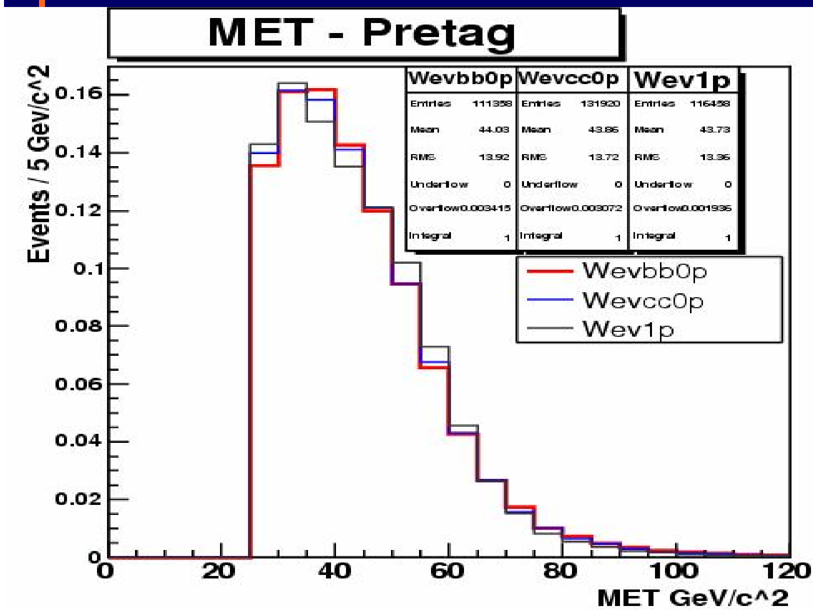


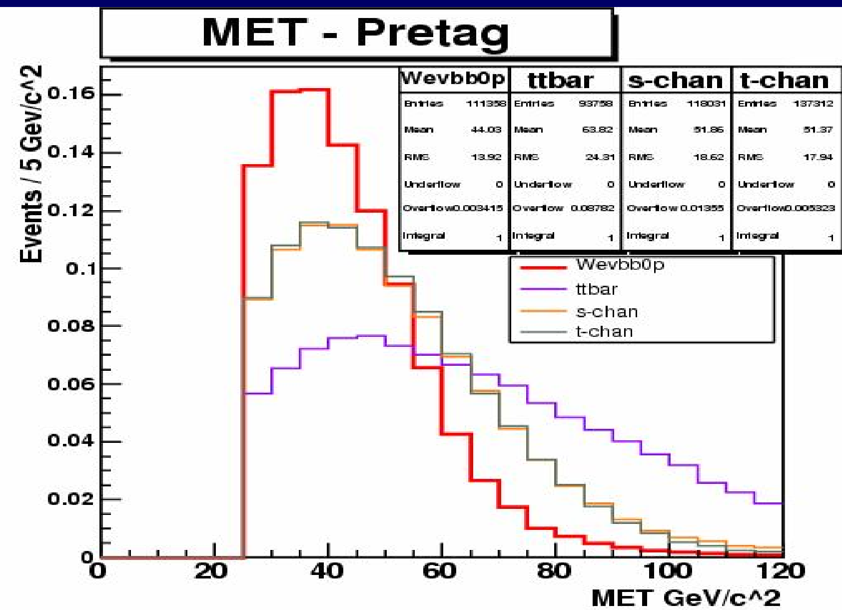


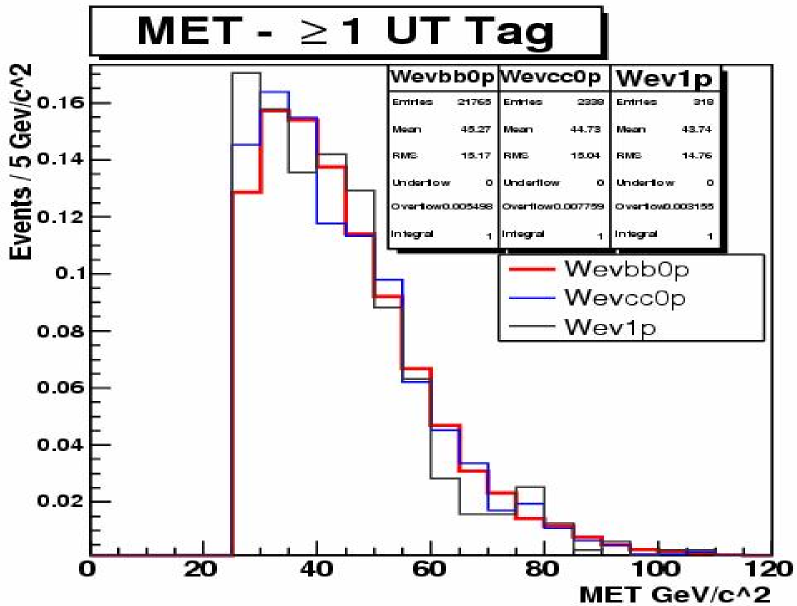


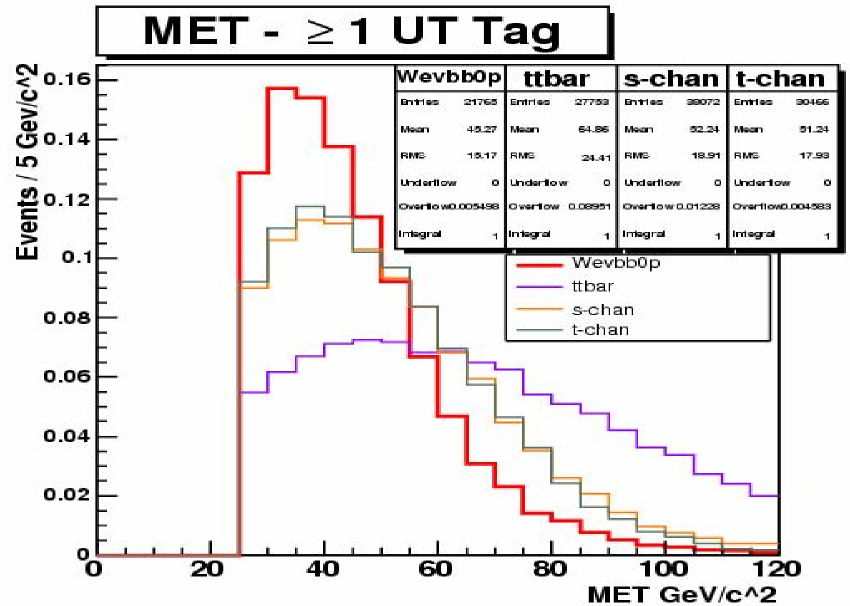


MET comparisons for W12j Contributors

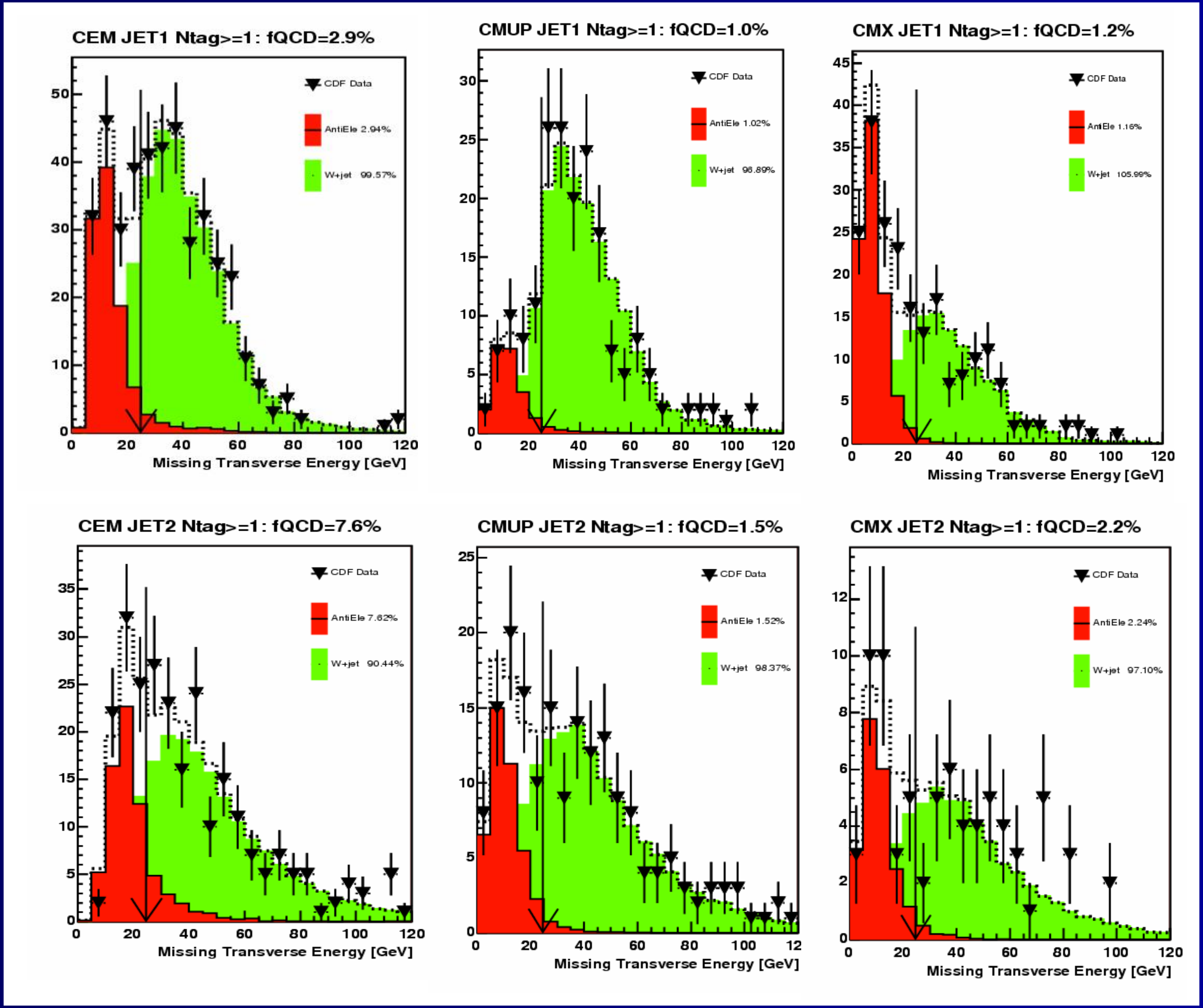








Background Sources of *b* Jets: Fake *W* Events



Acceptance Results

	A_{jet}	A_{sel}	ϵ_{UT}	w	$(A \times \epsilon_{tag})_{CEM}$
Wevbb0p	0.7871 ± 0.0008	0.4815 ± 0.001	0.1556 ± 0.001	0.7218	0.04256 ± 0.0003
Wevbb1p	0.6798 ± 0.0009	0.5629 ± 0.001	0.1606 ± 0.001	0.2037	$0.01252 \pm$ 9e-05
Wevbb2p	0.6811 ± 0.001	0.576 ± 0.002	0.1592 ± 0.002	0.02966	$0.001853 \pm$ 2e-05
Wevcc0p	1.306 ± 0.01	0.3438 ± 0.06	0 ± 0	0.0001566	0 ± 0
Wevcc1p	1.132 ± 0.01	0.3636 ± 0.05	0.03143 ± 0.03	7.927e-05	0 ± 0
Wevcc2p	0.9455 ± 0.03	0.4615 ± 0.07	0.03667 ± 0.04	2.348e-05	0 ± 0
Wevc0p	1.295 ± 0.01	0.1754 ± 0.05	0 ± 0	0.0004963	0 ± 0
Wevc1p	1.069 ± 0.01	0.3377 ± 0.05	0 ± 0	0.000165	0 ± 0
Wevc2p	1.2 ± 0.01	0.625 ± 0.06	0.01956 ± 0.02	2.123e-05	0 ± 0
Wevc3p	0.6786 ± 0.09	0.6316 ± 0.1	0 ± 0	1.707e-06	0 ± 0
Wev0p	1.538 ± 0.01	0.22 ± 0.04	0 ± 0	0.03081	0 ± 0
Wev1p	1.125 ± 0.01	0.3069 ± 0.03	0 ± 0	0.01062	0 ± 0
Wev2p	0.7209 ± 0.07	0.6452 ± 0.09	0.044 ± 0.05	0.002229	0 ± 0
Wev3p	0.7576 ± 0.07	0.56 ± 0.1	0 ± 0	0.0002845	0 ± 0
Wev4p	1.125 ± 0.01	0.4444 ± 0.2	0 ± 0	1.362e-05	0 ± 0
Total					0.0569 ± 0.0003

- A_{jet}, A_{sel} behavior different for elec, muon triggers – from tight jet+lep counting
- Ultratight tag efficiency stable across samples, triggers

Acceptance Results

	A_{jet}	A_{sel}	ϵ_{UT}	w	$(A \times \epsilon_{tag})_{CMUP}$
Wmvbb0p	0.8494 ± 0.0007	0.2559 ± 0.0009	0.1557 ± 0.001	0.7208	0.02439 ± 0.0002
Wmvbb1p	0.8867 ± 0.0006	0.2466 ± 0.0009	0.1579 ± 0.001	0.2041	$0.007043 \pm$ 7e-05
Wmvbb2p	0.9532 ± 0.0006	0.239 ± 0.001	0.1556 ± 0.002	0.02964	$0.001051 \pm$ 2e-05
Wmvcc0p	0.898 ± 0.04	0.1364 ± 0.05	0 ± 0	0.0001574	0 ± 0
Wmvcc1p	$1.052 \pm$ nan	0.2469 ± 0.05	0 ± 0	8.95e-05	0 ± 0
Wmvcc2p	0.9861 ± 0.01	0.09859 ± 0.04	0 ± 0	2.954e-05	0 ± 0
Wmvc0p	0.75 ± 0.06	0.2619 ± 0.07	0 ± 0	0.0006173	0 ± 0
Wmvc1p	$1.231 \pm$ nan	0.3438 ± 0.05	0 ± 0	0.0001761	0 ± 0
Wmvc2p	0.9726 ± 0.02	0.2676 ± 0.05	0 ± 0	2.56e-05	0 ± 0
Wmvc3p	$1.139 \pm$ nan	0.1707 ± 0.06	0 ± 0	1.325e-05	0 ± 0
Wmv0p	0.6528 ± 0.06	0.1915 ± 0.06	0 ± 0	0.0333	0 ± 0
Wmv1p	0.9926 ± 0.007	0.2985 ± 0.04	0 ± 0	0.008321	0 ± 0
Wmv2p	0.8261 ± 0.06	0.2105 ± 0.07	0 ± 0	0.00237	0 ± 0
Wmv3p	1 ± 0	0.1923 ± 0.08	0 ± 0	0.0002204	0 ± 0
Wmv4p	$1.25 \pm$ nan	0.3333 ± 0.1	0 ± 0	0.0001017	0 ± 0
Total					0.0325 ± 0.0002

Acceptance Results

	A_{jet}	A_{sel}	ϵ_{UT}	w	$(A \times \epsilon_{tag})_{CMX}$
Wmvbb0p	0.8494 ± 0.0007	0.1345 ± 0.0007	0.1543 ± 0.002	0.7208	0.01271 ± 0.0002
Wmvbb1p	0.8867 ± 0.0006	0.1339 ± 0.0007	0.1624 ± 0.002	0.2041	$0.003935 \pm$ 5e-05
Wmvbb2p	0.9532 ± 0.0006	0.1311 ± 0.001	0.1594 ± 0.003	0.02964	$0.0005902 \pm$ 1e-05
Wmvcc0p	0.898 ± 0.04	0.1591 ± 0.06	0 ± 0	0.0001574	0 ± 0
Wmvcc1p	$1.052 \pm$ nan	0.08642 ± 0.03	0 ± 0	8.95e-05	0 ± 0
Wmvcc2p	0.9861 ± 0.01	0.2113 ± 0.05	0 ± 0	2.954e-05	0 ± 0
Wmvc0p	0.75 ± 0.06	0.2143 ± 0.06	0 ± 0	0.0006173	0 ± 0
Wmvc1p	$1.231 \pm$ nan	0.1042 ± 0.03	0 ± 0	0.0001761	0 ± 0
Wmvc2p	0.9726 ± 0.02	0.1549 ± 0.04	0 ± 0	2.56e-05	0 ± 0
Wmvc3p	$1.139 \pm$ nan	0.122 ± 0.05	0 ± 0	1.325e-05	0 ± 0
Wmv0p	0.6528 ± 0.06	0.06383 ± 0.04	0 ± 0	0.0333	0 ± 0
Wmv1p	0.9926 ± 0.007	0.1119 ± 0.03	0 ± 0	0.008321	0 ± 0
Wmv2p	0.8261 ± 0.06	0.1316 ± 0.05	0 ± 0	0.00237	0 ± 0
Wmv3p	1 ± 0	0.1923 ± 0.08	0 ± 0	0.0002204	0 ± 0
Wmv4p	$1.25 \pm$ nan	0.1333 ± 0.09	0 ± 0	0.0001017	0 ± 0
Total					0.0172 ± 0.0002

Acceptance Systematics

Scale choice impact:

		$(A \times \epsilon_{tag})_i$	$(A \times \epsilon_{tag})_{CEM}$
	Wevbb0p	0.0467 ± 0.0006	
k =0.5	Wevbb1p	0.0128 ± 0.0002	0.0615 ± 0.0006
	Wevbb2p	$0.0020 \pm$ 4e-05	
	Wevbb0p	0.0447 ± 0.0003	
Default $k = 1$	Wevbb1p	0.0133 ± 0.0001	0.0600 ± 0.0003
	Wevbb2p	$0.0020 \pm$ 3e-05	
	Wevbb0p	0.0456 ± 0.0006	
k =2.0	Wevbb1p	0.0135 ± 0.0002	0.0611 ± 0.0006
	Wevbb2p	$0.0019 \pm$ 4e-05	

		$(A \times \epsilon_{tag})_i$	$(A \times \epsilon_{tag})_{CMUP}$
	Wmvbb0p	0.0257 ± 0.0003	
k =0.5	Wmvbb1p	$0.0076\pm$ 7e-05	0.034 ± 0.0003
	Wmvbb2p	$0.0011 \pm$ 2e-05	
	Wmvbb0p	0.0257 ± 0.0002	
Default $k = 1$	Wmvbb1p	$0.0076 \pm$ 6e-05	0.034 ± 0.0002
	Wmvbb2p	$0.0011 \pm$ 2e-05	
	Wmvbb0p	0.0266 ± 0.0003	
k =2.0	Wmvbb1p	$0.0076\pm$ 7e-05	0.035 ± 0.0003
	Wmvbb2p	$0.0011 \pm$ 2e-05	

		$(A \times \epsilon_{tag})_i$	$(A \times \epsilon_{tag})_{CMX}$
	Wmvbb0p	0.0133 ± 0.0003	
k =0.5	Wmvbb1p	$0.0038\pm$ 7e-05	0.018 ± 0.0003
	Wmvbb2p	$0.0007 \pm$ 2e-05	
	Wmvbb0p	0.0133 ± 0.0002	
Default $k = 1$	Wmvbb1p	$0.0038\pm$ 6e-05	0.018 ± 0.0002
	Wmvbb2p	$0.0006 \pm$ 1e-05	
	Wmvbb0p	0.0133 ± 0.0003	
k =2.0	Wmvbb1p	$0.0038\pm$ 7e-05	0.018 ± 0.0003
	Wmvbb2p	$0.0006 \pm$ 2e-05	

- Choice of scale, Q^2:
  - Choice could affect jet E_T and η distributions, which impacts jet counting
  - ALPGEN scale chosen via:

$$Q^2 = k \cdot (M_W^2 + \sum_p (m_p^2 + p_{T,p}^2))$$

$\delta(A \times \varepsilon)/(A \times \varepsilon) = 3\%$

- Choice of PDF:
  - Evaluated in 700/pb analysis, small

$\delta(A \times \varepsilon)/(A \times \varepsilon) = 2\%$

$$A_{W+b\ bjets} \cdot \varepsilon_{tag} = 0.057 \pm 0.005\,(\text{syst})\ \text{CEM}$$
$$= 0.031 \pm 0.003\,(\text{syst})\ \text{CMUP}$$
$$= 0.017 \pm 0.001\,(\text{syst})\ \text{CMX}$$

Acceptance Systematics

- Sources of systematic error on $A_{W+b\ bjets} \cdot \varepsilon_{tag}$
  - Tag efficiency
  - Jet Energy Corrections
  - Q^2 (event level and per-vertex level)
  - PDFs

- We know tag efficiency syst:
  - Comes from imprecise calibration of tag efficiency for data *b* jets, familiar to most from "the scale factor"

$\delta(A \times \varepsilon)/(A \times \varepsilon) = 6\%$

- Have quantified JES:
  - Look at $\pm 1\sigma$ variations on the L5 jet energy correction

$\delta(A \times \varepsilon)/(A \times \varepsilon) = 3\%$

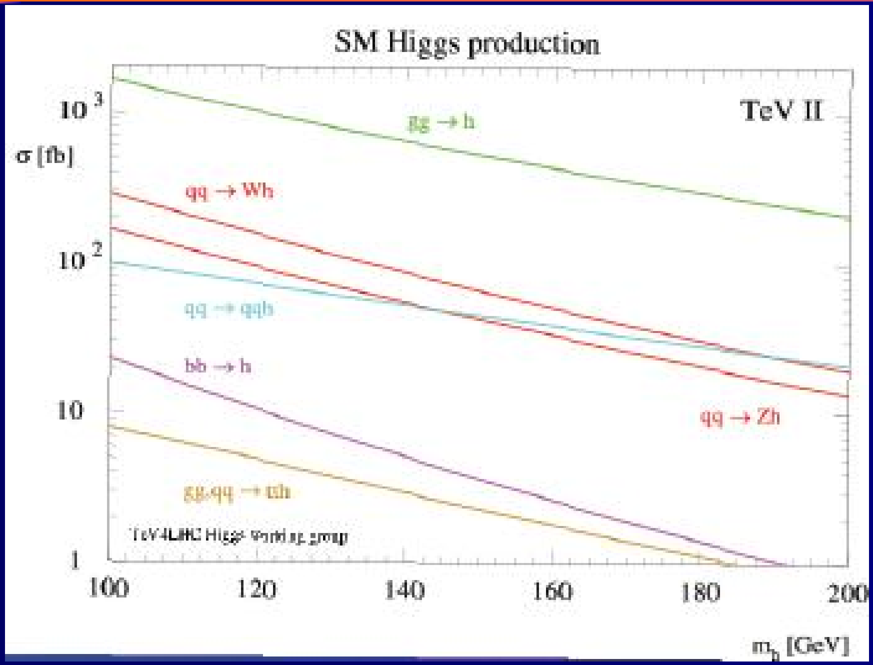
JES uncertainty impact:

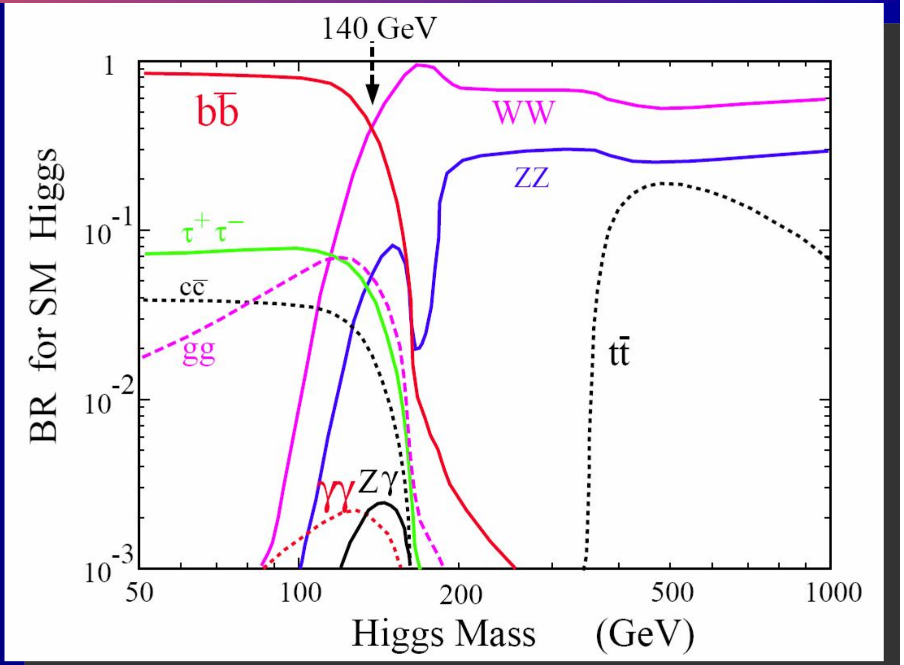
		$(A \times \epsilon_{tag})_i$	$(A \times \epsilon_{tag})_{CEM}$
	Wevbb0p	0.0466 ± 0.0004	
JES $+1\sigma$	Wevbb1p	0.0132 ± 0.0001	0.0616 ± 0.0004
	Wevbb2p	$0.0018 \pm$ 3e-05	
	Wevbb0p	0.0447 ± 0.0003	
Default	Wevbb1p	0.0133 ± 0.0001	0.0600 ± 0.0003
	Wevbb2p	$0.0020 \pm$ 3e-05	
	Wevbb0p	0.0434 ± 0.0004	
JES -1σ	Wevbb1p	0.0126 ± 0.0001	0.0581 ± 0.0004
	Wevbb2p	$0.0021 \pm$ 3e-05	

		$(A \times \epsilon_{tag})_i$	$(A \times \epsilon_{tag})_{CMUP}$
	Wmvbb0p	0.0271 ± 0.0003	
JES $+1\sigma$	Wmvbb1p	$0.0072 \pm$ 7e-05	0.035 ± 0.0003
	Wmvbb2p	$0.0011 \pm$ 3e-05	
	Wmvbb0p	0.0257 ± 0.0002	
Default	Wmvbb1p	$0.0076 \pm$ 6e-05	0.034 ± 0.0002
	Wmvbb2p	$0.0011 \pm$ 2e-05	
	Wmvbb0p	0.0241 ± 0.0003	
JES -1σ	Wmvbb1p	$0.0075 \pm$ 7e-05	0.033 ± 0.0003
	Wmvbb2p	$0.0012 \pm$ 3e-05	

		$(A \times \epsilon_{tag})_i$	$(A \times \epsilon_{tag})_{CMX}$
	Wmvbb0p	0.0133 ± 0.0003	
JES $+1\sigma$	Wmvbb1p	$0.0038 \pm$ 7e-05	0.018 ± 0.0003
	Wmvbb2p	$0.0006 \pm$ 2e-05	
	Wmvbb0p	0.0133 ± 0.0002	
Default	Wmvbb1p	$0.0038 \pm$ 6e-05	0.018 ± 0.0002
	Wmvbb2p	$0.0006 \pm$ 1e-05	
	Wmvbb0p	0.0123 ± 0.0003	
JES -1σ	Wmvbb1p	$0.0038 \pm$ 7e-05	0.017 ± 0.0003
	Wmvbb2p	$0.0007 \pm$ 2e-05	

The Search for the Higgs Boson

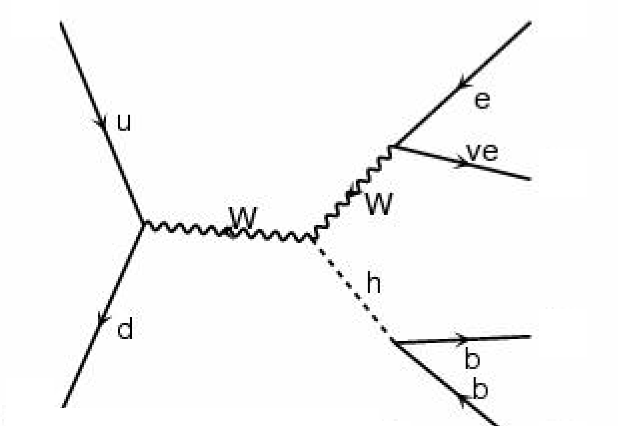




- Electroweak symmetry is broken in SM
  - Imposition of mass to fundamental particles
- EWSB in the Standard Model: Higgs Mechanism
  - Additional consequence: existence of Higgs boson
  - Not yet observed – a missing piece of the puzzle
- Promising Tevatron production mode:

  $p\,p \rightarrow W^* \rightarrow W^{\pm}\,H$: ~0.1-0.2 pb
- Higgs decays to *b* quarks if its mass is low:

  $H \rightarrow b\,b$



***WH* production**

Systematics Summary

Source	$\frac{\delta_{\sigma_{b-\text{jets}} \times BR}}{\sigma_{b-\text{jets}} \times BR}$ (%)
b shape modeling	8
c shape modeling	1
LF shape modeling	3
UT tag efficiency	6
Luminosity	6
Top Cross Sections	2
Fake $W^{\pm}$ $\not{E}_T$ fits	1
Tagged Fake $W^{\pm}$ b fraction	1
Jet Energy Scale	3
Q^2	3
PDF	2
$\lvert z_0 \rvert$ efficiency	<1
Trigger efficiency	<1
Lepton ID efficiency	<1