Searches for the Standard Model Higgs Boson at the Tevatron

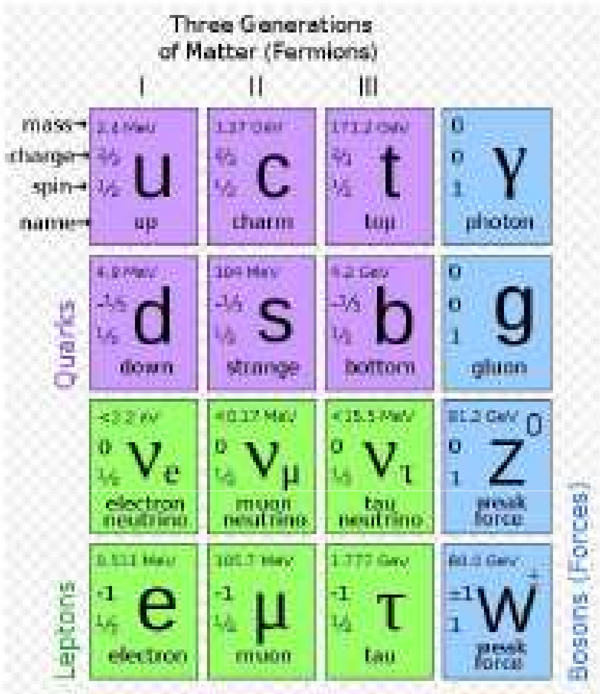
Jesús Vizán

Universite Catholique de Louvain

Outline

- Motivation
- Tevatron, and CDF/D0 Experiments
- Introduction to Tevatron Higgs Boson Searches
  - Low Mass Higgs Searches
  - $H \rightarrow bb$ modes
  - ***The WH→lνbb channel***
  - High Mass Searches
  - Other Channels
  - CDF and D0 Combinations
  - Tevatron Combination
- Summary

Standard Model of Particle Physics



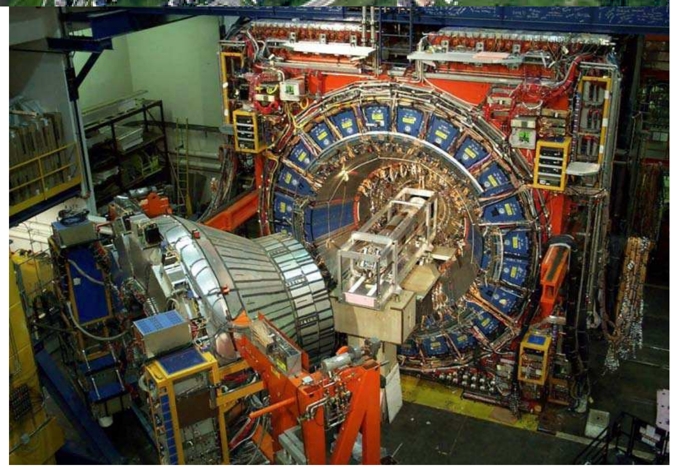
- ❑ Fundament particles:
  - ➢ 6 quarks, 6 leptons is enough
  - ➢ Bulk of visible matter: u, d, and e
  - ➢ Anti-particle for each particle
- ❑ Fundamental interactions:
  - ➢ Strong: gluons
  - ➢ Weak: $W^\pm$, Z^0
  - ➢ Electromagnetic: photon
  - ➢ Gravitational: graviton?
- ❑ Higgs field → Higgs particle
  - ➢ Higgs field is introduced to add mass to those particles
  - ➢ A new particle, the Higgs boson, is predicted coming from his own interactions

Direct searches at Colliders

- The Higgs boson is the only undiscovered “elementary” particle in the Standard Model
- Mayor goal at LHC (and Tevatron)
- The SM can not predict the Higgs boson mass
  - Need to be determined with an experiment
  - Very clear precision for production mechanisms, cross sections and branching ratios within the Standard Model
- Current direct searches SM Higgs boson
  - LEP II search exclude: $M_H < 114.4$ GeV/c^2 at 95% C.L.
  - Exclusion dominated by LHC experiments: approximately $110 < M_H < 115$ and $130 < M_H < 600$ GeV
  - Some excess of signal-like events in the region ~ 125 GeV/c^2

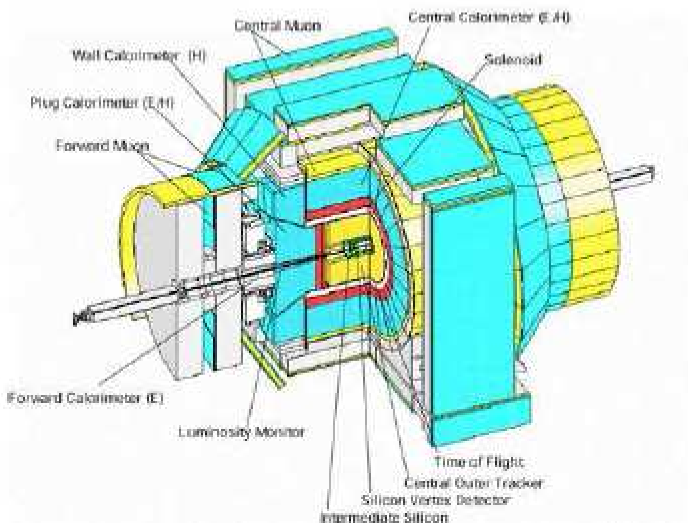
The Tevatron Collider

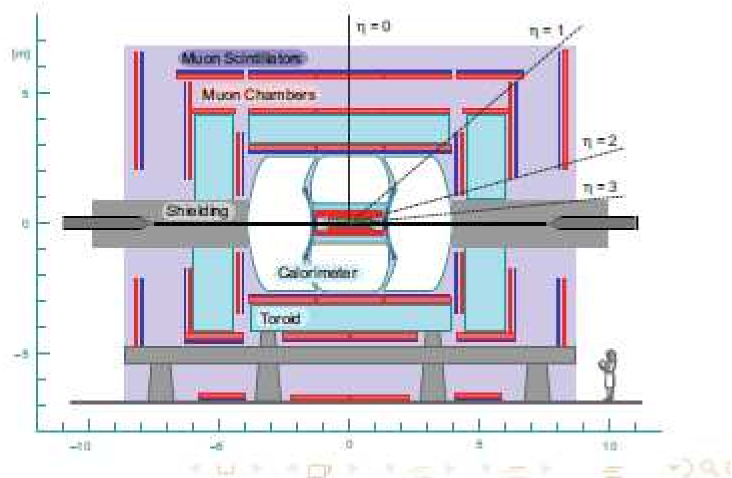
- Tevatron: proton-antiproton collisions at $\sqrt{s}$=1.96TeV
- It was closed on September 2011
- Analyses presented make use of full dataset
- ~12 fb^{-1} delivered data (~10 fb^{-1} acquired by the experiments)



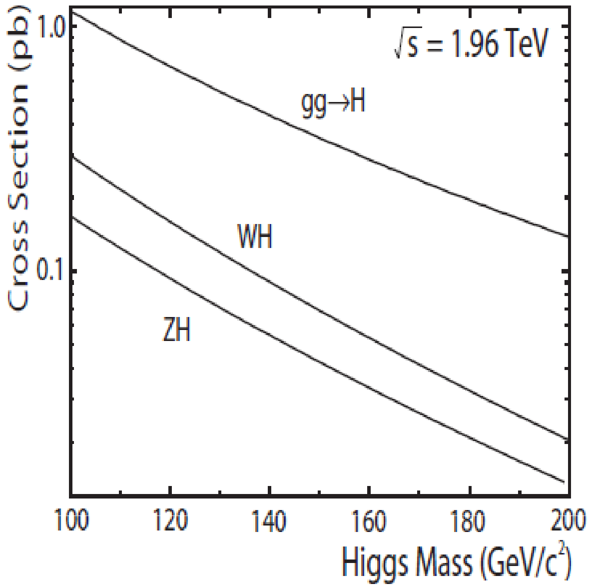
CDF and D0, two general purpose detectors

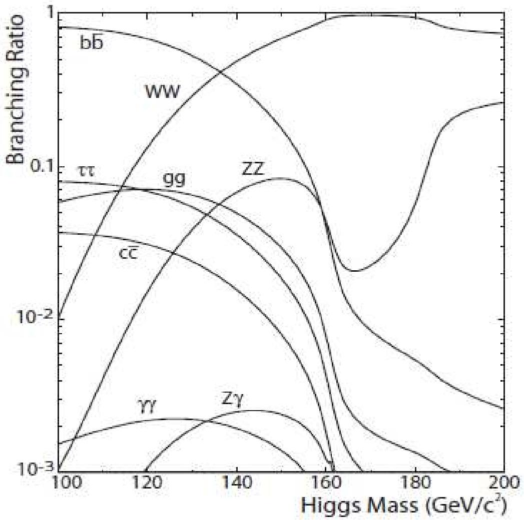
Tracking	Silicon	$\|\eta\| < 2 - 2.5$	Silicon	$\|\eta\| < 3$
	Drift cell	$\|\eta\| < 1.1$	Fiber	$\|\eta\| < 1.7$
Calorimetry	Scintillators	$\|\eta\| < 3.6$	LAr/DU	$\|\eta\| < 4$
Muon chambers	Drift Scintillators	$\|\eta\| < 1.5$	Drift Scintillators	$\|\eta\| < 2.0$





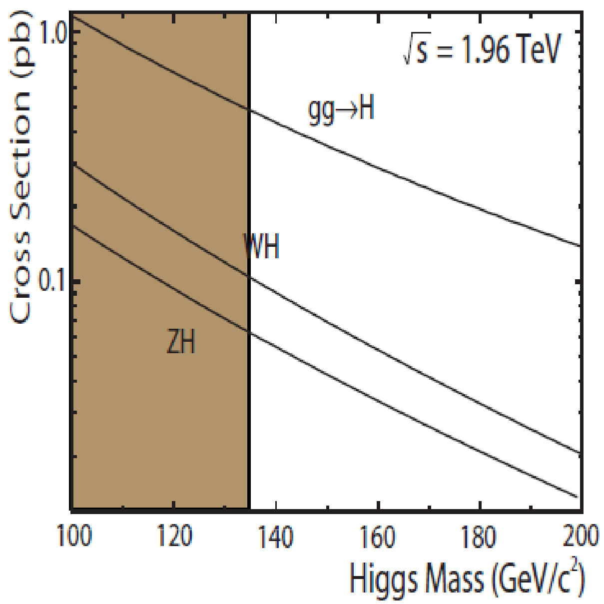
Higgs Searches at the Tevatron

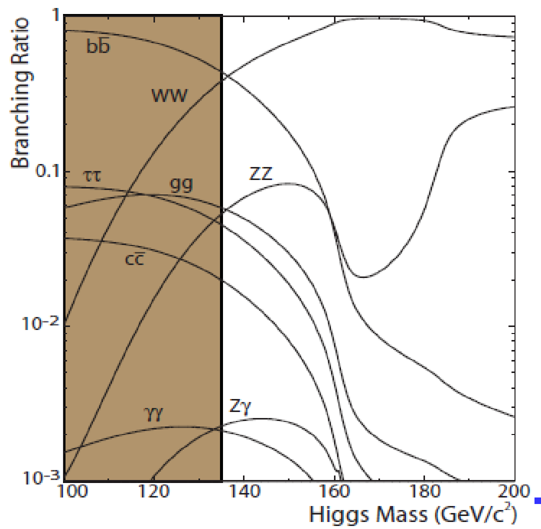




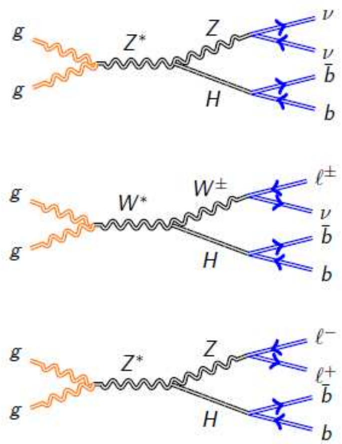
- **Direct production** dominates the mass spectrum
- **Associated production** (W/Z)~ 5 times smaller
- Two decay modes dominate: low / high mass
  - m_H < 135 GeV/c^2: H → bb
  - m_H > 135 GeV/c^2: H →WW
- Analyzing all decay channels for best sensitivity
  - Dedicated triggers to keep most of the Higgs events
  - No single analysis with sufficient statistical significance.
- No single channel can do. Need to combine all CDF Higgs searches and do the same also with DZero

Low mass Higgs searches





- H→bb dominant decay. Need efficient b-tagging algos
- σ (H) X B(H→bb) ~ 0.5 pb
  - Final state overwhelmed by QCD
- σ (VH) X B(H→bb) ~ 0.1 pb
  - Extra vector boson (in leptonic decay) helps reducing background
- Associated production: main low mass channels



$$VH \rightarrow \not{E}_T b\bar{b}$$

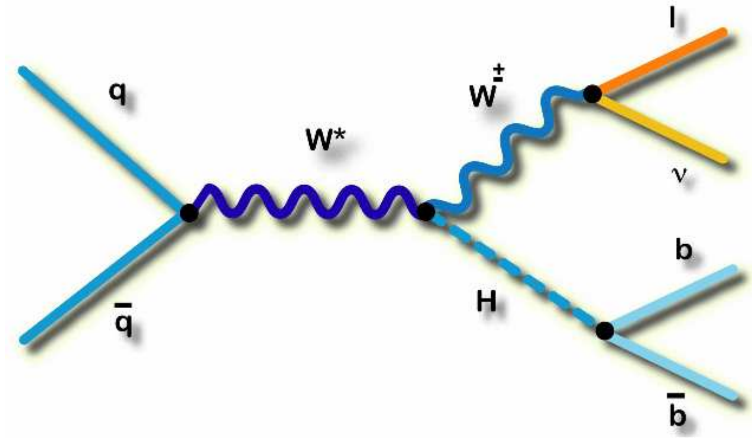
$$WH \rightarrow \ell\nu b\bar{b}$$

$$ZH \rightarrow \ell\ell b\bar{b}$$

- WH→lvbb will be described in detail in next slides
- Significant sensitivity for other channels at low mass
  - H→ τ τ , H→ γ γ , H→WW, ttH→lvqqbbbb

Why the WH→lνbb channel

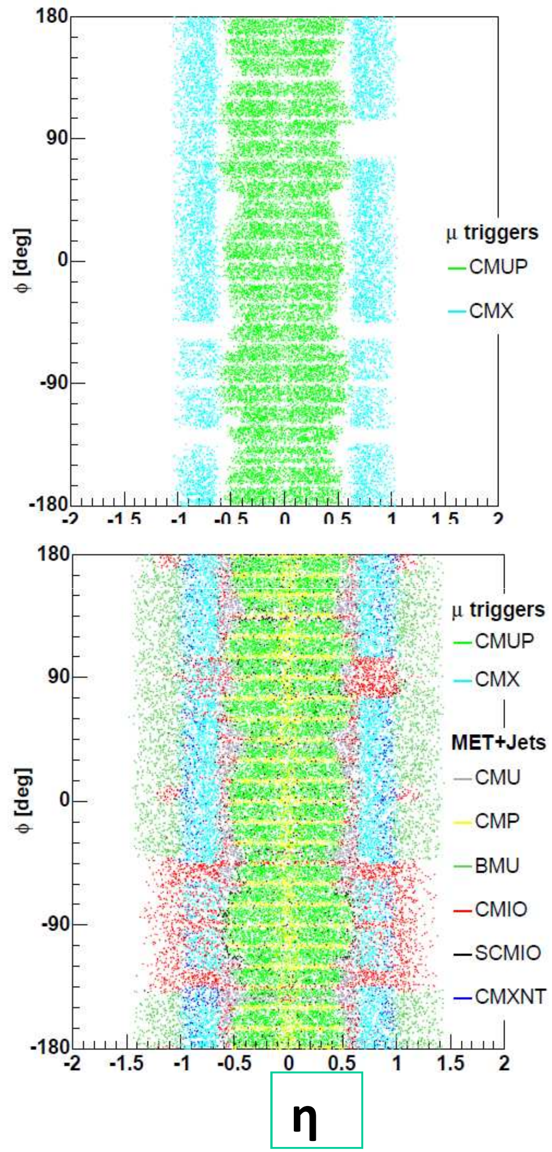
- qq → WH→lνbb: 2^{nd} highest cross section
- Reasonably clean signature
  - W→lν requirement: ***cleaner signature***
  - Presence of b-tagged jets
  - Relatively High MET (reject QCD)



- Many analysis techniques and strategies common to other analyses
  - Specially for H→bb final state
    - Use of Multiple "b-tagging categories"
  - Increase event acceptance by loosening standard lepton reco ID
  - Use of multivariate techniques
- CDF analysis will described. Similar strategies by D0

Event Selection: leptons

- High P_T isolated lepton (e/μ) or track: ***P_T > 20GeV***
  - Using different lepton categories depending on detector region and quality cuts applied.
    - ***Central leptons*** (CEM, CMUP, CMX): best S/B
    - ***Forward electrons*** (PHX)
    - ***Non-triggered muons*** (NTM)
      - Recovered relaxing the standard reconstruction from muons or from non-triggerable regions
    - ***Loose electron like leptons*** (LEL)
      - Non triggered reconstructed as isolated tracks matched to significant deposits of energy in the calorimeter
      - Triggered central electrons failing the standard reconstruction cuts but recovered using a multivariate electron likelihood

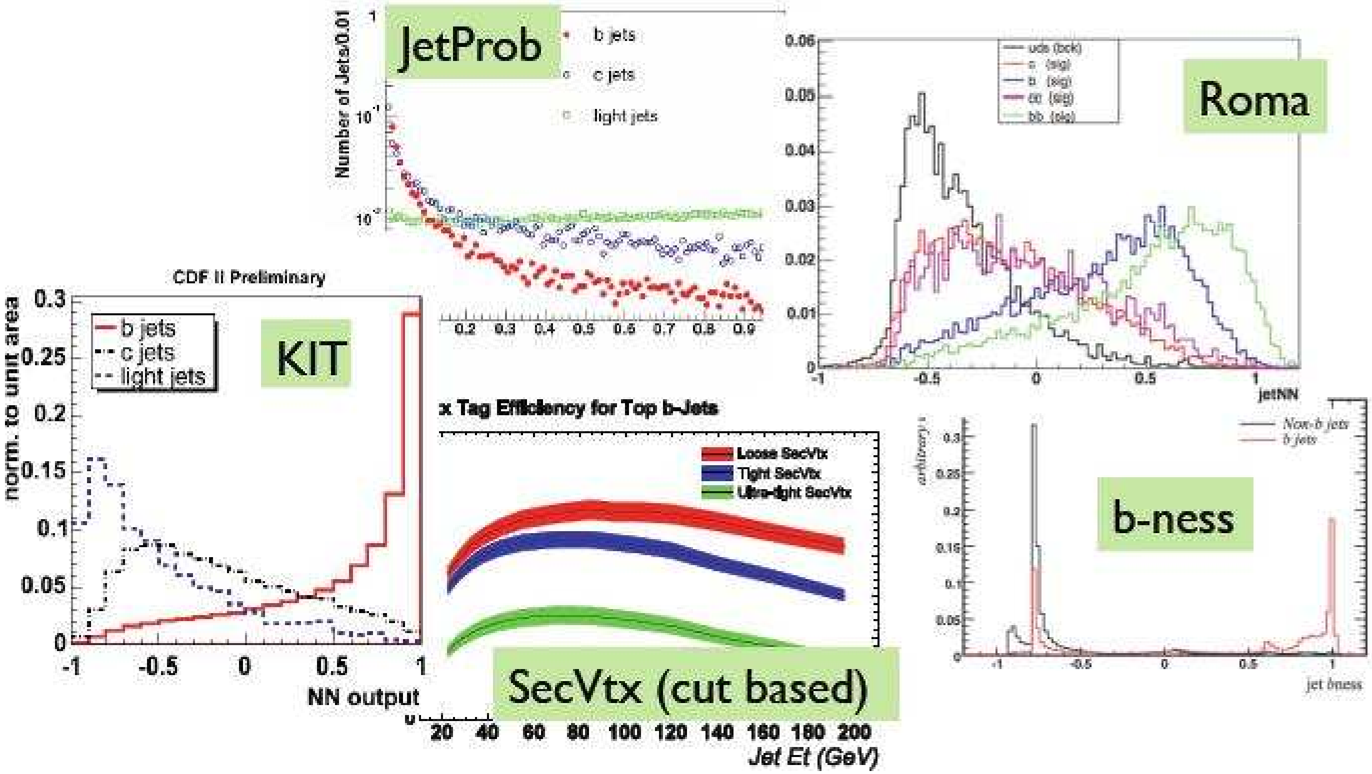


Event Selection: jet+met

- 2 (or 3) central high energy jets: ***E_T > 20GeV, $|\eta|$ < 2.0***
- Applying a QCD veto using a multivariate technique (Support Vector Machine) (***~6%*** improvement in sensitivity for single tag category)
  - E_T (second leading E_T jet)
  - MET
  - P_T (lepton), $\Delta\phi$ (lepton, MET^{raw})
  - …
  - Very good performance for muons. It allows to relax the MET cut for central muons
- Large missing transverse energy: ***MET> 20 GeV*** (***10 GeV*** for central muons)

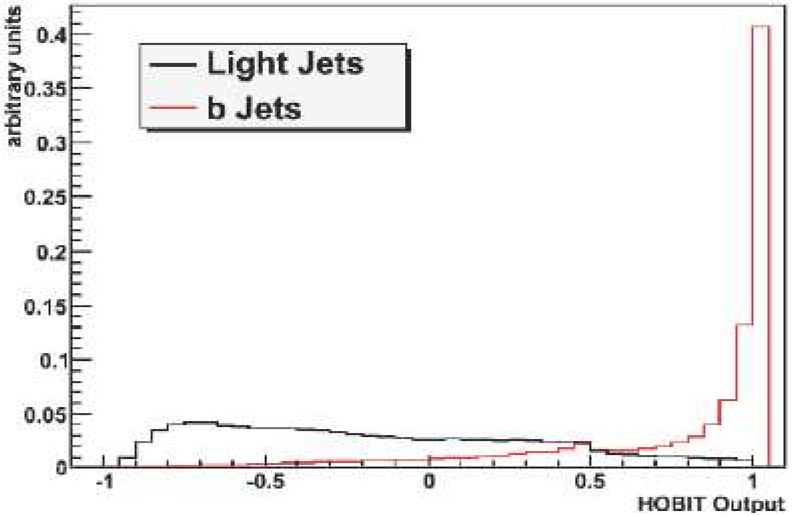
Event Selection: b-tagging

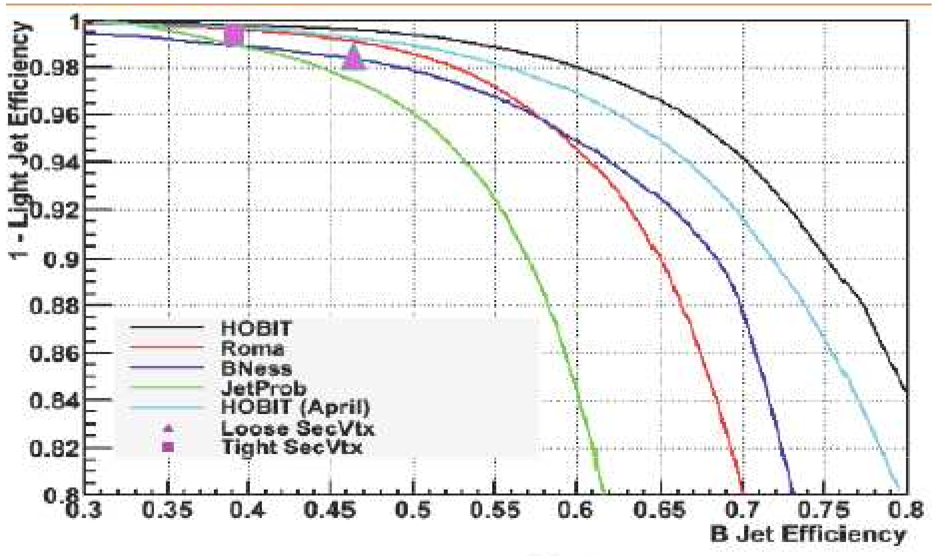
- Improvement in b-tagging benefits simultaneously all H→bb
- Important effor in CDF during last 1-2 years
- Create a new NN tagger using as inputs some of the old taggers used by CDF: Higgs Optimized b-Identification Tagger



- HOBIT
- 25 inputs
- Characterized (mistags/b-eff) using ttbar enriched and soft electron samples

Event Selection: HOBIT performance





- ❑ Better performance than any other previous tagger used at CDF

Event Selection: WH b-tagging

Summer 2011

Tagging Category	S/√B
SecVtx+SecVtx	0.228
SecVtx+JetProb	0.160
SecVtx+Roma	0.103
Single SecVtx	0.146
Sum	0.331

3 taggers

4 tagging categories

Winter 2012

Tagging Category	S/√B
Tight-Tight	0.266
Tight-Loose	0.200
Single Tight	0.143
Loose-Loose	0.053
Single Loose	0.044
Sum	0.369

2 Hobit Working Points

5 tagging categories

- ~12% improvement in signal significance by using
  - Similar improvement in expected exclusion limits
- Hobit included in 2/3 CDF $H \rightarrow bb$ analyses
- D0 is also working on improving the b-tagging

Background Estimation

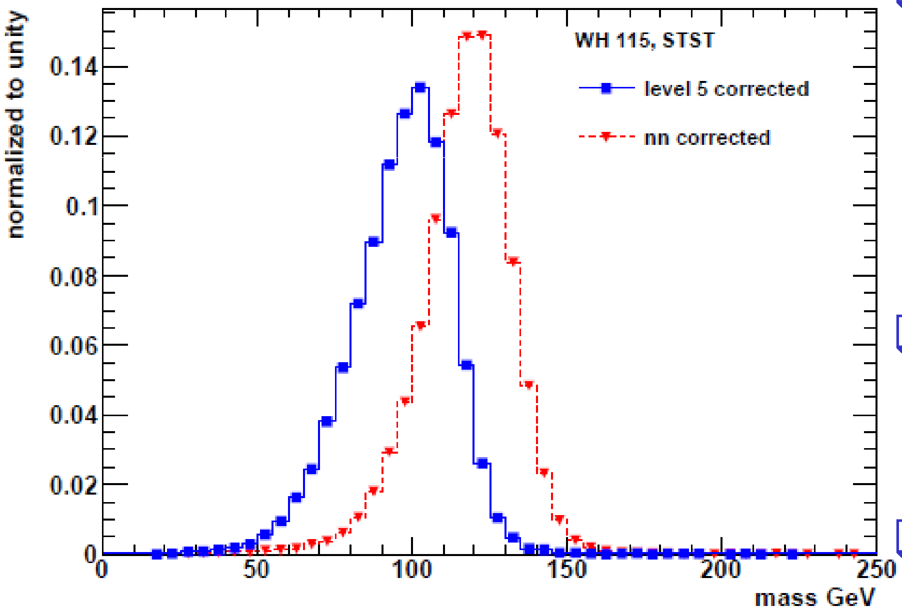
- Non-W: ***QCD*** multi-jet fake estimated ***from data***
- Mistag: falsely b-tagged ***W+LF jets*** ; W+HF: W+bb, W+cc/c -> ***Shapes from MC, Normalizations corrected using data***
- MC Total: top pair, single top, diboson(WW, WZ, ZZ), Z+jets: estimated ***from MC*** using theoretical/measured cross sections

Event Yield with events with 2 jets for central leptons

Number of Jets	2 jets				
Tagging categories	TT	TL	T	LL	L
DiTop	177.49±22.17	211.19±19.8	544.5±52.06	63.04±6.93	327.74±33.71
STopS	59.1±7.06	66.39±5.85	118.38±10.68	19.35±2.19	69.4±7.13
STopT	17.4±2.48	32.45±3.98	228.45±25.63	12.21±1.32	134.83±15.56
WW	1.9±0.48	15.54±3.13	217.47±27.09	29.26±4.5	719.24±70.55
WZ	21.86±2.63	25.97±2.28	63.3±6.23	11.8±1	115.13±10.59
ZZ	2.6±0.3	2.73±0.24	7.87±0.77	1.08±0.09	11.98±1.08
Zjets	11.94±1.29	23.24±2.47	184.32±19.71	30.93±3.46	815.82±85.61
Wbb	284.99±116.78	382.43±155.86	1372.45±559.67	129.59±52.89	948.67±387.04
Wcc	22.54±9.39	141.43±58.32	1379.5±564.72	196.66±80.63	3332.54±1360.1
Wlf	5.17±1.54	73.93±16.53	1179.09±191.85	293.49±47.08	9732.87±1094.5
QCD	12.35±7.94	101.82±41.71	680.92±272.42	125.62±50.72	2031.95±812.95
Bkg	617.34±172.05	1077.12±309.74	5976.25±1730.5	913.03±250.76	18240.17±3877.7
Obs	556	907	5737	865	18606
WH115	9.57±0.98	9.98±0.62	16.29±1.04	2.7±0.27	9.07±0.75

Improved Mass Resolution (Mjj)

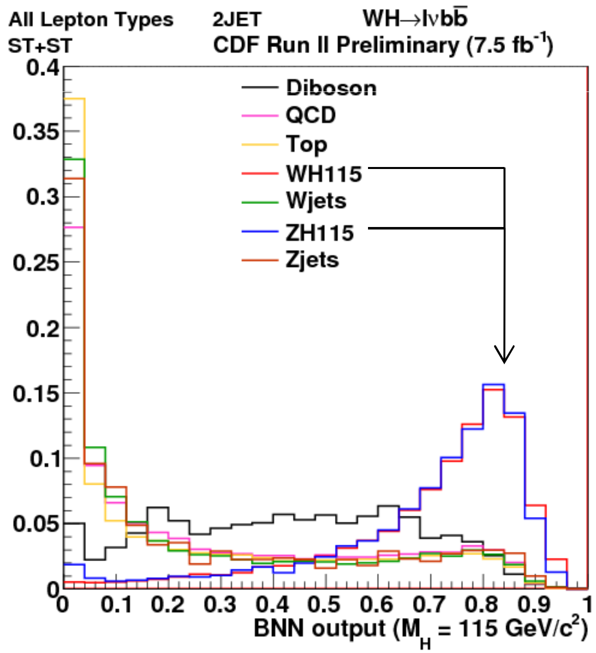
- Thet invariant mass of the bb pair is the most sensitive variable to the Higgs
- An improvement in resolution has direct impact on the sensitivity of the search



- Derive specific energy corrections for b-jets
  - Exploiting tracking and calorimenter informations with a NN
- Resolution in dijet mass improved from ~15% to ~11% for events with 2 b-tagged jets
- Similar techniques applied to other low mass analyses

Multivariate Analysis (BNN)

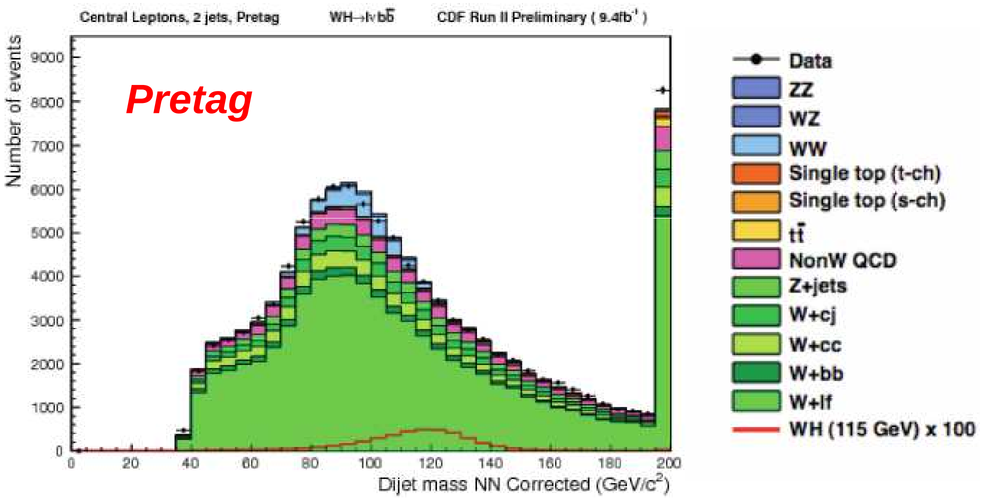
- Counting experiment is hopeless due to the large backgrounds
- Separate signal and background using multivariate techniques
- Bayesian Neural Network is employed
  - Trained separately for the 11 mass point M_H considered
  - The most important input variable is the NN-corrected invariant mass of the 2 jets
  - Other Inputs
    - P_T (jet1) + P_T (jet2) + P_T (lep)− MET
    - P_T (W)
    - $H_T = \Sigma E_T(\text{jets}) + P_T(\text{lep}) + \text{MET}$
    - MET
    - charge x η(lepton)
    - …

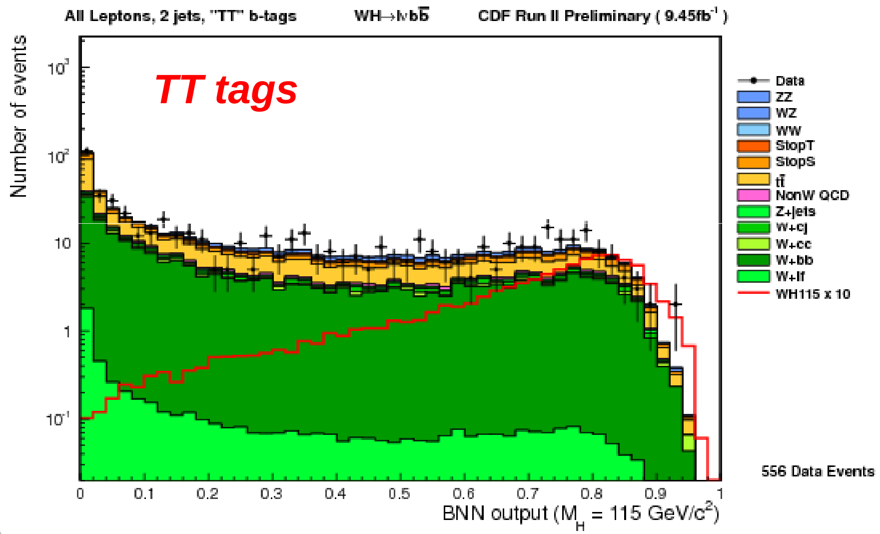


Multivariate Analysis (BNN)

- ❑ Counting experiment is hopeless due to the large backgrounds
- ❑ Separate signal and background using multivariate techniques
- ❑ Checks in control region to validate modelling of inputs and outputs

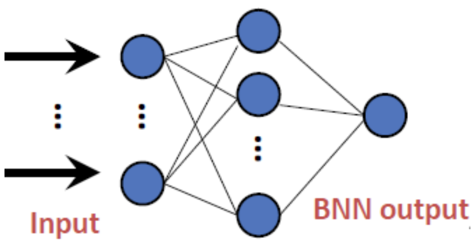
Bayesian Neural Network





+

Other kinematic variables

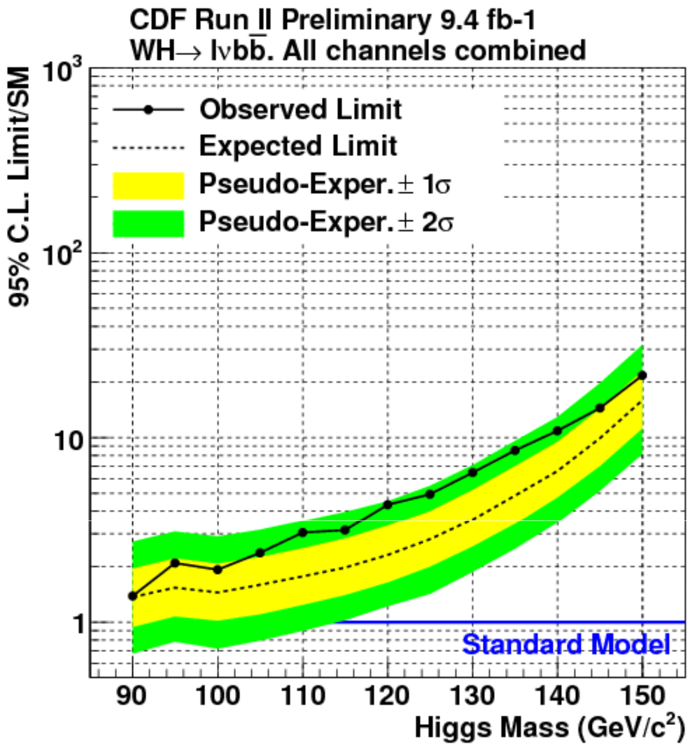


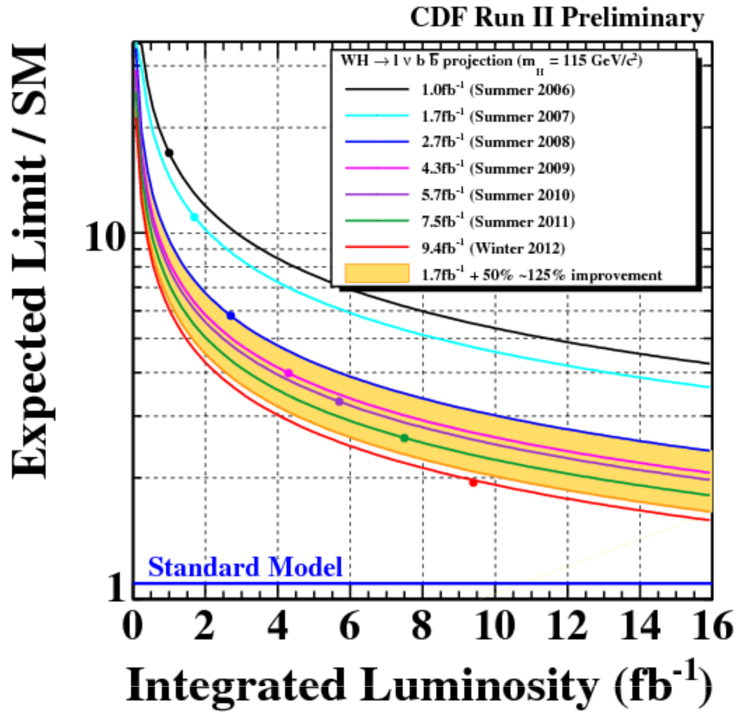
Systematic uncertainties

- Systematic Uncertainties on Signal Acceptance
  - Lepton identification: The difference of ID eff. between data and MC: ~2-4.5%
  - Trigger efficiency: The uncertainties of trigger efficiency: ~1-3%
  - Initial/Final state radiation (ISR/FSR):The difference between higher and lower ISR/FSR MC samples: 5-10%
  - Jet Energy Scale (JES): Estimated by JES 1σ shift from default value: ~2-10%
  - Parton distribution function (PDF): The difference among various PDFs (Estimated from MC) : ~1%
  - b-tagging: The difference of b-tagging efficiency between data and MC: 4-8%
- Shape systematics
  - JES (apply 1σ shift to MC samples)
  - Q^2 (apply 100% factor to default value for W+jets MC samples)

- Mistags: ~20 - 40%
- W+HF: ~30%
- QCD: 40%
- Luminosity ~6%

WH→lνbb Result





- Big improvement over time beyond luminosity beyond luminosity

Expected upper limit:

1.97 x σ(SM) (M_H= 115 GeV/c^2)

Observed upper limit:

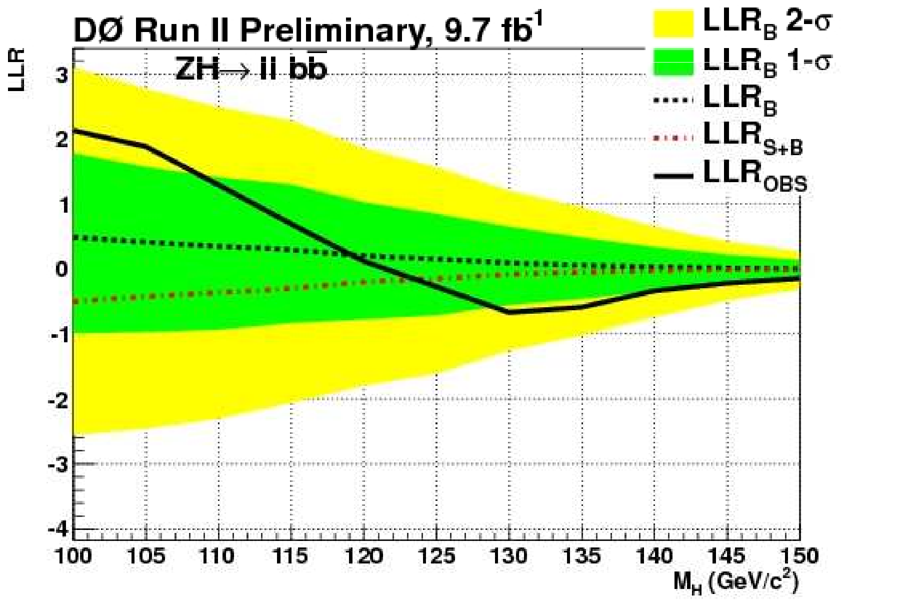
3.13 x σ(SM) (M_H= 115 GeV/c^2)

- Most sensitivity channel in Tevatron combination for M_H <130 GeV/c^2

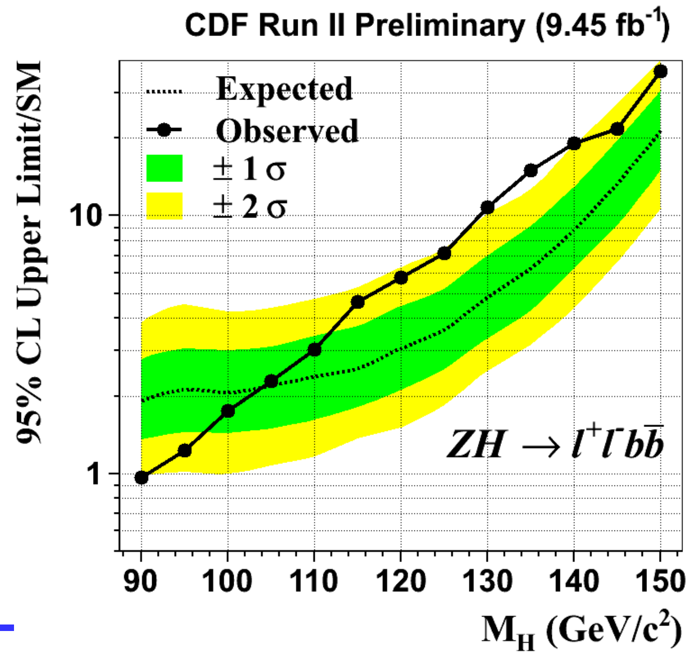
Other low mass channels: ZZ→llbb

- Lowest background, worse statistics
- 2 High PT muon/electrons + 2 or 3 jets
- Fully reconstructed final state: allows to improve Mjj resolution

- **D0** (*9.7 fb^{-1}*): ***Random Forest***
- ***Exp. Upper limit: 4.2x σ(SM)*** (M_H= 115 GeV/c^2)



- **CDF** (*9.4 fb^{-1}*): ***Combine 4 different NN***
- ***Exp. Upper limit: 2.5x σ(SM)*** (M_H= 115 GeV/c^2)

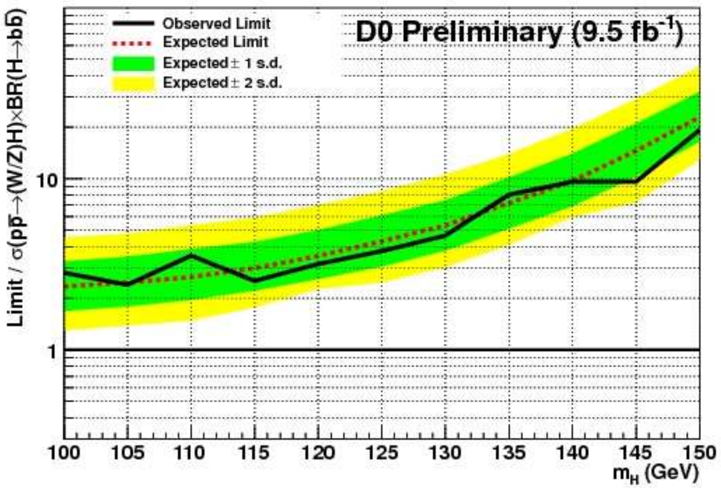


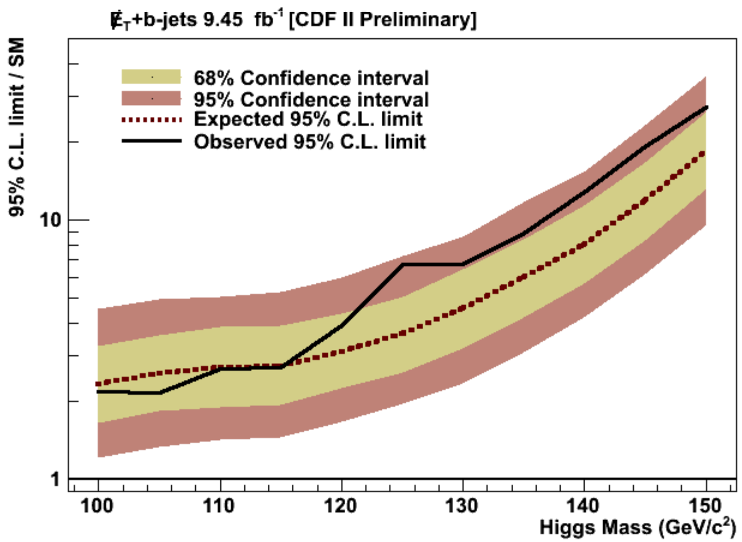
ZZ→ννbb

- ❑ Large MET + lepton veto + at least 2 jets
- ❑ Sensitive to WZ→lνbb with a missed lepton
- ❑ Large MET provides large background rejection

- ❑ **D0** (*9.5 fb^{-1}*): ***multijet DT + Final Boosted Decision Tree***
- ❑ Include b-ID ouput as part of BDT
- ❑ ***Exp. Upper limit: 3.0x σ(SM) @ 115 GeV***

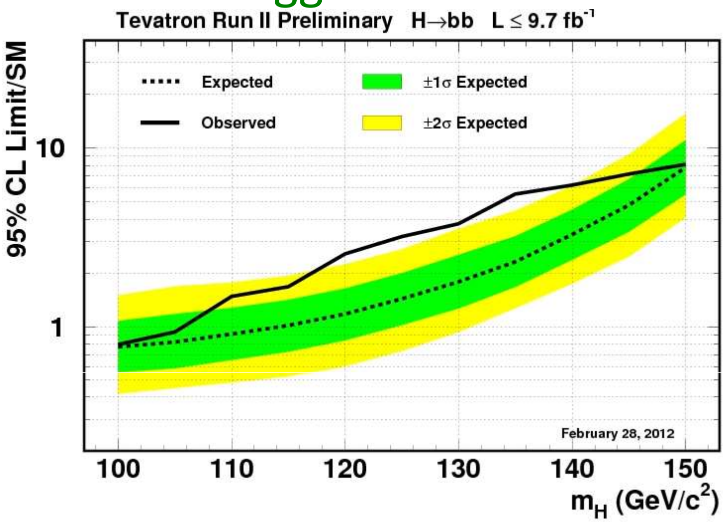
- ❑ **CDF** (*9.4 fb^{-1}*): ***multijet artificial NN + Multilayer perceptron***
- ❑ 2 or 3 jets
- ❑ ***Exp. Upper limit: 2.7x σ(SM)*** Does not make use of HOBIT yet ***yet***



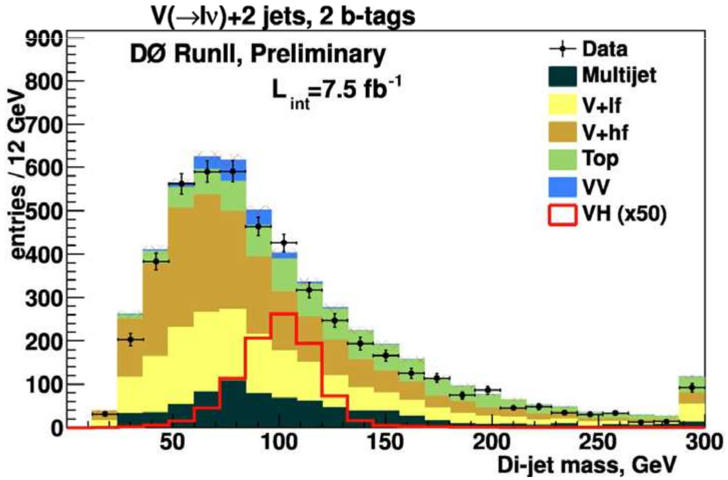


H→bb result and validation

- H→bb validation
  - WZ/ZZ →llbb, lνbb, ννbb same final states as most important channels for low-mass Higgs searches at Tevatron (WH/ZH)

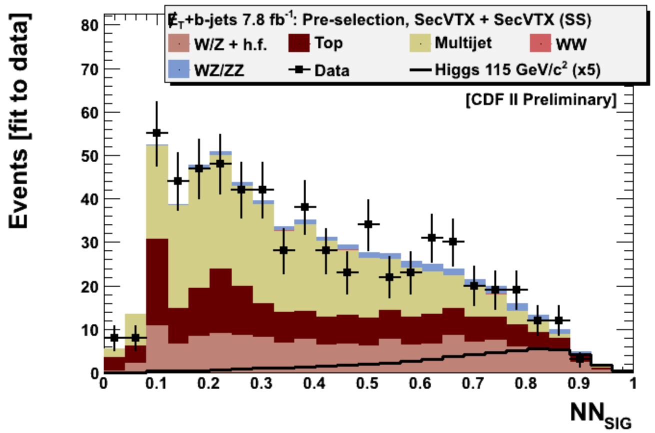


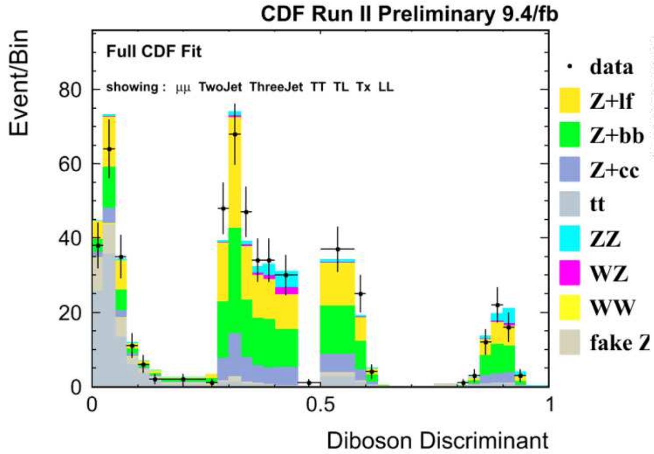
Final State	(W/Z)*H (115 GeV)*	(W/Z)*Z*
lνbb	**27 fb**	**105 fb**
llbb	**5 fb**	**24 fb**
ννbb	**15 fb**	**73 fb**

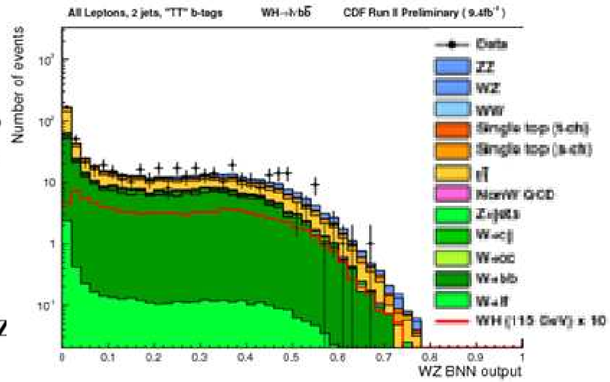


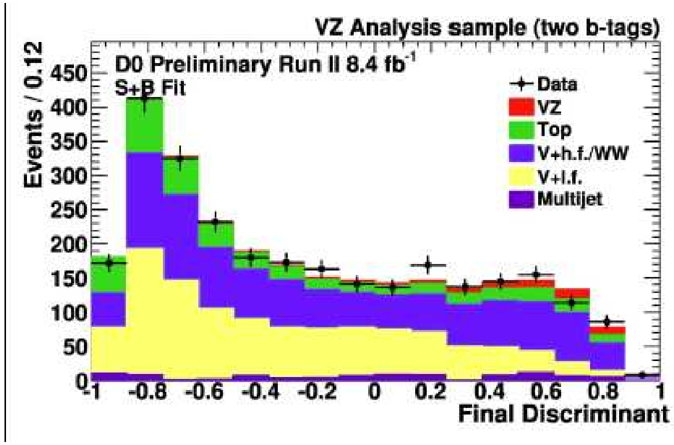
- WZ/WZ ~ 4-5 times larger cross section than WH/ZH
- WZ/ZZ harder to distinguish from large irreducible background (WW…)
- Train multivariate discriminants as for Higgs searches
- Some signal contribution from (W→cs, Z→cc)

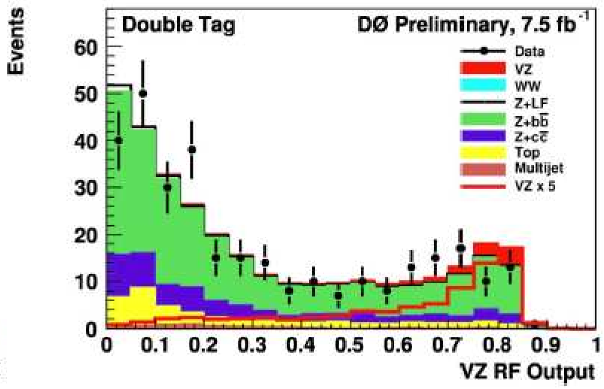
Discriminants optimized for diboson search

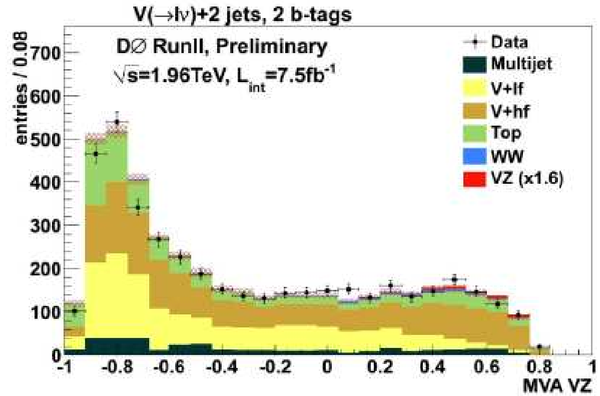






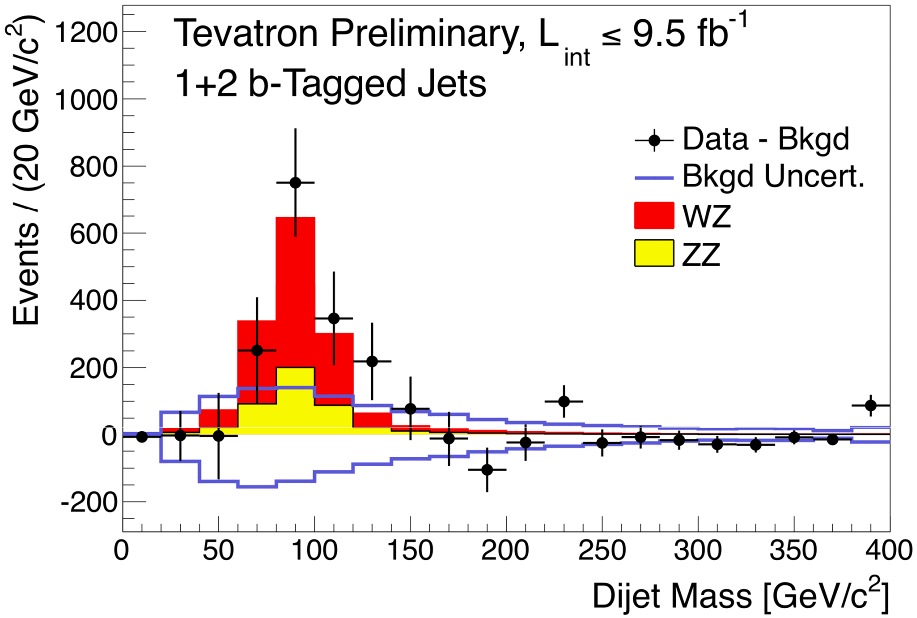


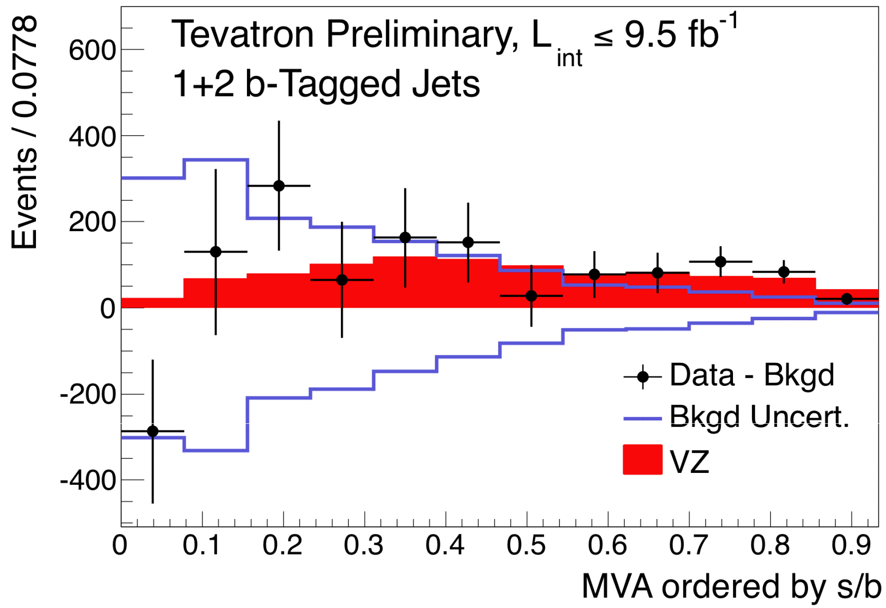




$\sigma(VZ)_{SM}$= 4.42 pb

Tevatron Combination for diboson search

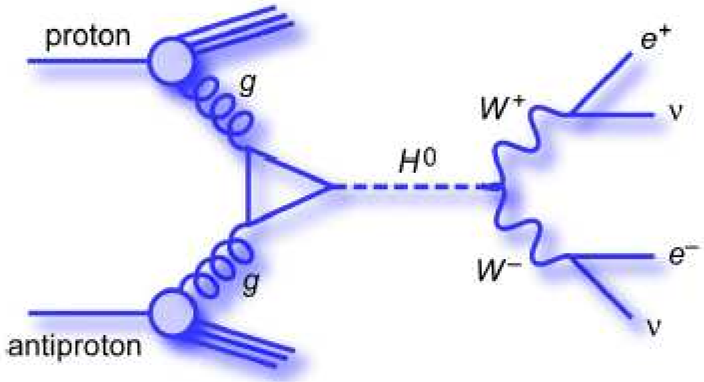




- ***$\sigma(VZ)$=4.47 ± 0.64 ± 0.73 pb***
- Compatible with the SM prediction
- Evidence with ***4.6σ***
- FERMILAB-CONF-12-063-E

High Mass Higgs boson Searches

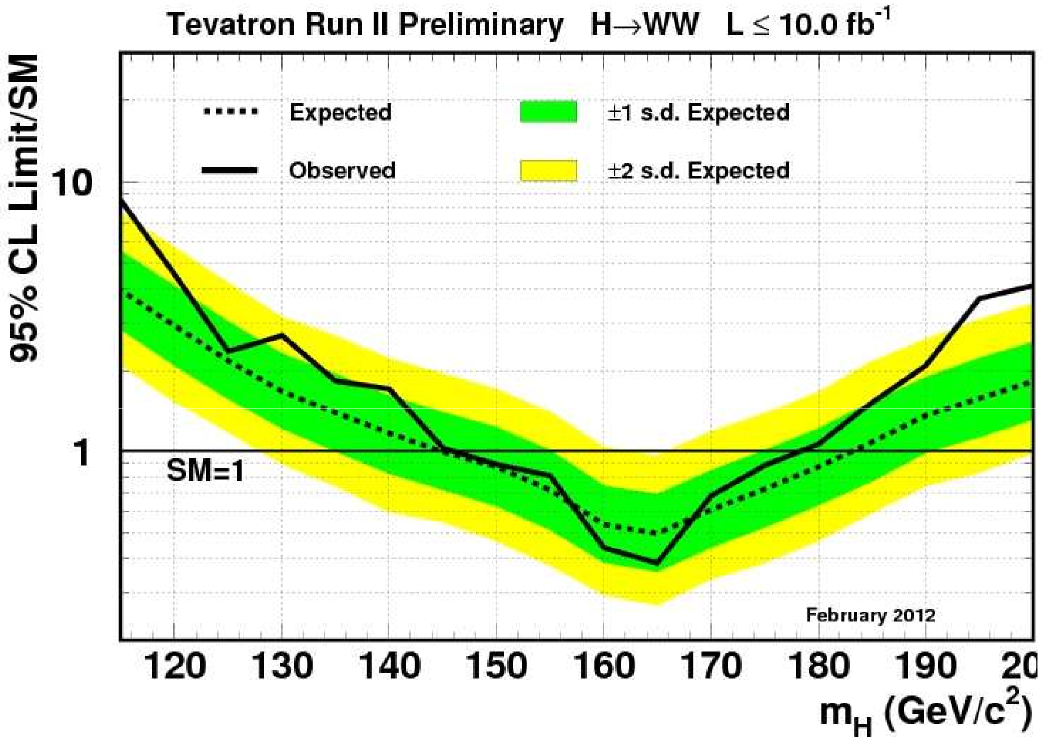
- Main Channel H→WW→lνlν



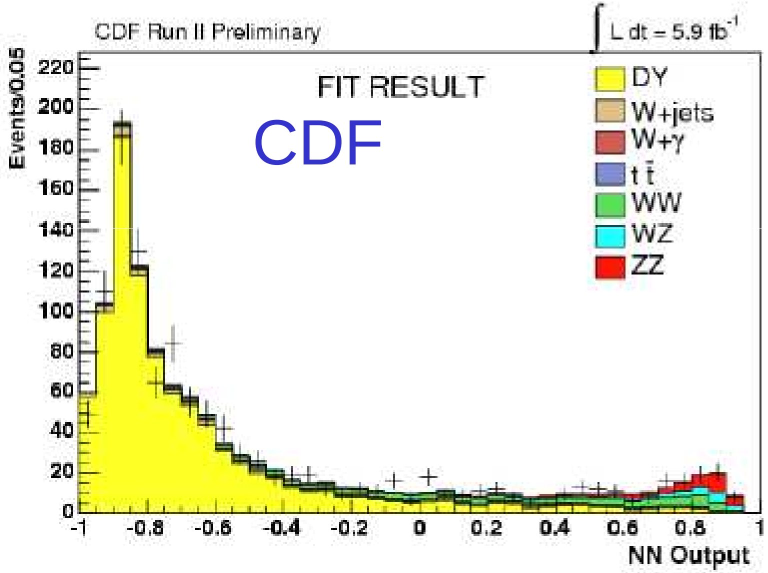
- Signature: 2 High E_T leptons and missing
- SM prediction
- Strategy
  - Maximize lepton and resonstruction and selection efficiencies
  - Separate events into multiple analysis channels
  - Best possible choice of kinematic event variables for separating signal and background

High Mass Higgs boson Searches

- Competitive with individual searches at m_H=125 GeV/c^2



NN optimized for ZZ→llνν



- Validated technique with σ(ZZ)

$$\sigma(ZZ) = 1.45^{+0.60}_{-0.51}\ \text{pb}$$
SM pred: 1.2±0.1 pb

Other Channels

CDF

Channel	Luminosity	95% CL limit M_H=125 GeV
$H \to \gamma\gamma$	10.0 fb^{-1}	10.8 x SM
$VH \to bb$+jets	9.45 fb^{-1}	11.0 x SM
$ttH \to l\nu$+jets	9.4 fb^{-1}	12.4 x SM
$H \to \tau\tau$+jets	8.4 fb^{-1}	14.8 x SM

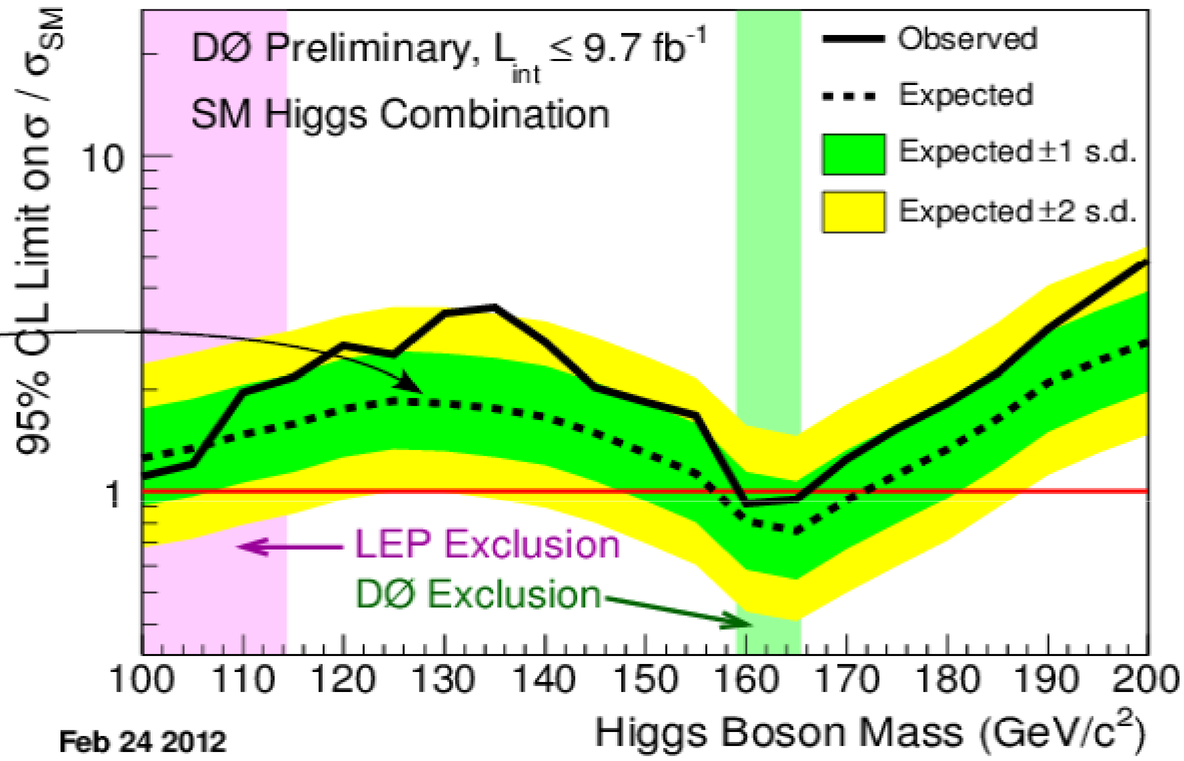
- No channel left behind
- Similar list for D0

Channel	Luminosity	95% CL limit M_H=150 GeV
$H \to ZZ \to llll$	9.7 fb^{-1}	9.4 x SM

D0 Combination

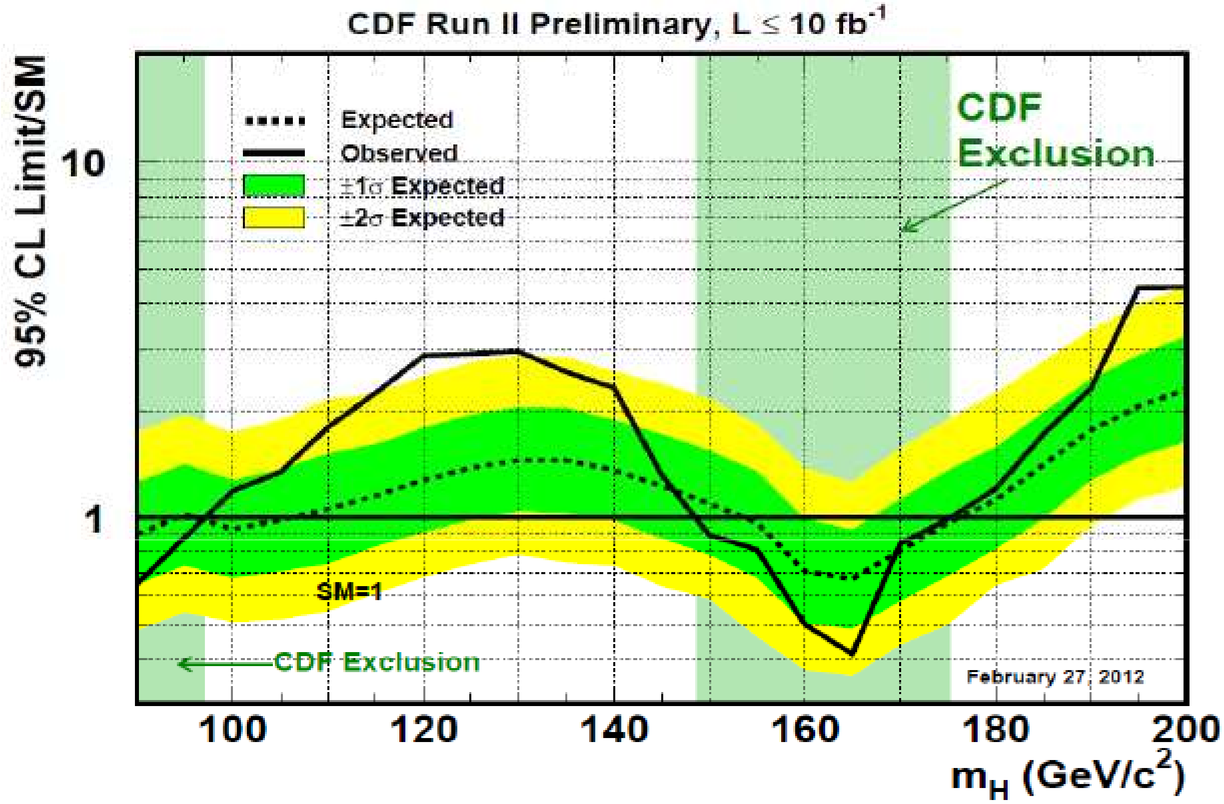
Combining all of the analyses:

Expected limit within $1.85\times\sigma_{SM}$



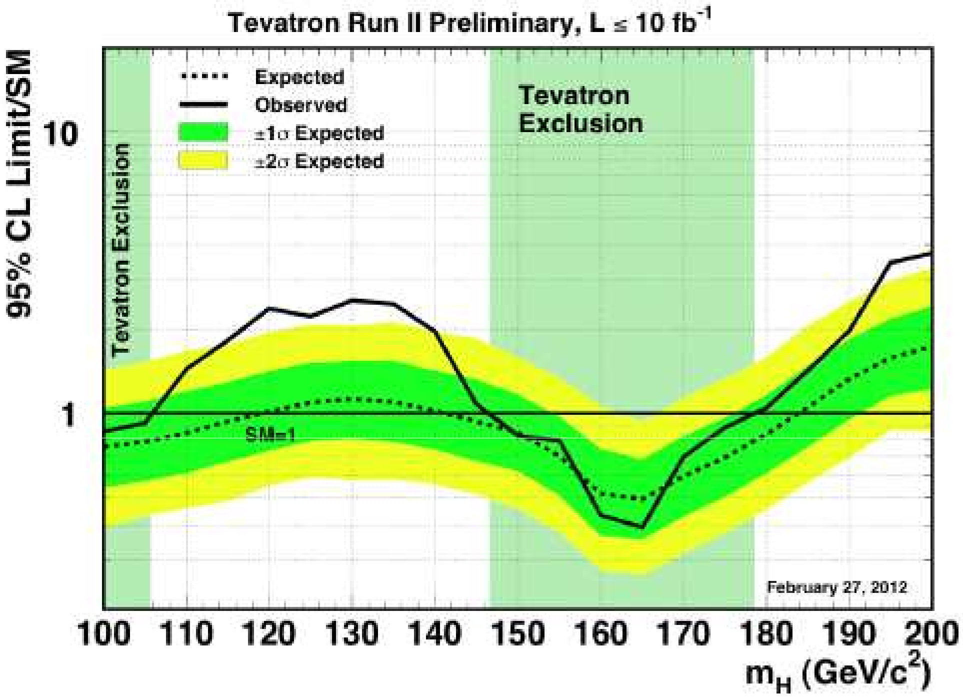
- Excess at the level of 1–2 standard deviations in the mass range 115–145 GeV
- More improvements to come

CDF Combination



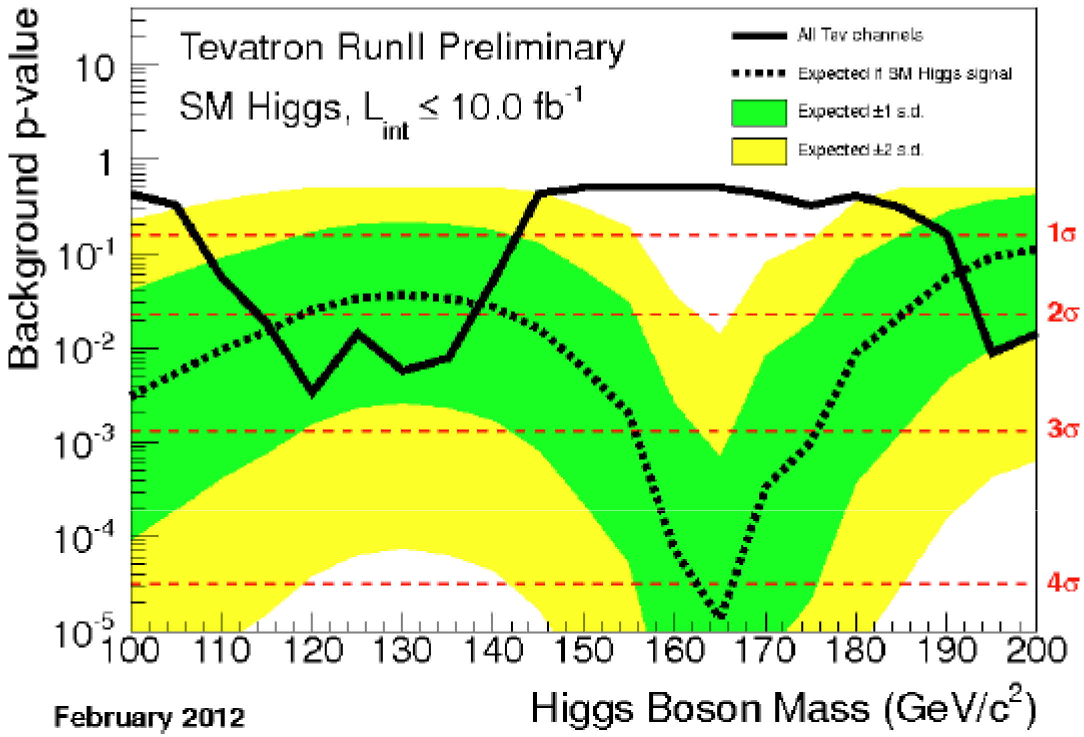
- Exclude 147 – 175 GeV/c^2
- Broad excess observed. Largest excess: 120 GeV

Tevatron Combination



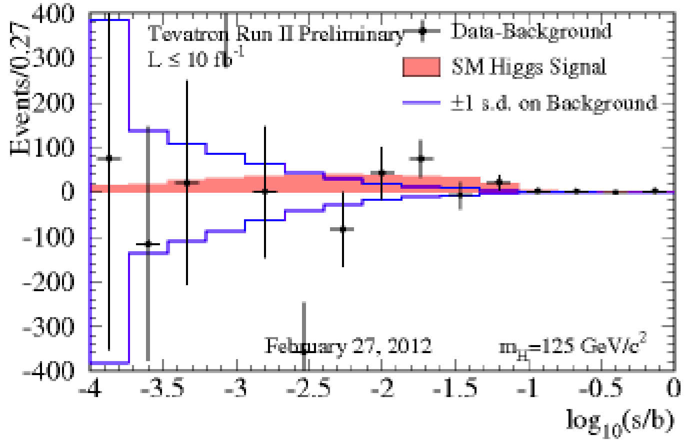
- ❑ Exclude SM Higgs at 95% C.L. for 147 < mH < 179 GeV/c^2
- ❑ Expect to exclude 100 < mH < 120 GeV/c^2 & 141 < mH < 184 GeV/c^2

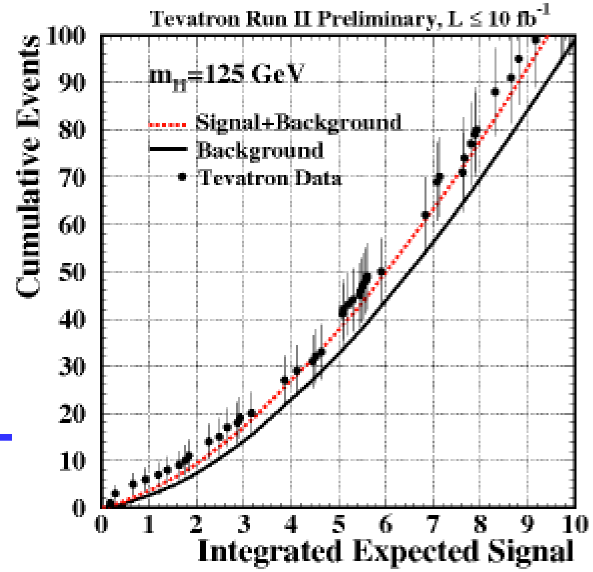
Global significance



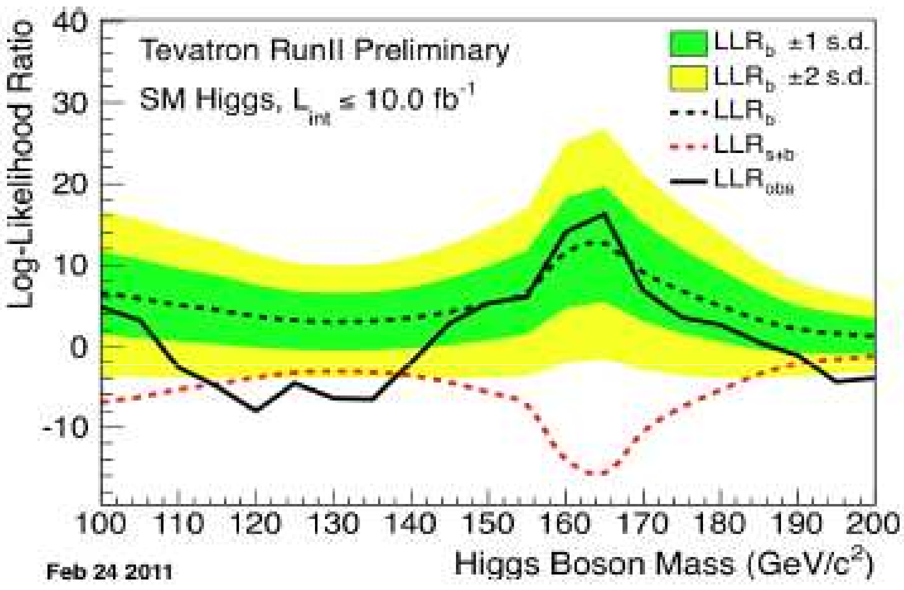
- Highest local p-value is found at mH=120 GeV/c2
- Same LEE of 4 for entire SM search range from 100 to 200 GeV/c2

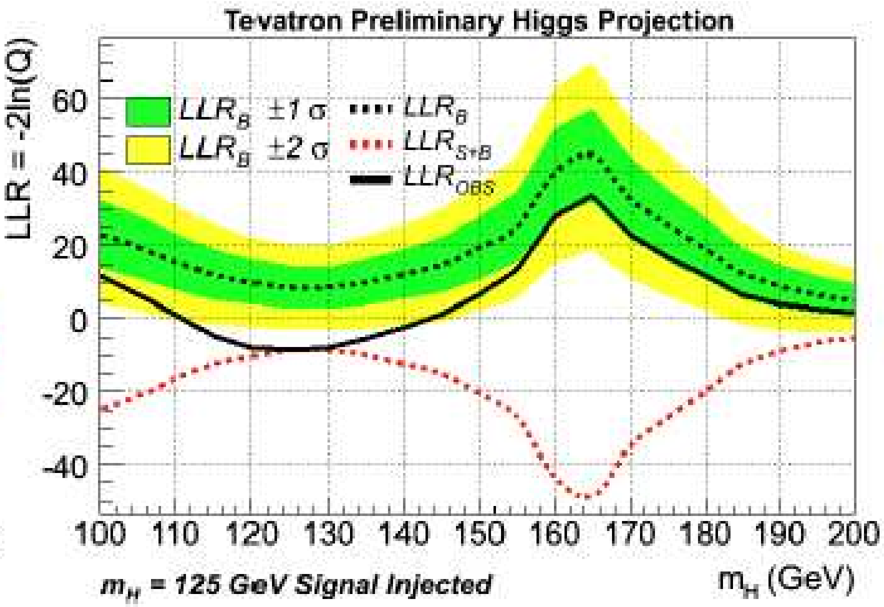
Minimum p-value	Mass (GeV)	global sig.	local sig.
Tevatron	**120**	**2.6 σ**	**2.2 σ**
CMS	**125**	**2.8 σ**	**2.1 σ**
ATLAS	**126**	**2.5 σ**	





Tevatron Combination





- Consistent with SM signal plus background hypothesis over Higgs mass range from 110 to 140 GeV/c2

Summary

- Tevatron: Expected upper limit in SM Higgs cross section < 2 x SM in the range 100 – 200 GeV
- At low mass sensitivity dominate by W/Z associated production in $H \rightarrow bb$ modes
- Described extensively $WH \rightarrow l\nu bb$ search
  - Use of sophisticated analysis techniques
- Measured $\sigma(WZ+ZZ)$ with a significance of 4.6 σ and a value compatible with SM
- Observed an excess of Higgs-like events consistent with SM Higgs production in the mass range from 115 to 140 GeV/c^2
- The global significance of this excess is 2.2 σ

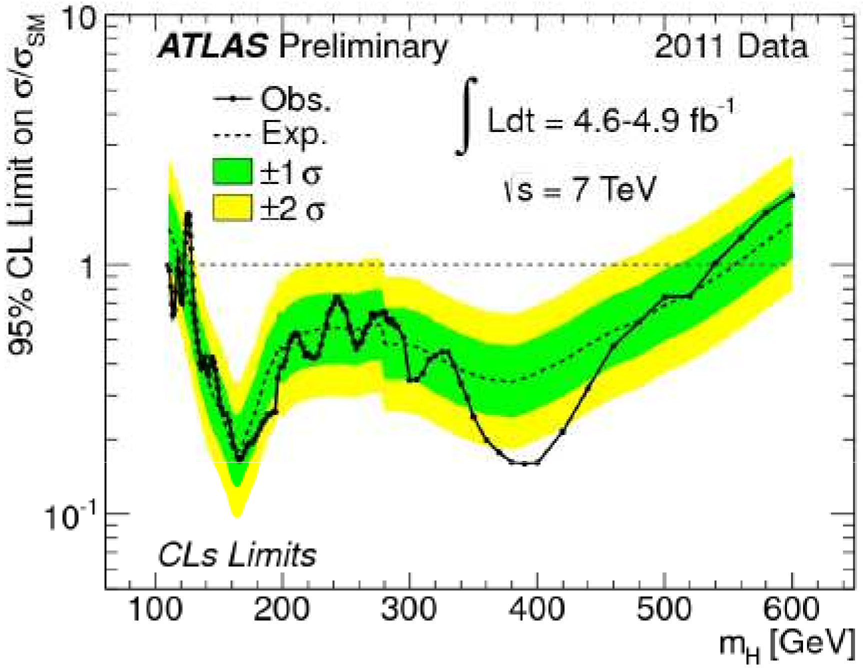
Back-up

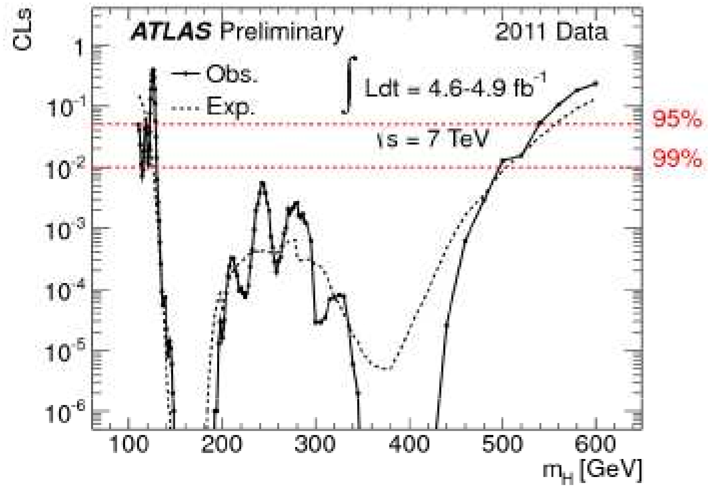
Event Selection: 2 Jet bin, 7.5 fb-1

Signal (WH M_H=115 GeV) event yields for different types of leptons

CEM	PHX	CMUP	CMX	NTM	LEL	All
10.3	1.68	5.89	3.01	7.22	3.33	31.4

ATLAS exclusion





Expected exclusion at 95% CL: 120-555 GeV

Observed exclusion at 95% CL: 110-117.5, 118.5-122.5, 129-539 GeV

Observed exclusion at 99% CL: 130-486 GeV

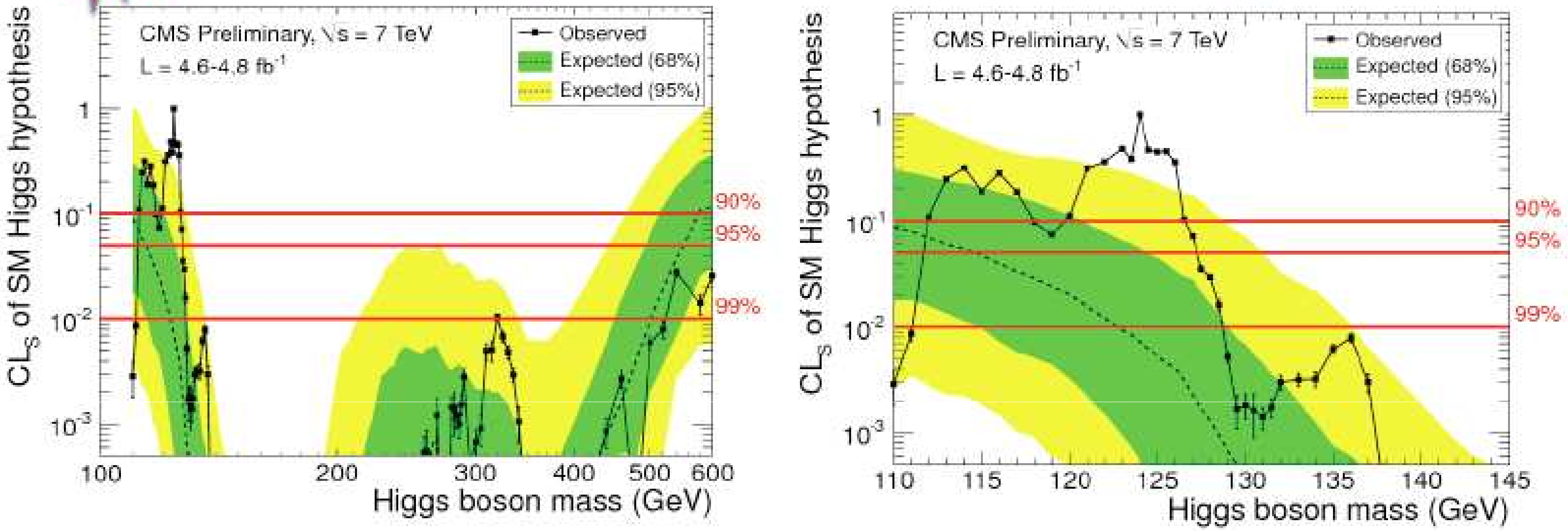
CMS exclusion

NEW

Full mass range

Low mass region

CMS document HIG-12-008

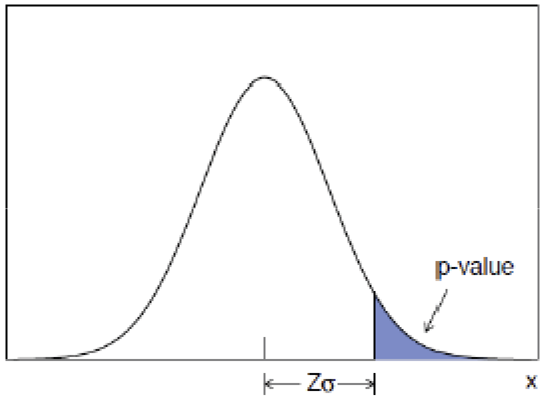


- Expected: 95% exclusion M_H in [114.5-543] GeV
- Observed: 95% exclusion M_H in [127.5-600] GeV
  99% exclusion M_H in [129-525] GeV
- 95% allowed mass range: 114.4-127.5 GeV
- Observed lower limit higher than expected because of excess in data at low mass

- ***p-value***: chance of bkgd fluctuating as high as or higher than what has been observed in data

$$p = \mathrm{P}(\ n \geq n_{\mathrm{obs}}\ |\ \mathrm{b}\ )$$

- ***significance* Z** is related to *p-value* via the tail probability of the normal distribution

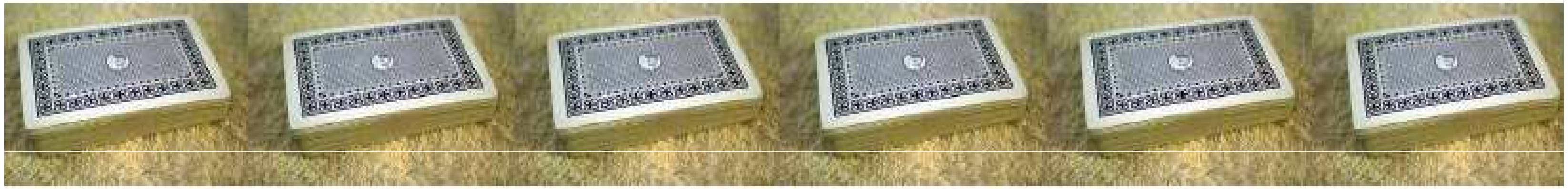


- *p-value* does not tell us whether the excess is consistent with the SM Higgs boson CS. Hence, we also quote the **best-fit value of the signal strength modifier $\hat{\mu}$**

Loke Elsewhere Effect

What is the chance that the top card in a deck is Queen of Hearts?

Local *p-value*



What is the chance that the top card in at least one of N decks is Queen of Hearts?

Global *p-value*

The fact that the two answers are not the same is known under a fancy name **Look-Elsewhere Effect.** The ratio of the two probabilities = **trials factor.**

From Discriminants to Limits

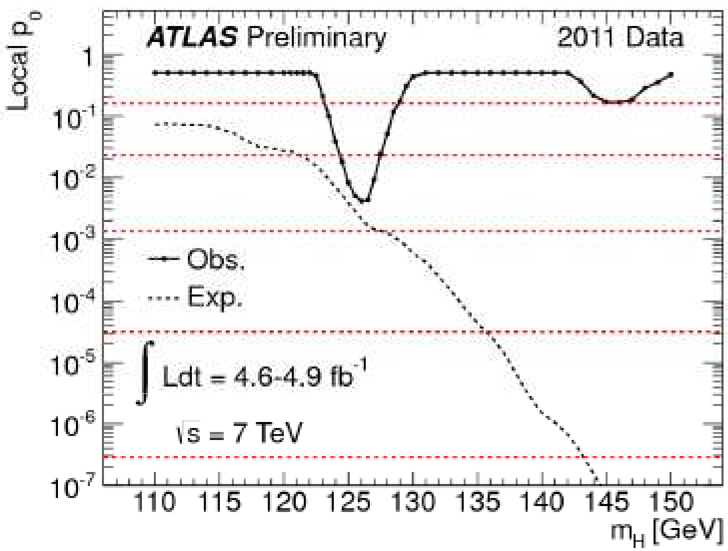
- Combined binned likelihood function

$$L(R, \vec{s}, \vec{b}|\vec{n}) = \prod_{i=1}^{N_{\text{channel}}} \prod_{j=1}^{N_{\text{bin}}} \frac{\mu_{ij}^{n_{ij}} e^{-\mu_{ij}}}{n_{ij}!}$$

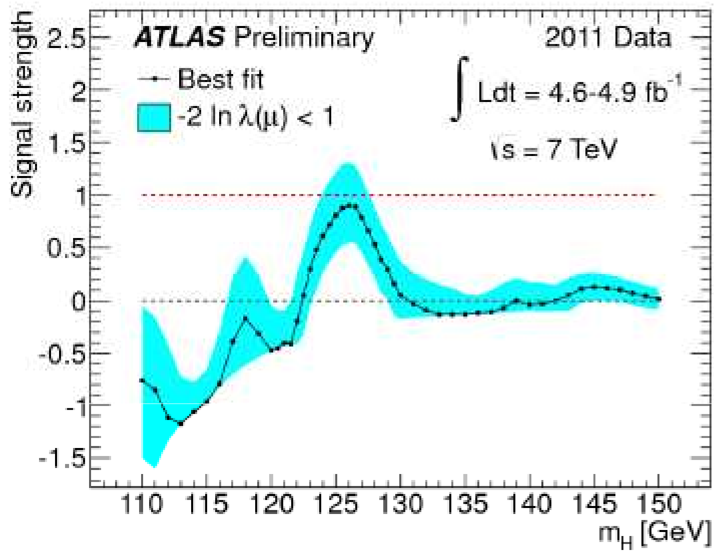
- Incorporate uncertainties as nuisance parameters
- Uncertainties taken on both the shapes and normalizations of signal & background templates
- Additional constraints on background model obtained directly from fit!!

ATLAS excess

Probability that the excess
is caused by a background fluctuation:



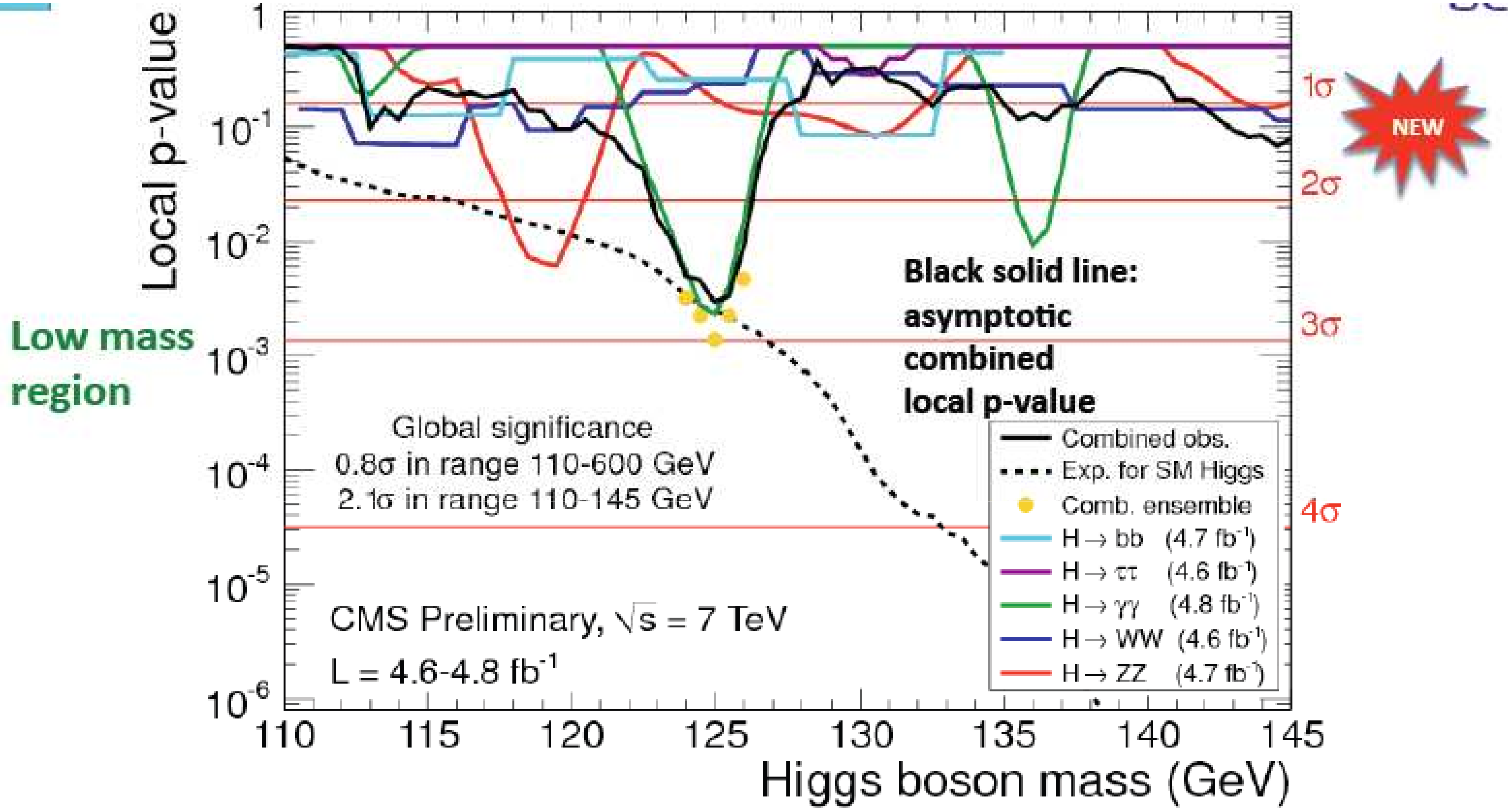
Best fit signal strength $\mu = \sigma/\sigma_{SM}$:



Excess of events observed at 126 GeV:

- Observed local significance 2.5σ (expected 2.9σ).
- Best-fit signal strength at 126 GeV: $\hat{\mu} = 0.9^{+0.4}_{-0.3}$.
- Global probability of such a background fluctation anywhere in the full explored mass range (110-600 GeV): 30%;
  in the mass range (110-146 GeV): 10%.

CMS excess



- Minimum p-value observed at 125 GeV with local significance: 2.8 σ
  - Similar significance expected from signal
- Estimated global significance:
  - 0.8 σ in [110-600] GeV, 2.1 σ in [110-145] GeV,