Standard Model Higgs Boson Search in the W(lν)H($b\bar{b}$) Channel at CMS

Analysis Overview

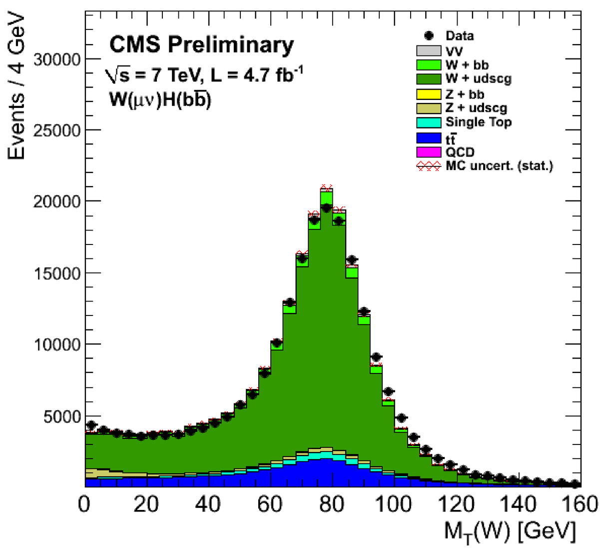
- We search for the Standard Model Higgs boson (mass range of 110-135 GeV) in the associated W/Z production channel with decay to $b\bar{b}$ pairs, using 4.7 fb^{-1} of CMS data collected from 7 TeV pp collisions at the LHC
- We perform cut-and-count analyses (using both dijet mass and multivariate analyzer output separately to define signal regions) on five modes – W(eν)H($b\bar{b}$), W(μν)H($b\bar{b}$), Z(ee)H($b\bar{b}$), Z(μμ)H($b\bar{b}$), and Z(νν)H($b\bar{b}$)
- Presented here are results for W(eν)H($b\bar{b}$) and W(μν)H($b\bar{b}$) modes, as well as combined results from all five modes

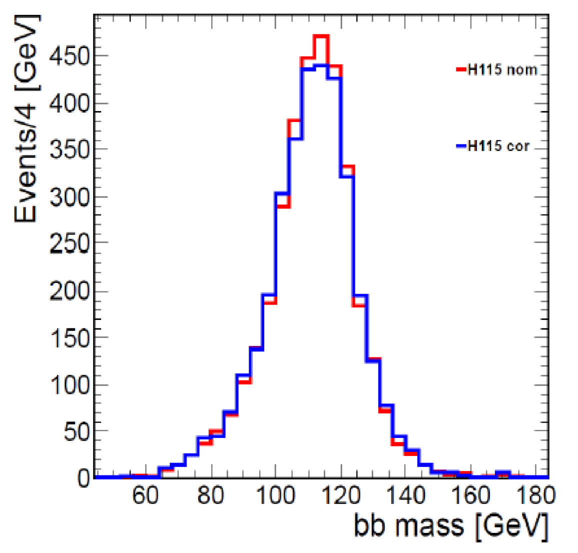
Triggers

- Primary triggers are single-lepton triggers with isolation; later data-taking makes use of jets and missing energy (electrons only)
- Tight tracker/calorimeter isolation for both muons and electrons to keep lepton E_T threshold low and to control trigger rate
- Trigger choices give offline E_T threshold of 20 GeV for muons and 30 GeV for electrons

Physics Objects

- Standard physics objects used with particle-flow algorithm, and Combined Secondary Vertex algorithm used for b-tagging
- Charged Hadron Subtraction and Fastjet Subtraction used in jet reconstruction and identification, also for lepton isolation
- Electron selection uses particle-flow algorithm and tight isolation to reject QCD
- Muon selection requires isolated global and tracker muon with |η| < 2.4 and at least one pixel hit, ten total tracker hits, one valid muon chamber hit, and two muon stations
- Anti-k_T jets with R = 0.5, p_T > 30 GeV used
- W boson selected with large boost (p_T >~ 150 GeV) which significantly reduces QCD, M_T cut unhelpful as Δφ(l,ν) ~ 0
- Higgs boson selected using dijet combination with highest p_T(jj) with large boost (p_T >~ 150 GeV), which improves dijet mass resolution over no boost



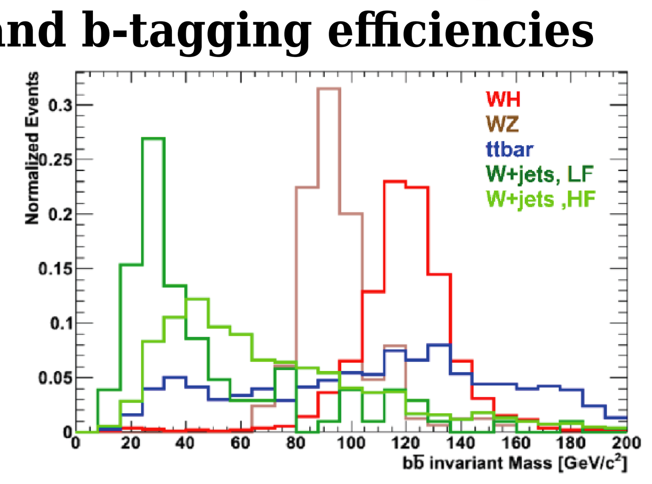


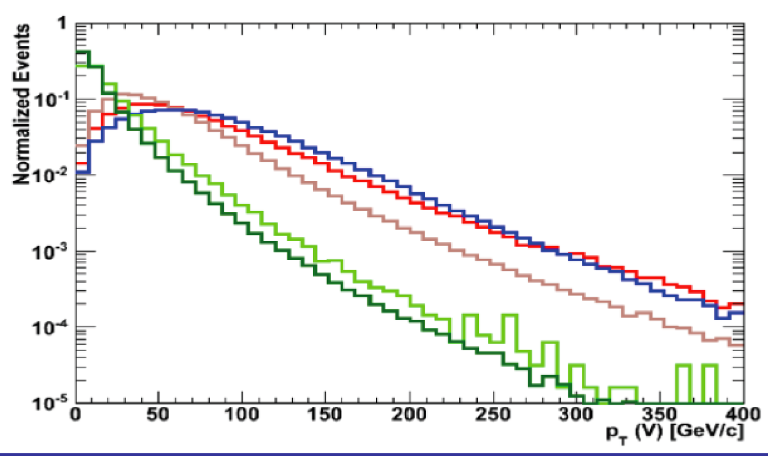
Event Selection

- Main backgrounds: W/Z+jets, $t\bar{t}$ (WW/WZ and Single Top lesser issue with small cross-section)
- W/Z+jets hard to reduce due to topological similarity to signal, use p_T(jj) cut and b-tagging
- Use topological cuts to reduce $t\bar{t}$ such as additional jet multiplicity and Δφ(W,H)
- Reweight simulated events to match data pileup and to account for lepton trigger/reconstruction and b-tagging efficiencies

M(jj) Analysis

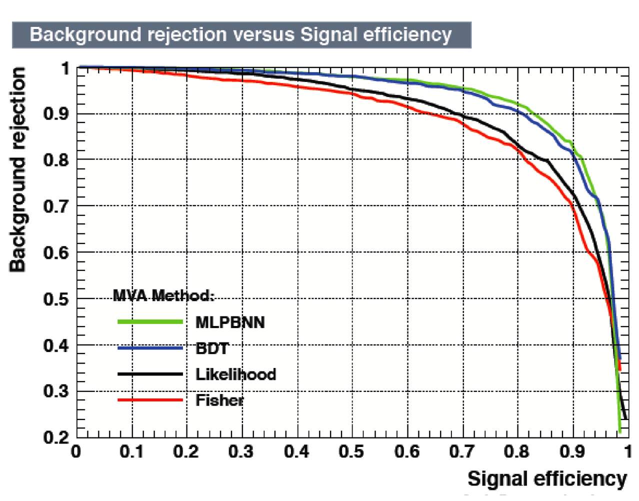
Variable	Cut
$p_T(b_1)$	> 30 GeV
$p_T(b_2)$	> 30 GeV
p_T(jj)	> 165 GeV
p_T(W)	> 160 GeV
CSV1	> 0.9
CSV2	> 0.5
Δφ(W, H)	> 2.95
N_{aj}	= 0
N_{al}	= 0
pfMET	> 30 GeV (W(eν)H)
M(jj) (110 GeV)	95–125 GeV
M(jj) (115 GeV)	100–130 GeV
M(jj) (120 GeV)	105–135 GeV
M(jj) (125 GeV)	110–140 GeV
M(jj) (130 GeV)	115–145 GeV
M(jj) (135 GeV)	120–150 GeV





Multivariate Analysis

- Makes full use of correlations between discriminating variables in signal and background
- Boosted Decision Tree (BDT) method used, implemented in TMVA framework
- Significance increase of ~15-25% over M(jj) analysis
- BDT training variables: Δφ(W,H), Δη(j1,j2), max jet CSV (CSV1), min jet CSV (CSV2), M(jj), p_T(jj), p_T(W), and N_{aj}



BDT Analysis

Variable	Cut
$p_T(b_1)$	> 30 GeV
$p_T(b_2)$	> 30 GeV
p_T(jj)	> 150 GeV
p_T(V)	> 150 GeV
CSV1	> 0.40
CSV2	> 0.40
N_{al}	= 0
pfMET/$\sqrt{\sum E_T}$	> 2 (W(eν)H)
BDT Output	> 0.050

Background Control Samples

- Define several control samples (CS) enriched in different background components: W+light, W+heavy, and $t\bar{t}$ ("light" - only u/d/s/c/g jets, "heavy" - at least one b jet)
- Find data/MC ratio in each control sample to obtain background scale factors for signal region (solve system of three equations, one for each CS, taking into account relative purity of backgrounds)
- Uncertainties on scale factors include statistical uncertainty and a systematic shape uncertainty determined by finding maximal variation of scale factor when moving CS cuts toward signal region

W+light

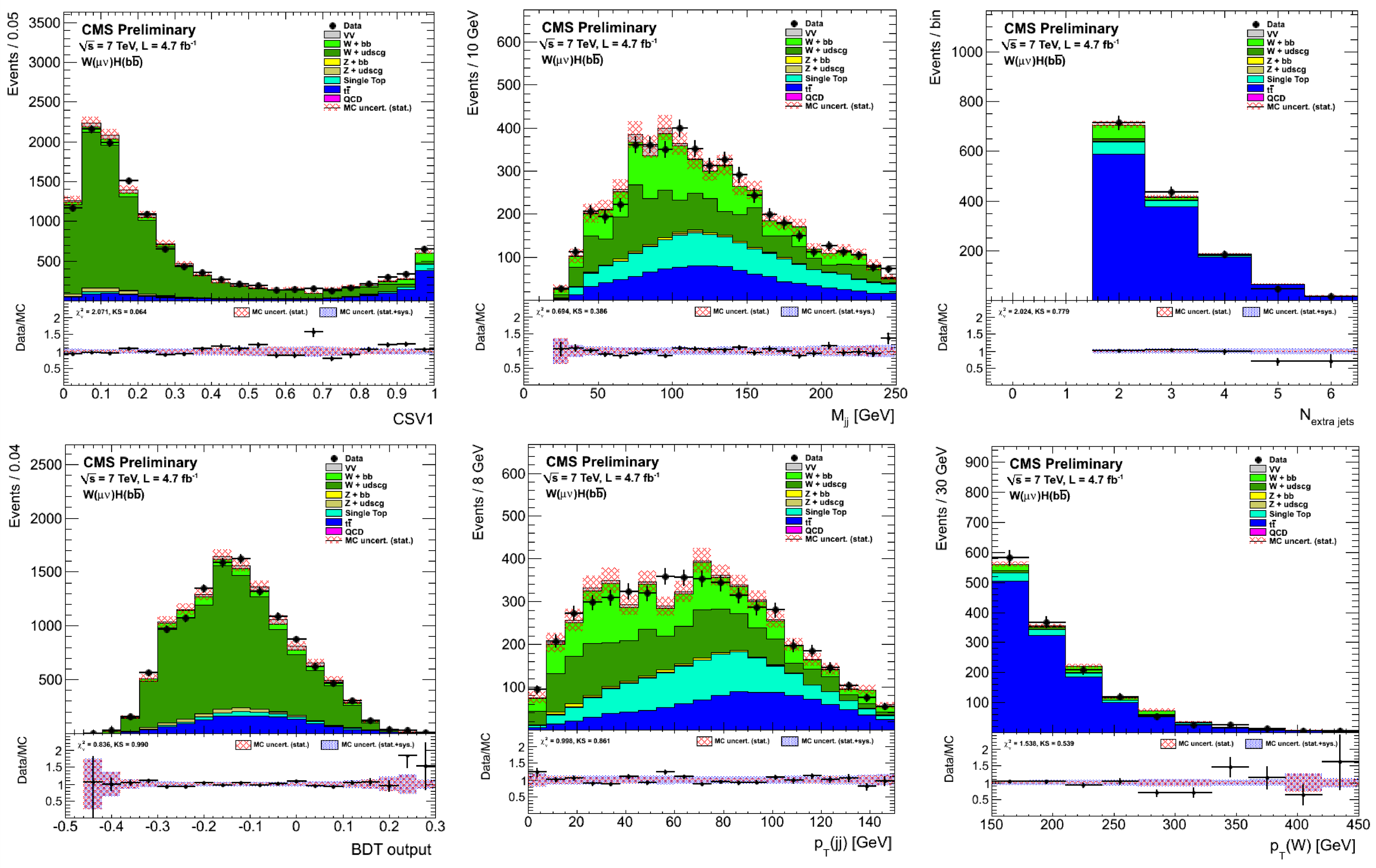
Variable	Cut
$p_T(b_1)$	> 30 GeV
$p_T(b_2)$	> 30 GeV
p_T(jj)	> 150 GeV
p_T(W)	> 150 GeV
N_{aj}	< 2
pfMET/$\sqrt{\sum E_T}$	> 2.5

W+heavy

Variable	Cut
$p_T(b_1)$	> 30 GeV
$p_T(b_2)$	> 30 GeV
p_T(jj)	< 150 GeV
p_T(W)	< 150 GeV
M_T	> 40 GeV, < 120 GeV
CSV1	> 0.9
N_{aj}	= 0
pfMET/$\sqrt{\sum E_T}$	> 2.5

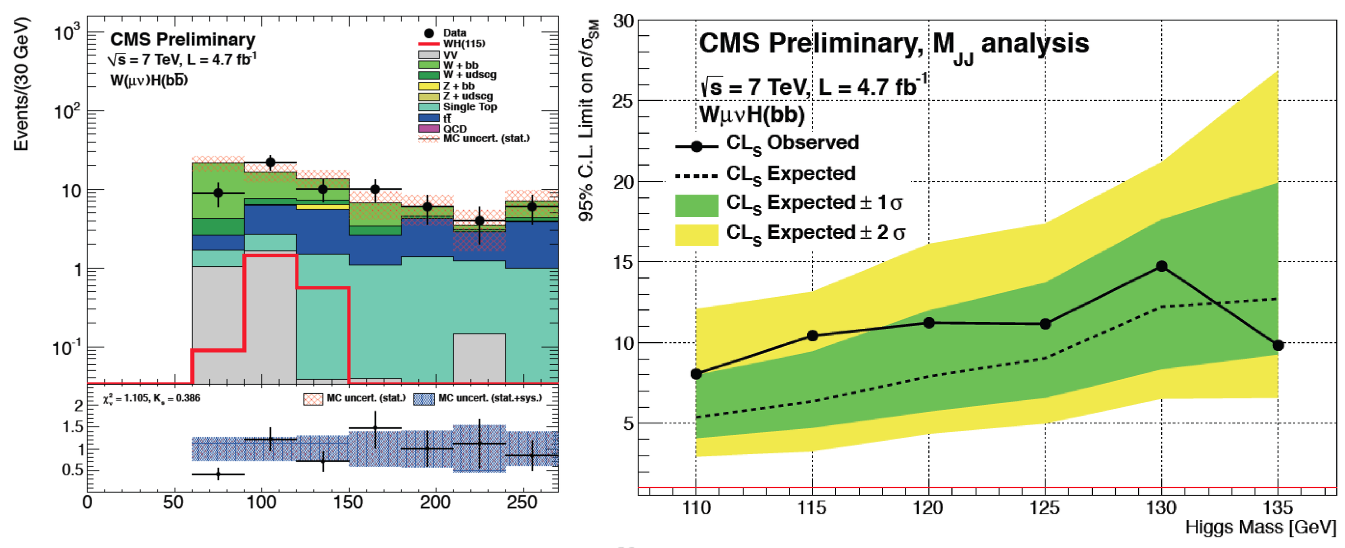
$t\bar{t}$

Variable	Cut
$p_T(b_1)$	> 30 GeV
$p_T(b_2)$	> 30 GeV
p_T(jj)	> 100 GeV
p_T(W)	> 100 GeV
CSV1	> 0.9
N_{aj}	> 1

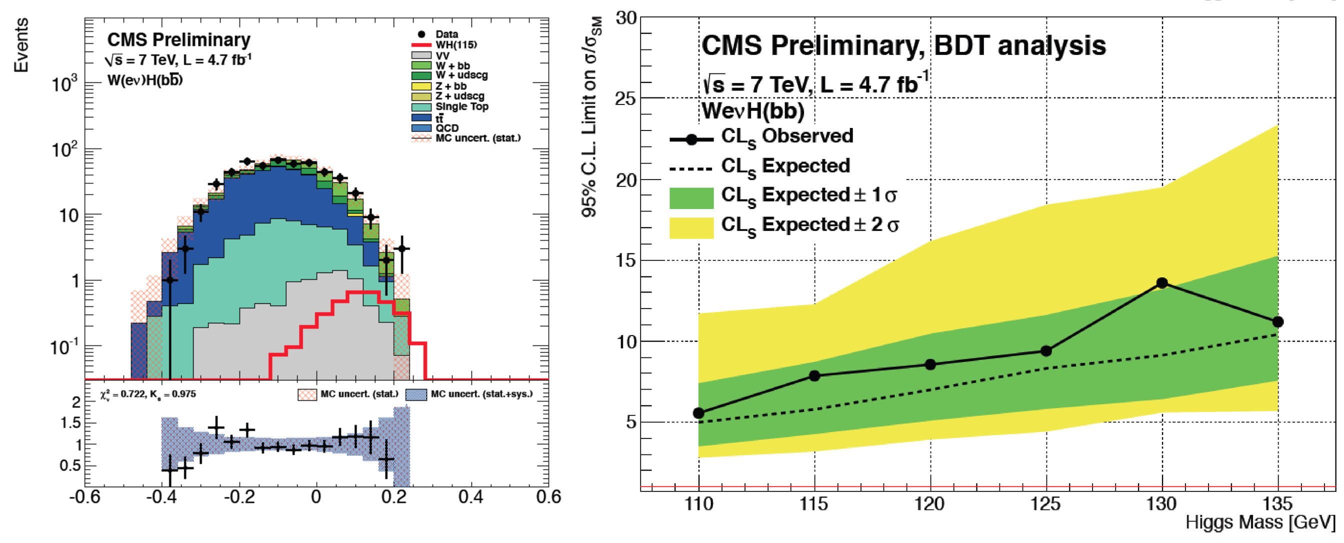


Results: W(lν)H($b\bar{b}$) Modes Only

W(μν)H($b\bar{b}$) M(jj) Analysis

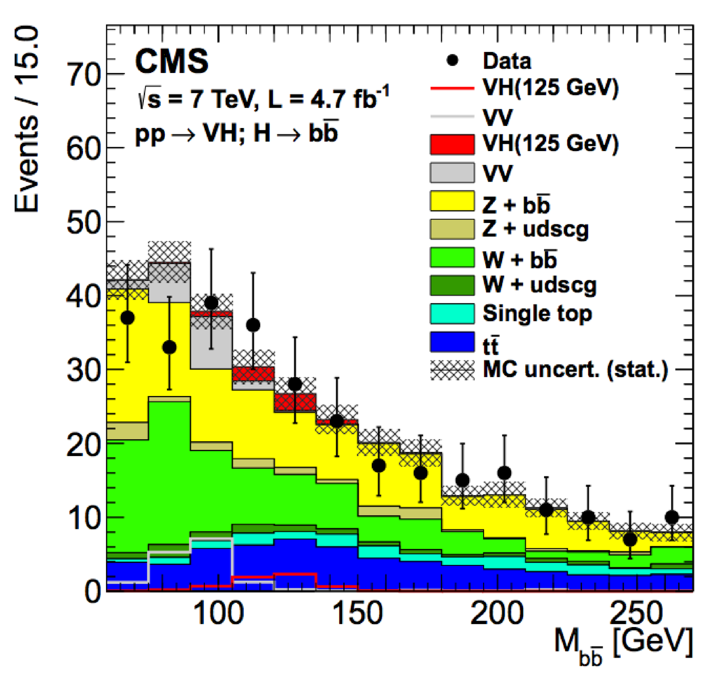


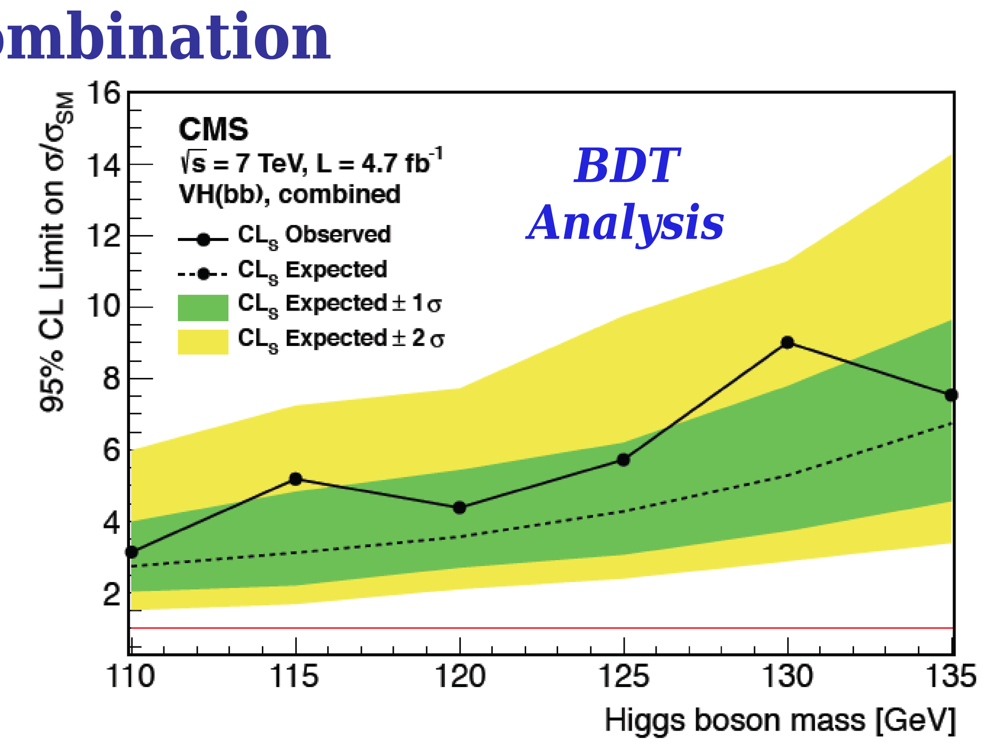
W(eν)H($b\bar{b}$) BDT Analysis



- Main systematics: b-tagging (20%), jet resolution (10%), signal p_T spectrum (10%)
- Find 95% C.L. upper limits using RooStats
- Obtain 5-20% gain in expected limit with BDT vs. M(jj)
- W(lν)H(bb) contributes most to expected limit: ~4*SM (115 GeV)

Results: Five Mode Combination





- Total expected limit: ~3*SM (115 GeV)
- Observed limit from combination consistent with expectations

Michael R. Mooney, Princeton University – LHCC Poster Session – March 21, 2012 – Geneva, Switzerland