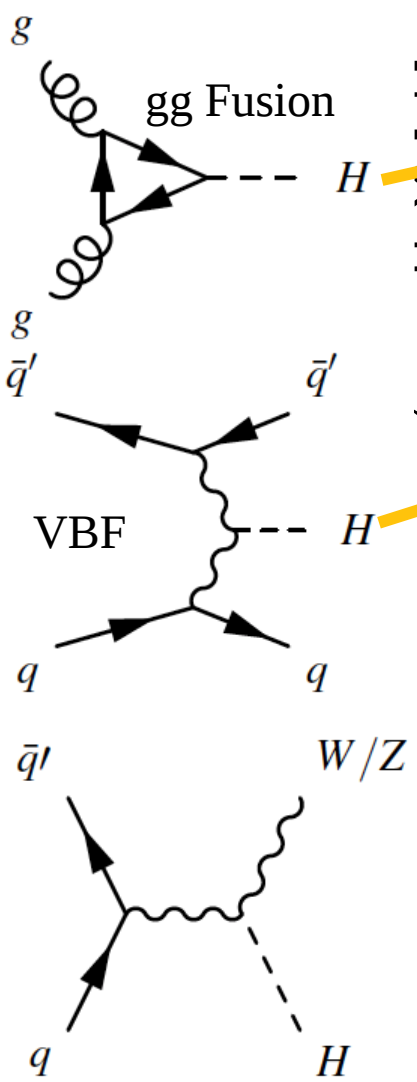
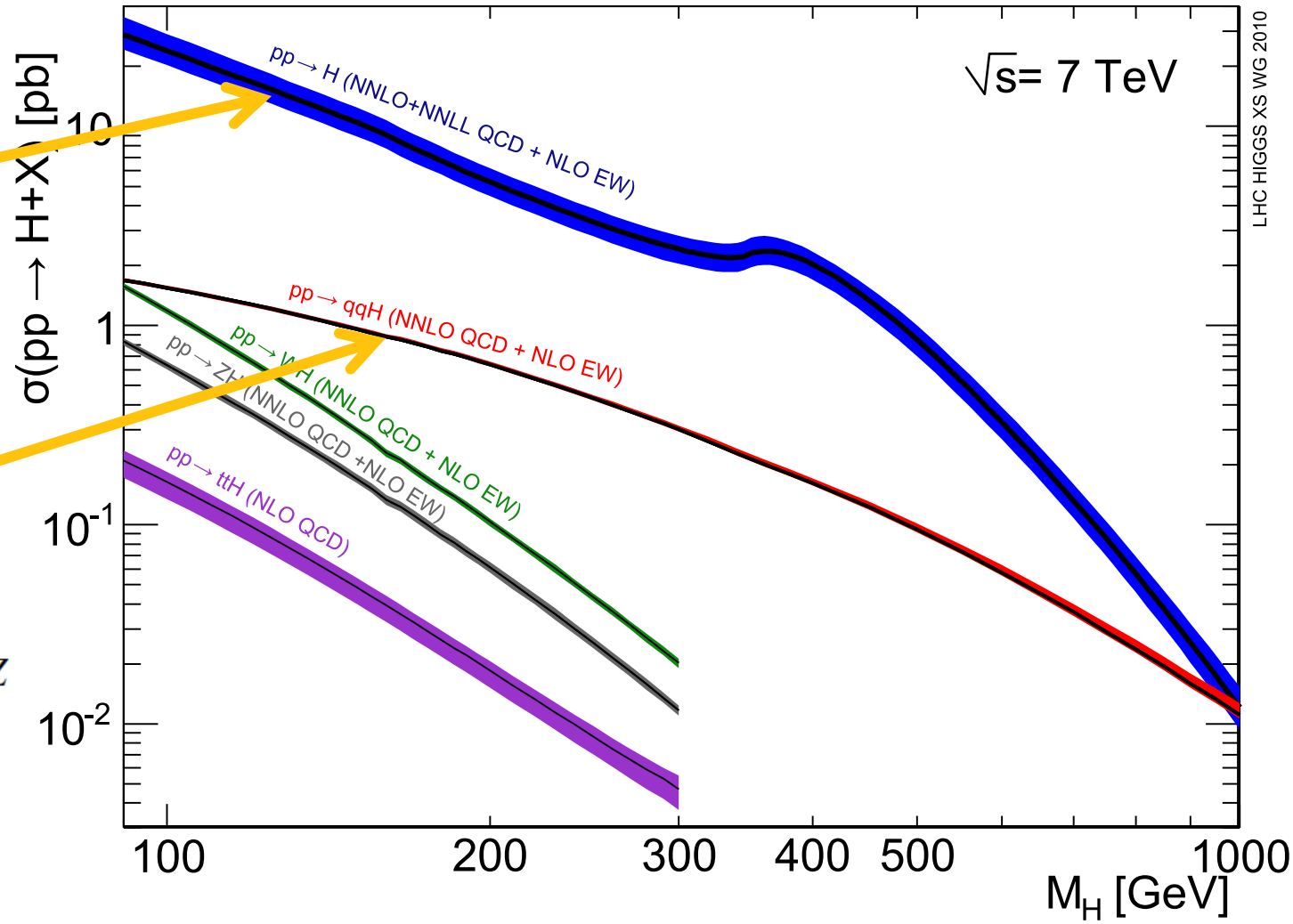
Search for the Standard Model Higgs boson in its decays to all leptonic WW in CMS

Jake Anderson

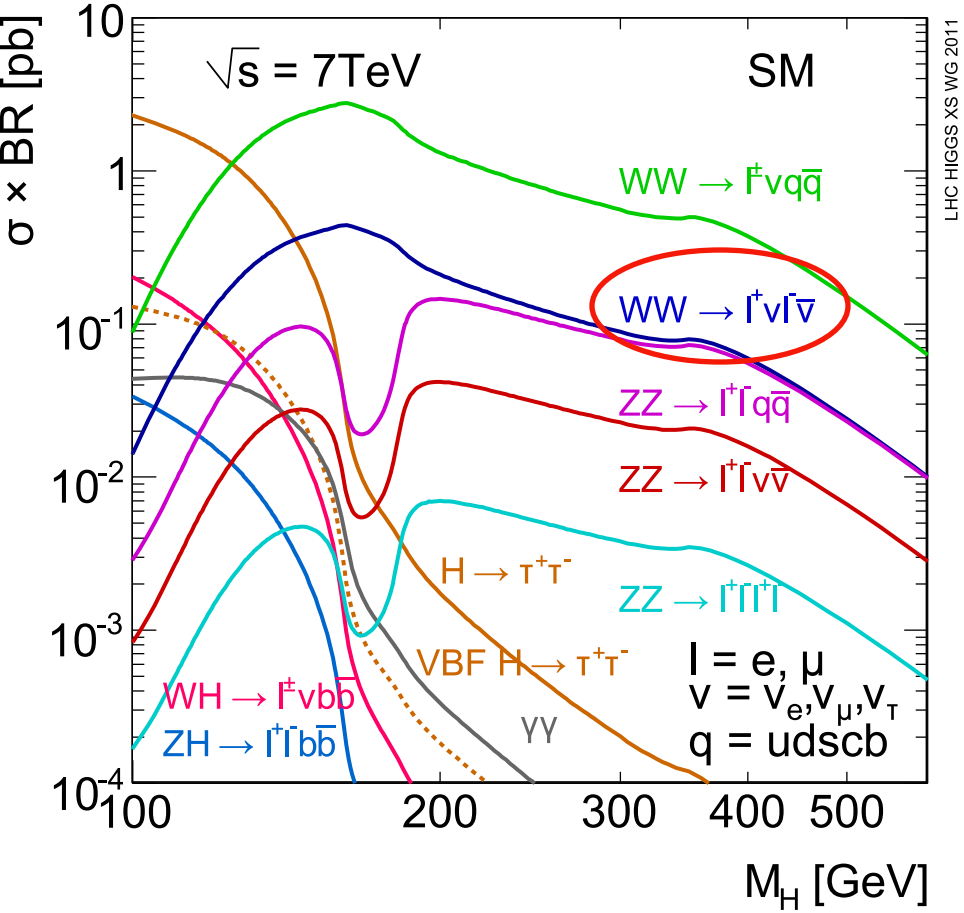
Fermilab, on behalf of the CMS collaboration

Higgs production



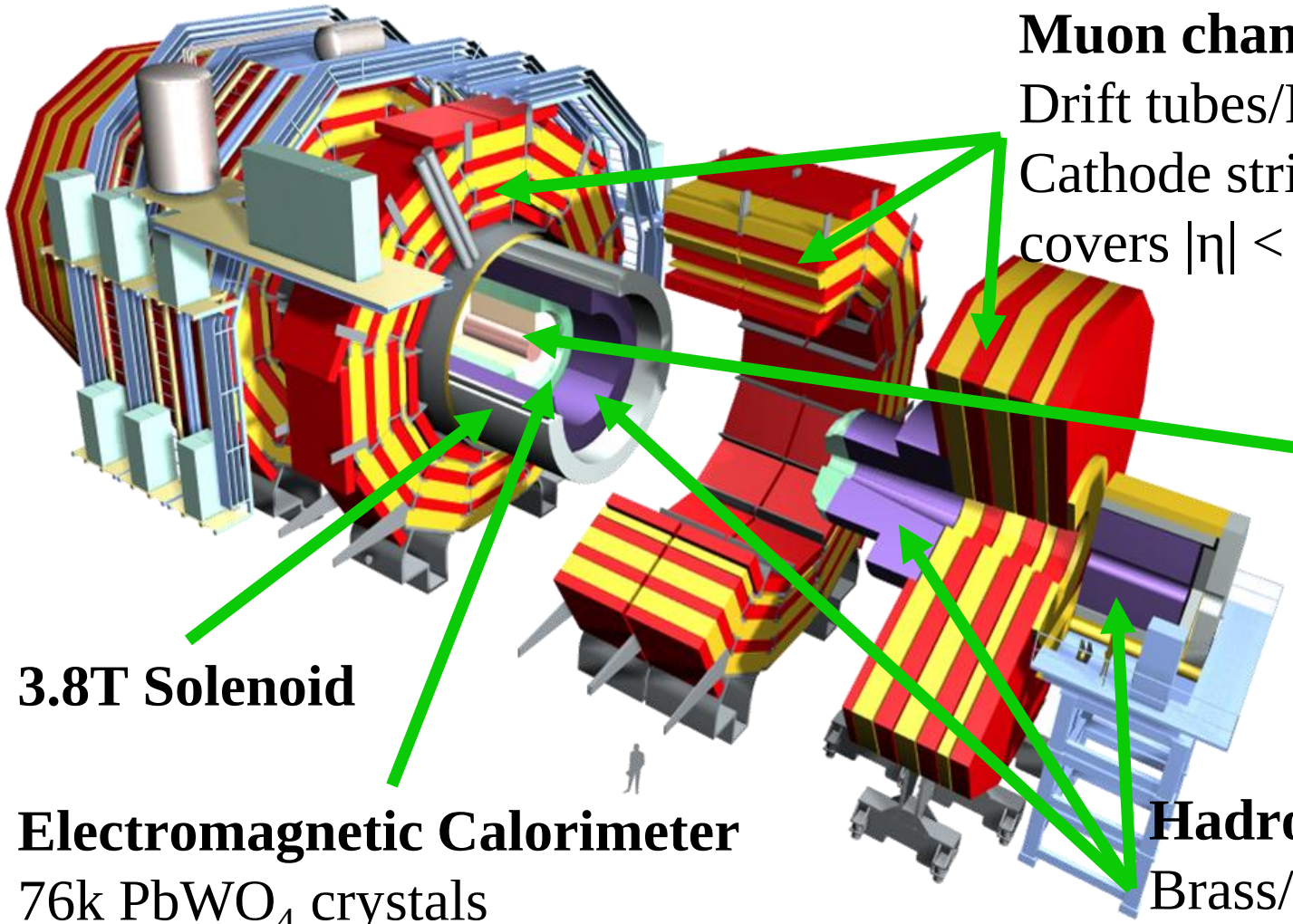


Higgs decays



- $H \rightarrow WW \rightarrow \ell\nu\ell\nu$ does the heavy lifting of any Higgs search.
  - large branching fraction all of the mass range.
  - clean, triggerable experimental signature
- Principal drawback is the lack of the definite mass peak.
- Without a strong WW search a Higgs discovery will be severely delayed.

The CMS detector



Muon chambers
Drift tubes/RPC in barrel
Cathode strip/RPC in endcaps
covers $|\eta| < 2.4$

Inner tracker
Silicon pixels
Silicon strips

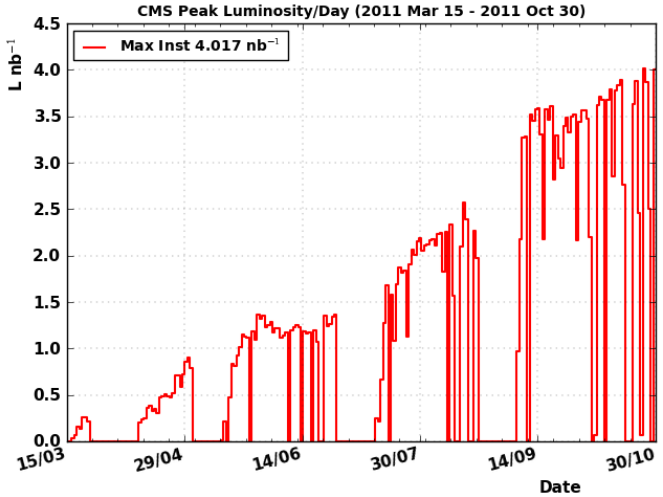
3.8T Solenoid

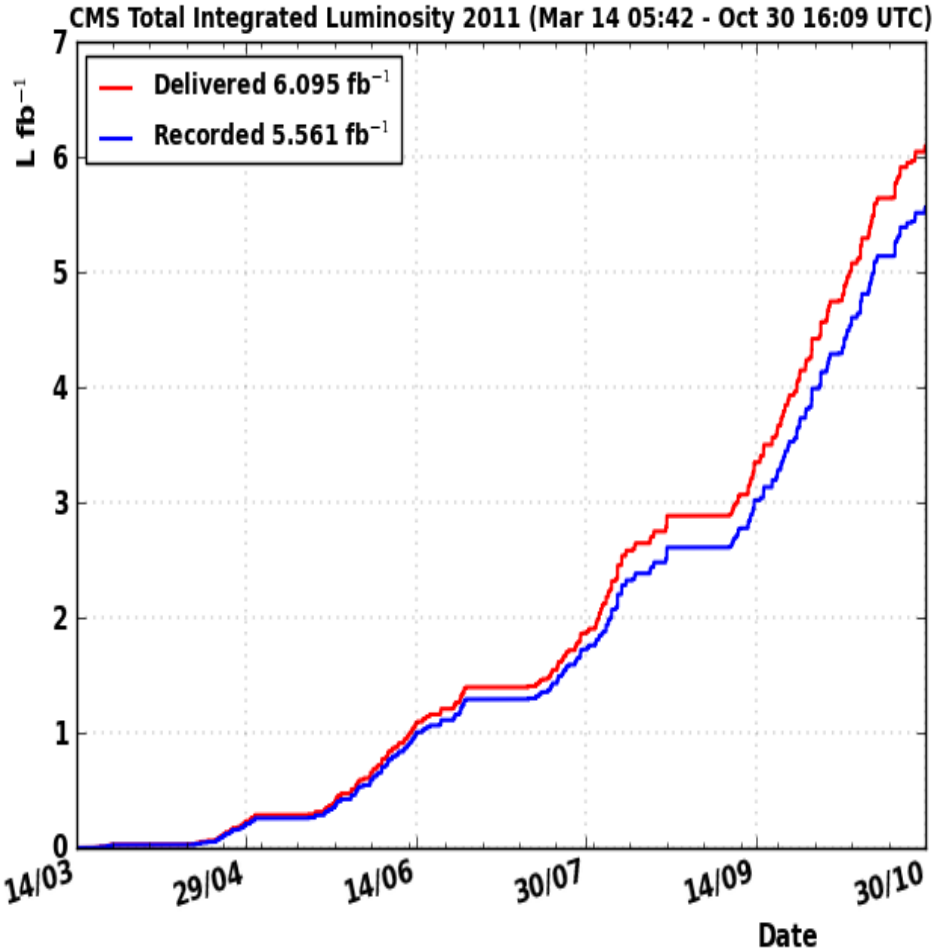
Electromagnetic Calorimeter
76k $PbWO_4$ crystals

Hadronic Calorimeter
Brass/scintillator
Iron/quartz fiber

The dataset

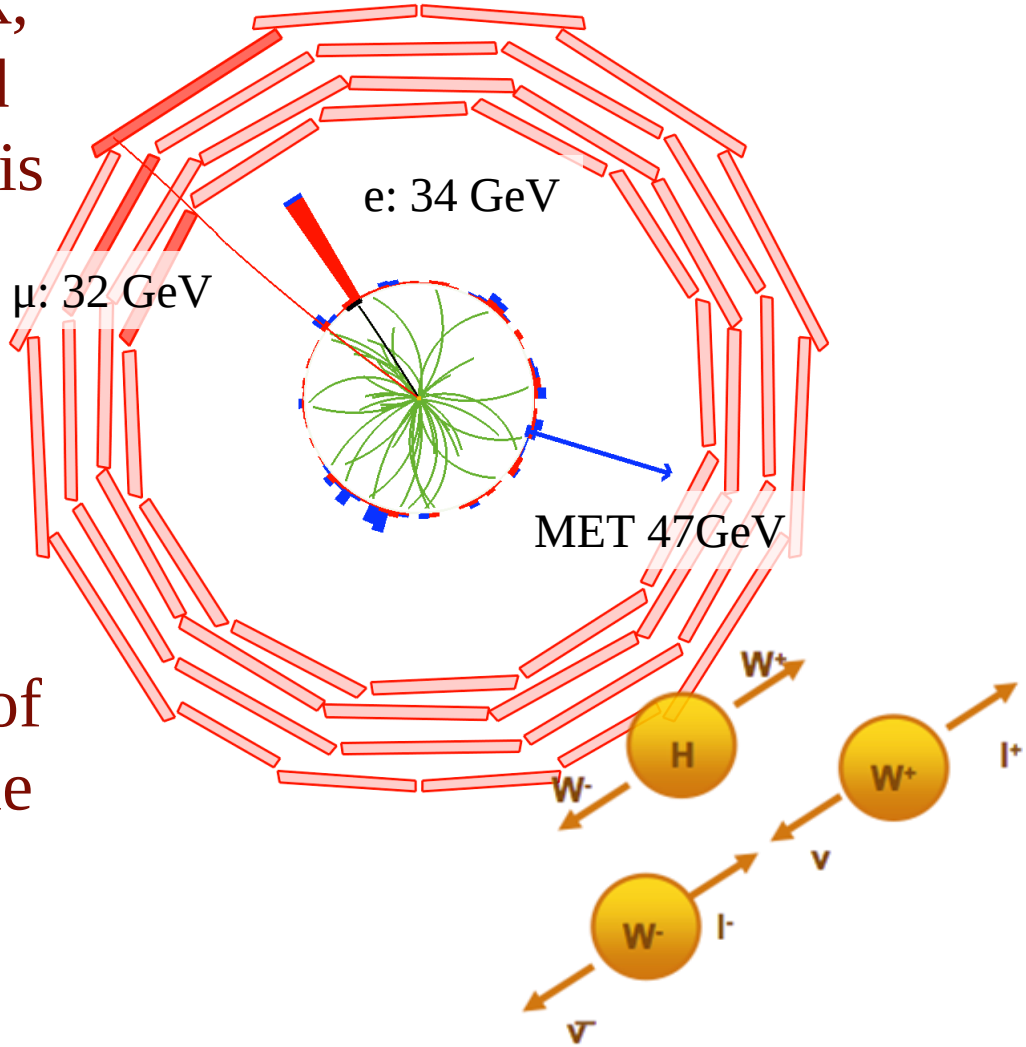
- We analyze the full dataset collected in 2011.
  - LHC delivered: 6.1 fb^{-1}
  - CMS recorded: 5.6 fb-1
- Peak instantaneous luminosity: 4.02 nb^{-1}/s





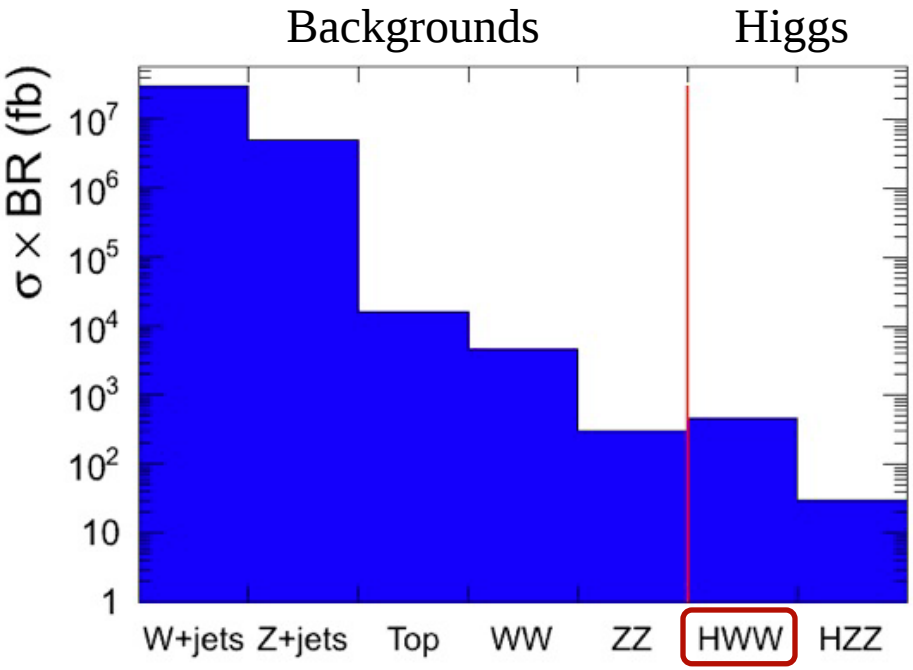
Strategy

- Without a definite mass peak, understanding, modeling and controlling the backgrounds is key to sensitivity.
- Key backgrounds:
  - V + jets
  - Top
  - EW diboson production
- Exploiting the scalar nature of the Higgs leads to best handle for irreducible non-resonant WW production.

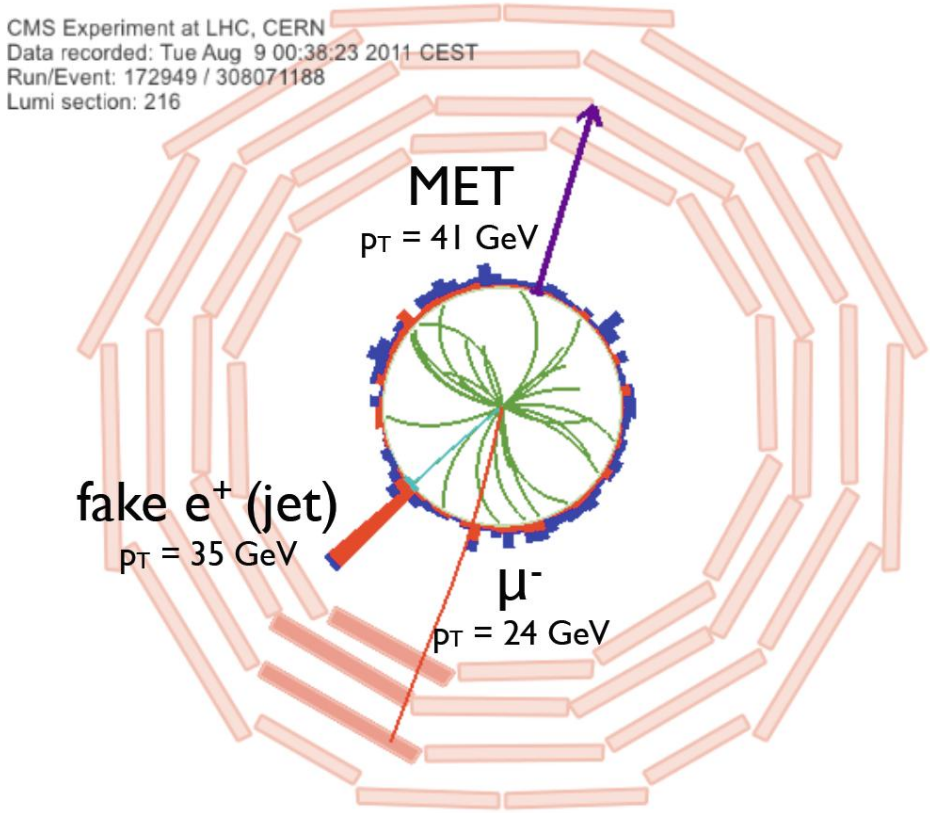


Strategy (II)

- Given the various backgrounds and their relative contributions the analysis is split into 5 separate sub-channels.
  - same-flavor/different-flavor
  - 0-jet/1-jet
  - 2 jets (VBF selections)
- The most sensitive is the different-flavor, 0-jet channel.

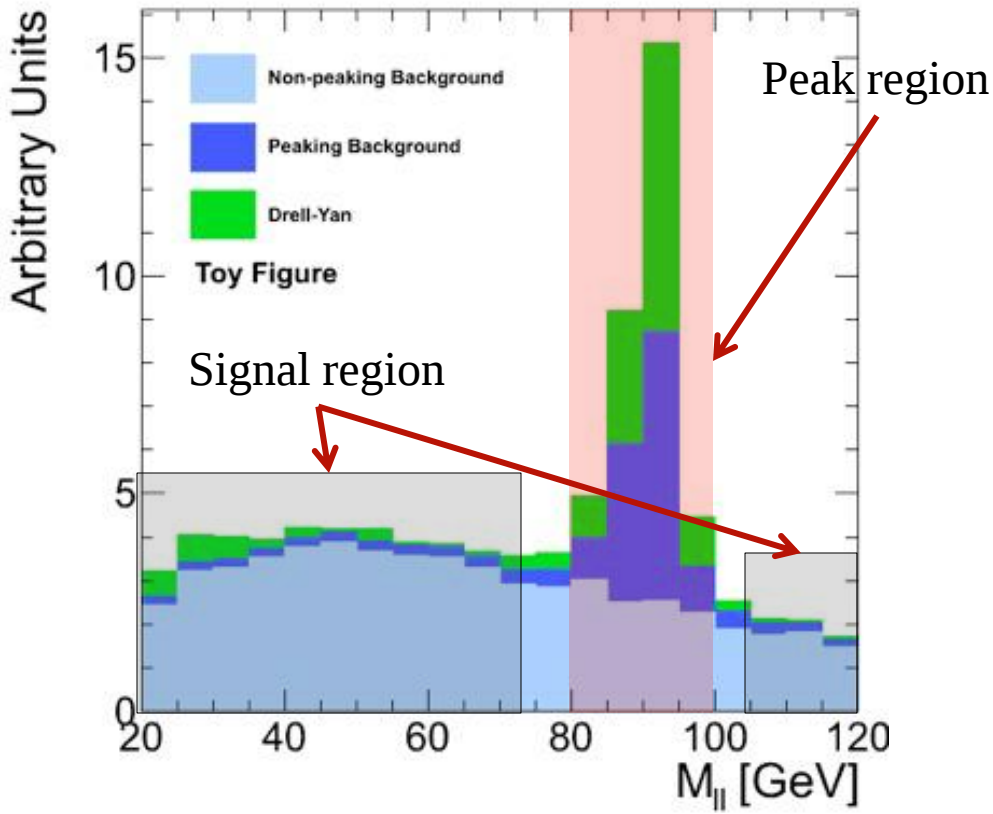


The background (W+jets)



- Tight lepton identification and isolation requirements keep this background in check.
- Normalization of the residual contribution is taken from the data.
  - Relax the lepton selection to define the control sample.
  - Use jet faking lepton rates to extrapolate to the signal sample.

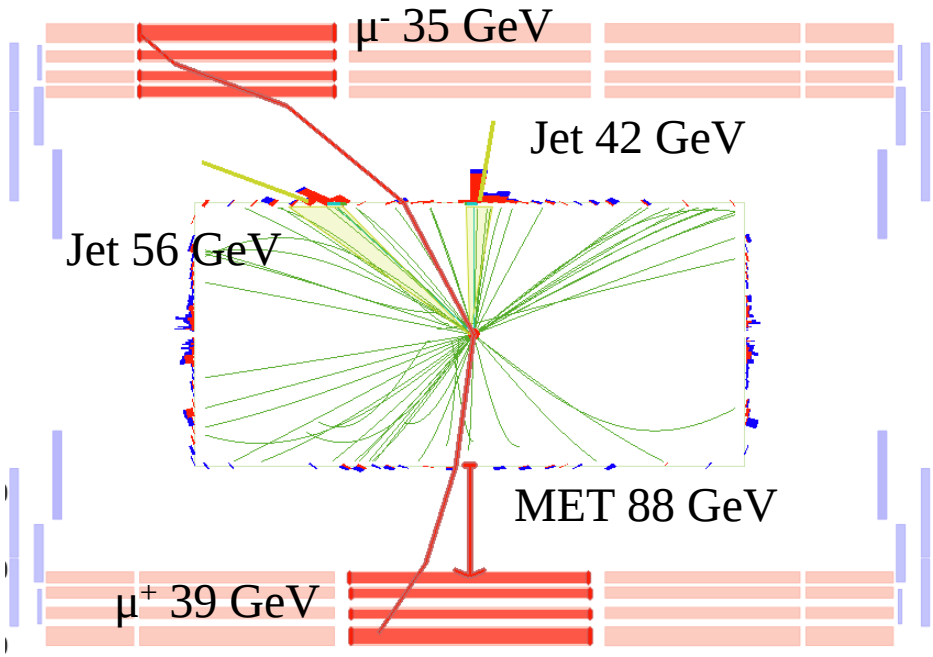
The background (Drell-Yan)



- Critical background in same flavor channels.
- Cut on relatively large amount of MET.
- Veto events with the same lepton flavor that fall near the Z peak.

- Residual background is determined from background subtracted data near the Z peak and extrapolated to the signal region using MC.

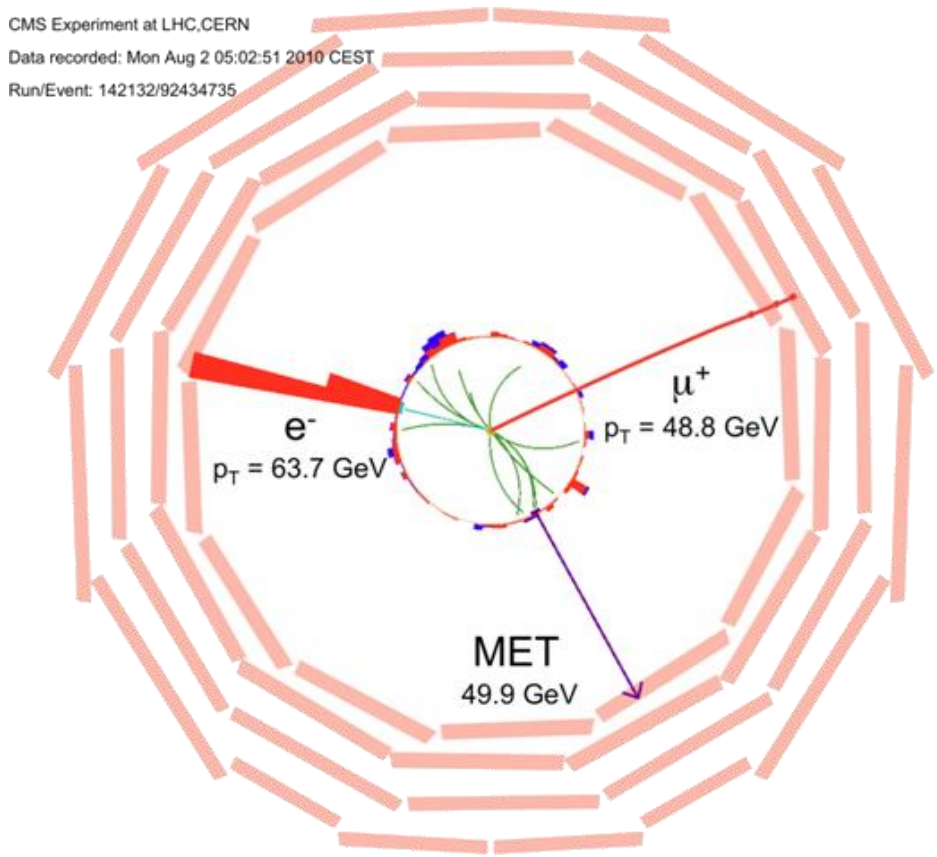
The background (top)



- Divide the data into jet bins (0, 1, and 2 extra jets).
  - 0 jet bin has much less top (ie more sensitivity)
- Veto events with b-tagged jets and soft muons (top-tagged)
- $\varepsilon_{\text{top-tag}}$ ~50% as measured with data.
- Residual background estimated as

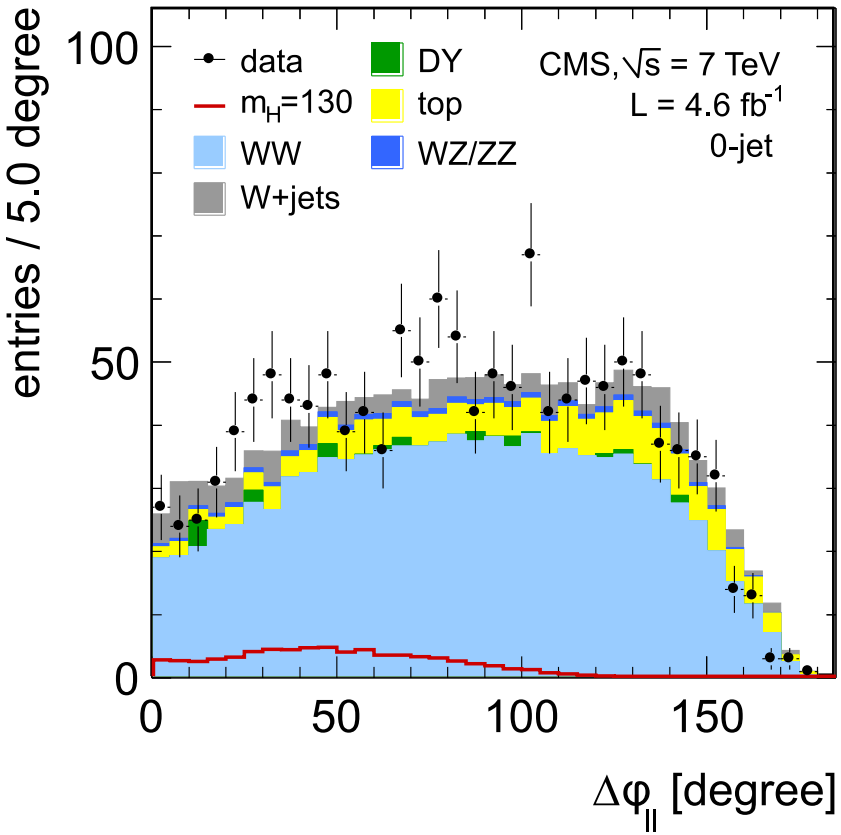
$$\frac{N_{\text{top-tag}}}{\varepsilon_{\text{top-tag}}} \times (1 - \varepsilon_{\text{top-tag}})$$

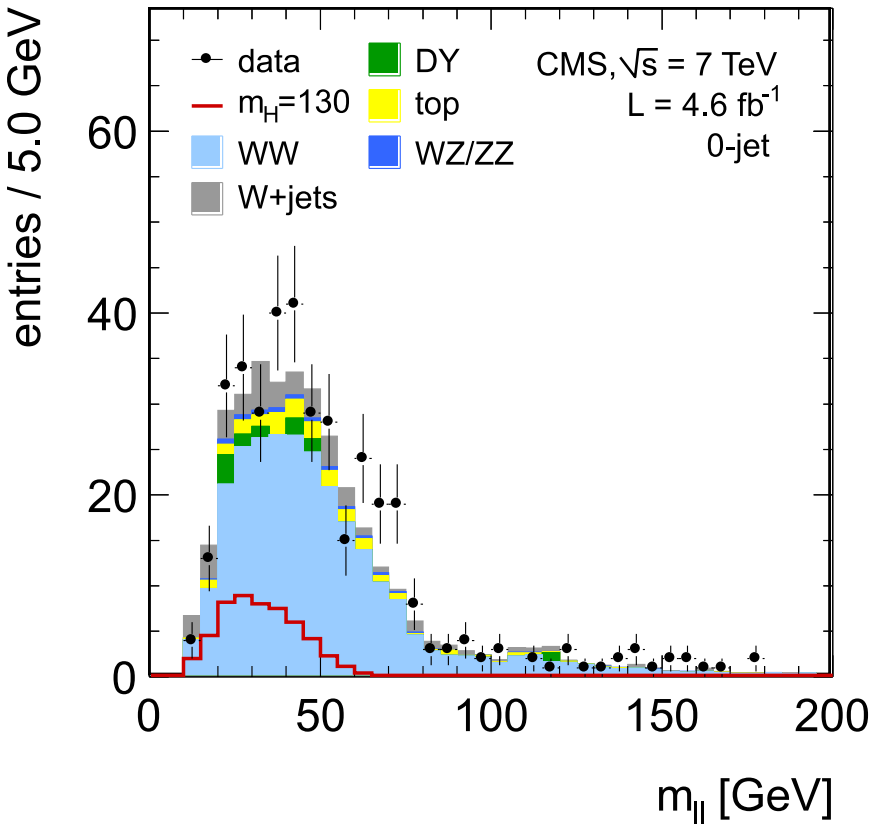
The background (WW)



- The non-resonant WW production is an irreducible background.
- We can still exploit the kinematics of the event to separate this from Higgs decays.
- The variables of choice: $\Delta\varphi_{\ell\ell}$ and $m_{\ell\ell}$.

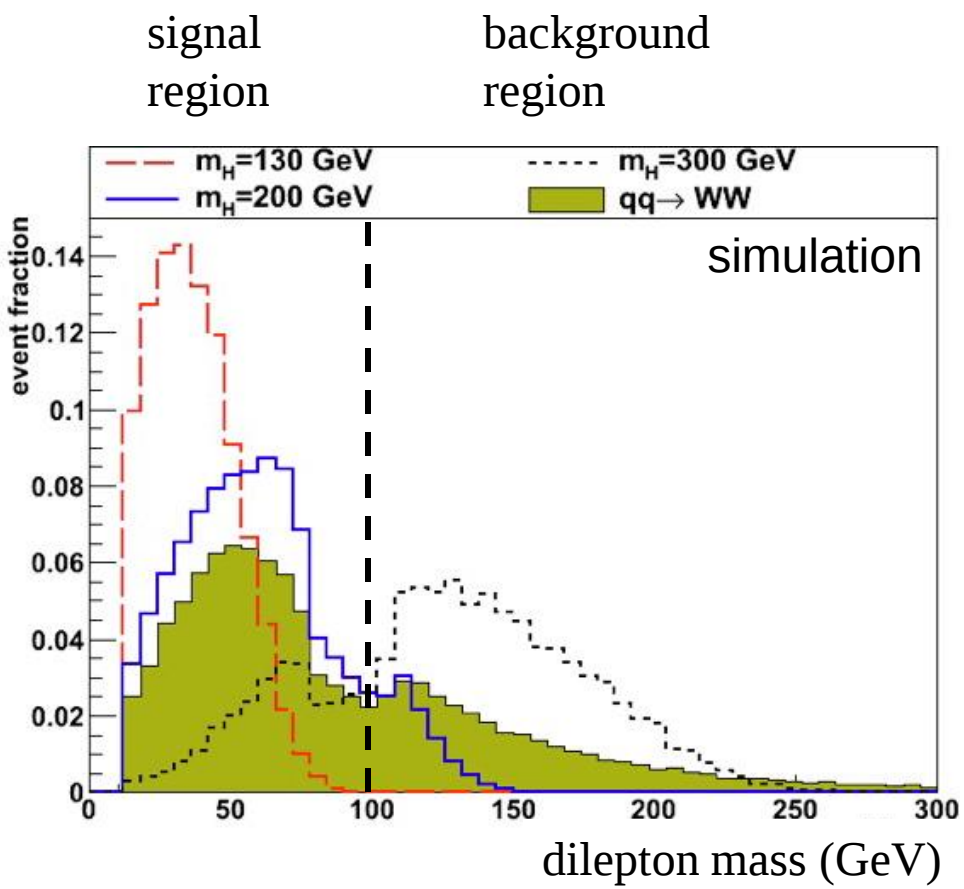
Kinematics





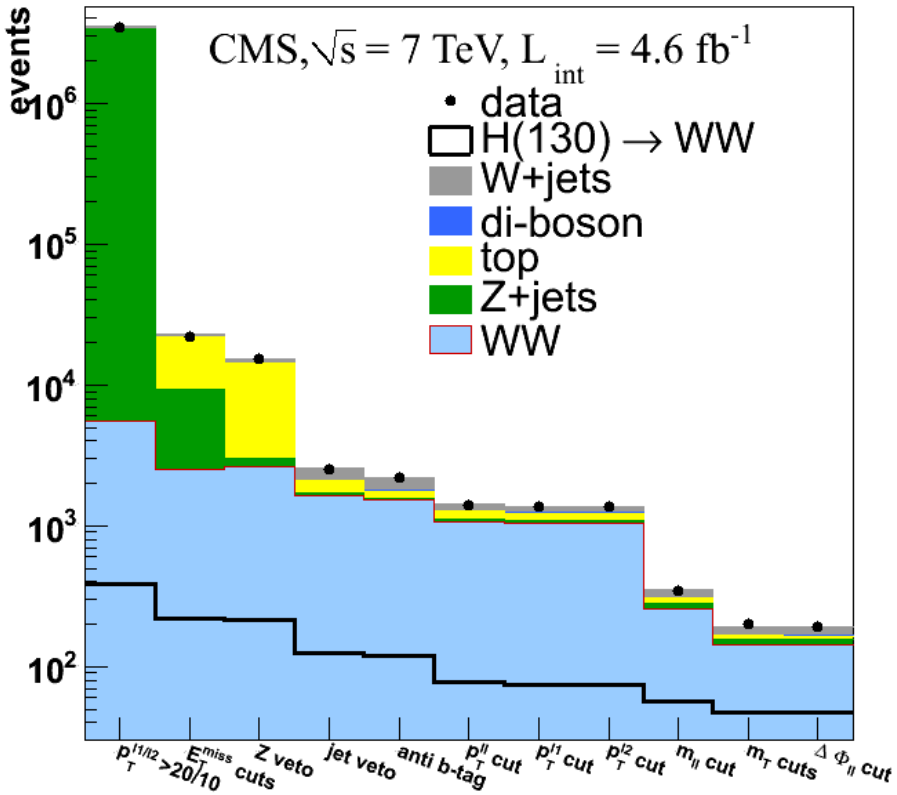
- Higgs decays have smaller opening angle for leptons.
- Higgs decays have smaller di-lepton mass.

The Di-boson background



- Normalization of the irreducible WW background is taken from data sidebands for low H mass and from simulation for larger masses.
- The remaining di-boson backgrounds are small and estimated from simulation.
- CMS has measured the electroweak WW production cross-section and it agrees very well with predictions.

Stacking up the background



- Extensive effort has been made to control and model the background components of this search.
- The agreement between the data and the MC is generally very good at every step of the analysis.
- Two complementary analysis techniques have been carried out
  - cut-based
  - multivariate based

Cut-based analysis

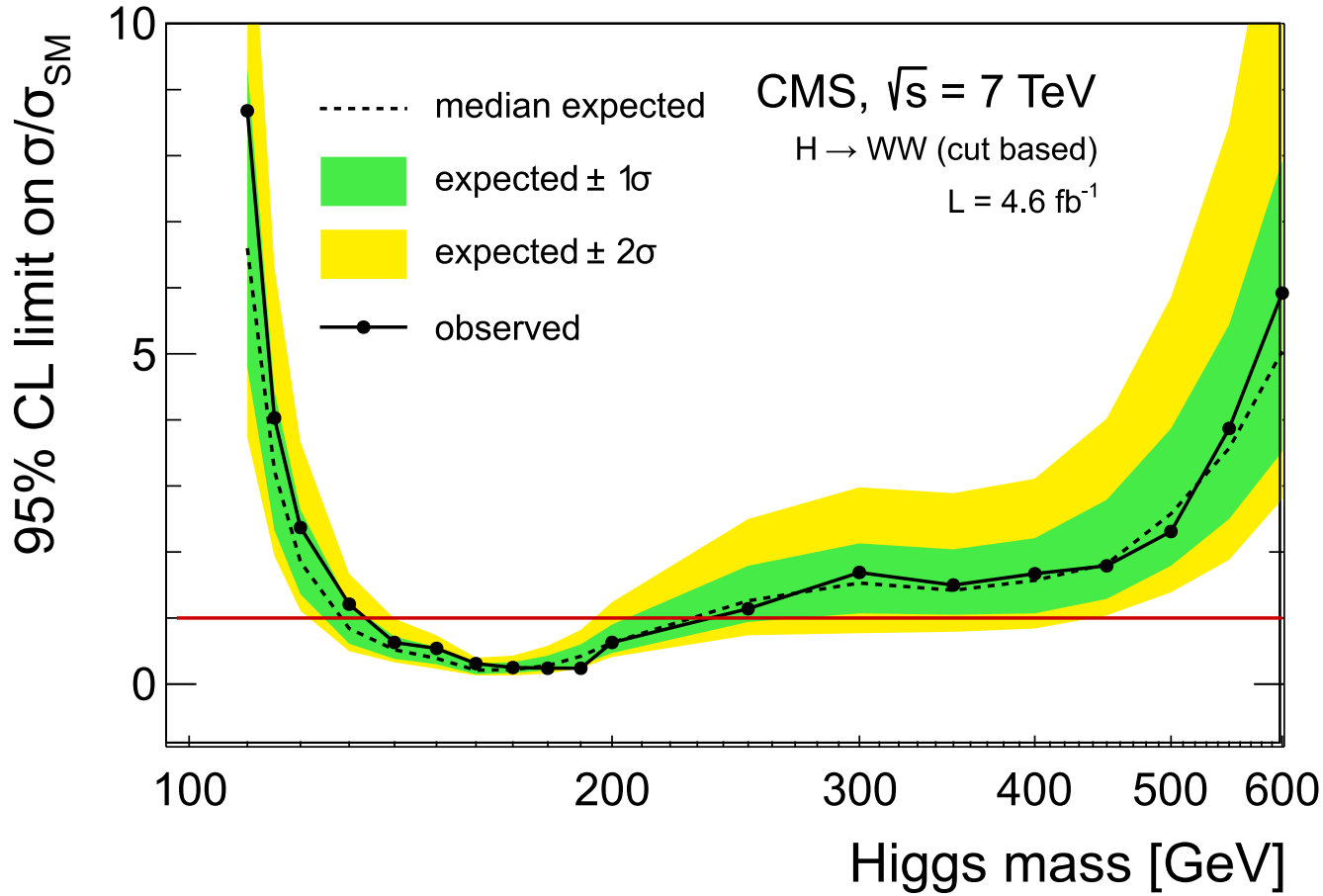
m_H	$p_T^{\ell,max}$	$p_T^{\ell,min}$	$m_{\ell\ell}$	$\Delta\phi_{\ell\ell}$	m_T
[GeV]	[GeV]	[GeV]	[GeV]	[°]	[GeV]
	>	>	<	<	[,]
120	20	10 (15)	40	115	[80,120]
130	25	10 (15)	45	90	[80,125]
160	30	25	50	60	[90,160]
200	40	25	90	100	[120,200]
250	55	25	150	140	[120,250]
300	70	25	200	175	[120,300]
400	90	25	300	175	[120,400]

- There are mass dependent cuts on the lepton p_T, $m_{\ell\ell}$, $\Delta\varphi_{\ell\ell}$ and the Higgs m_T.
- Events for the background and expected Higgs signal are counted and limits are set.

M_H	data	background	$H \rightarrow WW$
120	136	136.7 ± 12.7	15.7 ± 0.8
130	193	191.5 ± 14.0	45.2 ± 2.1
160	111	101.7 ± 6.8	122.9 ± 5.6
200	159	140.8 ± 6.8	48.8 ± 2.2
400	109	110.8 ± 5.8	17.5 ± 0.8

0 jet bin

Cut-based limit

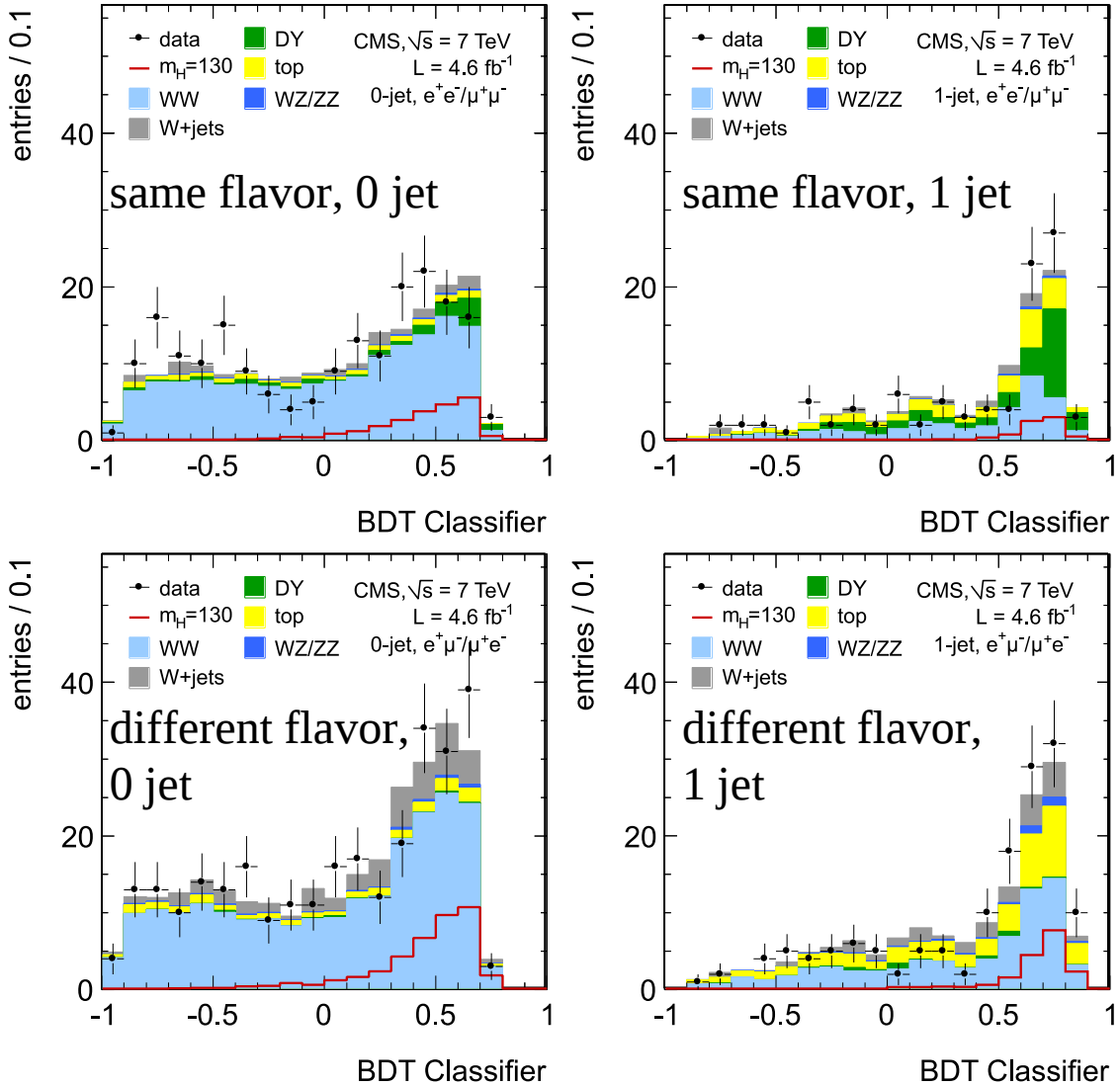


- expected exclusion: 129 GeV < m_H < 236 GeV
- observed exclusion: 132 GeV < m_H < 238 GeV

More advanced analysis

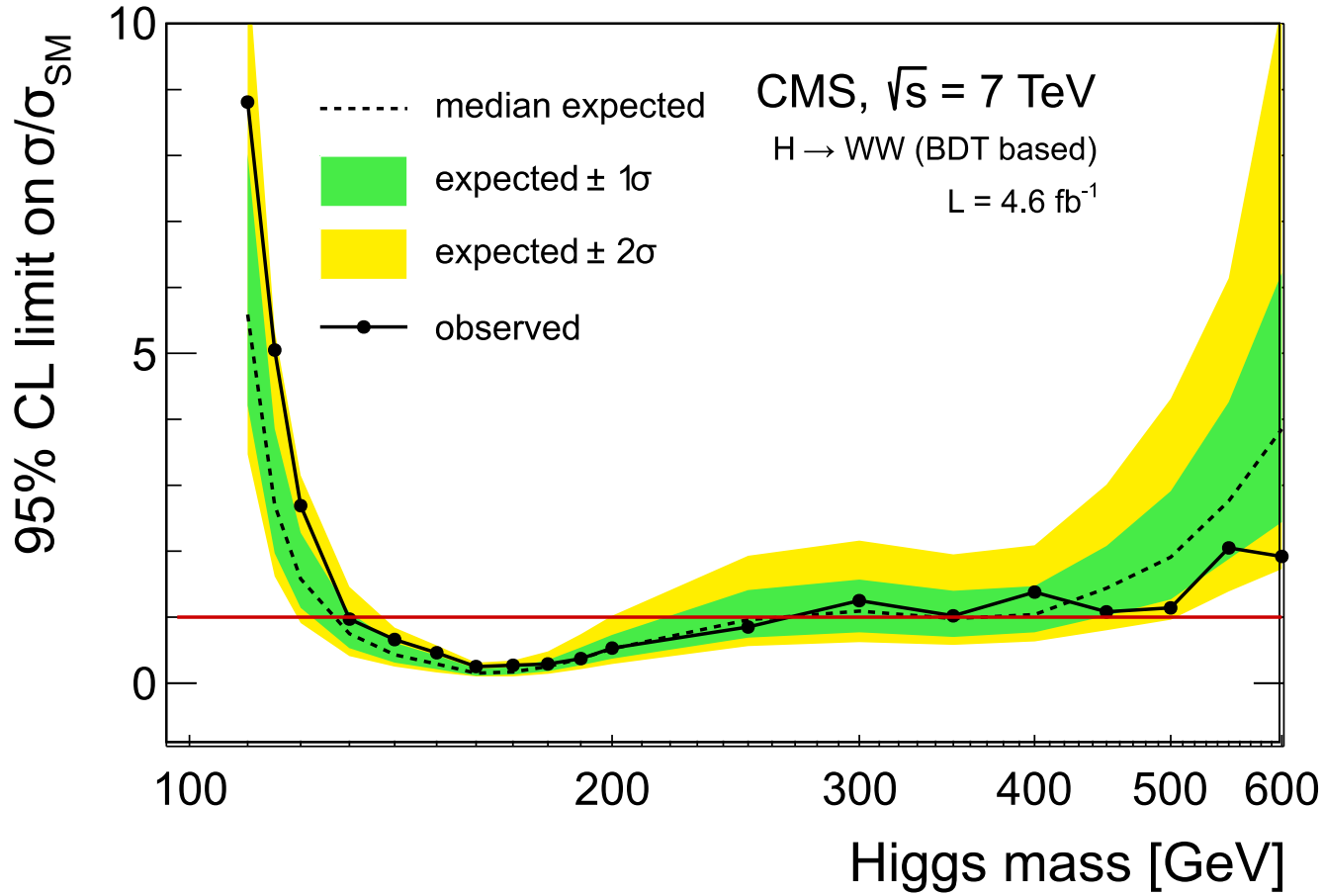
- To obtain more sensitivity, we employ a multivariate classifier based on boosted decision trees (BDT).
  - Uses kinematic observables from the cut based analysis and
    - $\Delta R_{\ell\ell} = [(\Delta\eta_{\ell\ell})^2 + (\Delta\varphi_{\ell\ell})^2]^{1/2}$
    - m_T for both ℓ-MET pairs
    - lepton flavors
  - trained to discriminate Higgs from non-resonant WW production.
  - Trained separately for each mass hypothesis and jet content bin.
  - **Shapes of the BDT output for both signal and background are used to set the final limit.**

BDT shapes



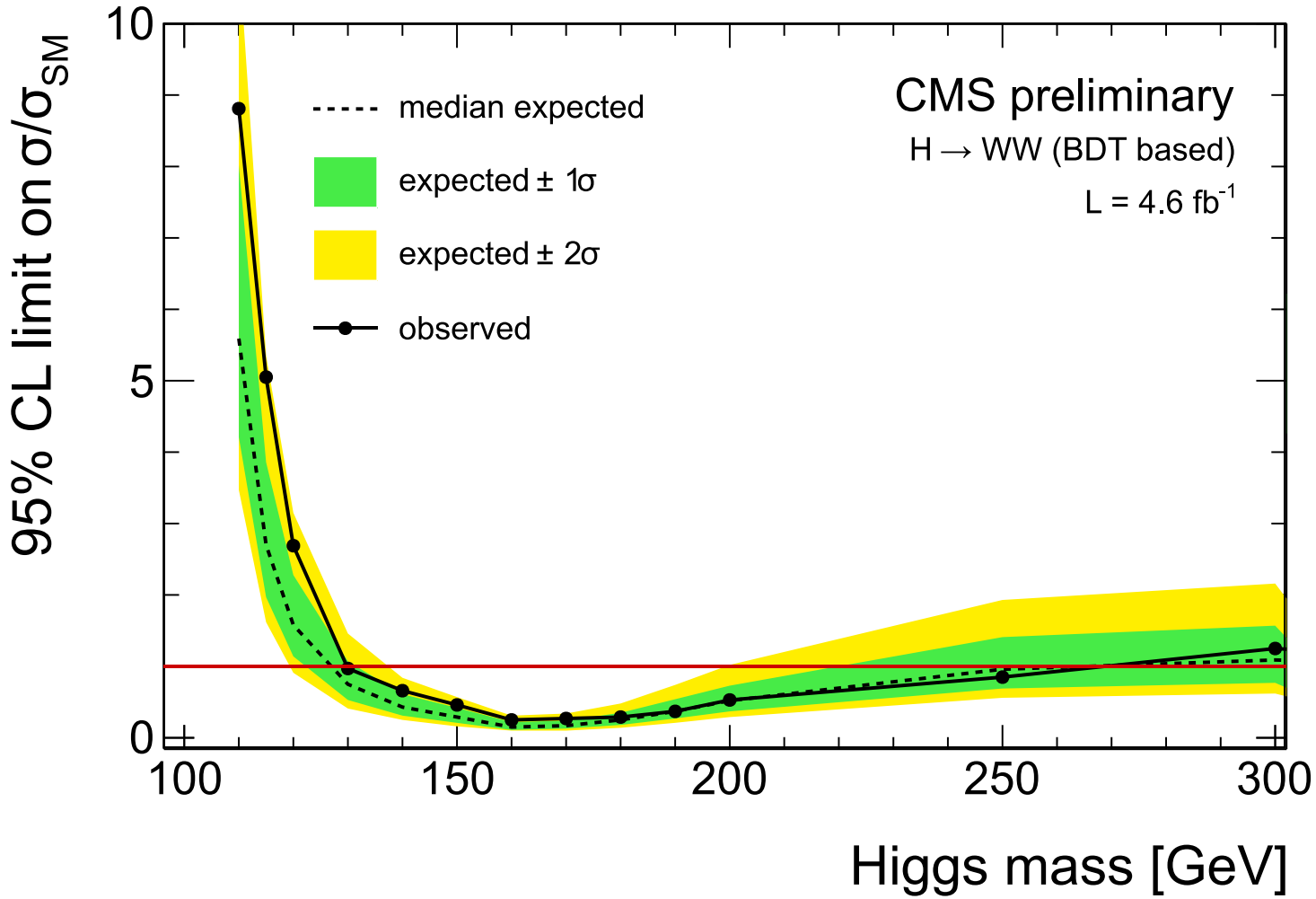
- The BDT analysis was cross-checked using both an invariant mass shape analysis and a matrix element analysis.
- Shape variations in the background components are fully studied and taken as a systematic error.

Final limit (BDT)

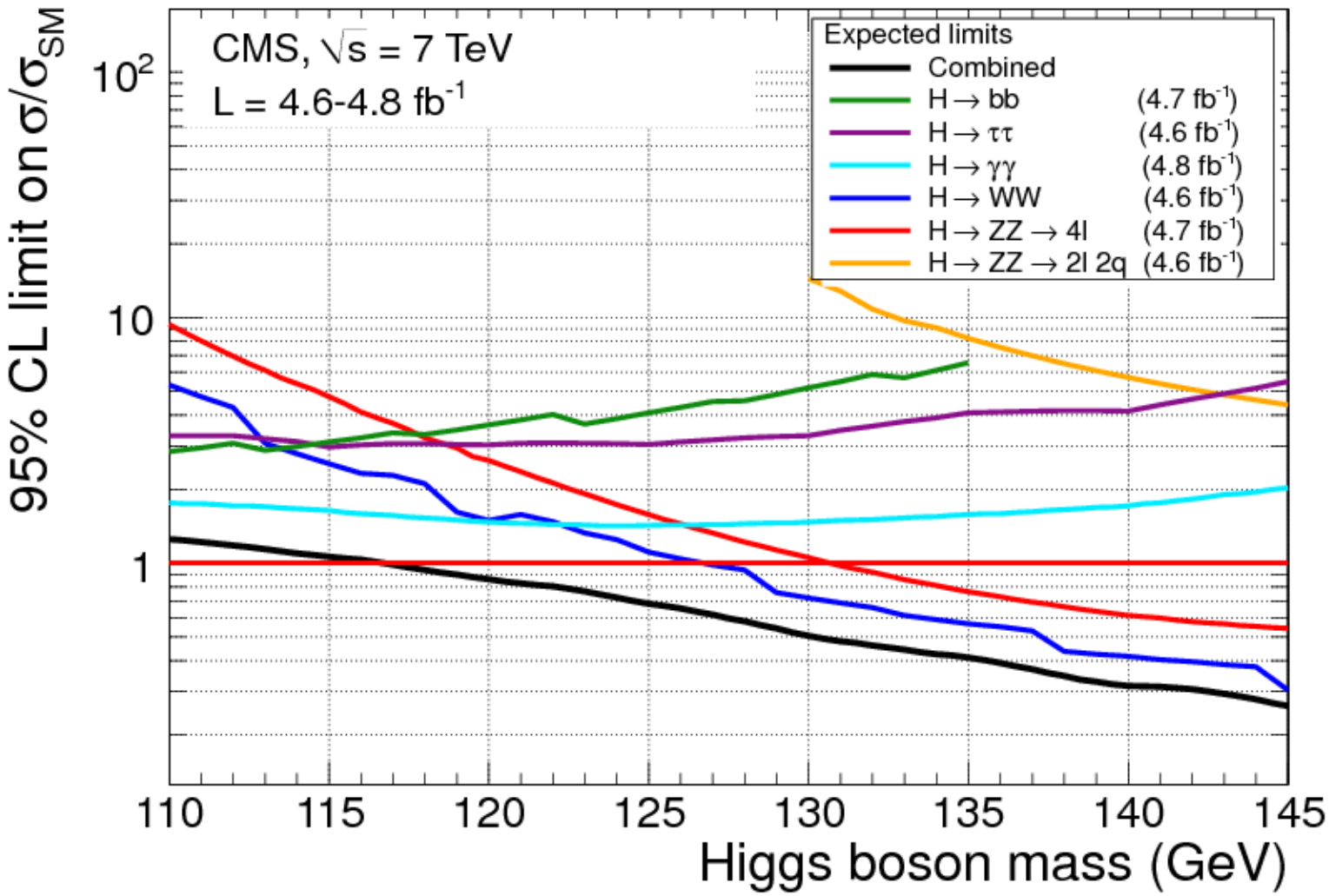


- expected exclusion: 127 GeV < m_H < 270 GeV
- observed exclusion: 129 GeV < m_H < 270 GeV

Limit (low mass view)



Contribution to the limits



Summary

- CMS has searched for the SM Higgs boson using the full dataset collected in 2011 in its leptonic WW decays, among other decay channels.
- We do not observe any significant excesses and are able to exclude the SM Higgs boson with masses from 127 GeV to 270 GeV.
- The data collected in 2012 will expand the reach of this decay mode.
- CMS also intends to include semi-leptonic $H \rightarrow WW$ to improve the sensitivity particularly at higher mass.

Backup

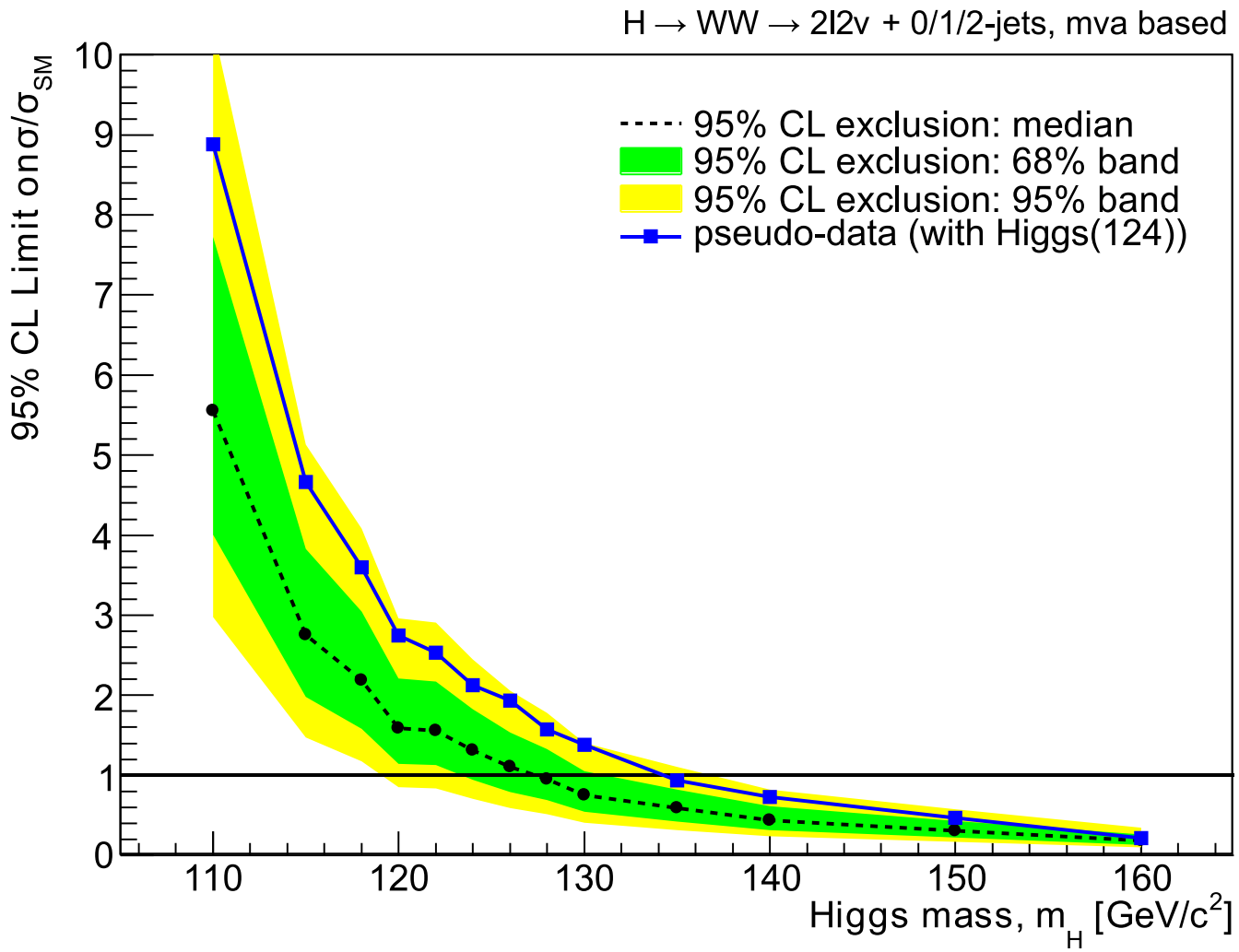
References

- https://twiki.cern.ch/twiki/bin/view/CMSPublic/Hig11024TWiki
- http://arxiv.org/pdf/1202.1489.pdf
- $WH \rightarrow WWW \rightarrow 3\ell 3\nu$
  - https://twiki.cern.ch/twiki/bin/view/CMSPublic/Hig11034TWiki
  - http://cdsweb.cern.ch/record/1429927?ln=en

Errors

- Errors on backgrounds
  - Rate and shape systematic errors on background processes via changing generators, parameters and smearing.
  - Normalization of most backgrounds taken from data control regions.
    - Errors based on both the control region and the extrapolation.
  - Effect of pileup on distributions are also included.
  - Total uncertainty on background is ~15% and dominated by statistics of the control regions.
- Errors on signal
  - Signal efficiencies taken from MC.
  - Errors on signal efficiencies fall into two categories
    - effects of signal migration between jet multiplicity bins – evaluated by comparing different generators.
    - effects of PDFs, α_s, higher-order corrections, etc. – evaluated using the PDF4LHC prescriptions.
  - Overall signal efficiency uncertainty ~20%.

Signal injection at 124 GeV



2-jet VBF selection

- 2 jets $p_T > 30$ GeV
- no jets with $p_T > 30$ GeV in the η region between the two jets.
- $|\Delta\eta_{jj}| > 3.5$
- $m_{jj} > 450$ GeV

Fermiophobic Higgs