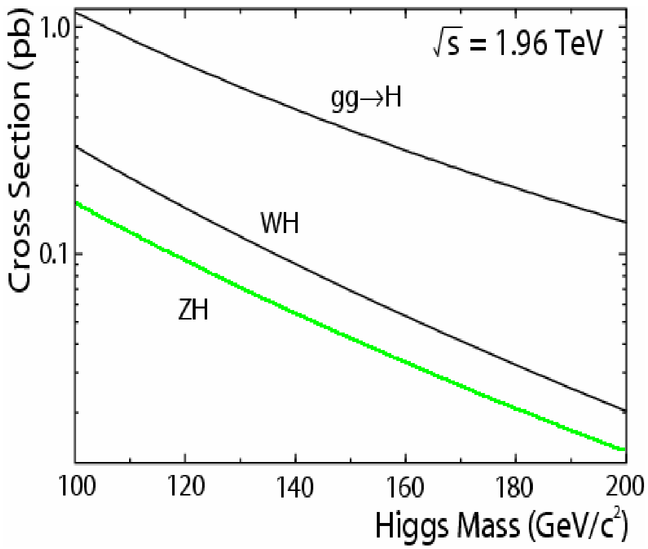
ZH->vvbb results from CDF

Viktor Veszpremi
Purdue University, CDF Collaboration

Tev4LHC Workshop, Oct 20-22 2005, Fermilab

Higgs Production at the Tevatron

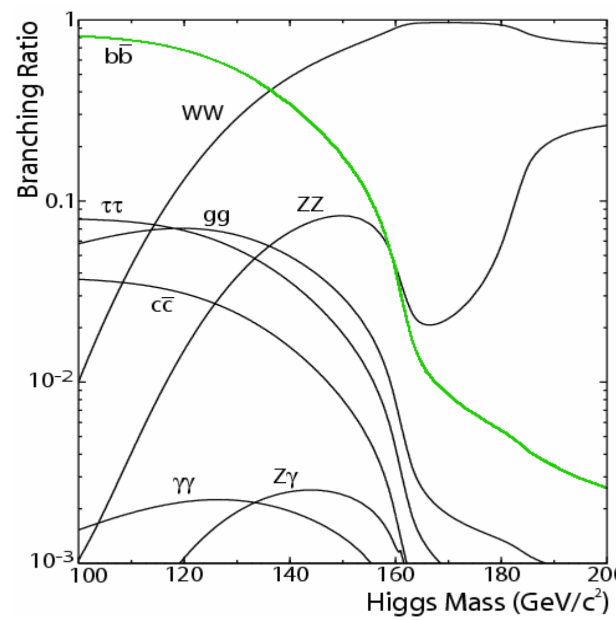


Production modes at Tevatron:

- Gluon fusion
  - Large cross section
- W/Z associated production
  - Smaller cross section, easier to trigger on the decay products of the vector boson

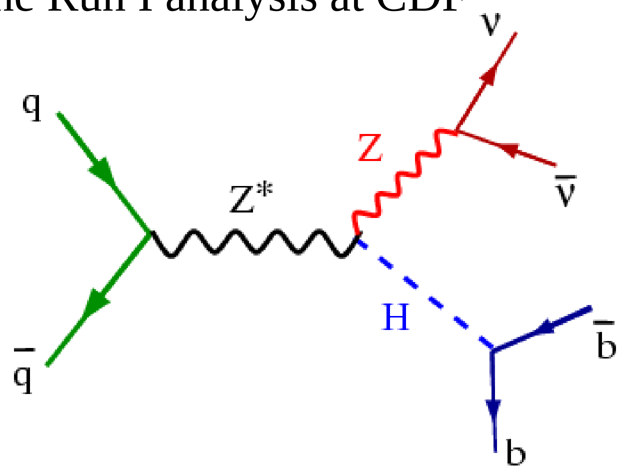
Decay modes:

- $H^0 \to b\bar{b}$ for Higgs mass below 125 GeV
  - The most promising channel at Tevatron
  - Requires good b-jet tagging
- $H^0 \to WW$ for Higgs mass above 130 GeV



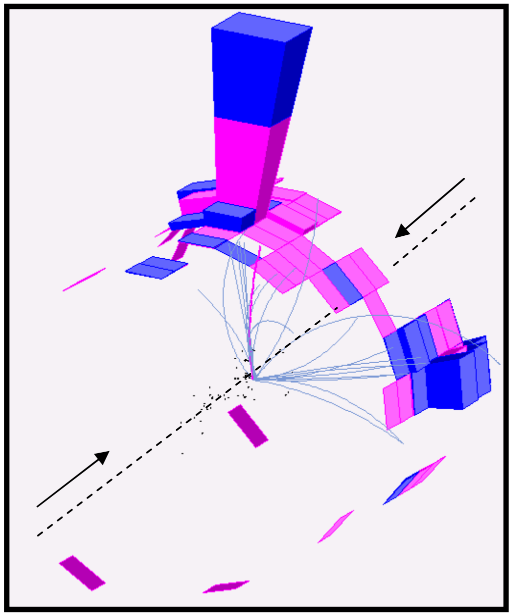
Select decay mode : Z -> ν ν, H ->b b

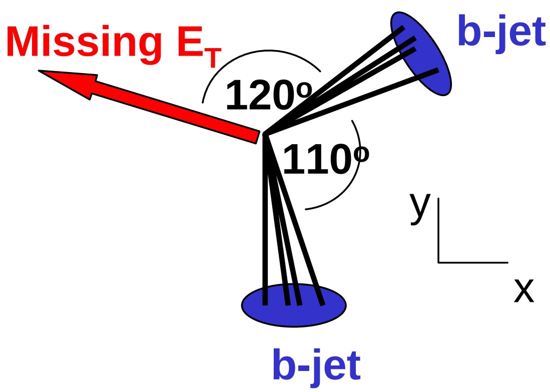
- This signature proved to be the most sensitive in the Run I analysis at CDF



- Signal has a distinctive topology
  - Large missing transverse energy
  - two jets (one is b-tagged)

- Trigger (MET35 + TWO JETS) on
  - Missing E_T > 35 GeV
  - Two jets E_T > 10 GeV





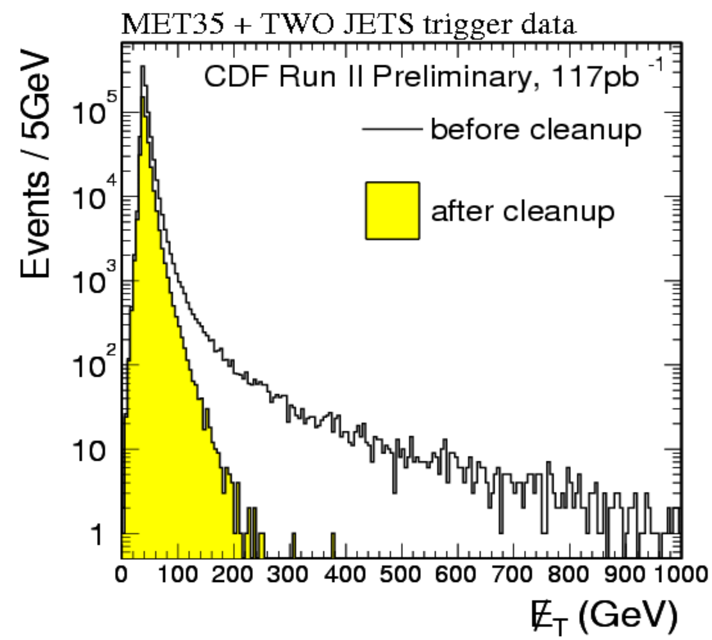
Origin of Missing E_T

Intrinsic Missing E_T – related to physics

- Produced by neutrinos
- Missing E_T is usually not aligned to jets in the event

Fake Missing E_T – detector related

- Beam/detector effects
  - A set of quality cuts are applied to eliminate beam effects
- Imbalanced, mis-measured jets
  - The primary source of $\not{E}_T$ in dijet events
  - Since the cross-section of QCD dijet production is large, most of the triggered events are like this
  - Makes QCD the main background



Missing ET before and after removal of beam and cosmic effects

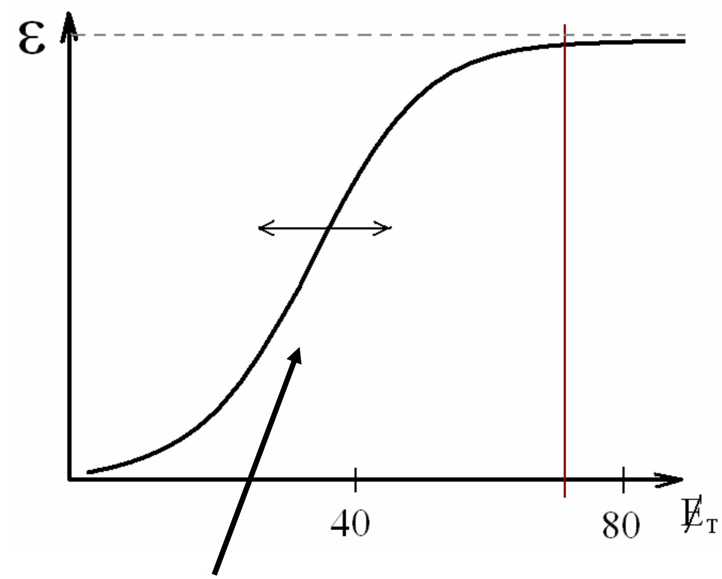
Triggering on Missing E_T

• The trigger efficiency depends on whether the Missing E_T is fake or real

• Efficiency for calculating event weight for Monte Carlo simulation was obtained using an inclusive jet data sample

• To avoid large systematic error from the trigger, we cut at 70 GeV in the Missing E_T, where the efficiency is ~ 1

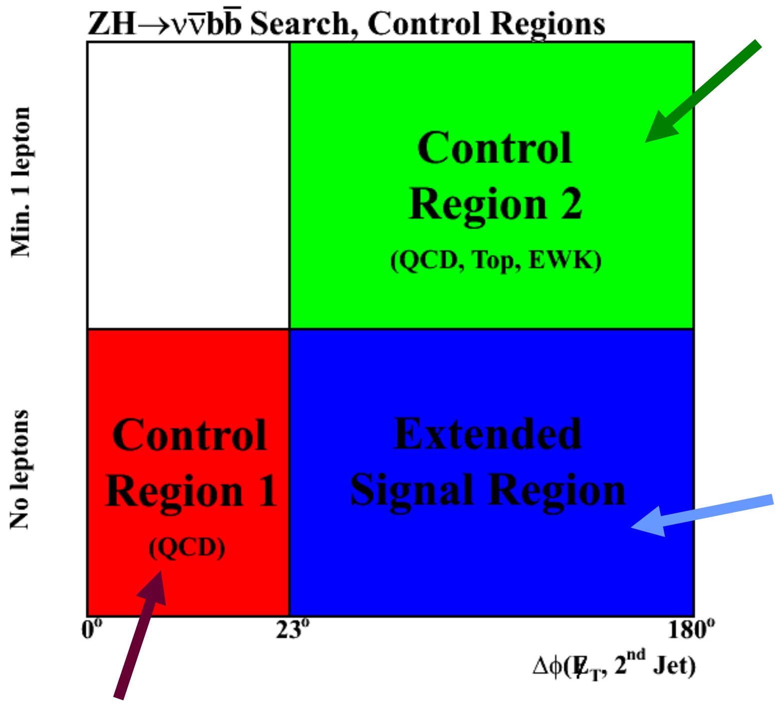
• Further trigger efficiency studies are needed taking into account the main source of the MET event by event



"Turn on" depends on whether it originates from a physical process (escaping neutrinos) or from detector effects (miss-measured jets)

Analysis Outline

We performed a **blind analysis**



Control Region 1 – QCD h.f.

- Veto events with identified leptons
- Missing E_T and 2nd leading jet are parallel

Control Region 2 – EWK

- Require min. 1 lepton
- Missing E_T and 2nd leading jets are not parallel
- Optimized cuts are tested in this region before looking at the real data in the Signal Region

Extended Signal Region (no optimization)

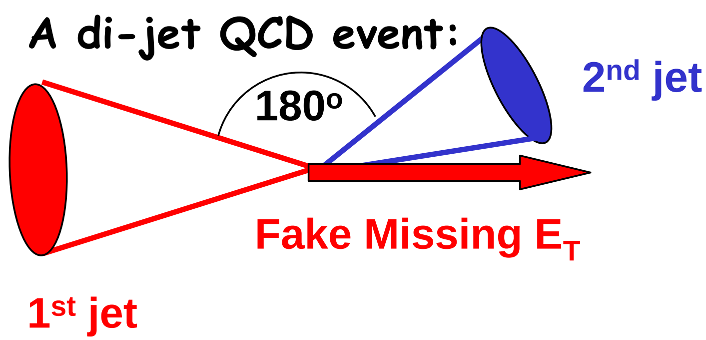
- Veto events with leptons
- Missing E_T and 2nd leading jet are not parallel
- Cut optimization is performed in this region based on Monte Carlo simulation before looking at the data

QCD Heavy Flavor Simulation

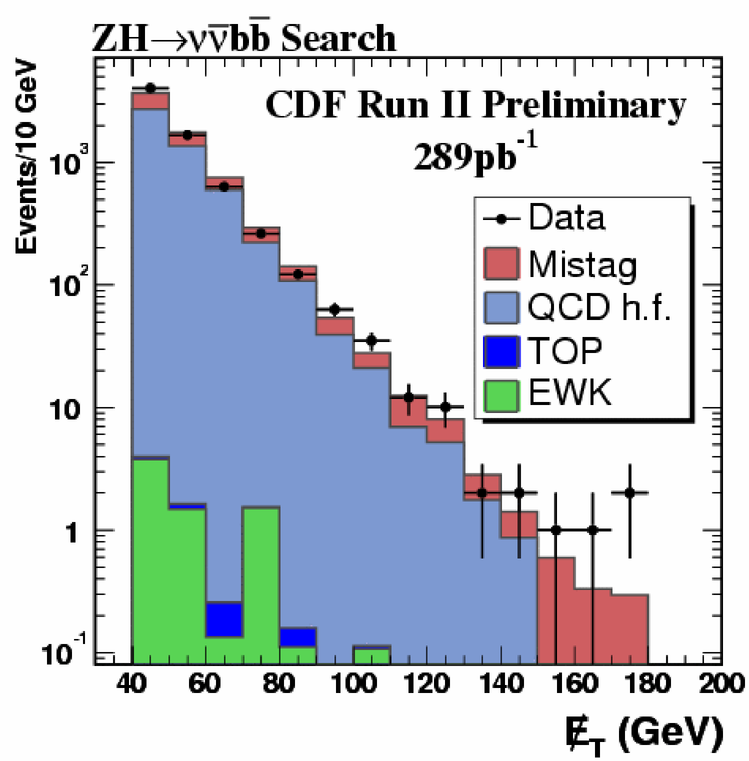
- Most of the triggered events are dijet events with fake Missing E_T
- After b-tagging, heavy flavor component is dominant
- Use Pythia to simulate it:
  - ~ 500 M events represent 1 fb^{-1} data
  - events are b-filtered (6%),
  - only events with Missing E_T >15 GeV are kept (32%)
  - few fb^{-1} data represents a great computational challenge
- Mistagged light flavor QCD is estimated from the data
- The simulation is normalized so that the sum of that and the estimated mistag is equal to the data in Control Region 1 with a 40 GeV MET threshold

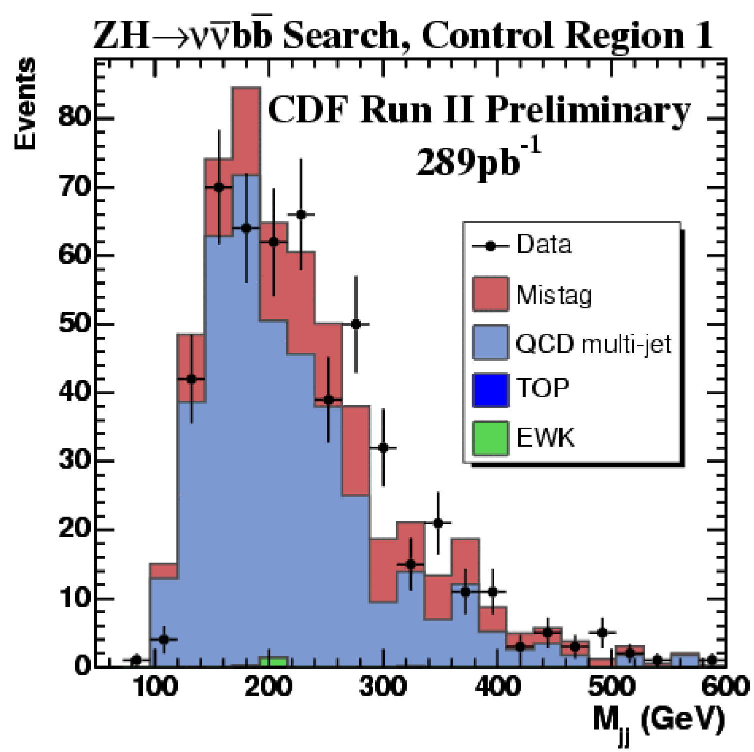
QCD events also have a particular topology:

- **jets are back-to-back**
- **Fake missing E_T points along the leading jet**



ZH→ννbb Control Region





Control Region 1: sum of the simulated heavy flavor and the mistag compared to data after normalization

Dijet invariant mass in Control Region 1 after normalization

Cut Optimization

The following optimized cuts were developed on a benchmark signal MC sample (M_h=120 GeV)

Selection cut	ZH 120 288.9 pb^{-1}	Acceptance (%)	S/sqrt(B)
Basic Cuts	**0.205±0.0035**	**5.92 ±0.1**	**0.026**
$\Delta\varphi(1^{st}\, Jet, \not{E}_T) > 0.8$	0.205 ±0.0035	5.92±0.1	0.027
$\not{H}_T$ significance*	0.183±0.0033	5.23±0.095	0.031
$1^{st}\, Jet E_T > 60 GeV$	0.161±0.0031	4.68±0.089	0.037
Di-jet mass cut	**0.126±0.016**	**3.64±0.079**	**0.062**

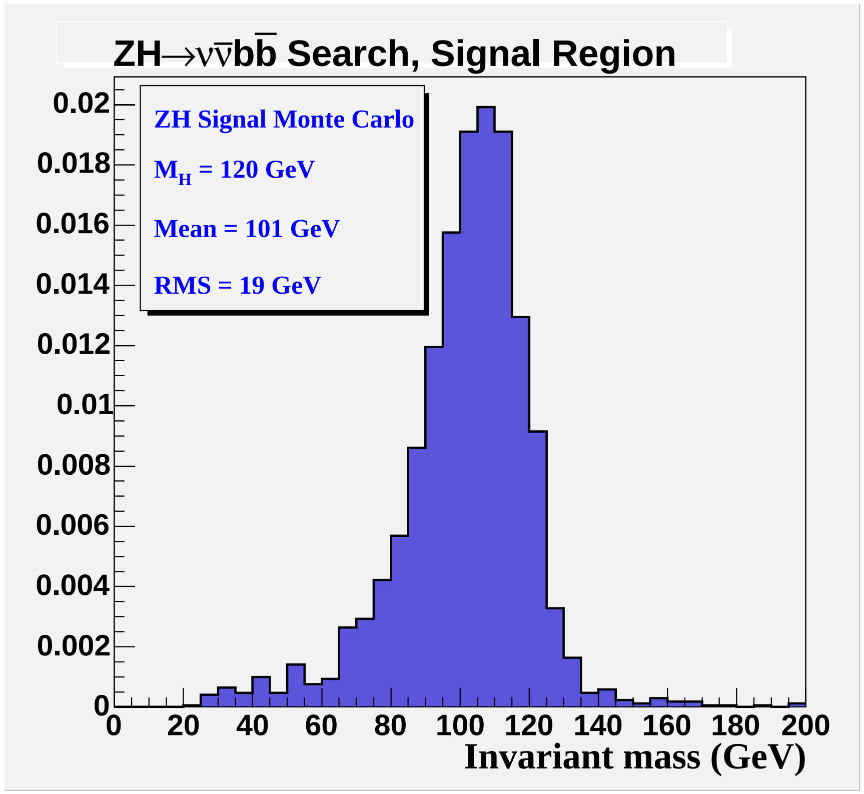
The last cut is on the di-jet mass. This is a "sliding" cut: a ±20 GeV mass window is applied to the data and Monte Carlo around the reconstructed Higgs mass peak

* H_T significance is the significance of the Missing E_T calculated from the two leading jets with respect to the scalar sum of the E_T of the jets

Choosing ZH→ννbb Mass Windows

- Last cut is on the dijet invariant mass
- A window of +20 GeV and –20 GeV is set around each of the mean of the mass peaks
- Dijet mass resolution is ~17 %

Invariant Mass (GeV)		$s / \sqrt{b}$
min.	max.	
60	140	0.043
70	130	0.047
80	**120**	**0.060**
90	110	0.056

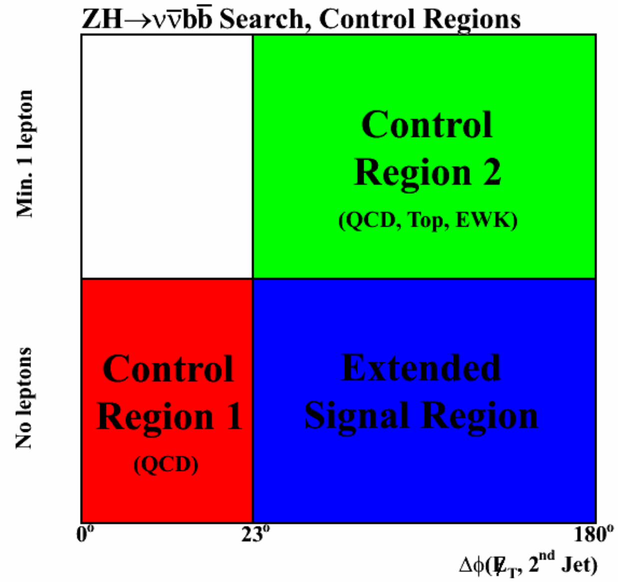


Cut Optimization

Expected number of events in the three regions

Before applying mass-window cut

Used L=289 pb^{-1} data

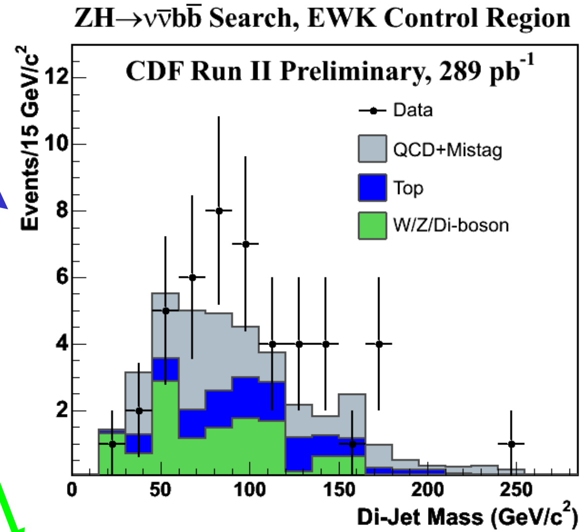


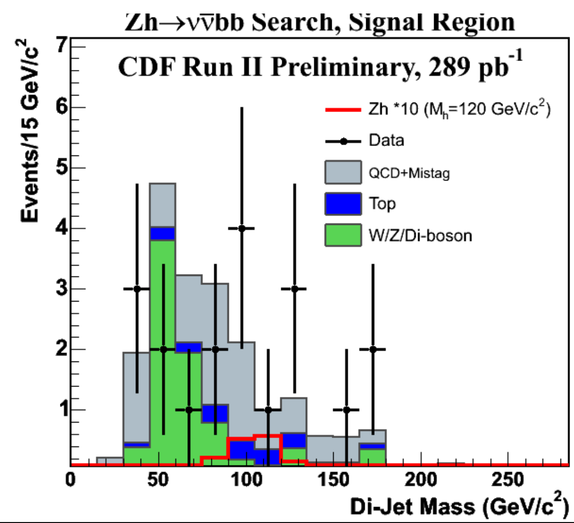
Process	Control Region 1	Control Region 2	Signal Region
QCD multi-jet	9.5 ± 4.3	5.2 ± 3	2.6 ± 1.7
TOP	0.01 ± 0.002	8.9 ± 2.3	2.1 ± 0.4
Di-boson	0 ± 1.2	1.5 ± 0.3	1.1 ± 0.2
W + h.f.	0 ± 1.2	9.7 ± 3.5	3.7 ± 2.6
Z + h.f.	0 ± 0.18	1.1 ± 0.3	3.2 ± 1.2
Mistag	2.9 ± 0.4	11.9 ± 2.3	7.0 ± 1.0
Total Expected BCK	12.4 ± 4.6	38.3 ± 5.7	19.7 ± 3.5
Observed	16	47	19

Results

The top plot : di-jet invariant mass in Control Region 2 after optimization, before the mass-window cut

The bottom plot is the same in the Signal Region





For the 120 GeV Higgs mass we apply a 80-120 GeV window cut:

SM background prediction: 4.36 ± 1.02 events

- QCD : 11.4%
- Top : 20.5%
- EWK : 18.2%
- Light flavor mistag: 50%

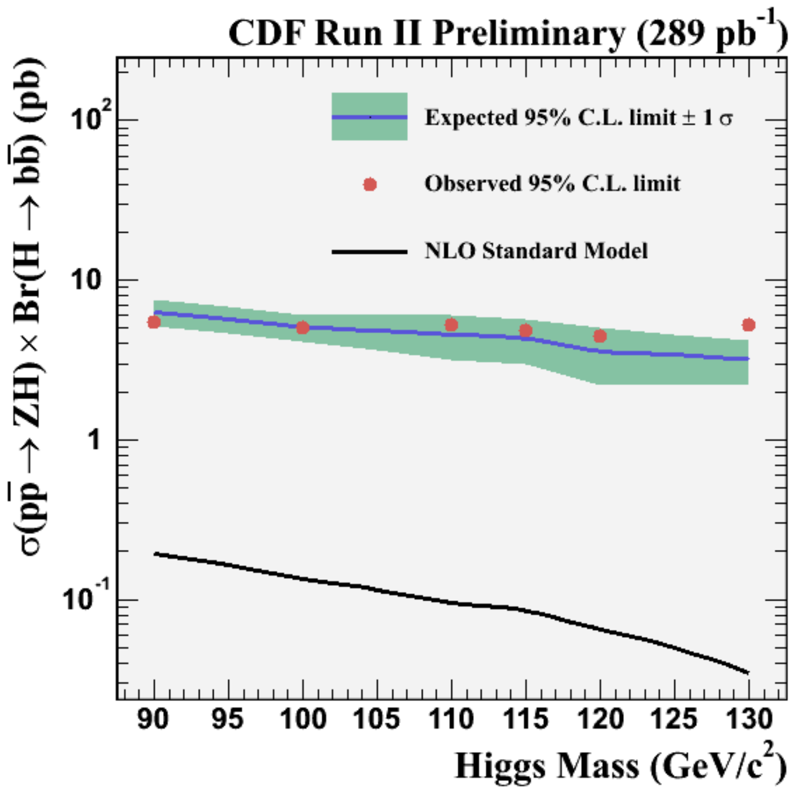
Observed: 6 events.

With a 95 % C.L. we expect the limit to be for the

$$\sigma(ZH) \cdot BR(b\bar{b}) < 3.6 \pm 1.4\, pb$$

Observed limit:

$$\sigma(ZH) \cdot BR(b\bar{b}) < 4.5\, pb$$



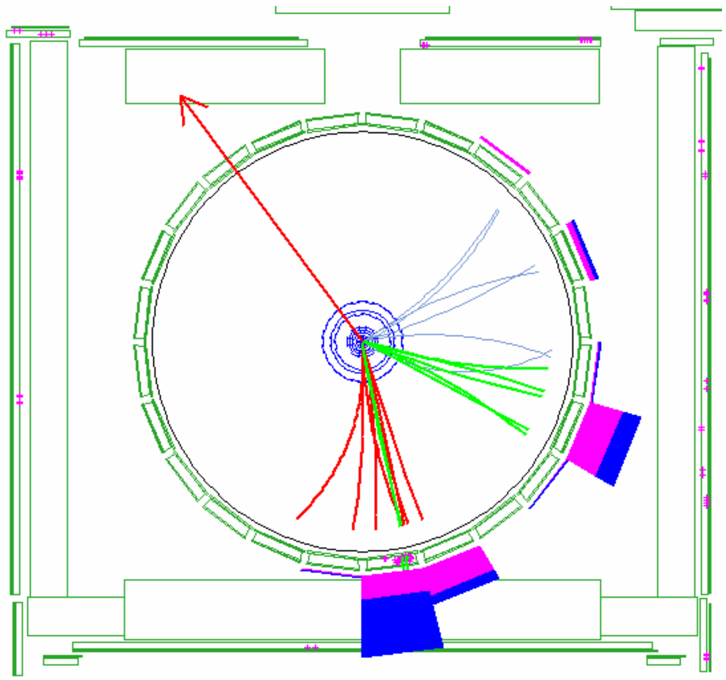
- Higgs mass dependence of the expected and observed limits

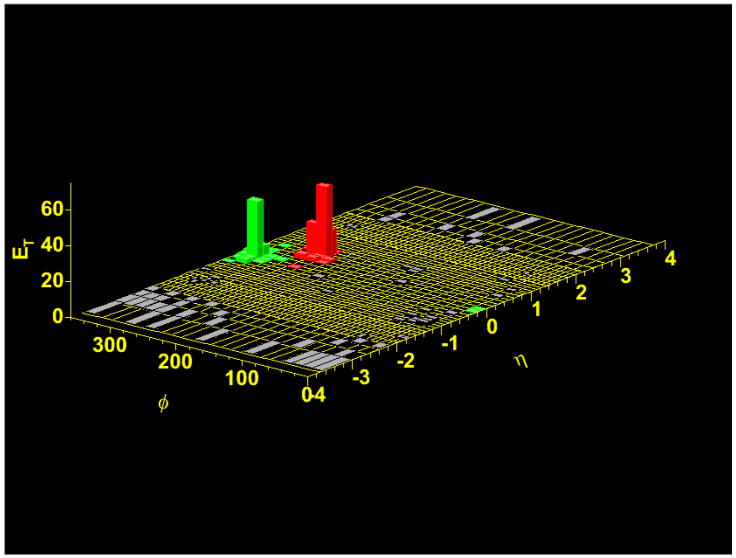
- Optimization was performed on a 120 GeV Higgs sample

- The largest systematic errors are the jet energy resolution (8%), luminosity (6%), b-tag efficiency (6%), and the statistical fluctuations

Mass (GeV)	Observed events	SM prediction	Higgs signal acceptance	Expected Limit (pb)	Observed Limit (pb)
90	6	7.18	0.45%	6.3 ± 1.2	5.4
100	7	7.07	0.55%	5.1 ± 1.0	5.0
110	7	5.9	0.64%	4.6 ± 1.4	5.2
115	7	5.9	0.67%	4.3 ± 1.4	4.8
120	6	4.36	0.73%	3.6 ± 1.4	4.5
130	8	4.11	0.77%	3.2 ± 1.0	5.2

Candidate Event





Two b-tagged jets

Jet_1 E_T= 100.3 GeV

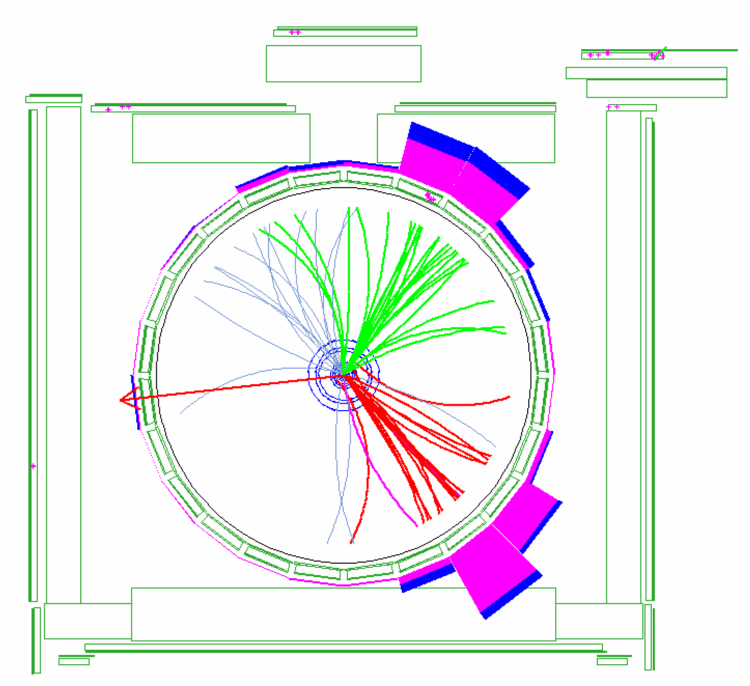
Jet_2 E_T= 54.7 GeV

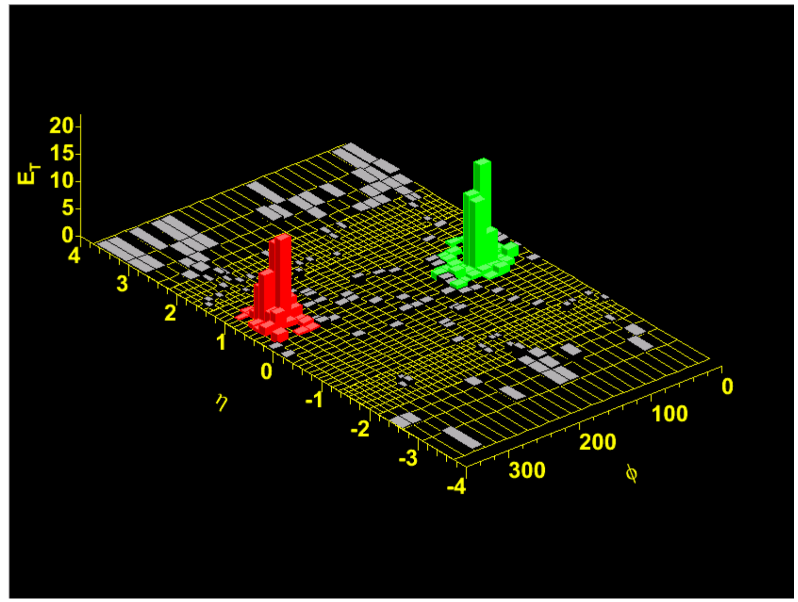
m_{jj}= 82 GeV

Missing E_T=145 GeV

Could be ZZ

Candidate Event





Jet$_1$ E_T=84.7 GeV
Jet$_2$ E_T=71.9 GeV -- Tagged

m_{jj} = 129 GeV

Missing E_T = 98 GeV

Summary and Future Improvements

Challenges in the analysis:

- Missing E_T trigger efficiency requires better understanding
- Monte Carlo generation is time costly

Further improvements:

- Taking into account the WH channel when the lepton is missed
- Changing (to looser) lepton veto
  - increases WH acceptance
  - increased amount of data with no requirements on the muon system
- Better dijet mass resolution
- Improving b-tag efficiency
- Develop better selection methods (Neural Network?)