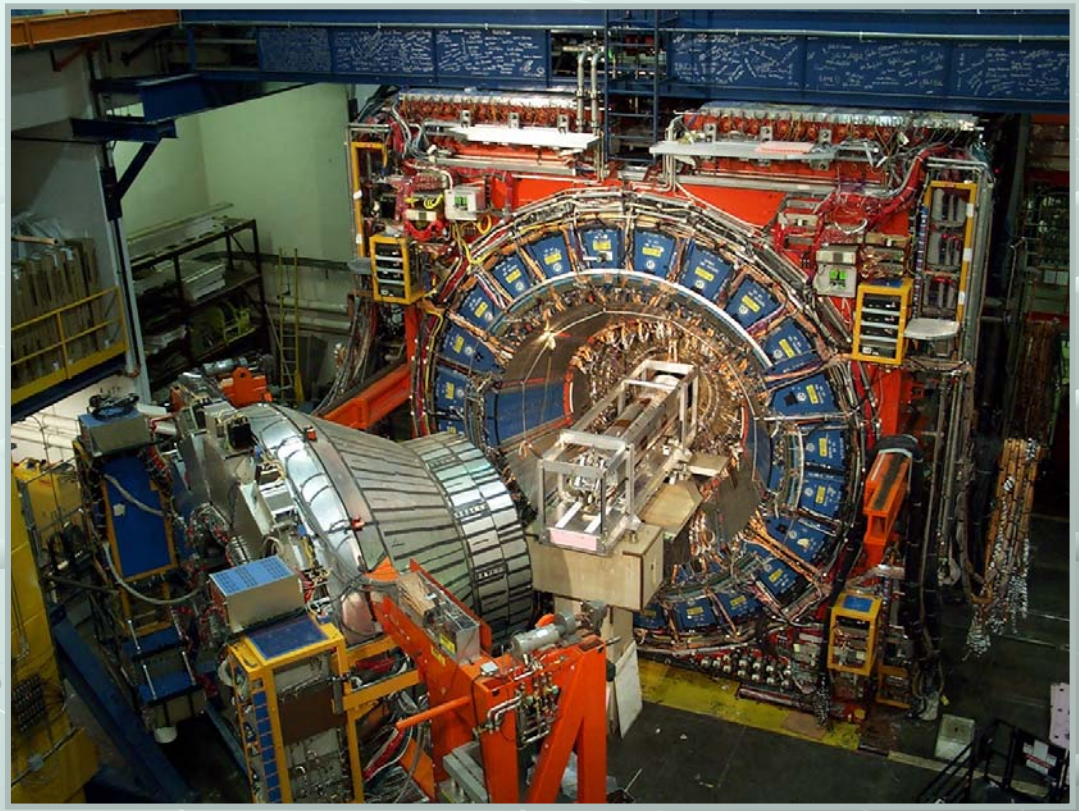
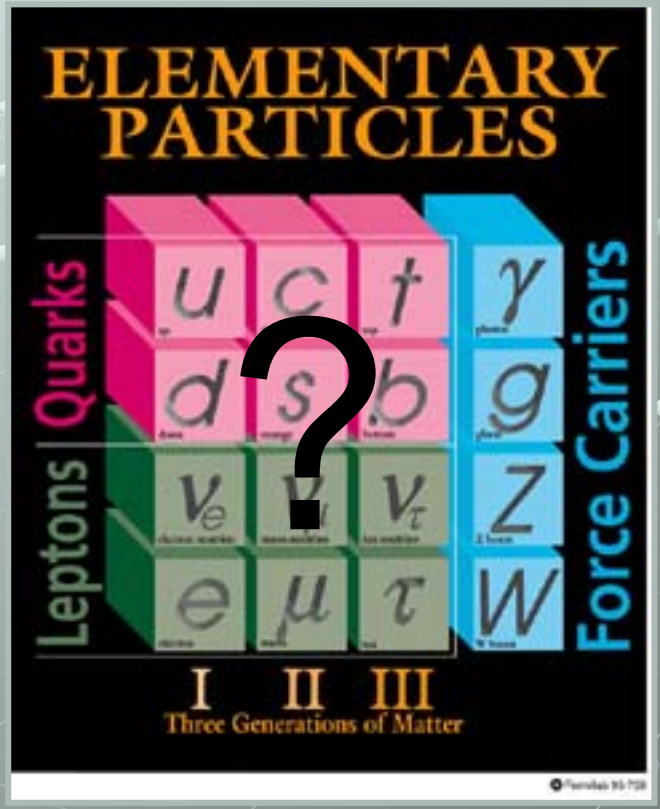
SEARCH FOR SM *HIGGS BOSON* USING ME+BDT WITH 2.7 fb^{-1} CDF DATA





Working group:

B. Álvarez, F. Canelli, B.Casal, J. Cuevas, P. Dong, R. Eusebi, C. Group, E. Palencia, A. Ruiz, B. Stelzer, R. Vilar, R. Wallny

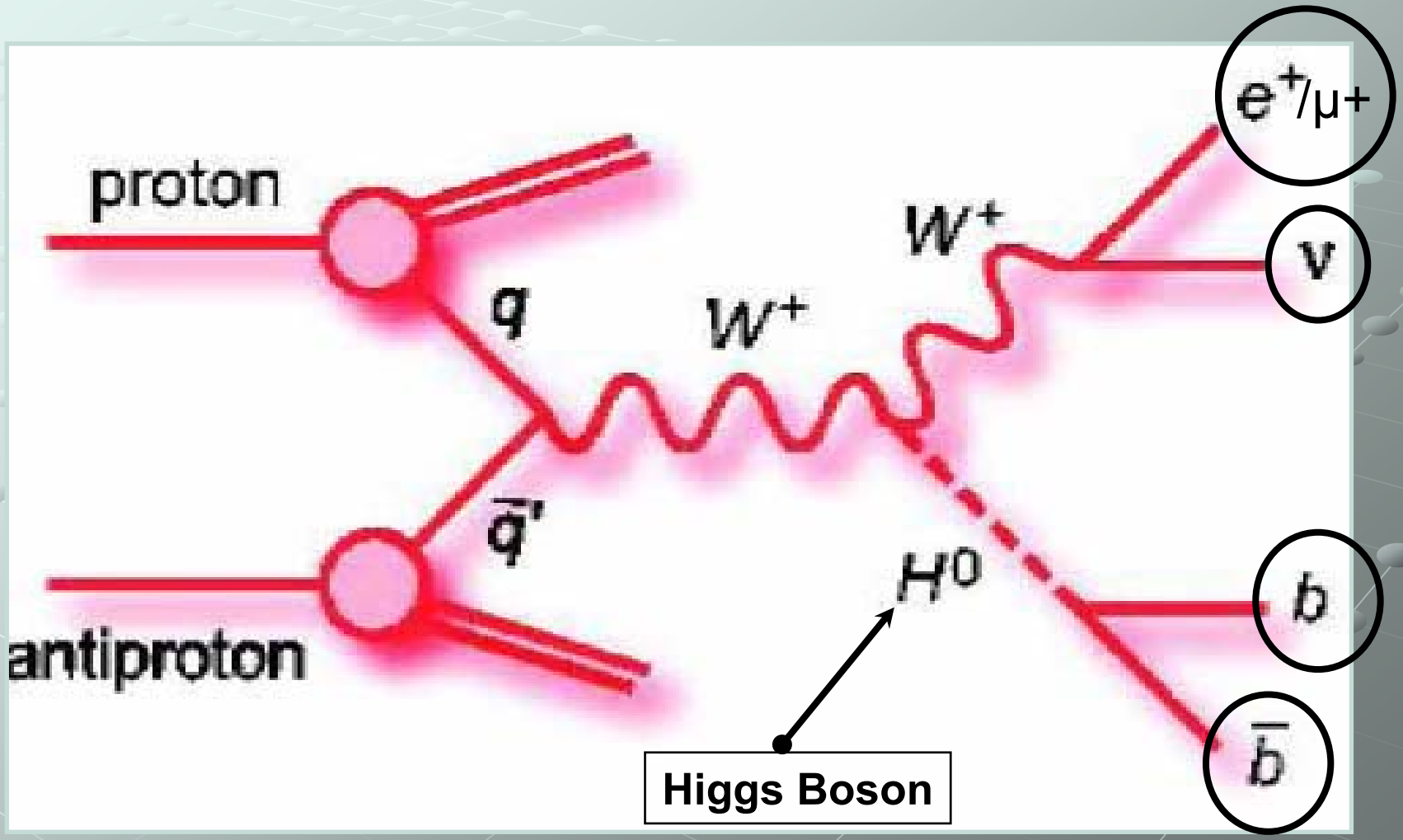
CDF Collaboration Meeting, October 16-17, 2008

OUTLINE

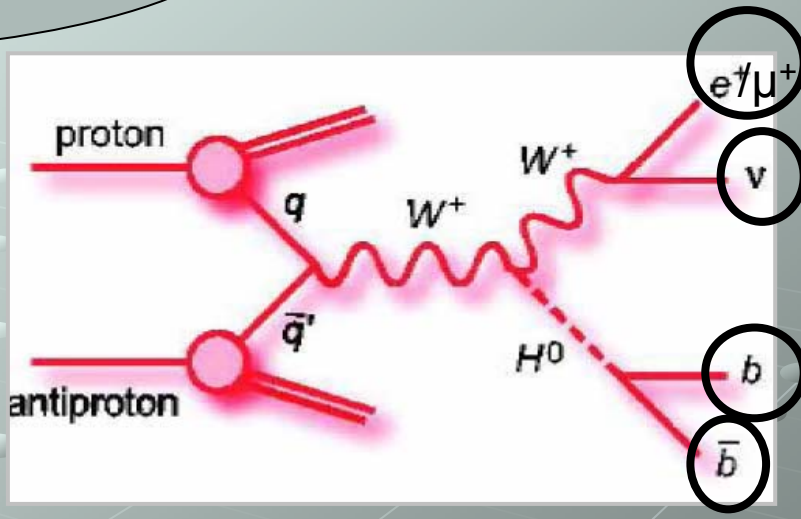
- Event Selection.
- Background Estimation.
- Matrix Element Method.
- Boosted Decision Trees Method.
- Systematics.
- Results (using MCLIMIT).
- Summary.

WH CHANNEL

At low mass m_H<135 GeV/c^2

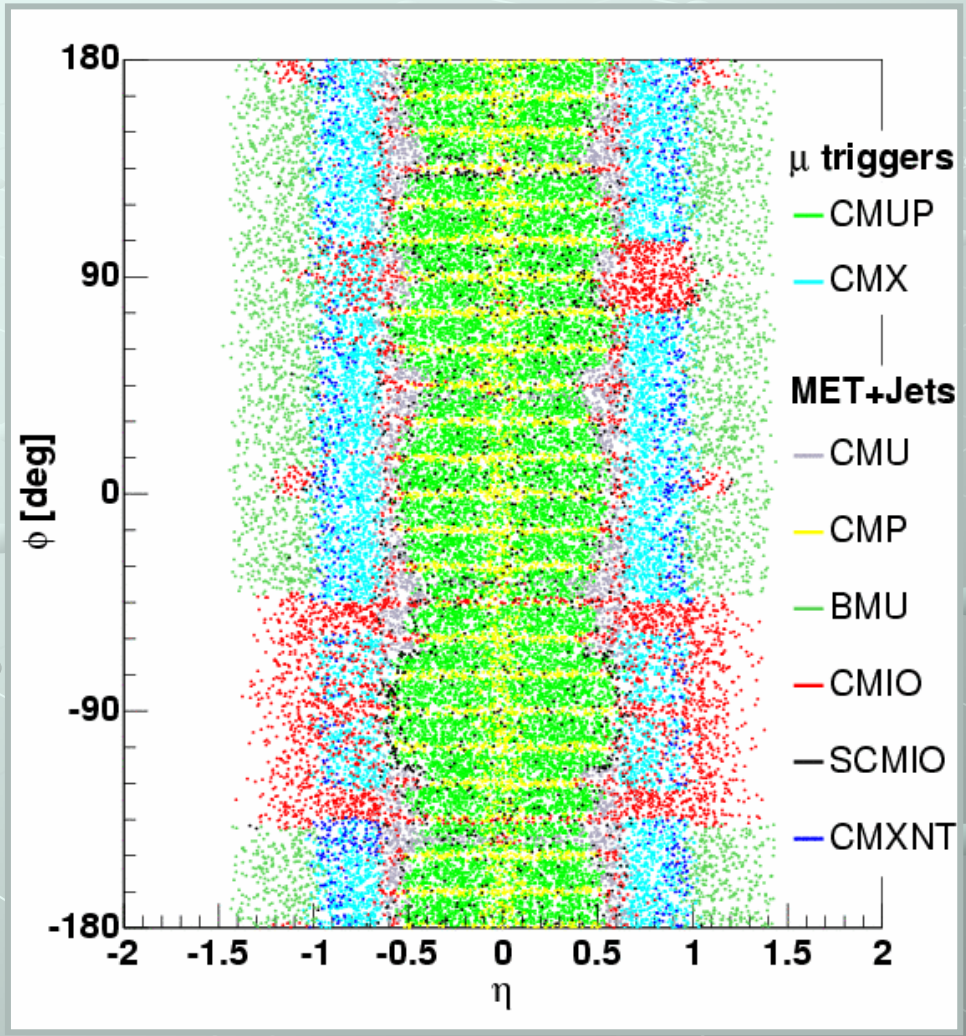


EVENT SELECTION



- One high p_T isolated lepton (e/μ): p_T> 20 GeV
- Two tight jets:
  - E_T(L5)>20 GeV
  - |η| < 2.0
- Cosmic, conversion, dilepton and Z vetos.
- QCD veto (as in the *single top* analyses).
- High missing transverse energy:
  - MET>20 GeV (25 GeV for PHX)
- At least one SECVTX b-tag.

EXTENDED MUONS



- Recover muons in the gaps left by the high p_T muon triggers.
- Acceptance gain of ~15%.
- Leptons used in the analysis: CEM, PHX, CMUP, CMX, CMU, CMP, CMIO, BMU, SCMIO, CMXNT

CDF Note 9105, see Bruno´s talk!!

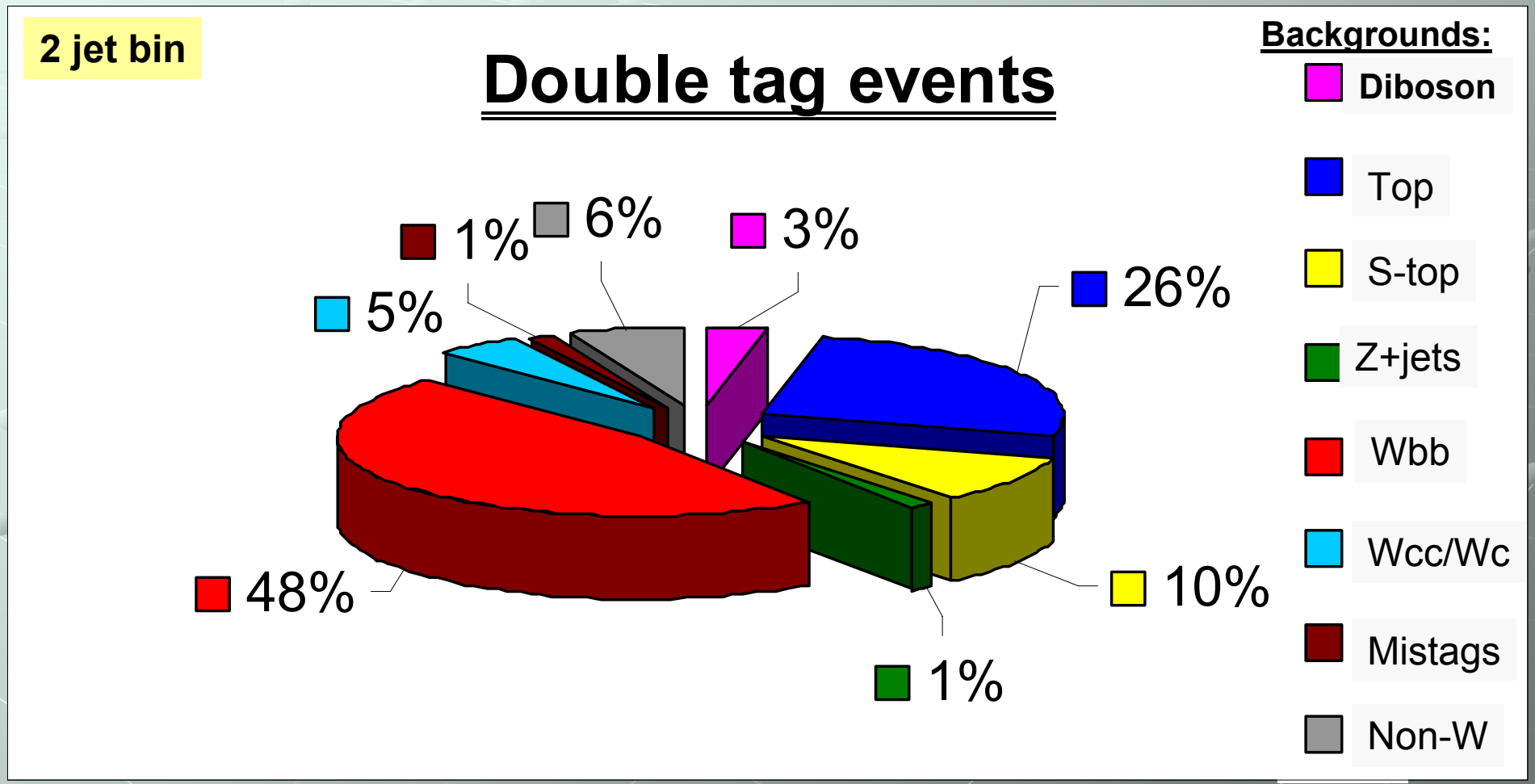
BACKGROUND ESTIMATION

- **W+jets**: Alpgen_v2 + Pythia, jet-based HF overlap removal (each process weigthed by x-section).
- **W+light**: pretag events (weighted by the mistag matrix).
- **tt and diboson**: standard Pythia samples (normalized to theorical x-section).
- **Single top**: Madgraph + Pythia (normalized to theorical x-sextion).
- **Z+jets**: Alpgen_v2 + Pythia, weighted the same as W+jets.
- **Non-W**: anti-electrons, jet-electrons, noniso-muons samples, pretag events.

2 jet bin

Channel	Single tag	Double tag
Diboson	80.0±8.0	5.4±0.6
TopLJ	121.3±17.1	23.8±3.9
TopDil	48.8±6.8	14.1±2.3
t-channel	64.0±9.3	1.8±0.3
s-channel	40.6±5.7	12.8±2.1
Z+jets	37.4±5.5	2.1±0.3
Wbb	538.7±162.5	70.3±22.5
Wcc/Wc	489.1±150.9	6.8±2.3
Mistags	458.0±57.9	2.2±0.6
Non-W	135.5±54.2	9.0±3.6
WH (115GeV)	6.3±0.5	2.0±0.2
Observed	1998.0	156.0

BACKGROUND ESTIMATION



MATRIX ELEMENT METHOD

- The probability of an event of being signal is proportional to the differential cross section of a given process:

$$P(x) = \frac{d\sigma}{\sigma}$$

- Fermi´s Golden Rule tells us the differential cross section for a given scattering process:

$$d\sigma = \frac{1}{2E_1 E_2 |v_1 - v_2|} \prod_i \left(\frac{d^3 p_i}{(2\pi)^3 2E_i} \right) |\mathcal{M}|^2 (2\pi)^4 \delta^{(4)} \left(p_1 + p_2 - \sum_i p_i \right)$$

EVENT PROBABILITIES

- Calculate probability density of an event resulting from a given process:

$$P(p_l^\mu, p_{j1}^\mu, p_{j2}^\mu) = \frac{1}{\sigma}\int d\rho_{j1} d\rho_{j2} dp_\nu^z \sum_{comb} \phi_4 \, |M(p_i^\mu)|^2 \frac{f(q_1) f(q_2)}{|q_1||q_2|} W_{jet}(E_{jet}, E_{part})$$

Phase Space Factor

Matrix Elements

Parton Distribution Functions

Transfer Functions

Inputs:
lepton and jet 4-vectors – no other information needed!

❖ **TRANSFER FUNCTIONS (TFs):**
We parameterize the Ep – Ej distribution as a sum of two Gaussian functions.
In this analysis we used a NNoutput as an estimation of Ej.

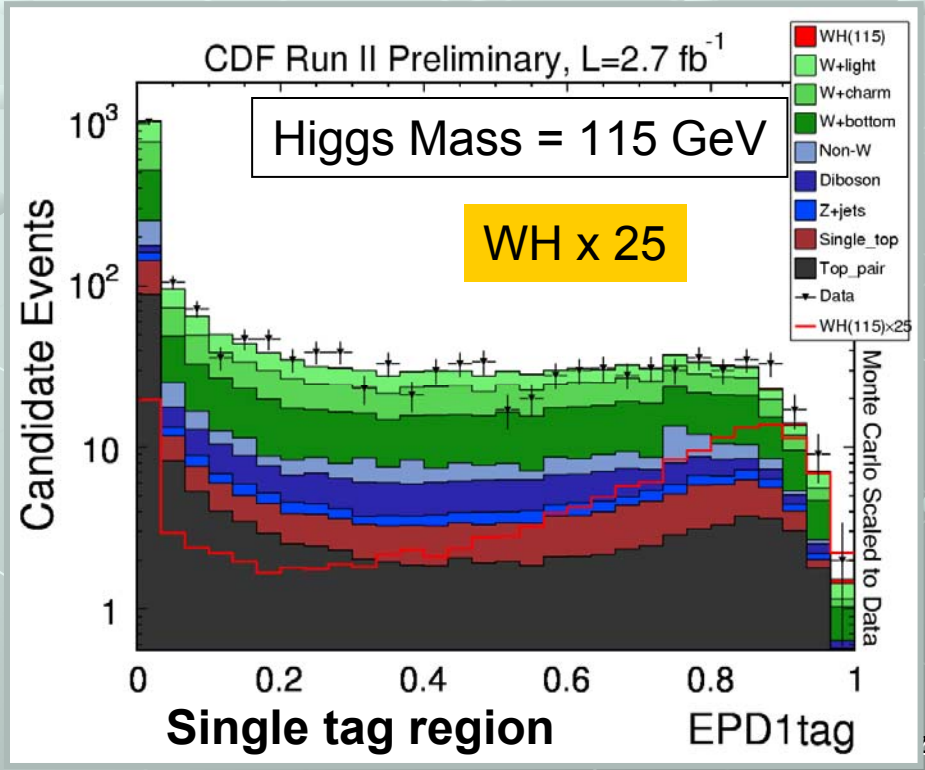
DISCRIMINANT

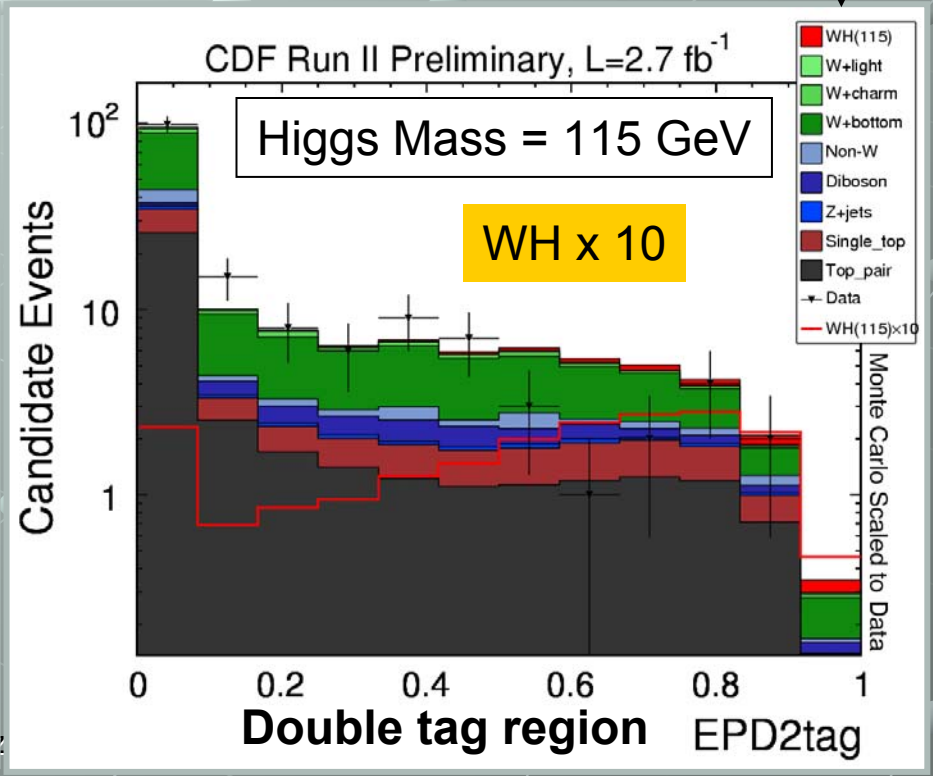
Most sensitive variable for the BDT

- **E**vent **P**robability **D**iscriminant (EPD) definition:

$$EPD = P_{signal} / (P_{signal} + P_{background})$$

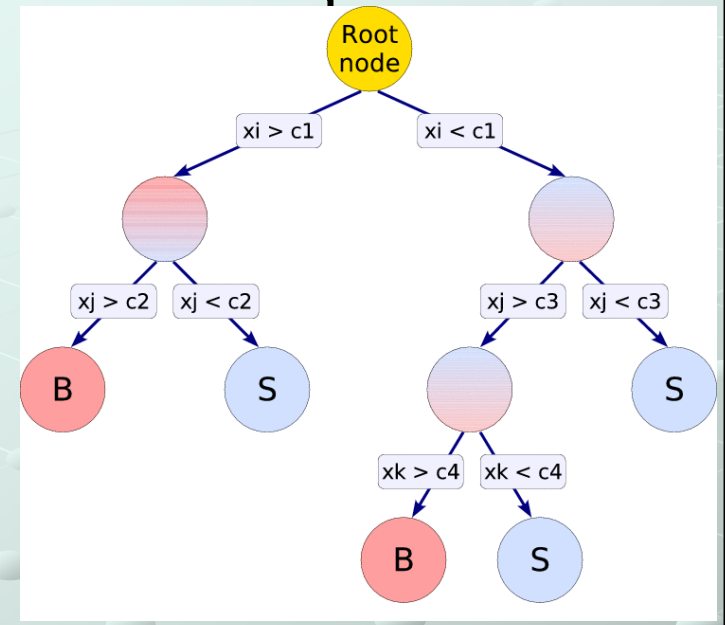
- Backgrounds peak to zero and signal peaks higher.





BDT

Overview



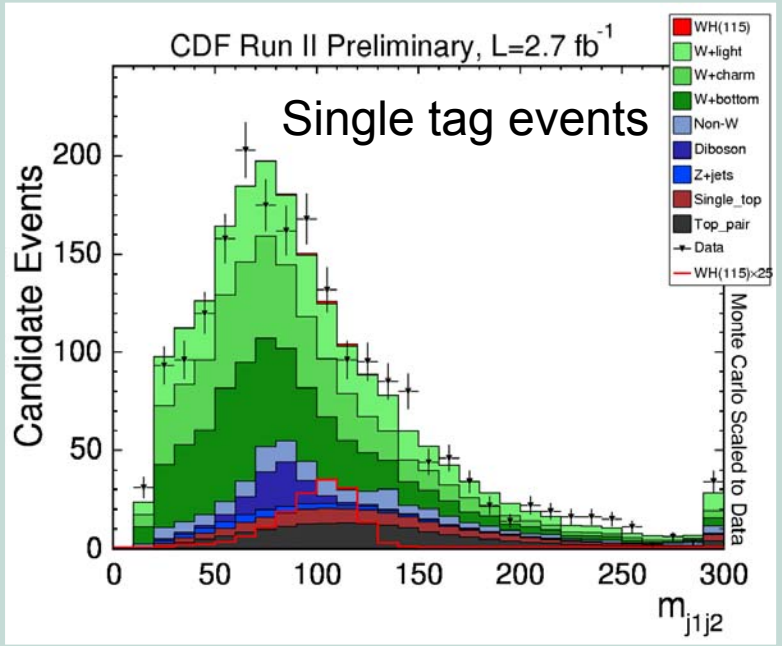
Training

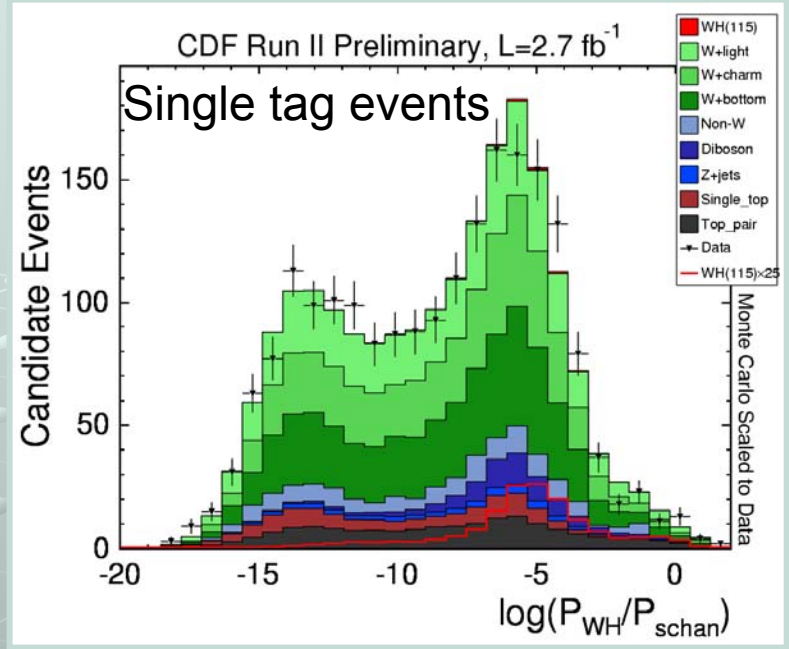
- ***Channels***
  - 2jet-1tag
  - 2jet-2tag
- ***Events***
  - Signal
    - WH
  - Background
    - Wbb
    - tt
    - Single top

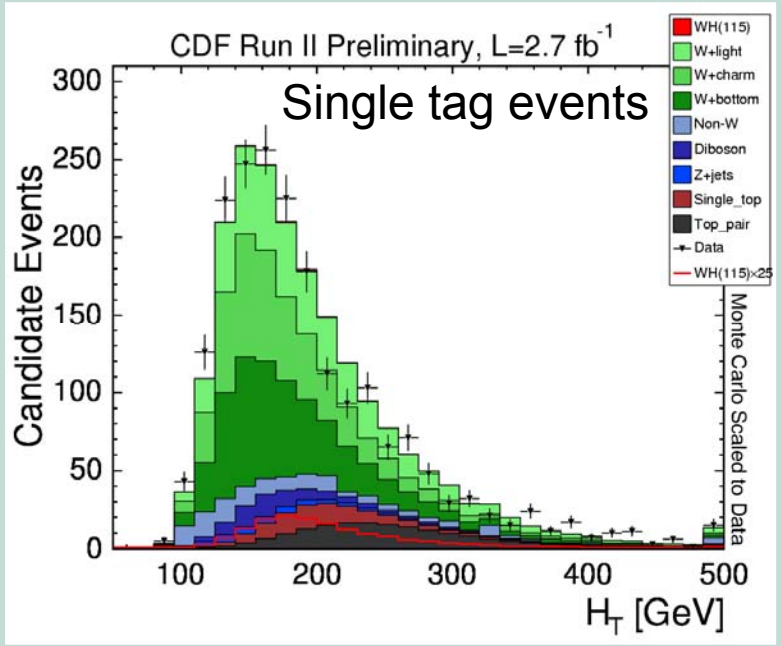
21 Input variables

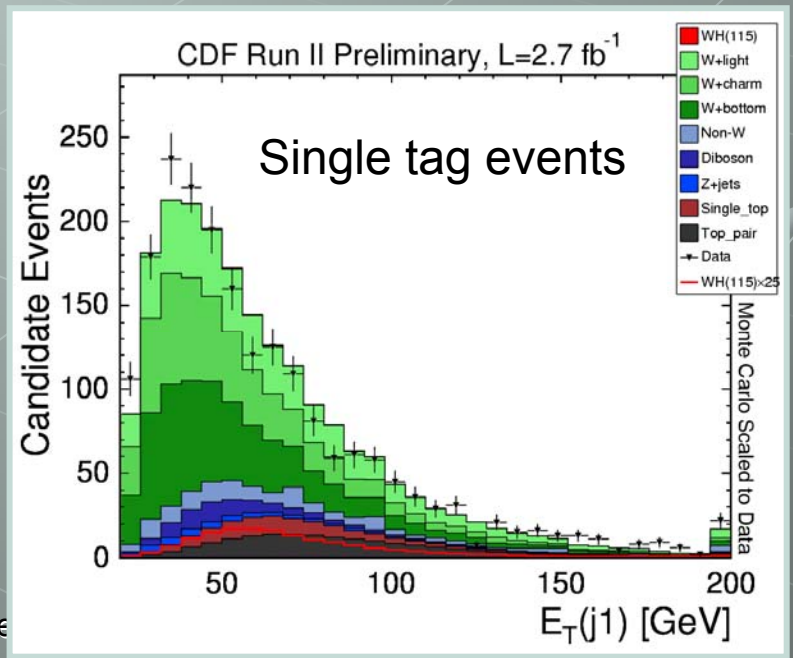
- 9 variables based on ME
- 12 kinematic variables

Some Input Variables for the BDT

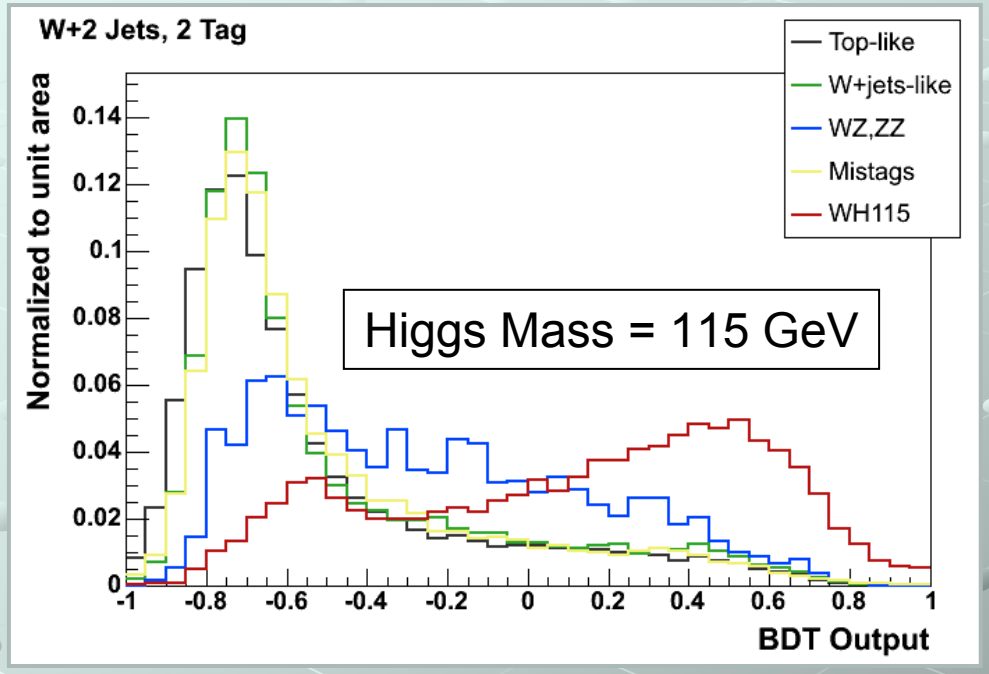


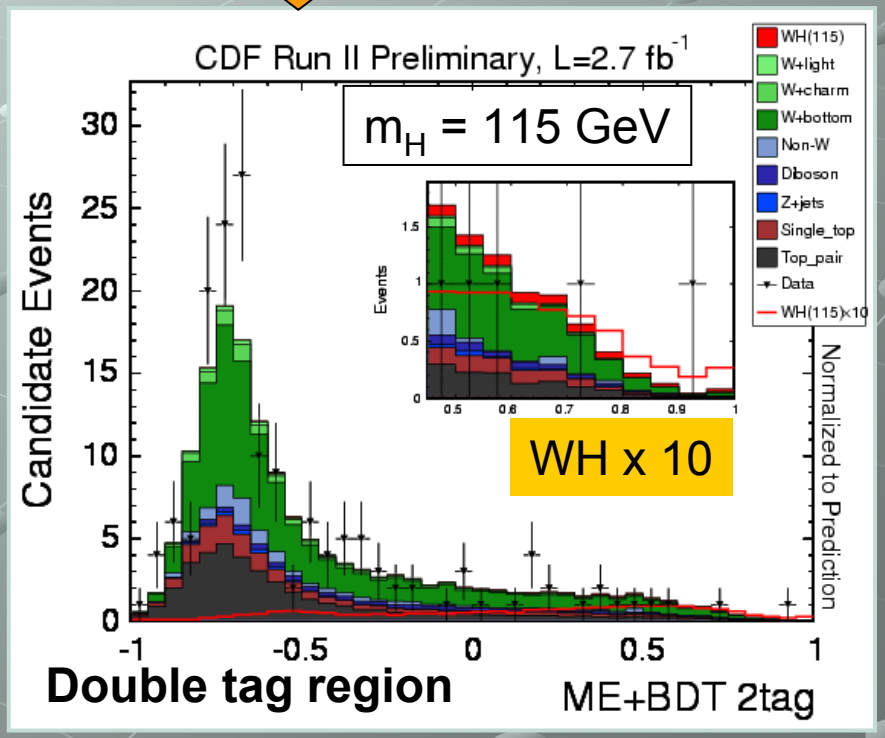






BDT OUTPUTS





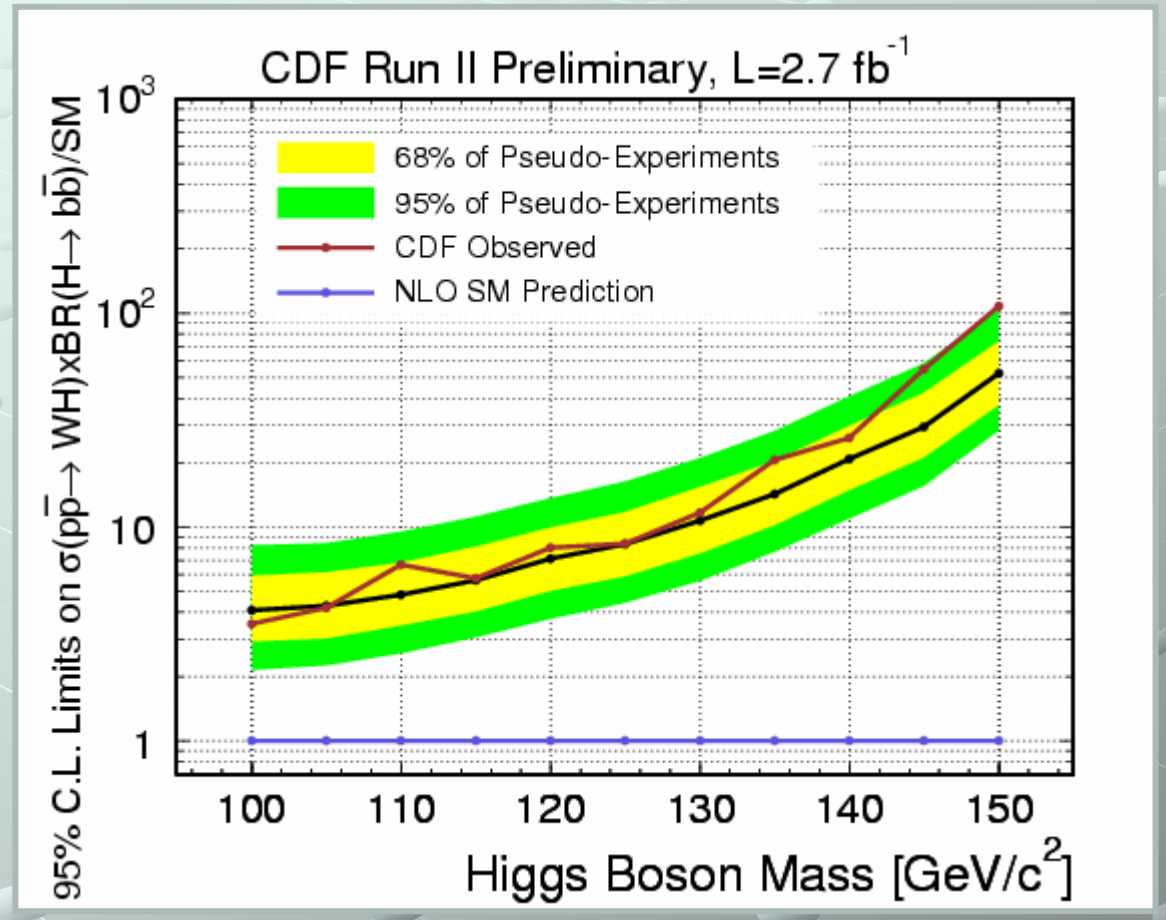
FINAL DISCRIMINANT!!

See BDT output for the *single tag region* in the back up slides!

RATE SYSTEMATICS

Systematic uncertainty	Single Tag (%)	Double Tag (%)
Jet energy scale	2.0	2.0
ISR/FSR + PDF	3.1	5.6
Lepton ID	~2.0	~2.0
Luminosity	6.0	6.0
b-tagging SF	3.5	8.4

RESULTS



m_H (GeV/c^2)	Expected (σ/SM)	Observed (σ/SM)
100	4.07	3.52
105	4.28	4.19
110	4.82	6.67
115	5.64	5.75
120	7.02	7.98
125	8.32	8.36
130	10.7	11.7
135	14.2	20.5
140	20.8	26.1
145	29.4	54.5
150	52.1	107.6

MCLIMIT used

NEAR FUTURE PLANS

- Combine this result with the ***NN WH*** 2.7 fb^{-1} result.
- Split into different tag categories: (SVSV, SVJP and SV).
- Study *loose* tag categories.
- Add *isolated tracks* and compare them with the extended muons categories.
- Optimize the transfer functions.
- Add new variables into the BDT training.
- ***New ideas***: add new variables from the ME calculation.

SUMMARY

- SM Higgs search with ME+BDT technique has been performed with 2.7 fb^{-1} of CDF data.
- Boosted decision trees use the ME output as input with some other kinematic variables.
- Our 95% C. L. expected upper limit is 5.64xSM for a Higgs of 115 GeV/c^2.
- We are working on combining this result with the NN WH group. See Jason´s talk!!

THANK YOU!

BACK UP SLIDES

DOCUMENTATION

- **Main note**: CDF Note **9404**.

- Related material:
  - ✓ WH search using ME in 2.2 fb^{-1}: CDF Note 9255.
  - ✓ Single Top search using ME in 2.2 fb^{-1}: CDF Note 9223.
  - ✓ Single Top search using BDT in 2.2 fb^{-1}: CDF Note 9263.
  - ✓ Extra muons: CDF Note 9105.
  - ✓ Method-II For You: CDF Note 9185.

QCD VETO

- QCD veto applied only in the single tagged sample.

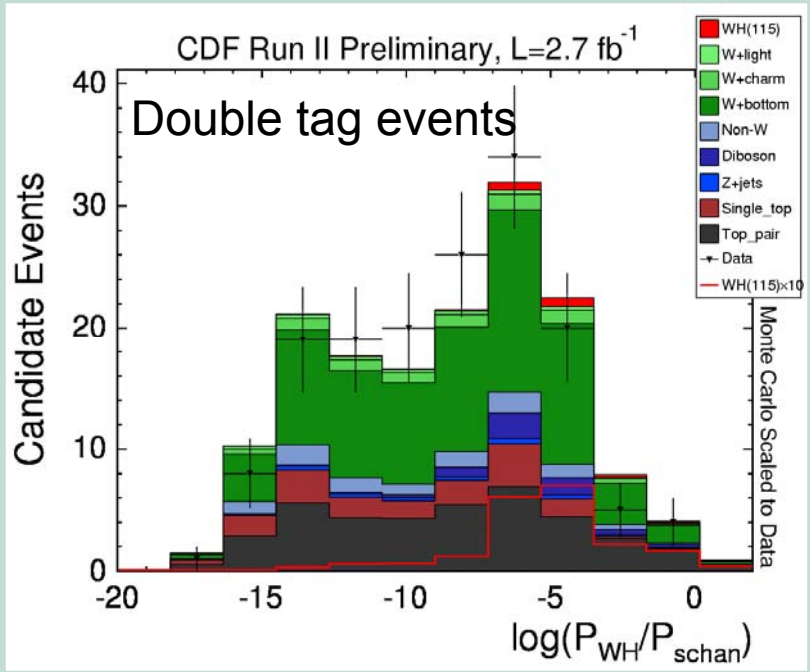
- Central electrons **(CEM)**:
  - ✓ $m_T(W) > 20$ GeV/c^2
  - ✓ METsig≥ -0.05 x $m_T(W)$ + 3.5
  - ✓ METsig≥ 2.5 - (2.5/0.8) x Δφ(MET,jet2)

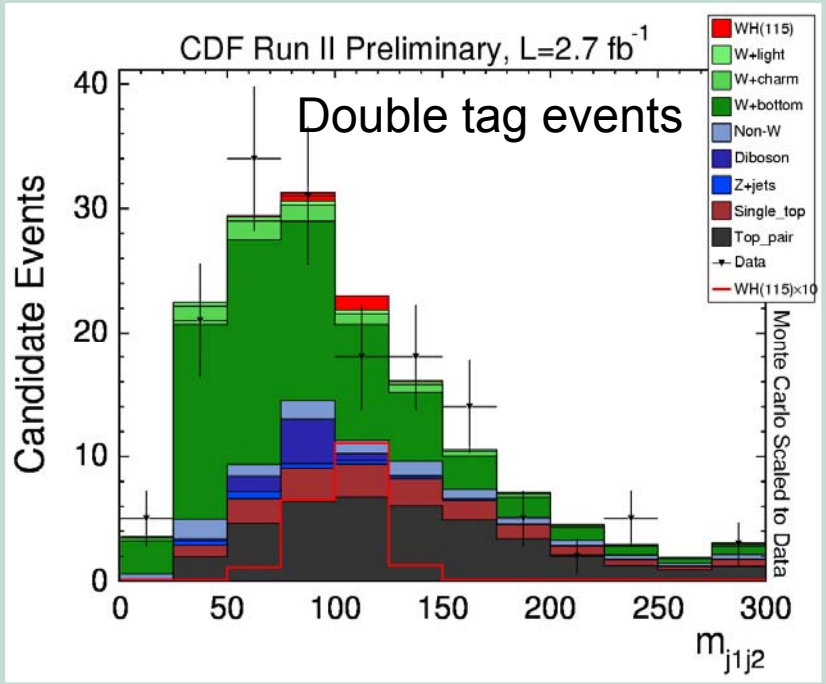
- Muons:
  - ✓ $m_T(W) > 10$ GeV/c^2

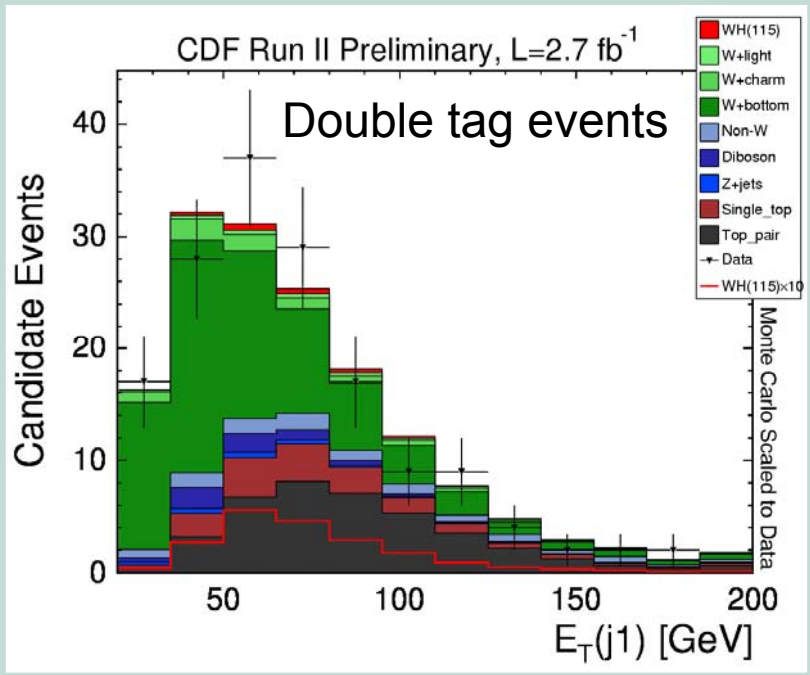
- Plug electrons **(PHX):**
  - ✓ MET > 25 GeV
  - ✓ $m_T(W) > 20$ GeV/c^2 METsig> 2
  - ✓ MET > 45 - 30 x Δφ(MET,jet1)
  - ✓ MET > 45 - 30 x Δφ(MET,jet2)

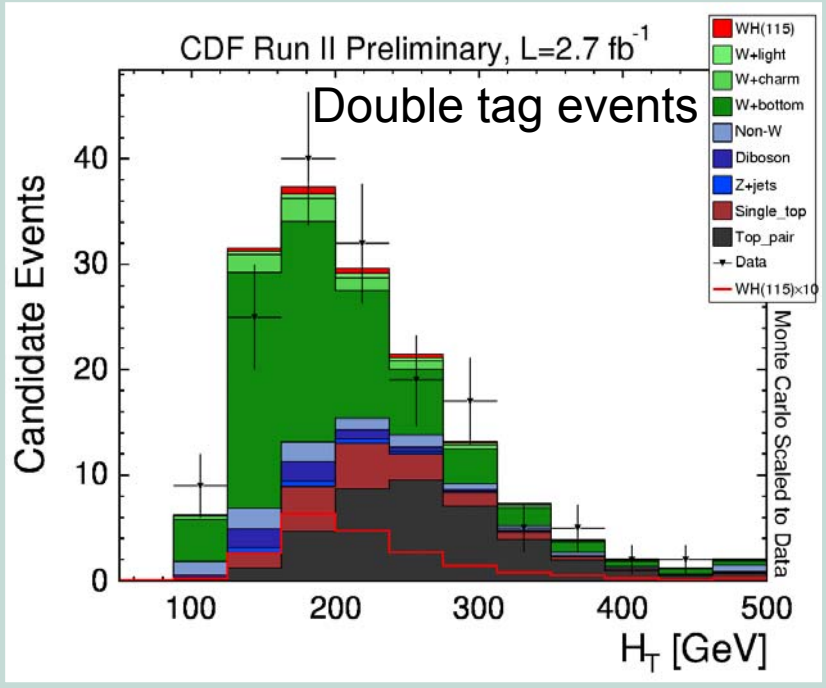
Process	Theoretical x-sections
s-channel	0.884 ± 0.11
t-channel	1.980 ± 0.25
WW	12.4 ± 0.25
WZ	3.96 ± 0.06
ZZ	1.58 ± 0.05
tt	6.7 ± 0.8

m_H (GeV/c^2)	BR(H ->bb)	σ	BR x σ pb
100	0.812	0.286	0.232
105	0.796	0.253	0.201
110	0.770	0.219	0.169
115	0.732	0.186	0.136
120	0.679	0.153	0.104
125	0.610	0.136	0.083
130	0.527	0.120	0.063
135	0.436	0.103	0.045
140	0.344	0.086	0.030
145	0.256	0.078	0.020
150	0.176	0.070	0.012


CDF Run II Preliminary, L=2.7 fb^{-1}
Double tag events
Candidate Events
$\log(P_{WH}/P_{schan})$
WH(115)
W+light
W+charm
W+bottom
Non-W
Diboson
Z+jets
Single_top
Top_pair
Data
WH(115)×10
Monte Carlo Scaled to Data


CDF Run II Preliminary, L=2.7 fb^{-1}
Double tag events
Candidate Events
m_{j1j2}
WH(115)
W+light
W+charm
W+bottom
Non-W
Diboson
Z+jets
Single_top
Top_pair
Data
WH(115)×10
Monte Carlo Scaled to Data

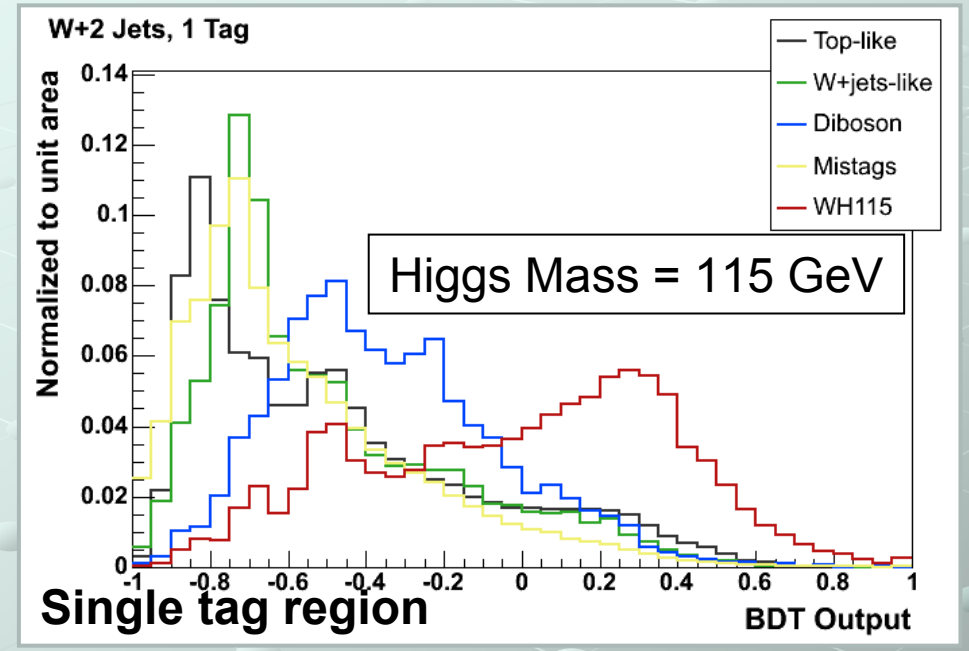

CDF Run II Preliminary, L=2.7 fb^{-1}
Double tag events
Candidate Events
$E_T(j1)$ [GeV]
WH(115)
W+light
W+charm
W+bottom
Non-W
Diboson
Z+jets
Single_top
Top_pair
Data
WH(115)×10
Monte Carlo Scaled to Data

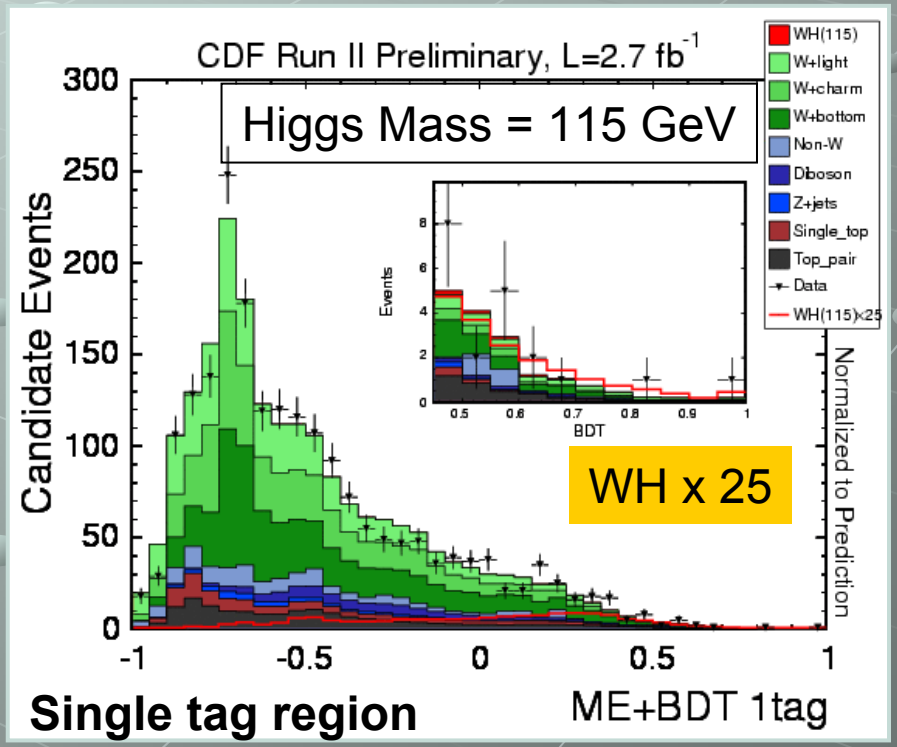

CDF Run II Preliminary, L=2.7 fb^{-1}
Double tag events
Candidate Events
H_T [GeV]
WH(115)
W+light
W+charm
W+bottom
Non-W
Diboson
Z+jets
Single_top
Top_pair
Data
WH(115)×10
Monte Carlo Scaled to Data

INPUT VARIABLES FOR THE BDT

Rank	Variable (1 TAG)	Importance	Variable (2tag)	Importance
1	EPD	2.428e-01	EPD	9.952e-02
2	Ht	8.334e-02	log(ProbWH/Wc)	9.199e-02
3	J2Et	6.797e-02	log(ProbWH/tchan)	8.354e-02
4	log(ProbWH/Wjg)	5.456e-02	log(ProbWH/ttbar)	7.425e-02
5	log(ProbWH/schan)	5.202e-02	log(ProbWH/schan)	7.367e-02
6	log(ProbWH/Wc)	4.960e-02	log(ProbWH/tchan2)	6.198e-02
7	met	4.872e-02	log(ProbWH/Wgg)	5.751e-02
8	J1Et	4.373e-02	log(ProbWH/Wjg)	4.353e-02
9	log(ProbWH//WccWbb)	4.006e-02	MetLepDPhi	4.174e-02
10	mJ1J2	3.932e-02	mJ1J2	4.145e-02
11	LepPt	3.901e-02	LepEta	4.134e-02
12	log(ProbWH/tchan)	3.723e-02	J1Et	3.968e-02
13	log(ProbWH/ttbar)	3.461e-02	Ht	3.460e-02
14	MetLepDPhi	2.893e-02	KaNN	3.358e-02
15	log(ProbWH/tchan2)	2.842e-02	MetJ1DPhi	3.062e-02
16	log(ProbWH/Wgg)	2.726e-02	J2Et	2.821e-02
17	KaNN	2.715e-02	log(ProbWH//WccWbb)	2.664e-02
18	wmt	1.785e-02	met	2.656e-02
19	LepEta	1.284e-02	cosLepJ1	2.587e-02
20	cosLepJ1	1.246e-02	LepPt	2.431e-02
21	MetJ1DPhi	1.209e-02	wmt	1.939e-02

BDT OUTPUTS





Single tag region!!