A Search for the Higgs Boson in $H \to WW$

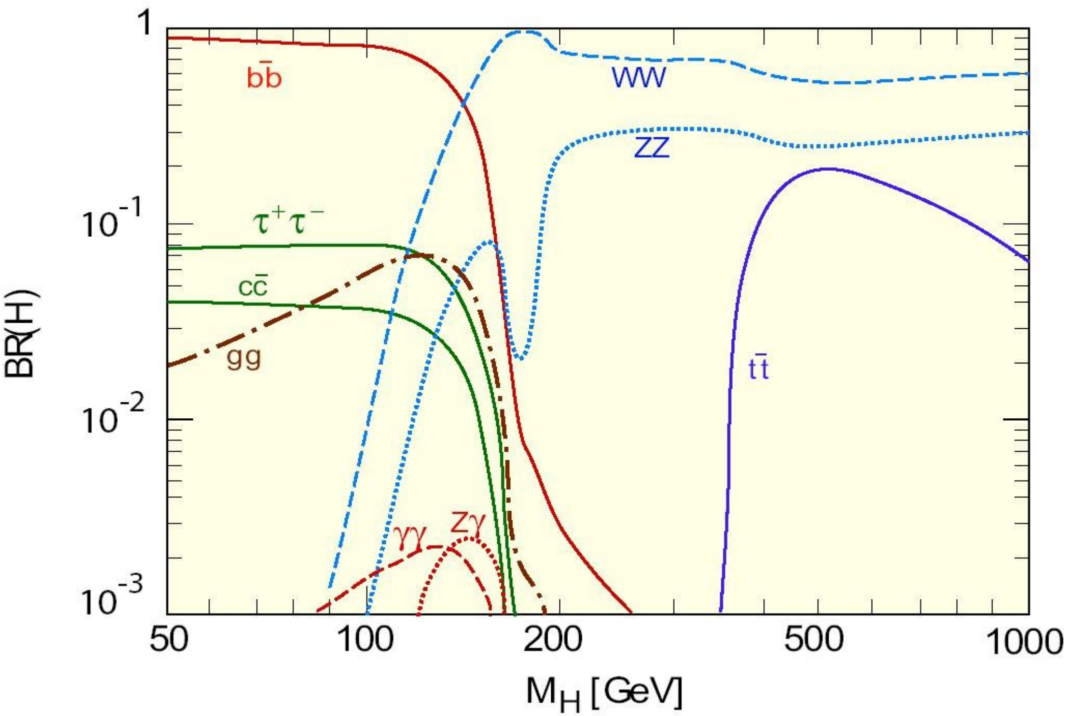
Kevin Sung (MIT)

CMS Collaboration

Meeting of the Division of Particles and Fields
of the American Physical Society, 2011

Introduction

- *H→WW*
  - High sensitivity to the presence of Higgs over a broad mass range
- 2010 CMS results:
  - *H→WW→lνl'ν'* with *36 pb⁻¹*
  - Set limits on models with four fermion generations
  - Phys. Lett. B699 (2011) 25-47, ePrint: arXiv:1102.5429 [hep-ex]



Today: preliminary 2011 results from CMS with *1.1 fb⁻¹*

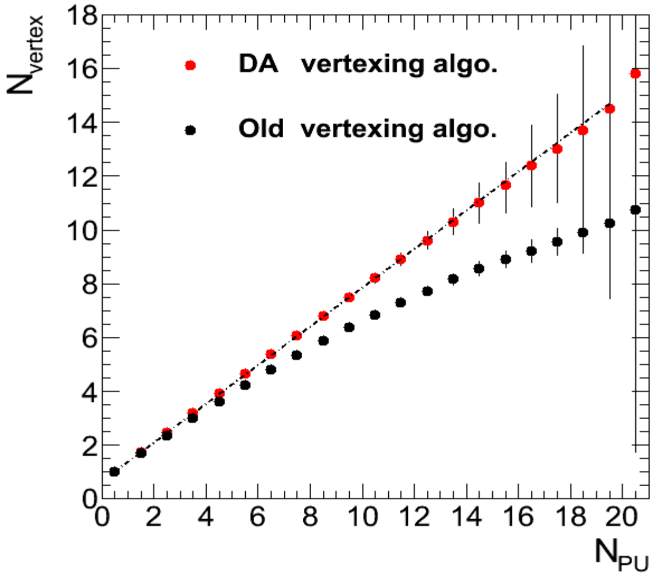
– CMS PAS HIG-11-003

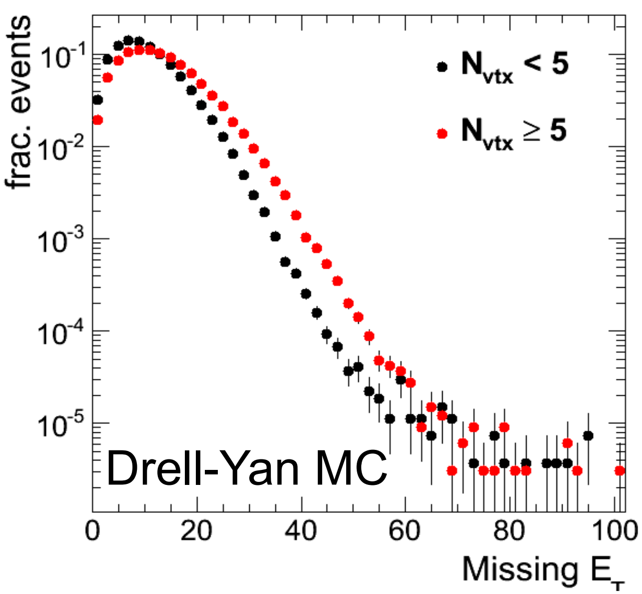
Analysis Outline

- WW Event Selection
- Categorize by 0-jet, 1-jet, 2-jets (VBF)
  - Count jets with $p_T > 30\ GeV$
- Background Estimation
  - Data driven: Drell-Yan (ee, $\mu\mu$), W+jets, Top, WW
  - MC-based: Drell-Yan ($\tau\tau$), $W\gamma$, WZ, ZZ
- Higgs Signal Extraction
  - Cut-Based Analysis
    - Additional cuts on p_T, $\Delta\varphi_{ll}$, m_{ll}, m_T depending on m_H hypothesis
  - Multivariate Analysis (MVA)

Pile-Up

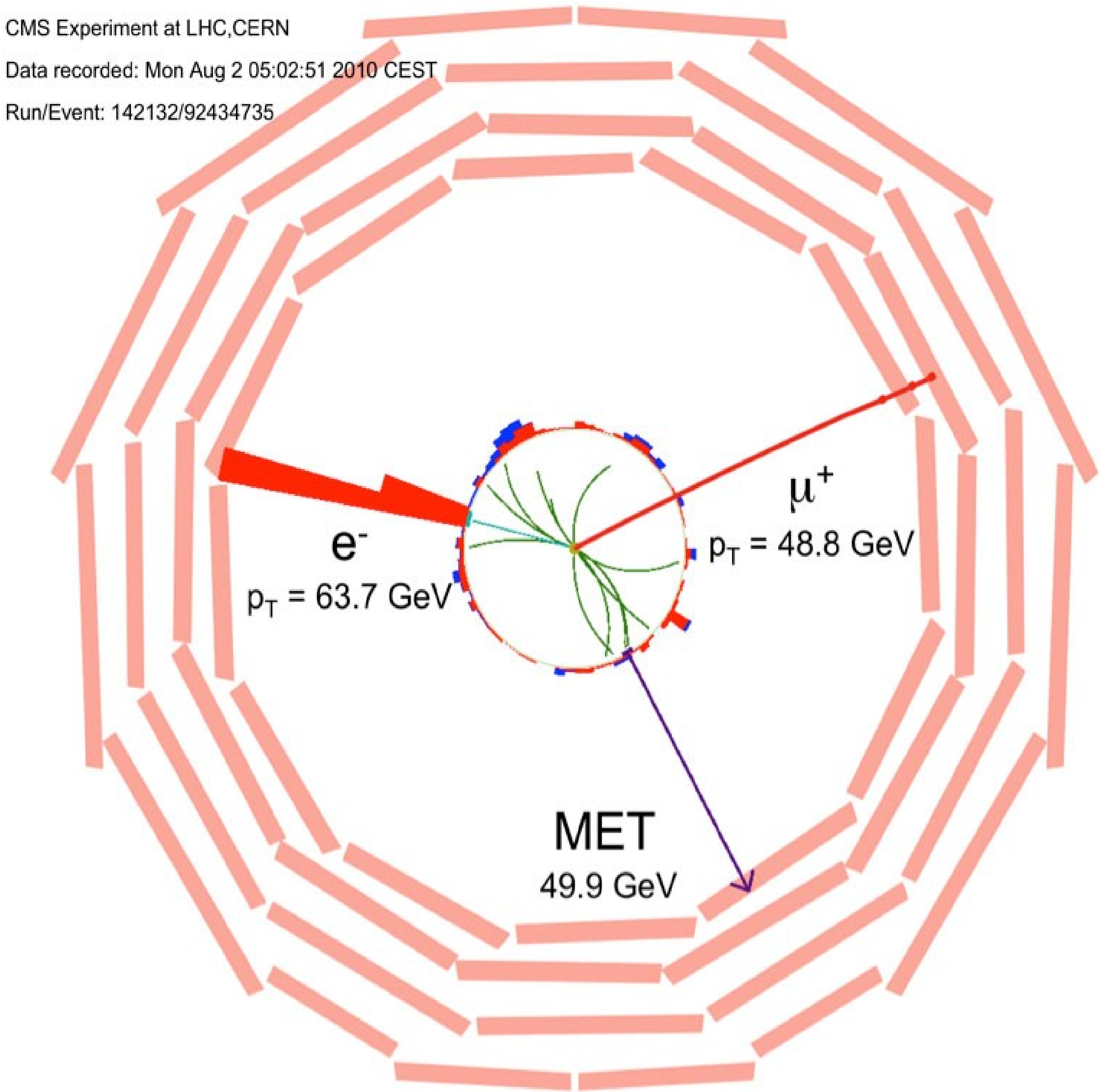
- LHC is a busy environment:
  - Large number of bunches and protons per bunch results in multiple pp interactions per crossing: pile-up
  - Number of interactions is Poissonian: with mean of ~2 interactions in 2010, ~5 interactions in 2011
- Challenges for analyses:
  - Multiple interaction merged into one vertex
  - Dependence of isolation quantities on number of pile-up interactions
  - Larger MET tails





Understanding/mitigating the effects of pile-up is a major aspect of the $H\to WW\to l\nu l'\nu'$ analysis

WW Event Selection



- Acquire events with double and single lepton triggers: $\sim 99\%$ efficiency w.r.t. selection
- Exactly two well identified, isolated leptons ($ee, \mu\mu, e\mu$)
  - $p_T^{leading} > 20\ GeV/c$
  - $p_T^{trailing} > 10\ GeV/c$
- Large missing energy
- Reject events with b-decays identified to reduce top bkgd
- ee, $\mu\mu$ not consistent with Z decay: $|m_{ll} - m_Z| > 15\ GeV/c^2$
- Reject events when ee, $\mu\mu$ is back-to-back with leading jet to reduce Z/γ^*+jets

Top and W+jets Backgrounds

- Similar approach to estimate these 2 backgrounds:
  - Given the efficiency of a cut and the number of events that failed this cut, we can deduce the number of events that passed
    - Measure (in)efficiency in background enriched sample
    - Extrapolate to signal region from sample of rejected events

- Top $\rightarrow$ failure to identify b-decays
  - In a top-enriched sample, measure the b-tagging efficiency
  - Extrapolate from a sample of b-tagged events
- W+jets $\rightarrow$ mis-identify a jet to be a good lepton
  - In a QCD-enriched sample, measure the efficiency of loosely identified leptons to pass full selection
  - Extrapolate from a sample where one lepton fails selection

Drell-Yan Background

- Residual Drell-Yan contamination from large “fake” MET:
  - Pile-up
  - Lepton mis-measurement
  - Hadronic recoil mis-measurement
- Use data to normalize Drell-Yan MC

Assume mass shape is well described in MC and consider the ratio of yield **out** of the Z window (i.e. signal region) to yield **in** the Z window:

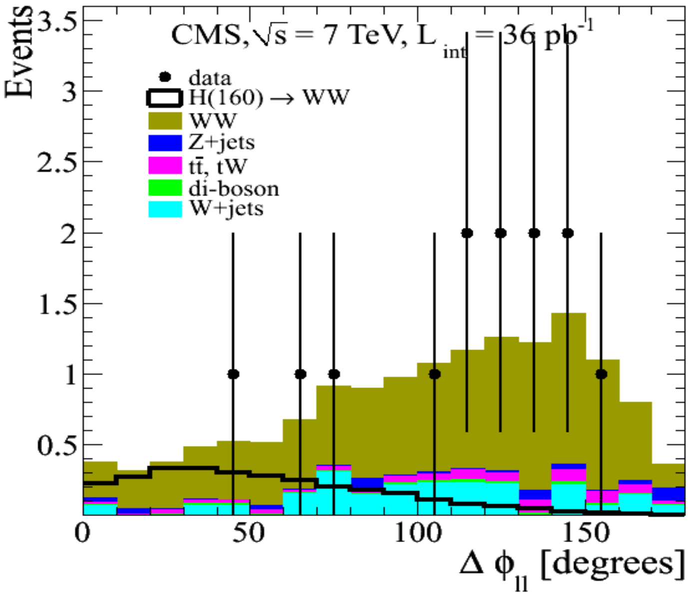
$$R^{\ell\ell}_{\mathrm{out/in}} = N^{\ell\ell,\mathrm{loose}}_{\mathrm{out}} / N^{\ell\ell,\mathrm{loose}}_{\mathrm{in}}$$

← Looser cuts than full selection to gain events

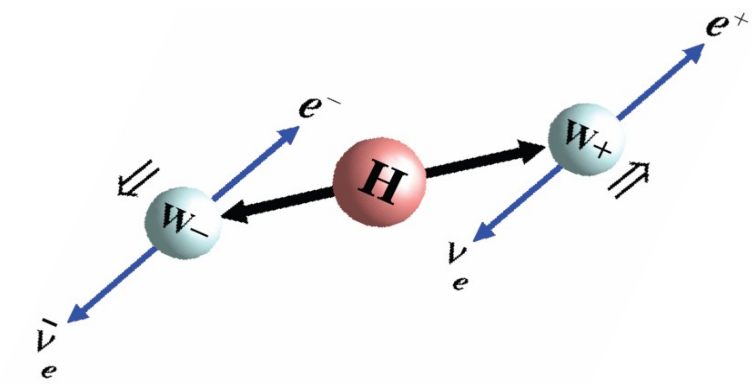
Estimate background from Drell-Yan yield **in** the Z window from data after all m_H-dependent cuts:

$$N^{\ell\ell,\mathrm{exp}}_{\mathrm{out}} = R^{\ell\ell}_{\mathrm{out/in}}(N^{\ell\ell}_{\mathrm{in}} - N^{\mathrm{non-Z}}_{\mathrm{in}} - N^{ZV}_{\mathrm{in}})$$

WW Background

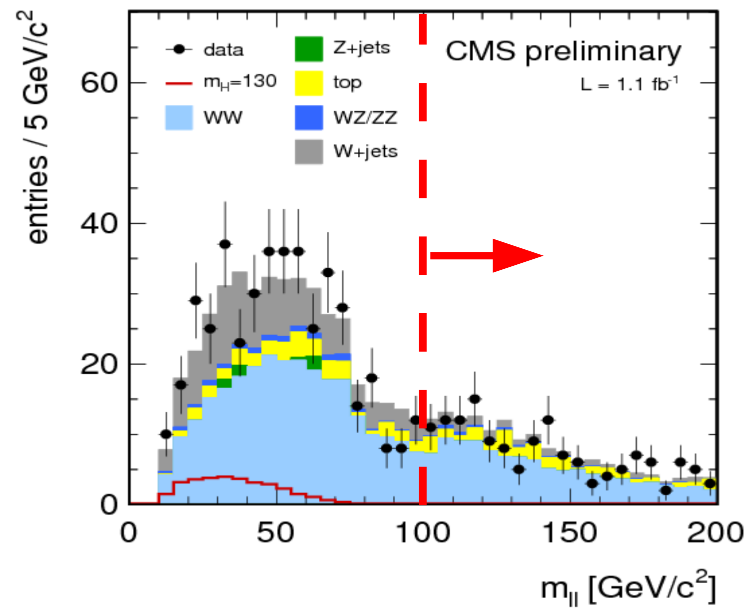


Two W's from decay of scalar Higgs must have opposite spin → favors final state leptons to be more collinear



Exploit spin correlation to separate $H\to WW$ signal from WW background: require small $\Delta\varphi_{ll}$

- Background estimation method:
  - For $m_H < 200\ GeV$, the dilepton mass region above $100\ GeV$ is relatively signal-free: *use data to normalize WW MC*
  - For higher m_H, the overlap is too large; *rely on MC*



WW Event Yields for $1.1\ fb^{-1}$

- Yields after *WW* event selection before Higgs signal extraction
  - Data yields in agreement with predicted background yields
- Dominant contributions:
  - 0-jet: *WW*, *W*+jets
  - 1-jet: Top, *WW*
  - 2-jet: Top

	0-jet	1-jet	2-jet
Top	63.8 ± 15.9	141.1 ± 14.1	99.3 ± 9.9
W+jets	106.9 ± 38.9	36.9 ± 13.8	16.4 ± 6.4
Z/γ^*	13.8 ± 5.3	21.1 ± 11.5	23.1 ± 13.5
WW	366.9 ± 30.3	107.3 ± 9.3	23.2 ± 2.0
WZ, ZZ	8.5 ± 0.9	7.2 ± 0.8	1.5 ± 0.2
$W+\gamma$	8.7 ± 1.7	2.4 ± 0.8	1.1 ± 0.5
All Bkg	568.6 ± 52.2	316.0 ± 24.7	164.6 ± 18.0
$m_H=130$	36.5 ± 0.3	13.8 ± 0.2	4.8 ± 0.1
data	626	334	175

Cut-Based Analysis

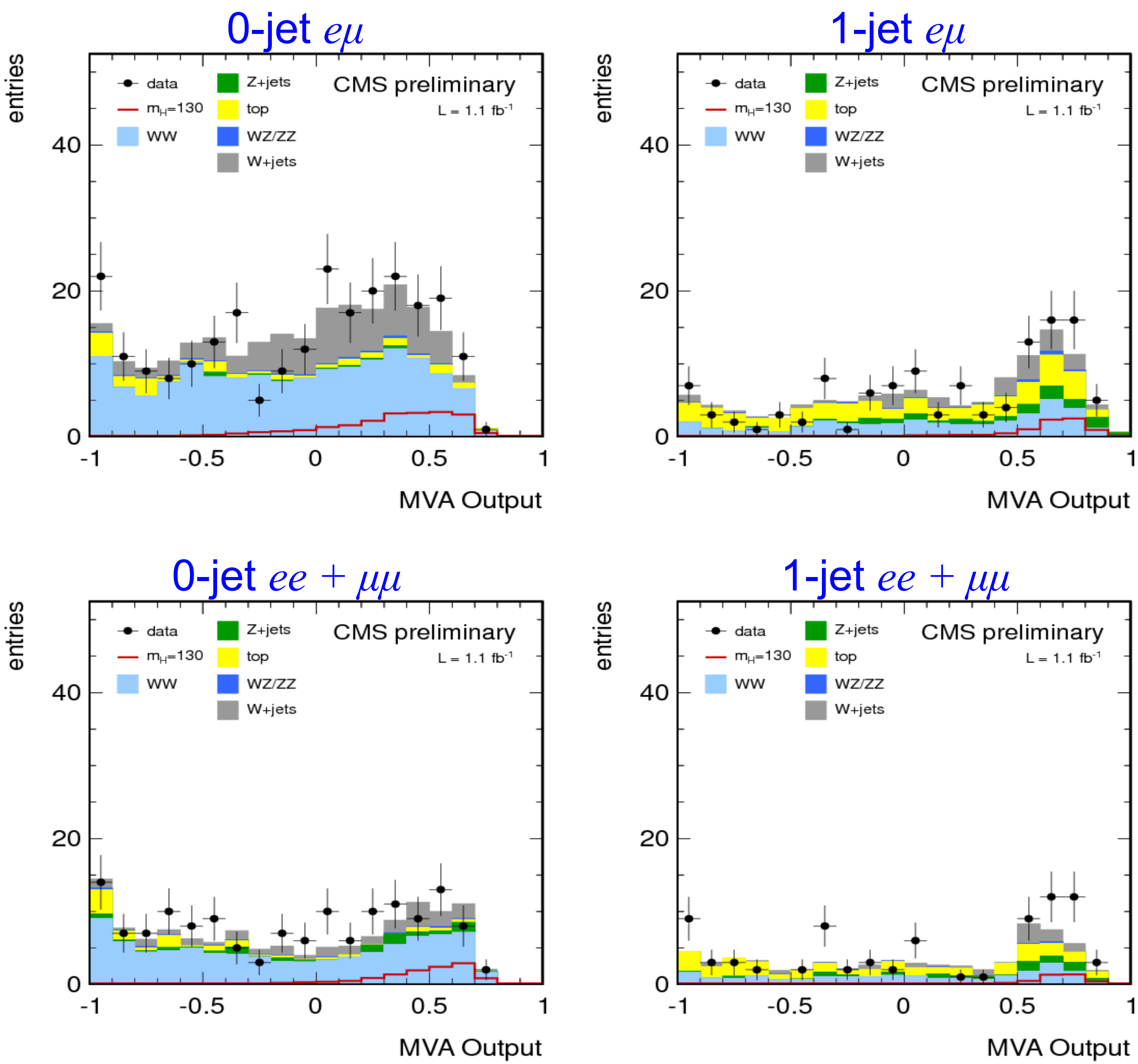
- Additional cuts that vary with m_H hypotheses to optimize for signal
- Mass points: m_H = 115, 120, 130, 140, ..., 200, 250, 300, ..., 600
- In the 0-jet and 1-jet categories:
  - Increase minimum lepton p_T thresholds with higher m_H
  - Increase maximum m_{ll} with higher m_H
  - Adjust maximum $\Delta\phi_{ll}$ depending on expected boost of W's
  - Select events in a window of transverse mass:

$$m_T^{\ell\ell E_{\mathrm{T}}^{\mathrm{miss}}} = \sqrt{2p_{\mathrm{T}}^{ll} E_{\mathrm{T}}^{\mathrm{miss}}(1 - cos(\Delta\phi_{\ell\ell - E_{\mathrm{T}}^{\mathrm{miss}}}))}$$

- In 2-jet category:
  - Increase maximum m_{ll} and with higher m_H
  - Requirements on the two jets to select VBF topology:

$$\Delta\eta(j_1 - j_2) > 3.5 \qquad m_{j_1 j_2} > 450\ \mathrm{GeV/c}^2$$

Multivariate Analysis



MVA algorithm:

- Boosted Decision Tree
- Trained against *WW* bkgd
- Involve the same variables used in cut-based approach plus a few more
- Use shape of BDT output

Expected exclusion range for *1.1 fb*$^{-1}$

- Cut-based: *[140,190] GeV*
- MVA-based: *[130, 200] GeV*

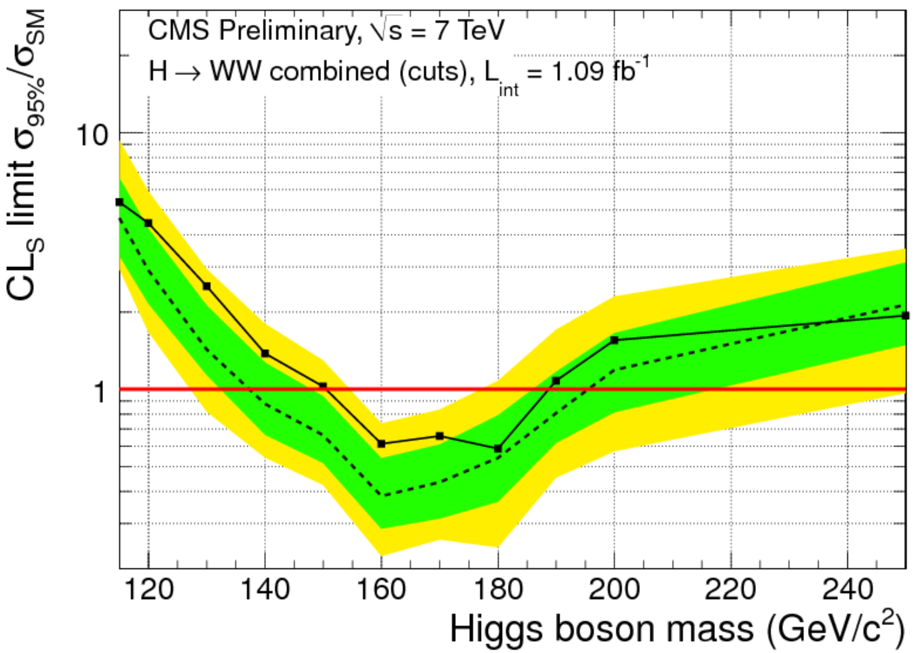
Systematic Uncertainties

- Theoretical uncertainties: *5% - 20%*
- Systematic uncertainties on data-driven background estimation are mostly statistical in nature, will decrease with more data

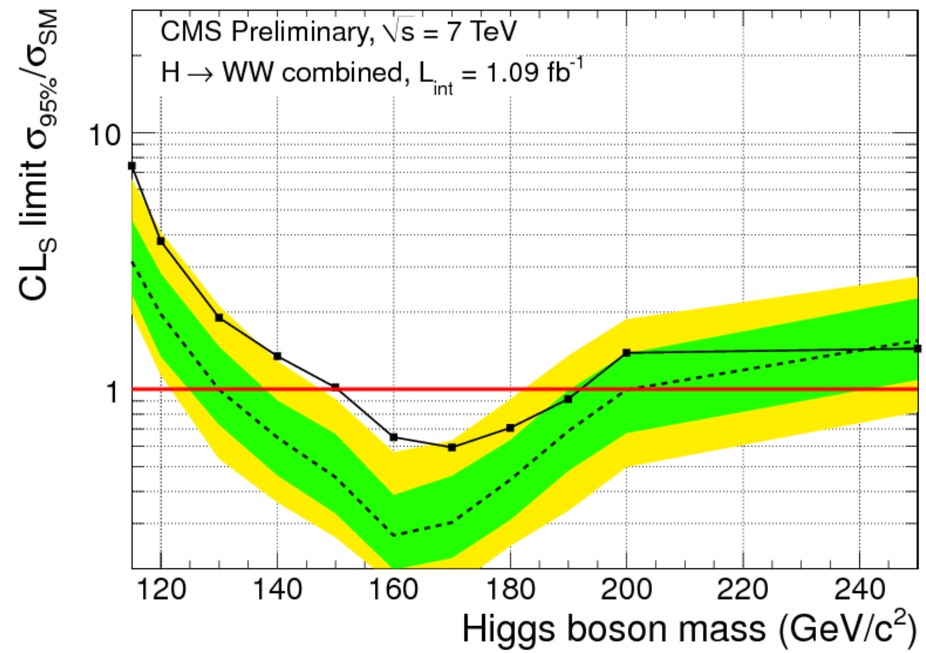
Process	Relative Systematic Uncertainty	0-jet Yield
Drell-Yan (*ee/μμ*)	60%	1.48 ± 0.26 (stat) ± 0.89 (syst)
***W*+jets**	36%	18.74 ± 2.40 (stat) ± 6.75 (syst)
Top	25%	3.66 ± 0.82 (stat) ± 0.92 (syst)
WW	15%	52.56 ± 0.71 (stat) ± 7.88 (syst)

Systematic uncertainties on data-driven background estimates in 0-jet bin at $m_H=130$ cut-based selection

Limits with $1.1\ fb^{-1}$



Limit with cut-based analysis



Limit with MVA-based analysis

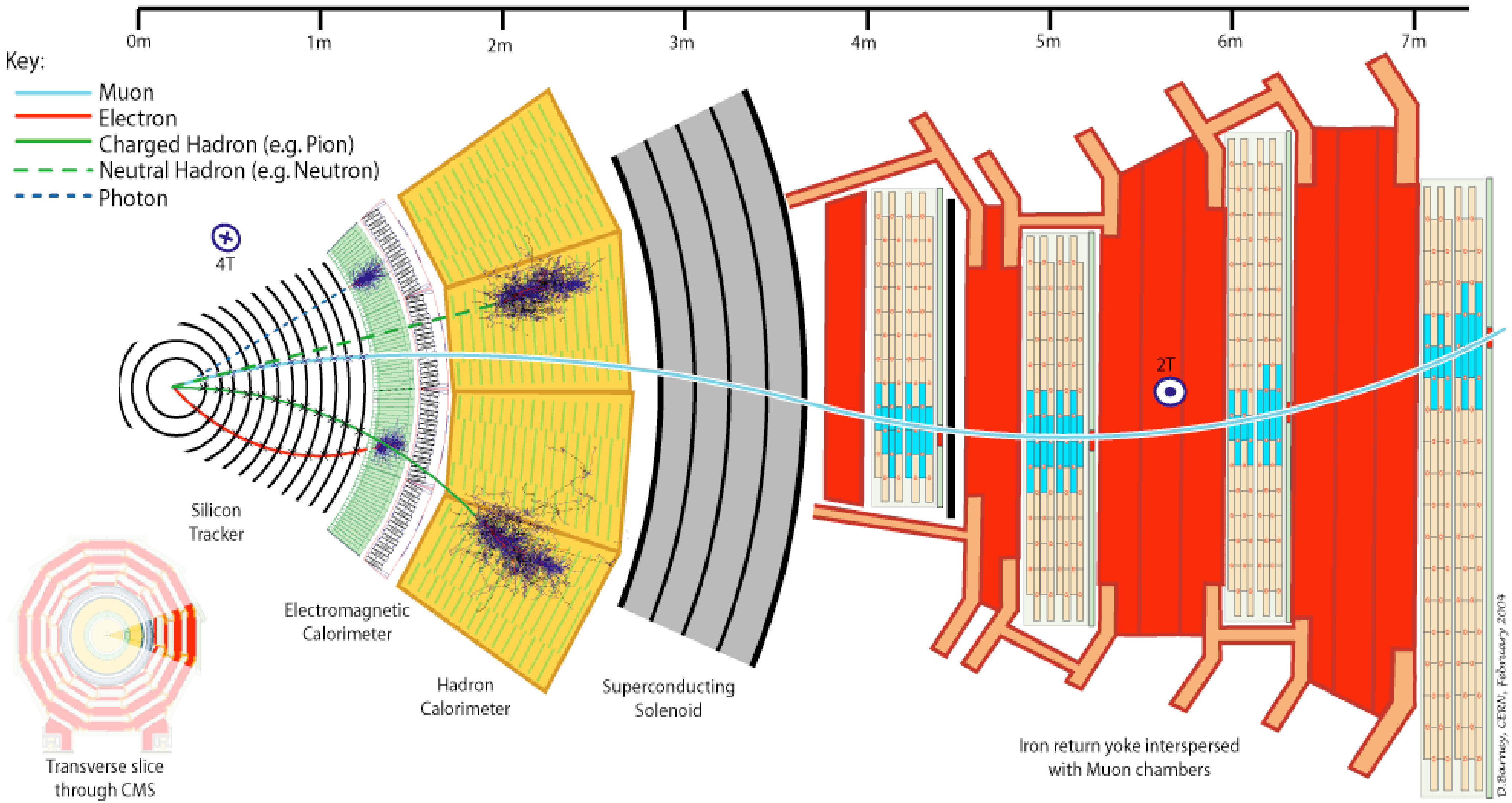
- Observed exclusion region is *[150, 193] GeV* at *95%* C.L. with MVA-based analysis, which is smaller than expectation
- $\sim 1\sigma - 2\sigma$ excess in low mass region is consistent with background fluctuation or presence of signal
- $H \rightarrow WW \rightarrow l\nu l'\nu'$ has poor mass resolution; yields are correlated between mass points

Summary

- CMS has performed a search for the Higgs boson in $H \to WW \to l\nu l'\nu'$ channel with $1.1\ fb^{-1}$ of data
- Observed exclusion region of $[150,193]\ GeV$ at 95% C.L.
- The observed exclusion region is smaller than expectation due to an excess observed in the low mass region
- The excess is consistent with background fluctuation or signal building up
  - The focus of current on-going studies
  - Getting more data!

Backup Slides

The CMS Detector



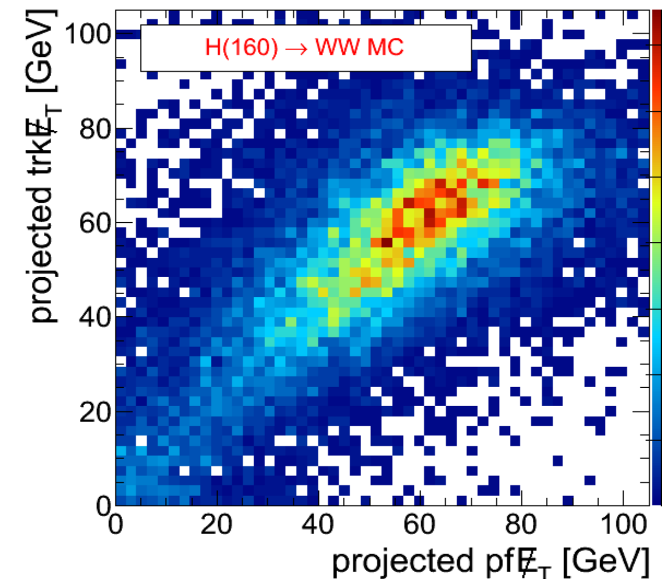
Missing Transverse Energy

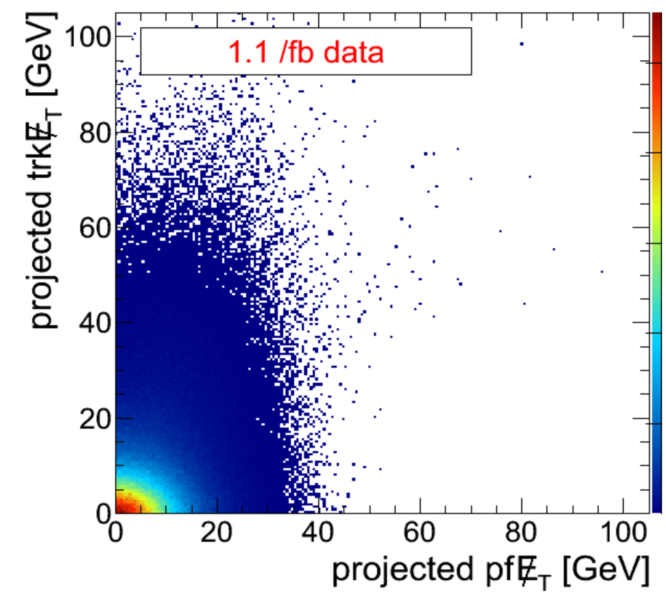
- Large MET in signal events from neutrinos
- Useful variable to further reject Drell-Yan
- Two algorithms:
  - Based on Particle Flow (PF) algorithm
  - Based on tracks associated with PV
- Use projected MET (helps reduce $Z \to \tau\tau$):

$$\Delta\phi_{min} = min(\Delta\phi(\ell_1, E_T^{miss}), \Delta\phi(\ell_2, E_T^{miss}))$$

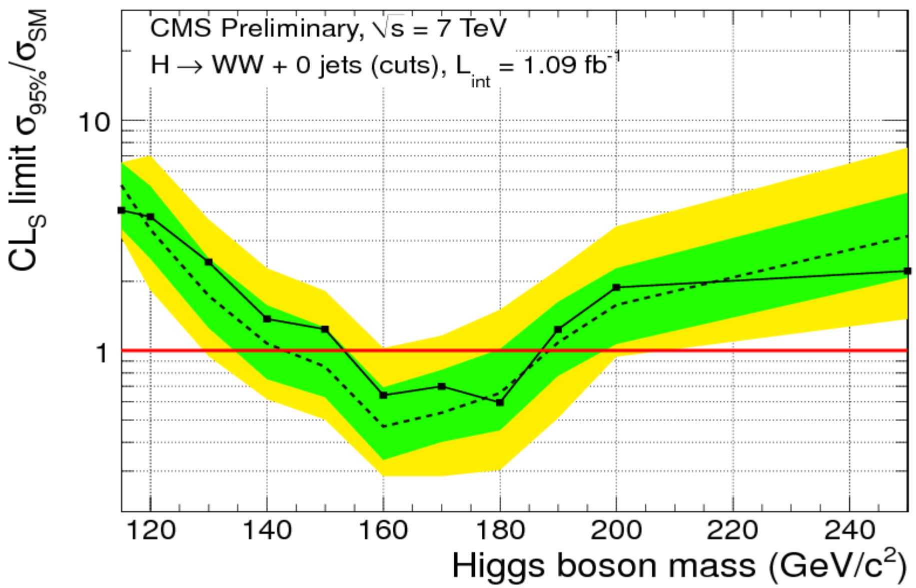
$$\text{projected } E_T^{miss} = \begin{cases} E_T^{miss} & \text{if } \Delta\phi_{min} > \frac{\pi}{2}, \\ E_T^{miss}\sin(\Delta\phi_{min}) & \text{if } \Delta\phi_{min} < \frac{\pi}{2} \end{cases}$$

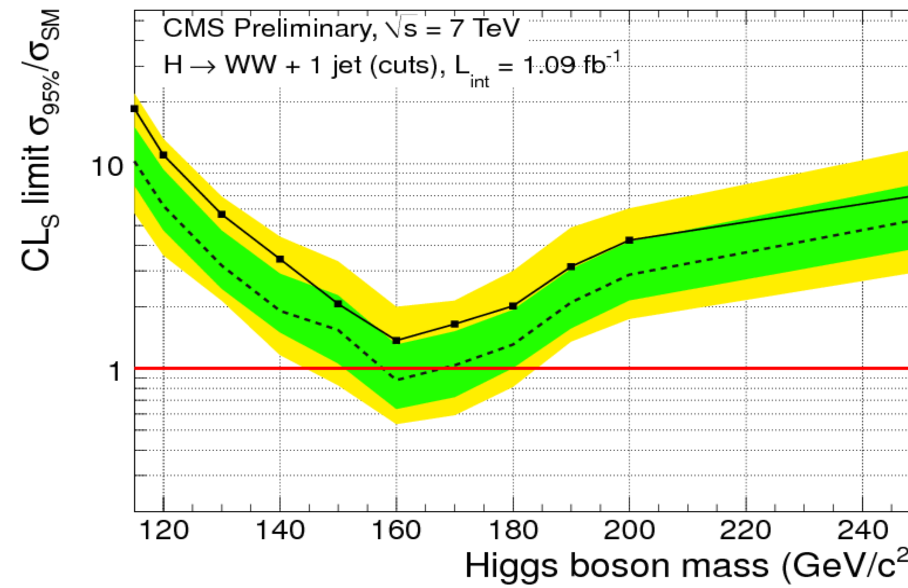
- The two MET types are strongly (weakly) correlated in events with (without) neutrinos, require large Min(proj-PF-MET, proj-trk-MET)
  - Larger than *40 GeV* for *ee, μμ*
  - Larger than *20 GeV* for *eμ*

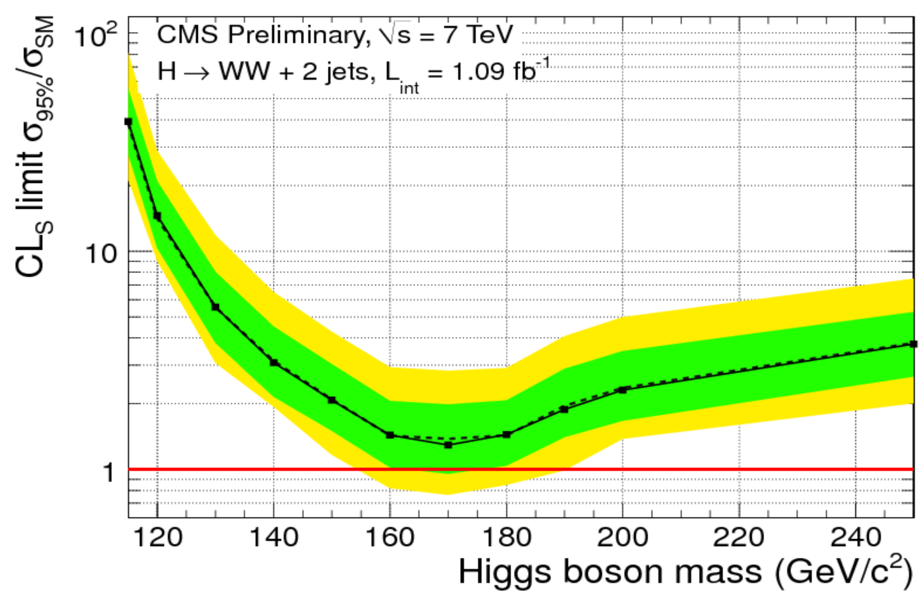




Limits in N_{jet} (cut-based)







Limits by N_{jet} (MVA-based)