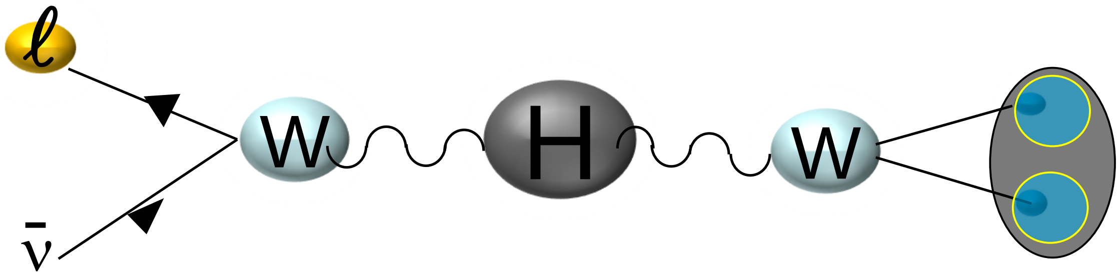


Search for a SM-like Higgs boson decaying into $WW \rightarrow l\nu q\bar{q}'$ in CMS

Zijun Xu

Peking University

ISHP, Beijing

Aug 15, 2013

Outline

- Introduction
  - $H \to WW \to$ Lep + MET + 2 jets[1]: *160-600 GeV*
  - $H \to WW \to$ Lep + MET + Fat-Jet[2]: *600-1000 GeV*
- Event Selection, and Optimization
  - $l\nu jj$: Kinematic Fit, MVA
  - $l\nu j$: Jet Grooming, Jet Substructure
- Analysis Strategy
- Statistical interpretation on SM and BSM
- Summary

[1]: CMS PAS HIG-12-046(HCP 2012), PAS HIG-12-021, PAS HIG-12-003, PAS HIG-12-034, Eur. Phys. J. C 73 (2013) 2469
[2]: CMS PAS HIG-13-008

SM Higgs Search

- H → WW→ $l\nu q\bar{q}'$ [1,2]

  <1>: $q\bar{q}'$ → 2 jets

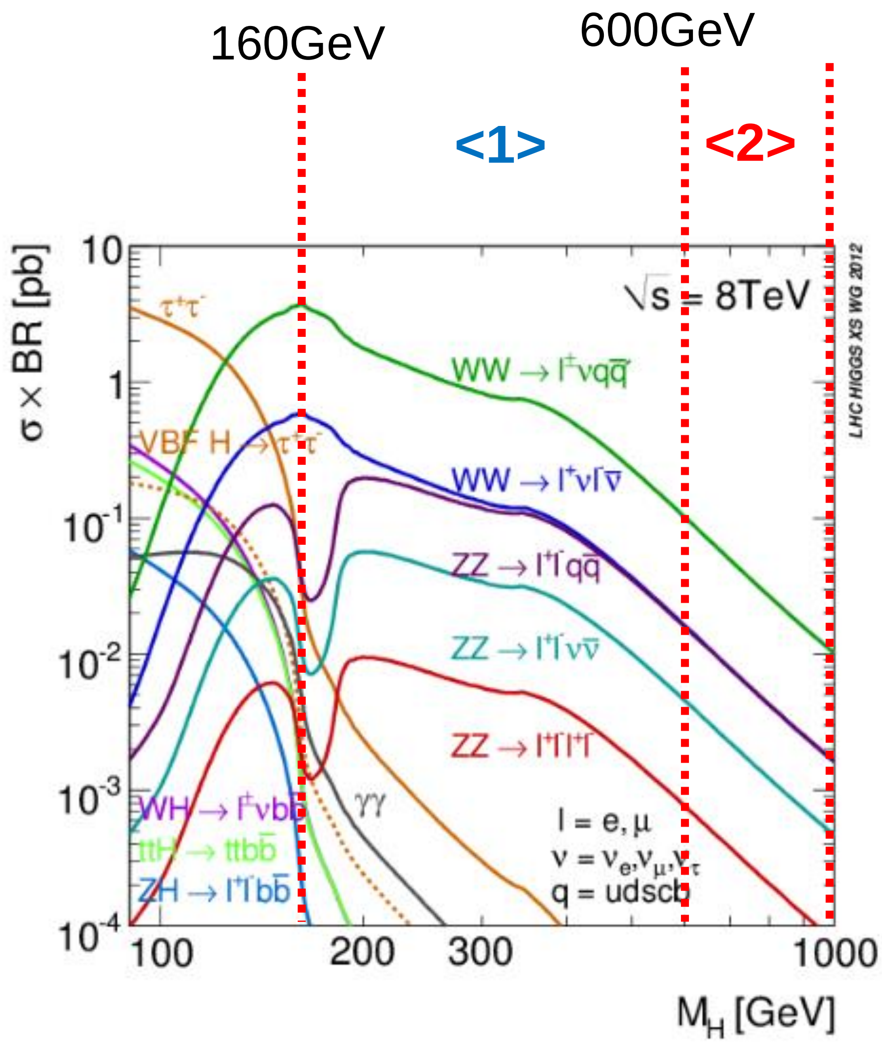
  <2>: $q\bar{q}'$ → a Fat-Jet

  Normal W→ 2 jets search loses statistics at high mass regime (~>600GeV).

- PROs:
  - High branching ratio
  - Full Higgs mass reconstruction

    Neutrino p_Z from m_W constraint

- CONs:
  - Only one lepton: huge W+jets background, QCD contamination
  - Broad detector resolution from jets and missing energy



[1] Han, Tao et al. Phys.Rev.Lett. 82 (1999) 25-28 hep-ph/9807424 MADPH-98-1067 ;
[2] B. A. et al. JHEP 1004 (2010) 083, doi:10.1007/JHEP04(2010)083, arXiv:0912.3543.

Heavy Higgs search in the era of H126

It's important to keep searching for the SM Higgs boson in the high mass regime:

- Higgs boson is one of the most essential particles in SM, need to check everything.
- It's possible that the 126GeV particle is not fully responsible for the electroweak symmetry breaking mechanism .
- Several popular BSM scenarios predict additional resonances at high mass.

This analysis is a benchmark for future di-boson final state searches and measurements in the high mass region

- Di-boson scattering measurement is meaningful: TGC, unitarization of the WW scattering …
- Jet substructure techniques are studied, which will be more widely used with LHC CM energy increasing.

Analysis Selections

An isolated, high-pT **lepton**:

pT > 30 (30) [30 (35)] GeV/c for electrons (muons)

High E_T^{miss} from a neutrino:

E_T^{miss} > 30 (25) [50 (70)] GeV for electrons (muons)

Leptonic W: mT(lepton+ E_T^{miss}) > 30 GeV

*: Red number is for heavy higgs search

- **Two** high pT **jets** with m_{jj} ~ 80 GeV

  pT > 30 GeV/c,

  |η| < 2.4,

  ΔR(jet-lepton) > 0.3

  #extra-jets = 0,1

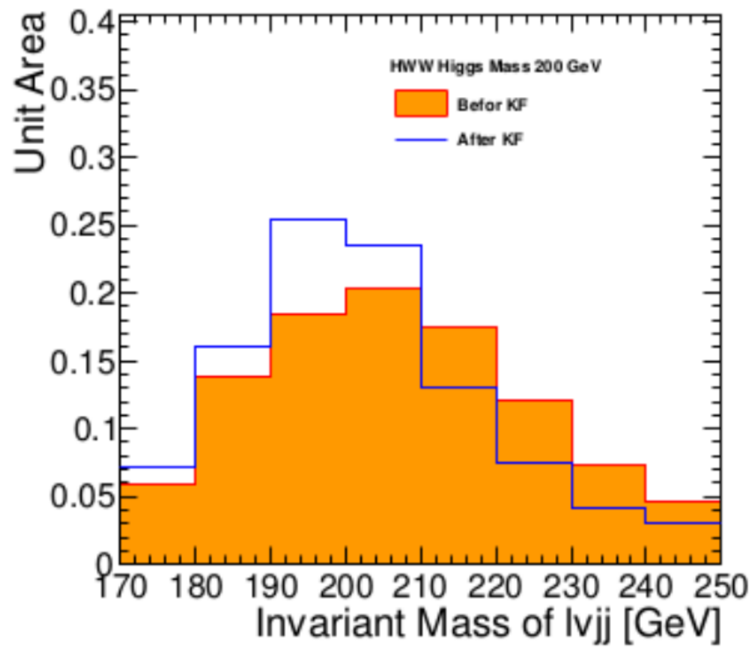
- A highly boosted leptonic **W**: pT_W > 200 GeV
- A High pT **fat-jet:** CA8 jet with pT > 200 GeV
- A back-to-back topology:

  $\Delta R_{l,j} > 1.57 = \pi/2$, $\Delta\Phi_{MET,j} > 2.0$, $\Delta\Phi_{V,j} > 2.0$

- Top events veto: none b-tagged jet

lvjj Optimization: KF+MVA

Kinematic Fit on lepton, E_T^{miss}, 2 jets to improve Higgs mass resolution

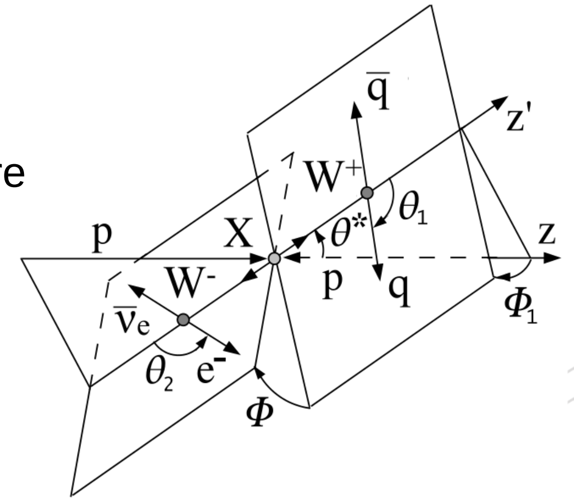


MVA to improve significance: a simple **likelihood** discriminator:

a different likelihood is built for each different final state

$\{\theta_1, \theta_2, \theta^*, \phi, \phi_1, (p_T)_{WW}, y_{WW}$, lepton charge$\}$

The lepton charge is a good variable since signals are charge-symmetric, while W+jets production is not.



lvj Optimization: Jet grooming

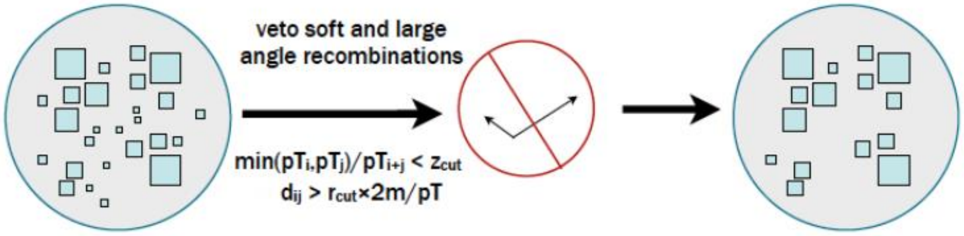
- W+jets is dominant background. Jet mass is correlated with jet pt and R, and easily be effected by soft QCD and pile-up.

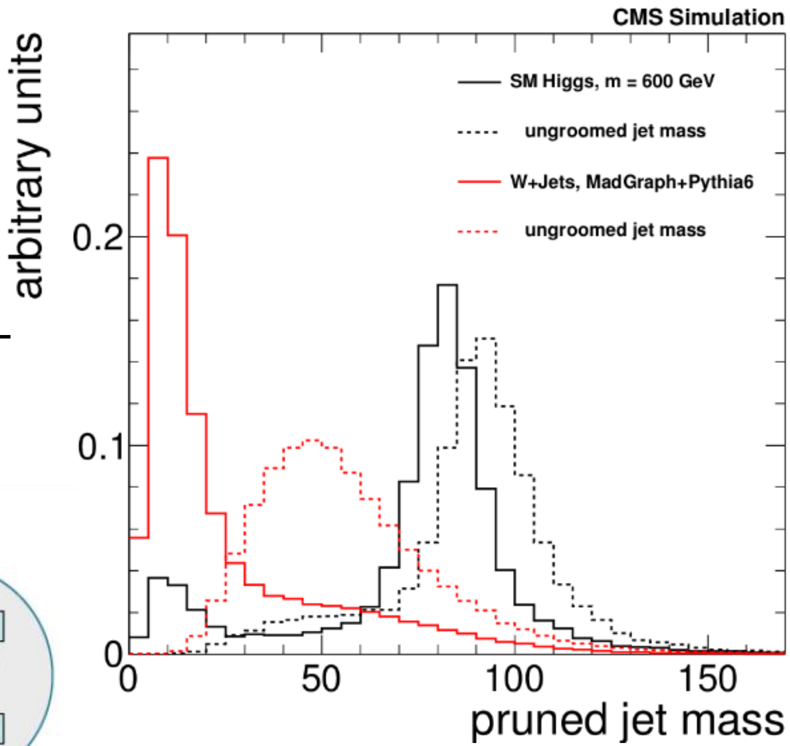
- **Jet grooming**:

Filtering[1] / Trimming[2] / **Pruning[3]**

Removing soft radiation and pileup contributions to jets, which gives better signal-to-bkg separation.

Pruning parameters: $z_{cut} = 0.1$, $r_{cut} = 0.5$





[1] Butterworth, Jonathan M. et al. Phys.Rev.Lett. 100 (2008) 242001 arXiv:0802.2470
[2] David, Jesse, Lian-Tao Wang. JHEP 1002 (2010) 084 arXiv:0912.1342
[3] Ellis, Stephen D. et al. Phys.Rev. D80 (2009) 051501 arXiv:0903.5081

lvj Optimization: Jet substructure

- **Jet substructure**:
  - Variables: Mass drop, N-subjettiness, Qjets, Cores and planar flow, Subjets kinematics…
  - **N-subjettiness** is the most sensitive variable, MVA is a small improvement in performance.

$$\tau_N = \frac{1}{d_0}\sum_i p_{T,i}\min\{(\Delta R_{1,i})^{\beta},(\Delta R_{2,i})^{\beta},...,(\Delta R_{N,i})^{\beta}\}$$

$$d_0 = \sum_i p_{T,i}(R_0)^{\beta} \qquad R_0=0.8,\ \beta=1$$

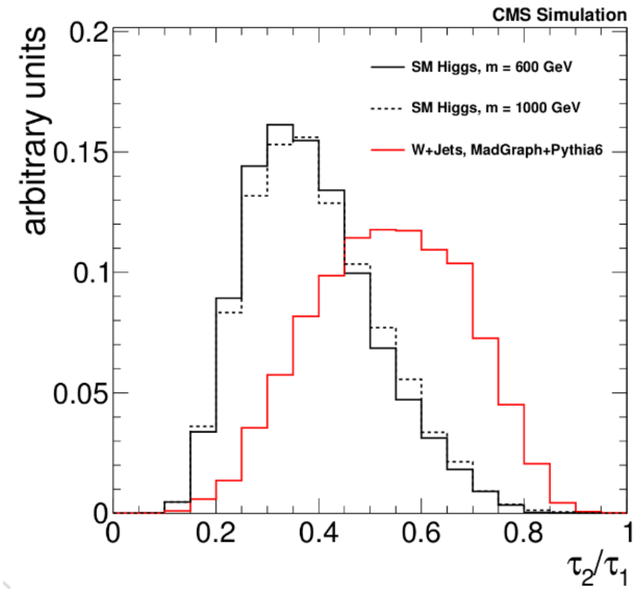
τ_N tends to be zero as the jet becomes more consistent with N Subjets.

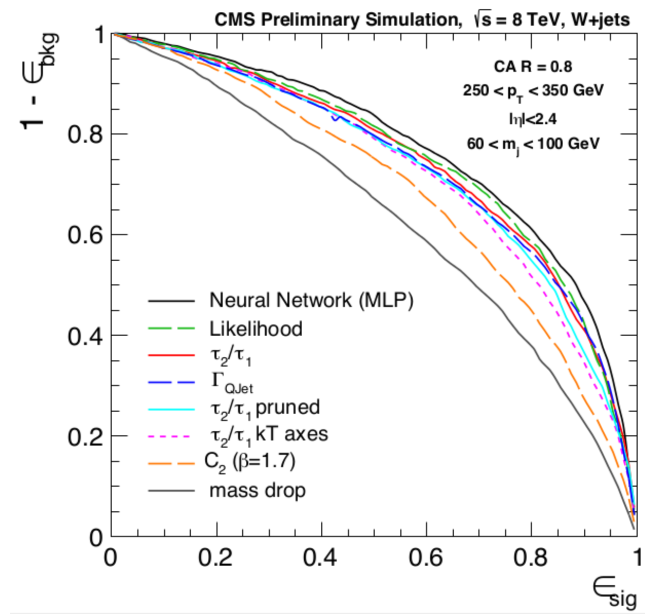
- **Wtagger**: identify the W-jet from Huge QCD

CA8 jet Pt cut before grooming +

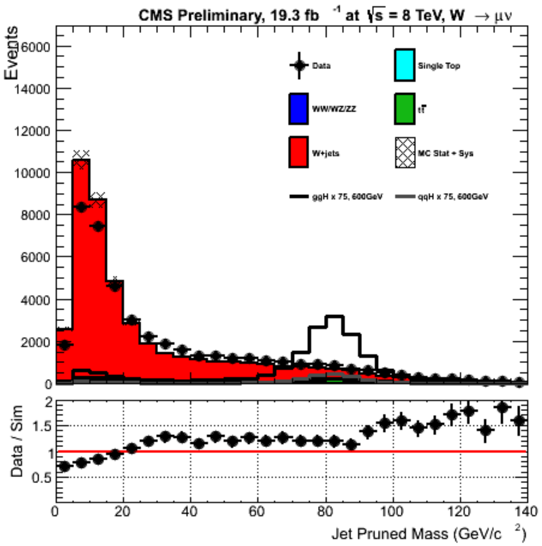
pruned jet mass cut + N-subjettiness τ_2/τ_1 cut

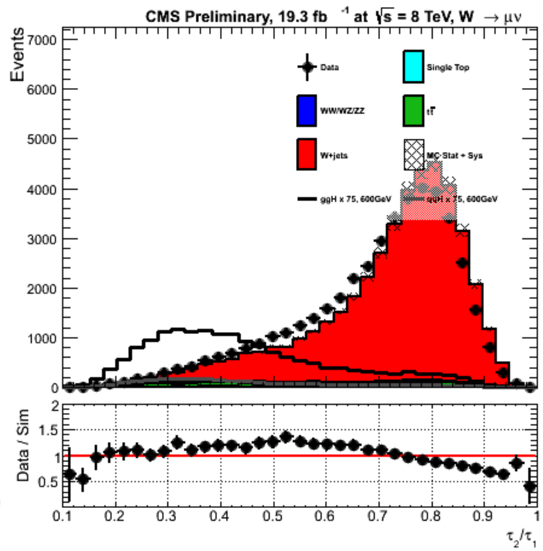
(cut value depends on desired signal efficiency and background fake rate)



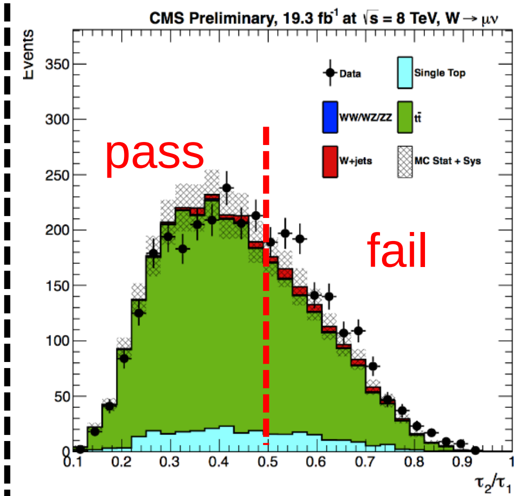


W-jet MC Correction





TTbar, after Jet Pt and mass cut



Selecting Top-enriched samples with a merged W

- inverting the Top veto: #bjets >0
- not requiring a back-to-back topology

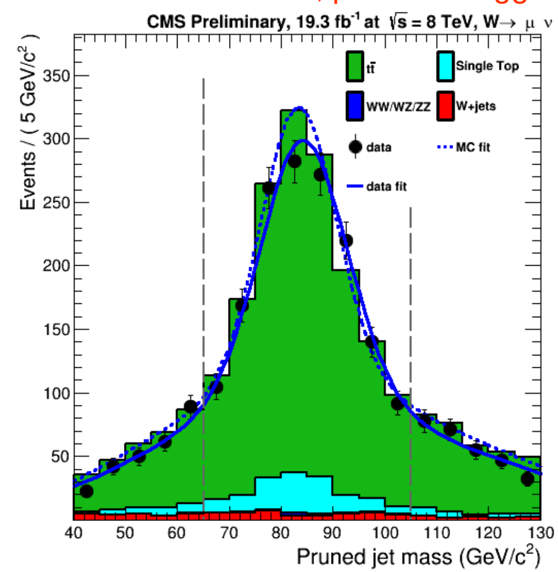
W-tagger efficiency SF:

- Fit background subtracted W-jet peak for W-tagging efficiency SF;
- Use signal window scale factor for top events.

Jet mass scale corrections:

- Gaus-like function for describing the peak;
- Corrections on the mean and sigma value.

TTbar, passed wtagger

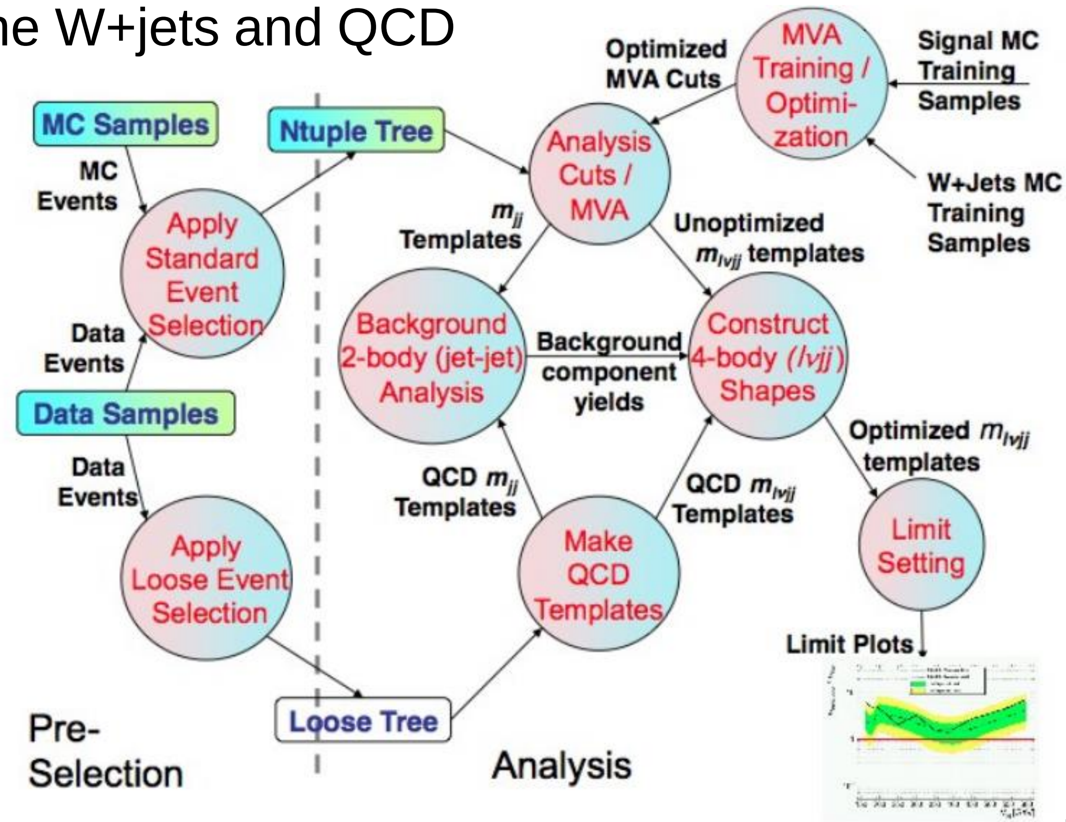


Analysis strategy

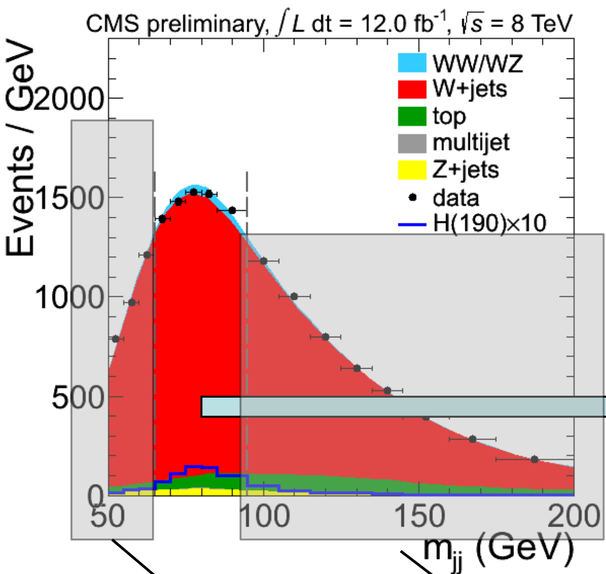
Main steps for signal and Bkg estimation:

- The $\mathbf{m_{jj}}$ or $\mathbf{m_{fat\text{-}jet}}$ distribution used to get the **bkg normalization**
- The $\mathbf{m_{lvjj}}$ or $\mathbf{m_{lvj}}$ distribution used to extract the **limit**
- Data-driven method for the W+jets and QCD
- Minor bkgs from MC:

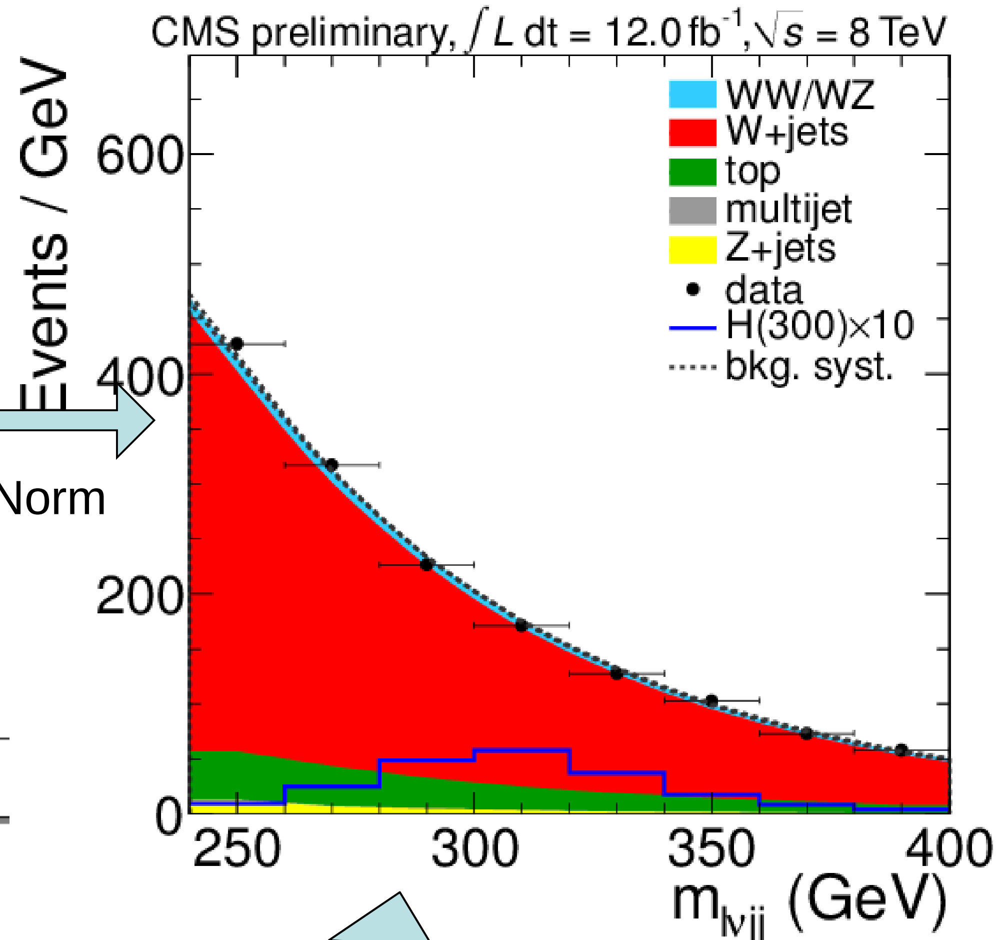
  WW/WZ/ZZ, Top, Z+jets



Analysis of L+ν+JJ



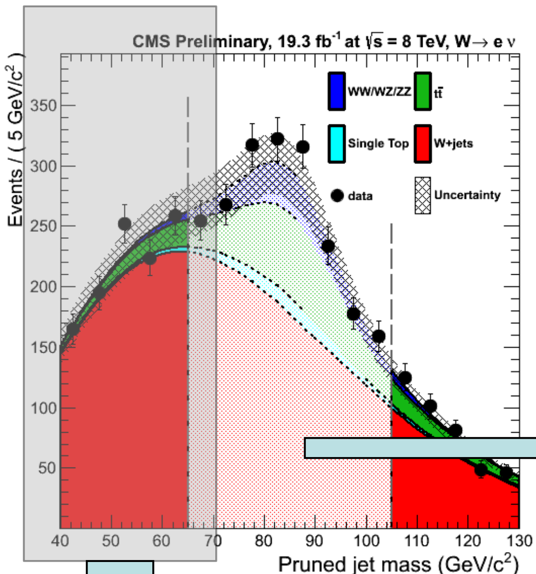
W+jets Norm

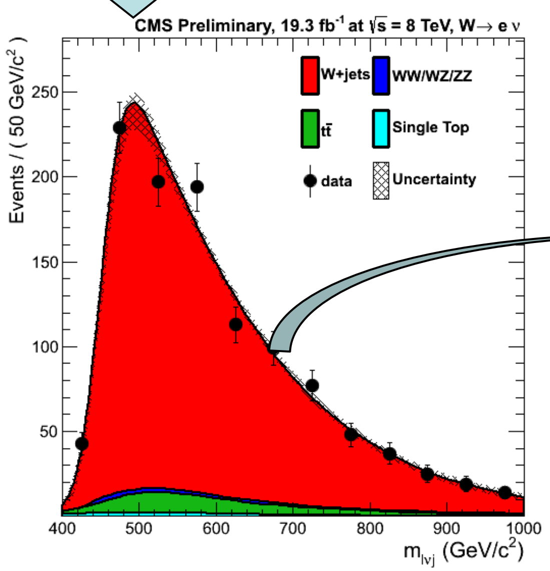


$$m_{l\nu jj}^{i} = \alpha \cdot m_{SBL}^{i} + (1-\alpha) \cdot m_{SBH}^{i}$$

W+jets Shape

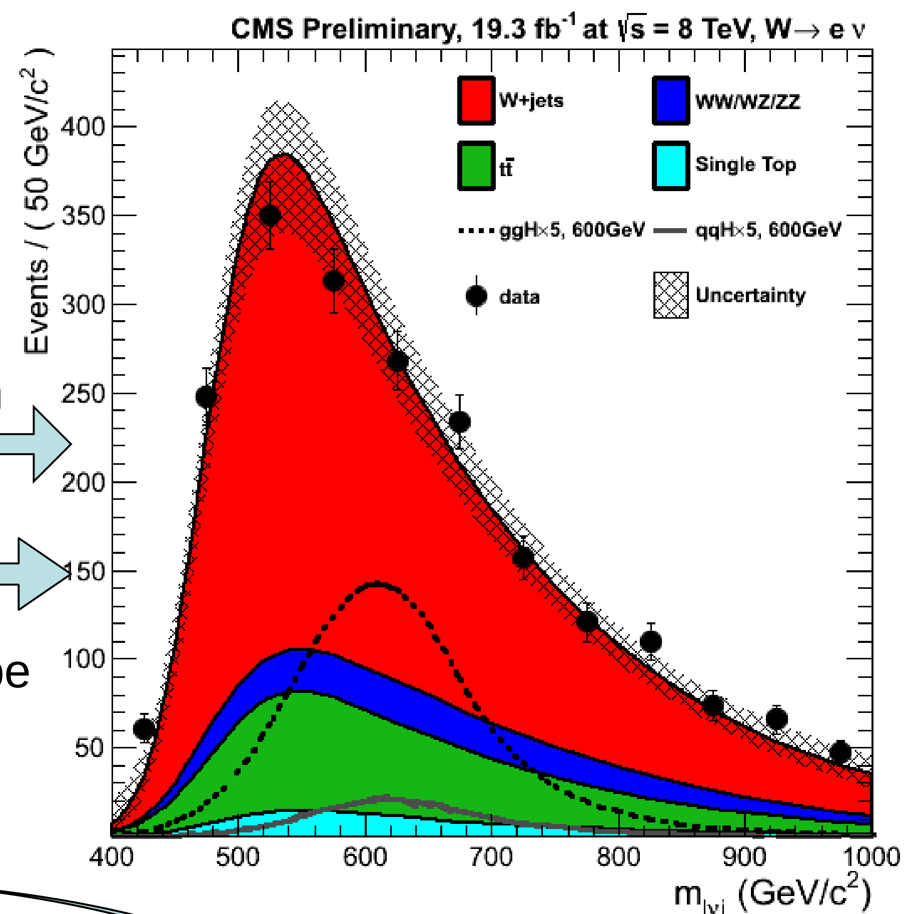
Analysis of L+ν+Fat-Jet





W+jets Norm

W+jets Shape



$$\boxed{F_{\text{data,SR}}(m_{l\nu j}) = \alpha_{\text{MC}}(m_{l\nu j}) \times F_{\text{data,SB}}(m_{l\nu j})}$$

Systematic Uncertainties List

normalization uncertainty	<= 2 %
W+jets fit uncertainty	shape
inclusive cross-section	13-15%
scale uncertainties from jet binning	4-28%
acceptance uncertainty due to pdfs	1-2%
luminosity	4.4%
jet energy scale and MET	1%
lepton efficiencies	2%
lepton trigger efficiencies	1%
pile-up	<1%
likelihood selection	10%
interference	shape

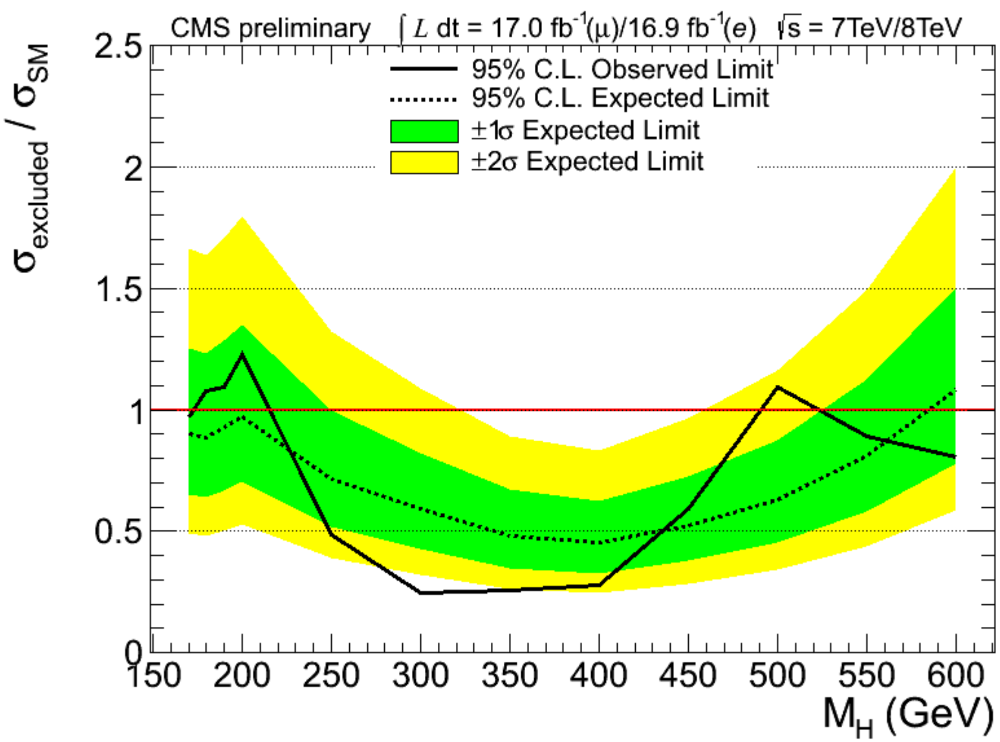
background systematics

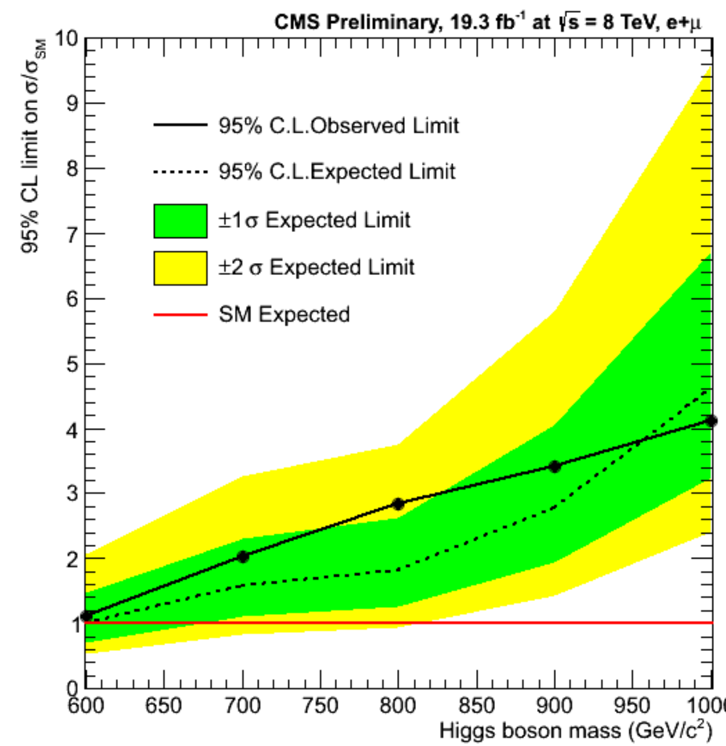
signal systematics

Syst. uncertainty	sig, ggH	sig, VBF	W+jets	$t\bar{t}$	single t	WW/WZ
lumi	4.4%	4.4%	-	4.4%	4.4%	4.4%
Higgs cross-section	13.7% †	0.7% †	-	-	-	-
Higgs PDF/scale	9.9% †	3.6% †	-	-	-	-
Intf (sig/bkg)	10.0%	50.0%	-	-	-	-
Bkg cross-section	-	-	-	-	30.0%	30.0%
W+jets norm.	-	-	8%	-	-	-
W-tagging	10.0%	10.0%	-	-	-	10.0%
$t\bar{t}$ norm.	-	-	-	6.0%	6.0%	-
Jet mass/energy scale	2%	2%	-	2%	2%	2%
W+jets shape	-	-	see Section. 6	-	-	-
b-tagging	2.5%	2.5%	-	-	2.5%	2.5%
Trigger (e & μ)	1%	1%	-	-	1%	1%
Selection Eff. (e & μ)	2%	2%	-	-	2%	2%

Table 1: Summary of systematic uncertainties related to normalization of expected signal and background yields. The symbol † denotes a mass-dependent uncertainty.

SM limits





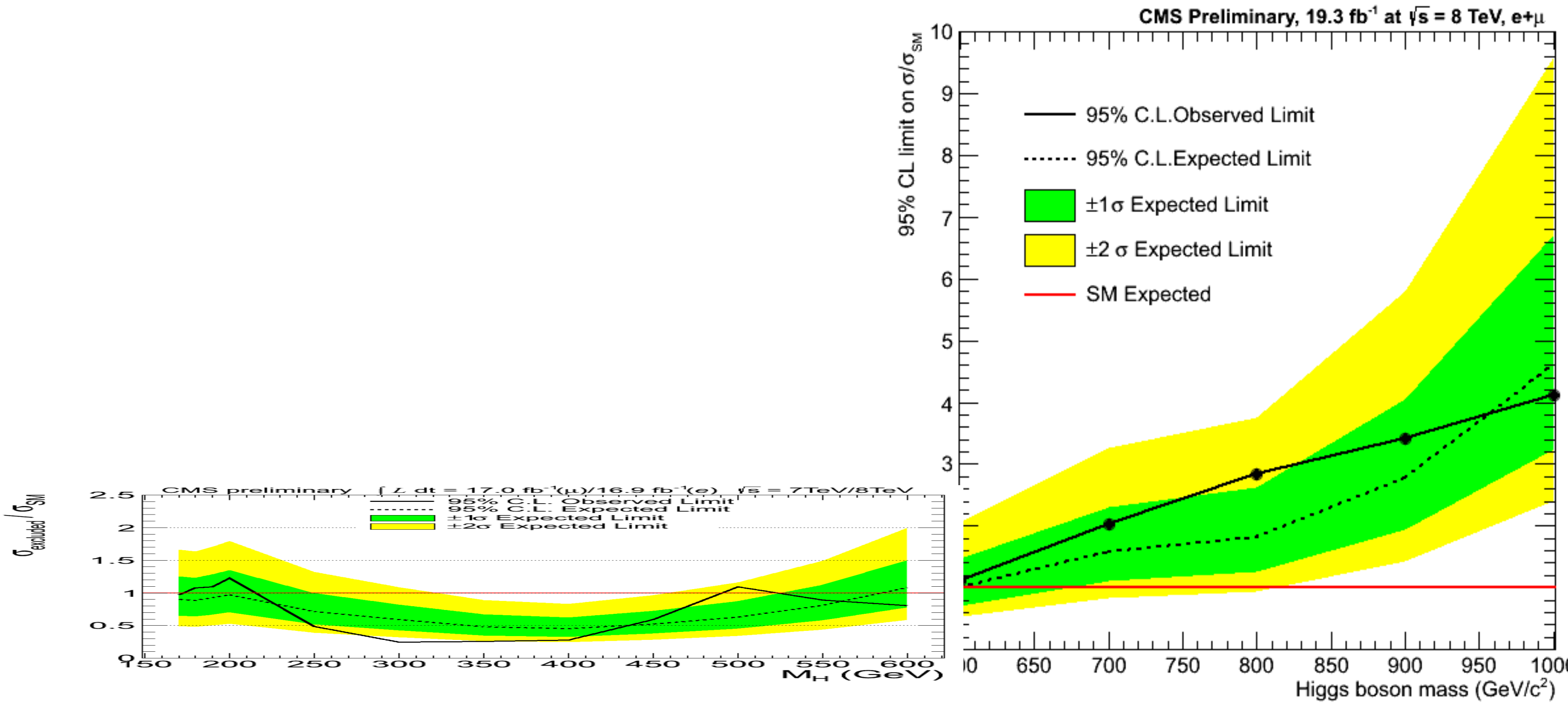
H → WW → Lep + MET + 2 jets:

- observed exclusion 95% CL: **215–490 GeV** and **525–600 GeV**

H → WW → Lep + MET + Fat-Jet

- No significant excess is observed: exclude at 1.1 (4.1) times the SM Higgs cross-section for a mass of 600 (1000) GeV hypothesis.

SM limits



H → WW → Lep + MET + 2 jets:

- observed exclusion 95% CL: **215–490 GeV** and **525–600 GeV**

H → WW → Lep + MET + Fat-Jet

- No significant excess is observed: exclude at 1.1 (4.1) times the SM Higgs cross-section for a mass of 600 (1000) GeV hypothesis.

Lvjj and lvj have same sensitivity at 600 point!

BSM Limits:σ

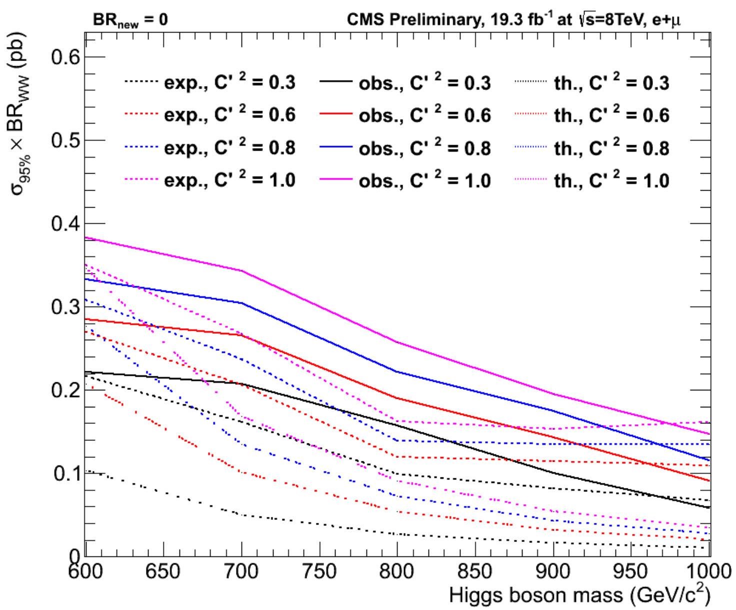
A simple model for BSM: A heavy Higgs mixes with the new boson at 126GeV:

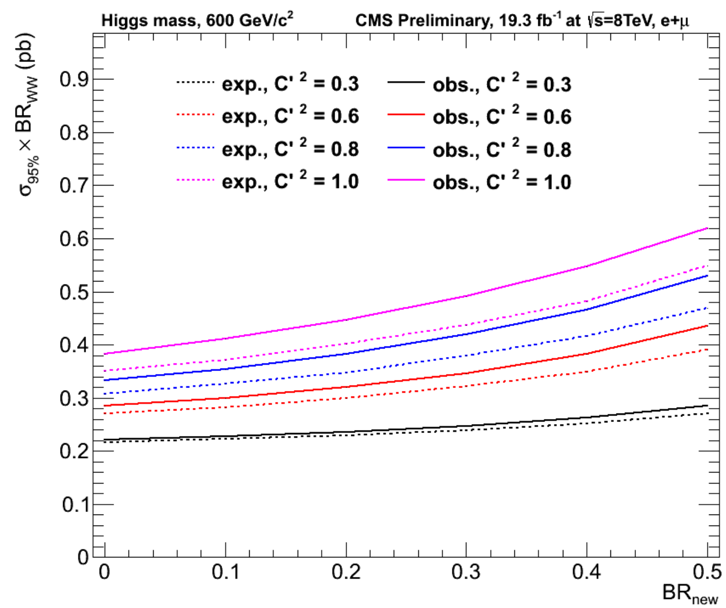
The heavy Higgs, coupling scaled by C', completes unitarization with H(126), coupling scaled by C, such that $C'^2 + C^2 = 1$;

The heavy Higgs has a non-SM-like decay modes: BR_{new}

The cross-section and width are modified in the following way:

$$\mu' = C'^2 \times (1 - BR_{new}) \;,\;\; \Gamma' = (C'^2/(1-BR_{new})) \times \Gamma_{SM}$$





The typical upper limit on the $\sigma_{95\%} \times BR_{WW}$ ranges from ~60 to 400 fb when BR_{new}=0 and C'^2 ranges from 0.3 to1.0.

*more plots in additional materials

Summary

- $H \to WW \to$ Lep + MET + 2 jets:

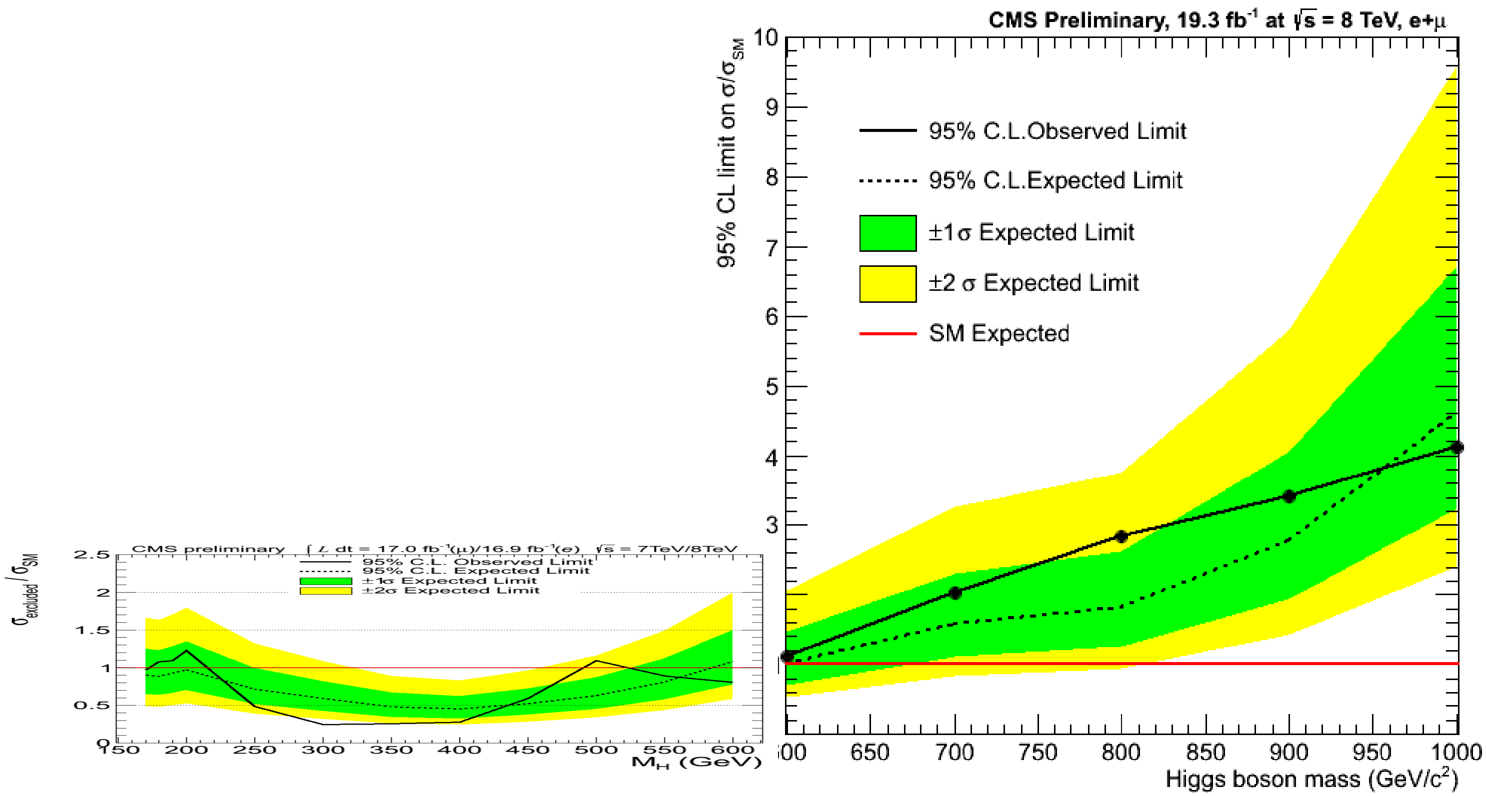
  The SM higgs mass ranges 215-490 GeV and 525-600 GeV be excluded at 95% confidence level;

- $H \to WW \to$ Lep + MET + Fat-Jet
  - No significant excess is observed in 600-1000GeV,
  - Upper limit on the $\sigma_{95\%} \times BR_{WW}$ be set for a simple BSM scenario;
- Jet **grooming** and **substructure** will be important for searches in high mass region in the future.

  EXO-12-021(EXO->WW (800GeV-2.5TeV)) and JME-13-006(detailed W tagging study) just be approved by CMS this week, see talk by **John Paul Chou**

additional material

SM limits



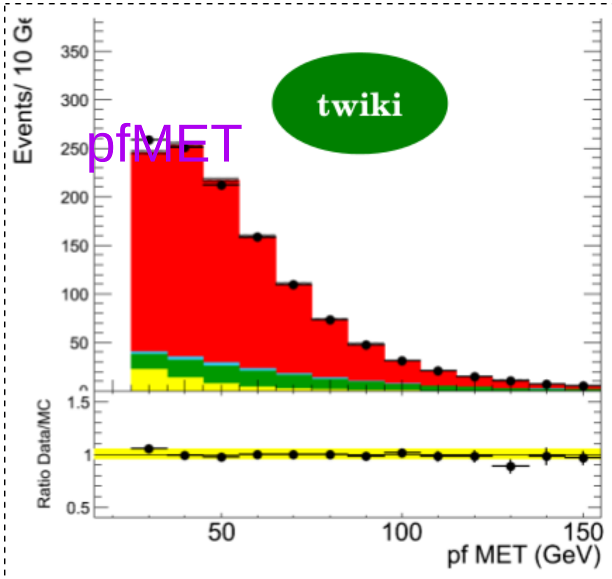
H → WW → Lep + ν + 2 or 3 jets:

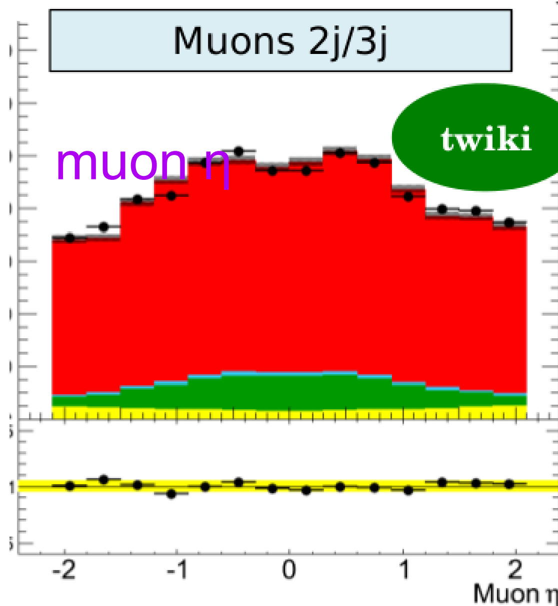
- observed exclusion 95% CL: **225–485 GeV** and **550–600 GeV**

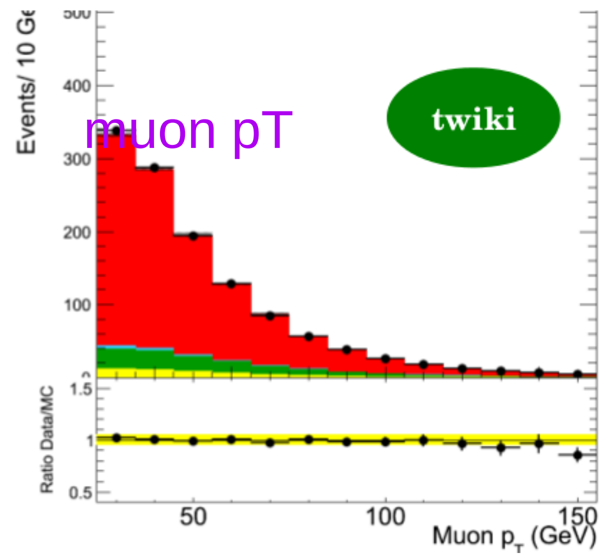
H → WW → Lep + ν + Fat-Jet

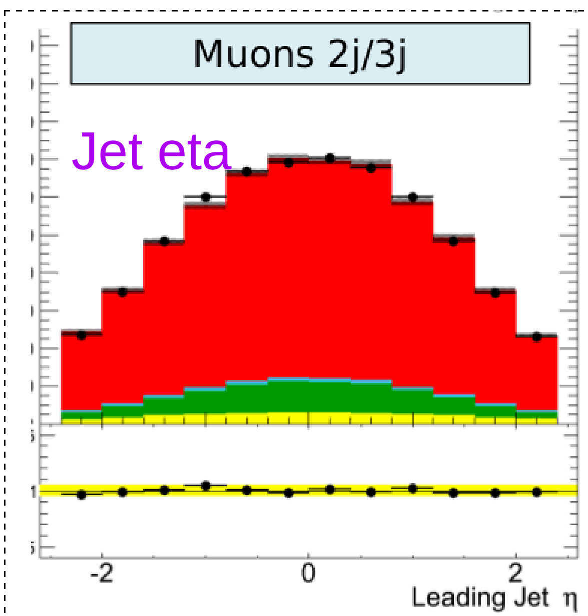
- No significant excess is observed: exclude at 1.1 (4.1) times the SM Higgs cross-section for a mass of 600 (1000) GeV hypothesis.

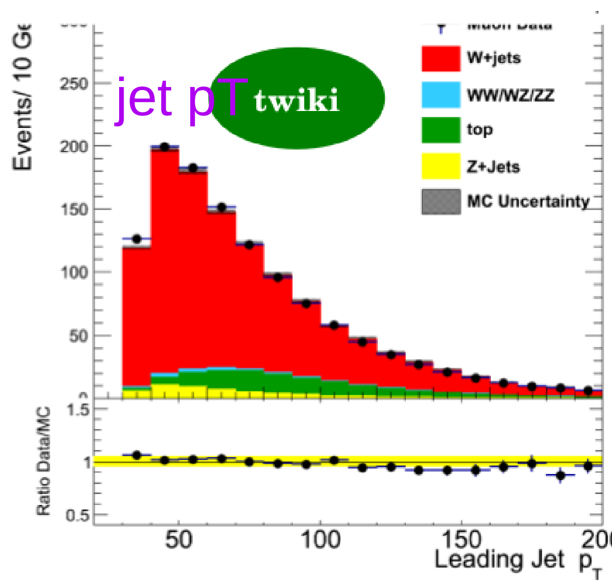
Control Plots

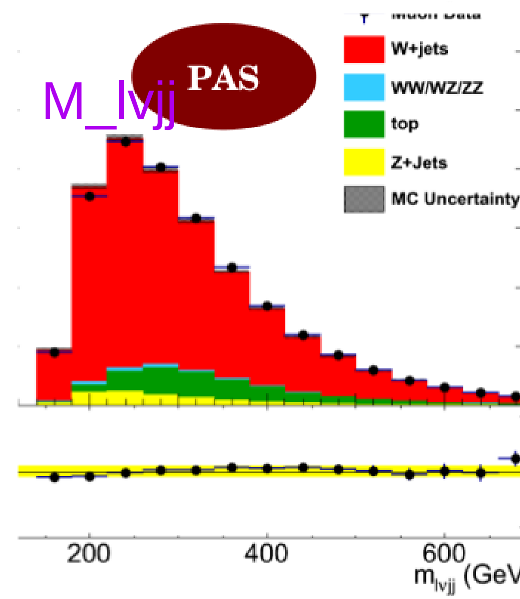








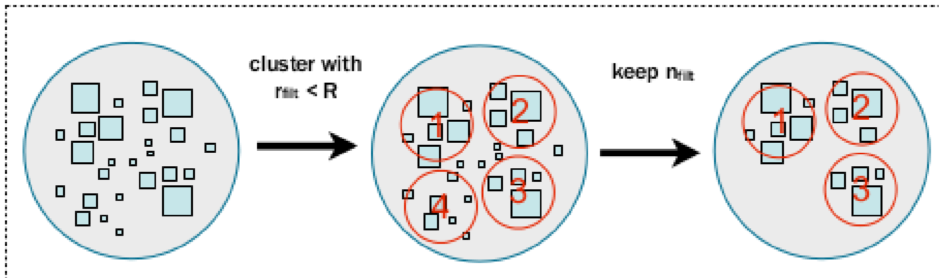




Grooming

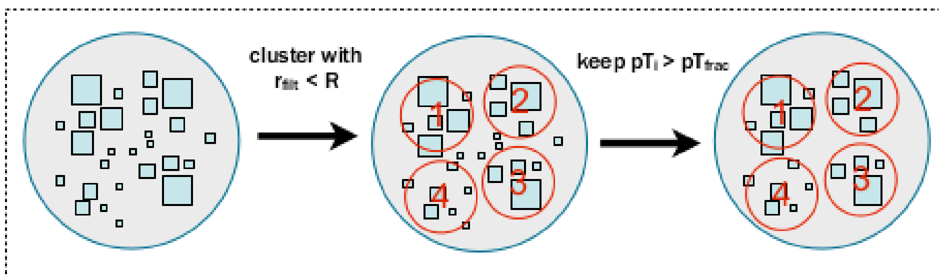
Grooming algorithms can be used to remove soft QCD and pileup contributions to the jet

filtering
[0802.2470]



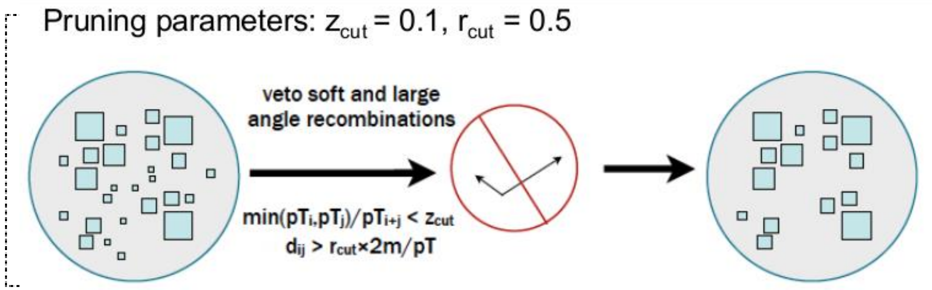
default:
$r_{filt} = 0.3$
$n_{filt} = 3$

trimming
[0912.1342]



default:
$r_{filt} = 0.2$
$pT_{frac} = 0.03$

pruning
[0903.5081]



default:
$z_{cut} = 0.1$
$r_{cut} = 0.5$

Data/MC samples, Triggers

Data:

Full 2012 8 TeV sample, 19.2-19.3 fb^{-1}(A,B,C,D)

Trigger:

SingleElectron PD
SingleMu PD

Monte Carlo:

Signal

SM ggH and qqH, Powheg [mH = 600-1000 GeV]

Background

W+jets (GEN W pT > 100 GeV), Herwig
W+jets (GEN W pT > 100 GeV), MG + Pythia
TTbar, Powheg + Pythia
WW/WZ/ZZ, Pythia
single top, Powheg
Z+jets, MG + Pythia

* Dataset names and run ranges given in additional material

Signal Reweighting: SM and BSM

- SM Higgs lineshape at high mass (mH > 400 GeV) requires reweighting from Powheg **fixed width BW** to **complex pole scheme** (CPS)
- Additional reweighting is needed for the **interference** effects between the gg→H→WW with SM continuum background gg→WW
- No LHC XS WG recommendation exists for the interference effects reweighting of the VBF signal, currently assign large uncertainty

- A simple benchmark model for BSM*:
  1) A heavy Higgs mixes with the new boson at 125GeV:
     The heavy Higgs, coupling scaled by C', completes unitarization with H(125), coupling scaled by C, such that $C'^2 + C^2 = 1$;
  2) The heavy Higgs has a non-SM-like decay modes: BR_{new}

  The cross-section and width are modified in the following way:

$$\mu' = C'^2 \times (1 - BR_{new})$$

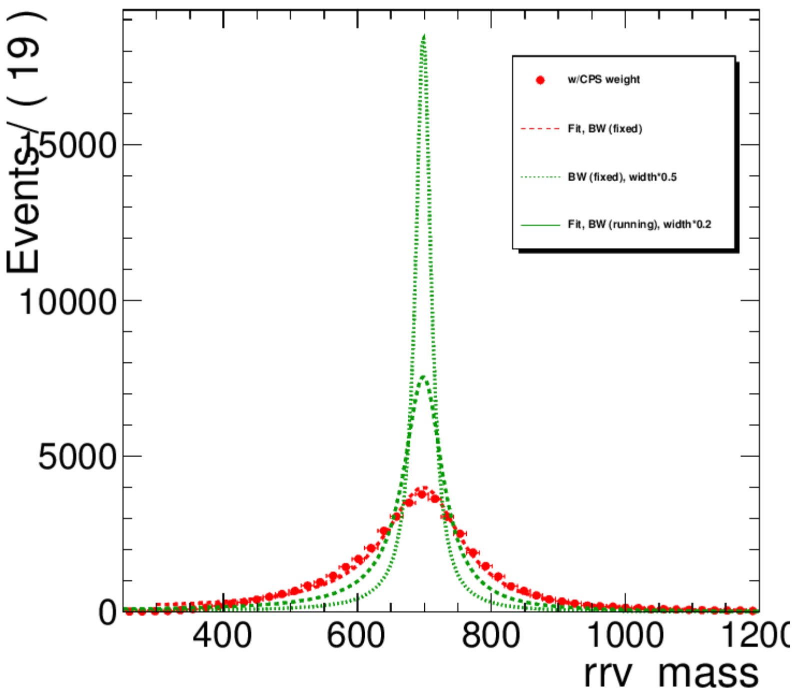
$$\Gamma' = (C'^2/(1-BR_{new})) \times \Gamma_{SM}$$

*is going to be included in the Yellow Report 3

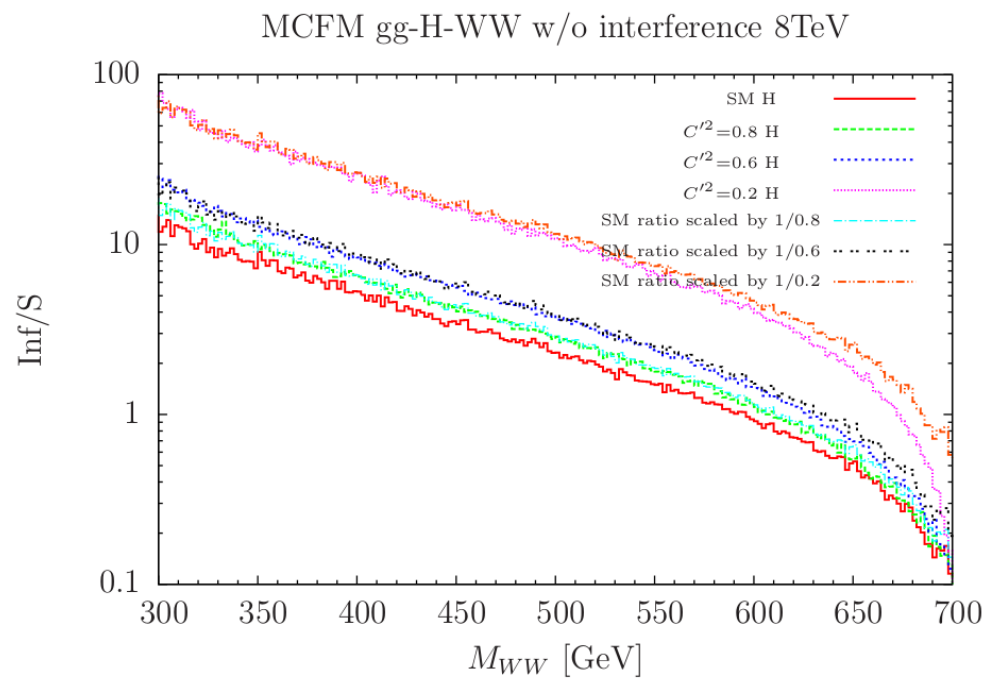
Signal Reweighting: SM and BSM

Two steps for signal reweighting in this BSM scenario:

1) Reweighting the CPS lineshape, before the inclusion of interference reweighting with gg → WW, to a narrower width:
2) Applying the interference reweighting which has been modified to include effects from a narrower width resonance with modified couplings.



fitting by running width Breit-Wigner function, then the fitted width scaled by a factor of $C'^2 \times (1 - BR_{new})$

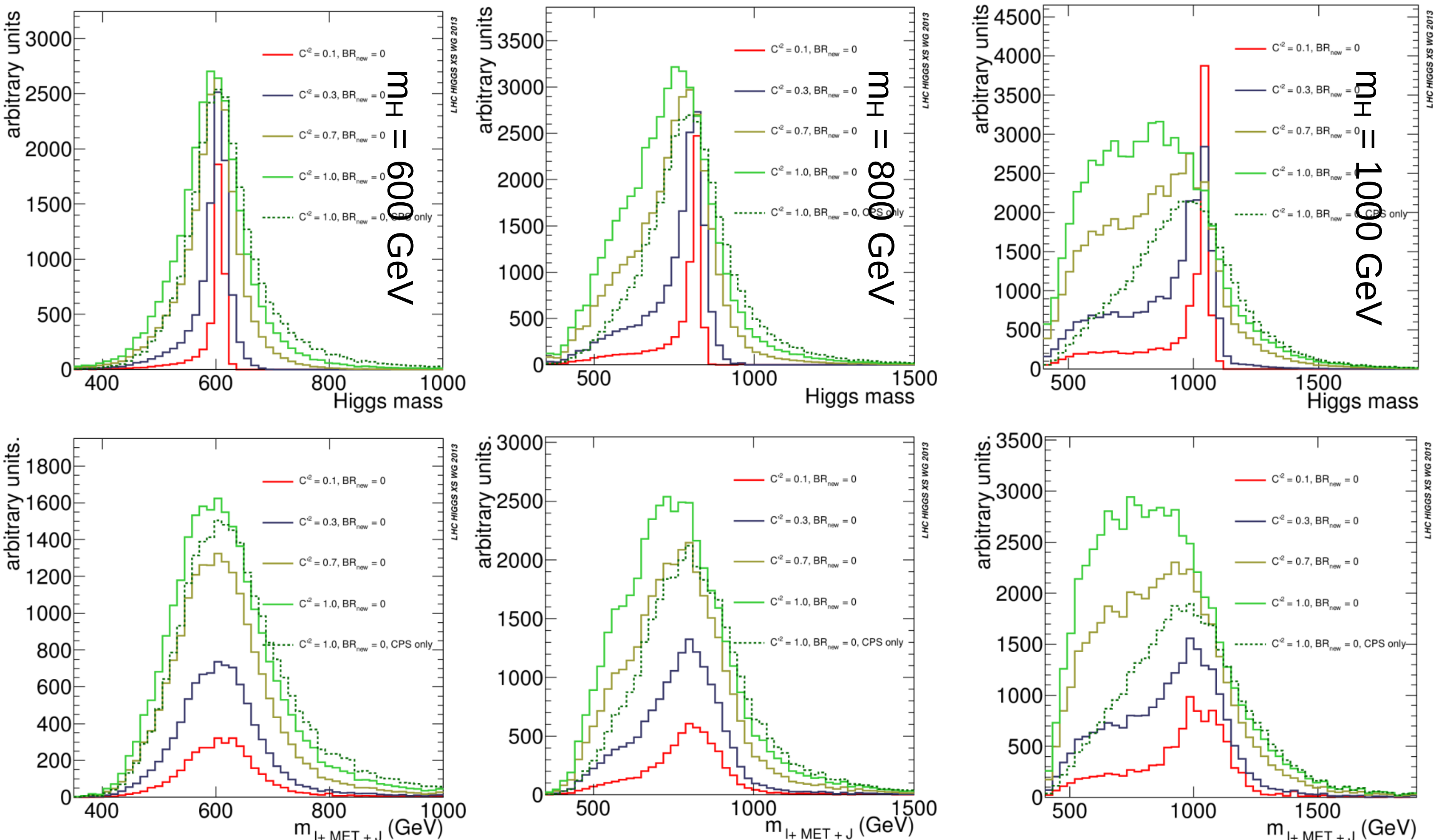


The agreement is not perfect, but acceptable.

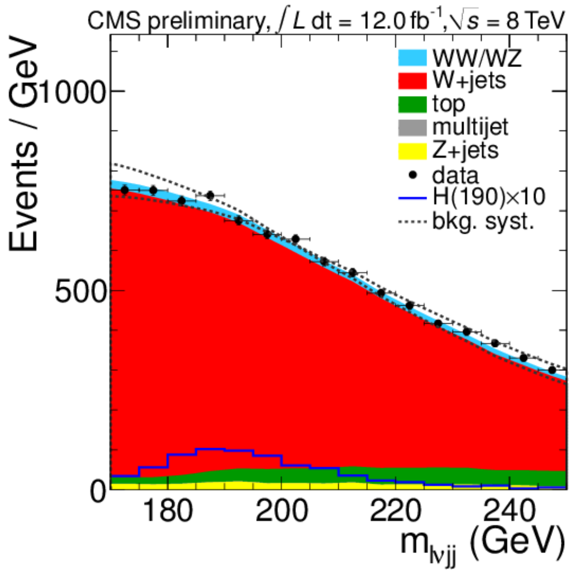
* Nhan Tran: https://indico.cern.ch/getFile.py/access?contribId=1&resId=0&materialId=slides&confId=230465

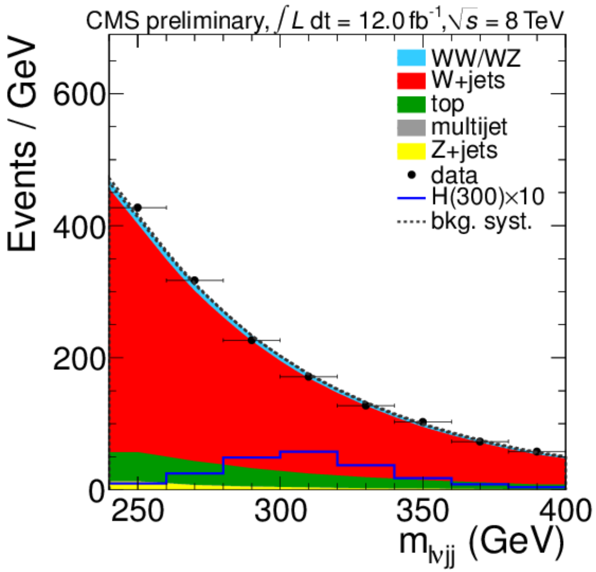
BSM Lineshapes

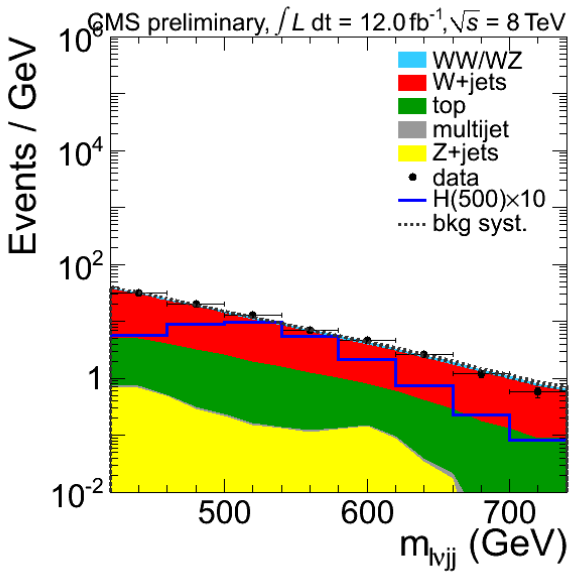
BSM model lineshapes at GEN (top) and RECO (bottom)



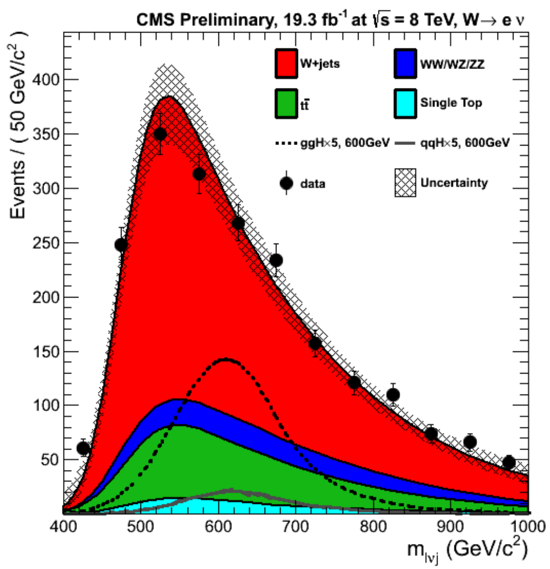
Final m_{WW} spectra

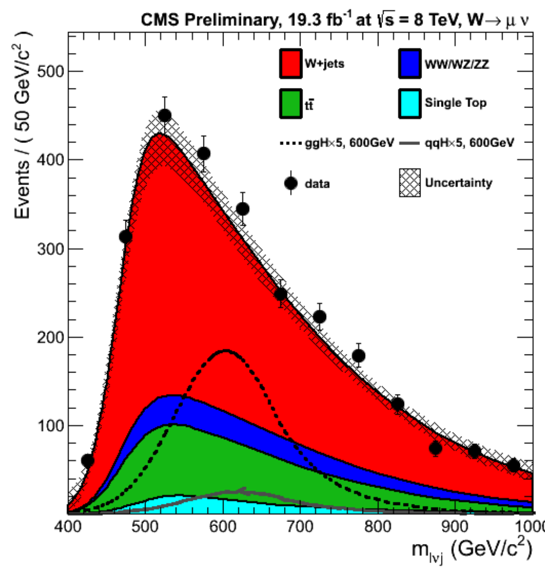


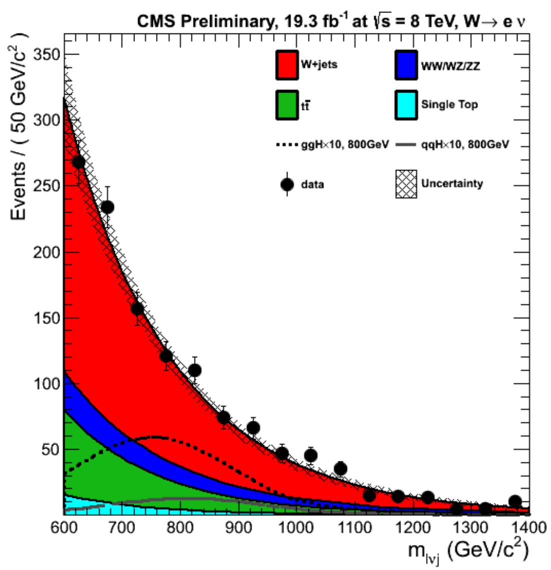


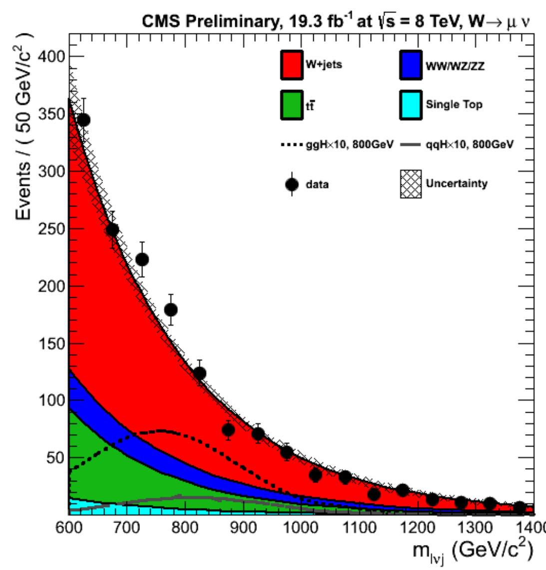


Final Distributions: mH = 600, 800 GeV









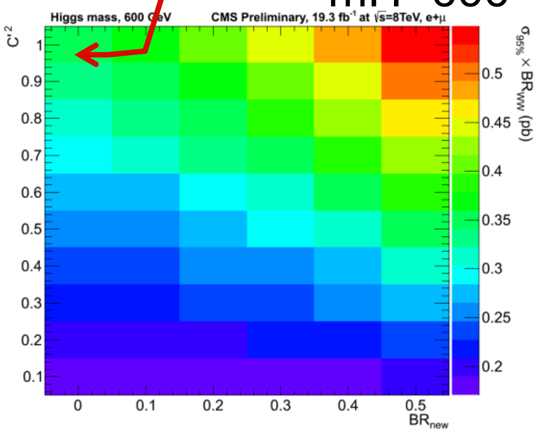
BSM: σ

BR_{new} vs. C'^2

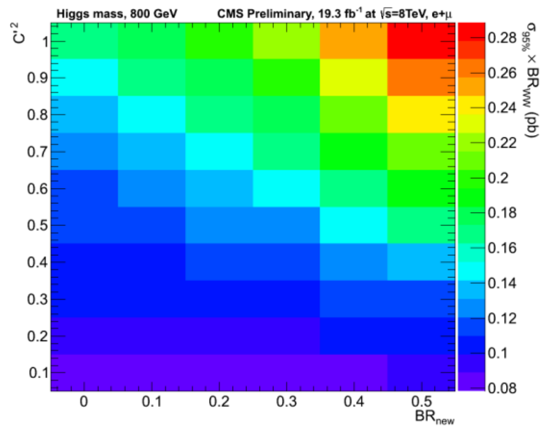
Exp

SM case

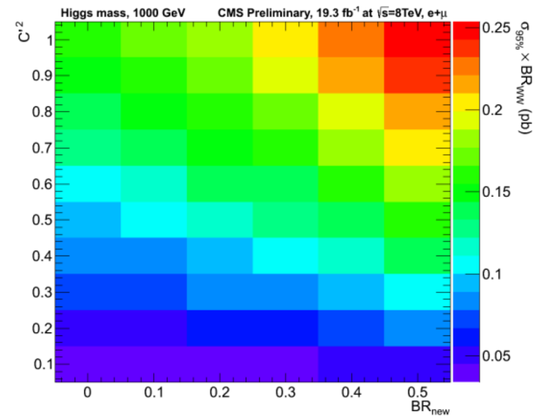
mH=600



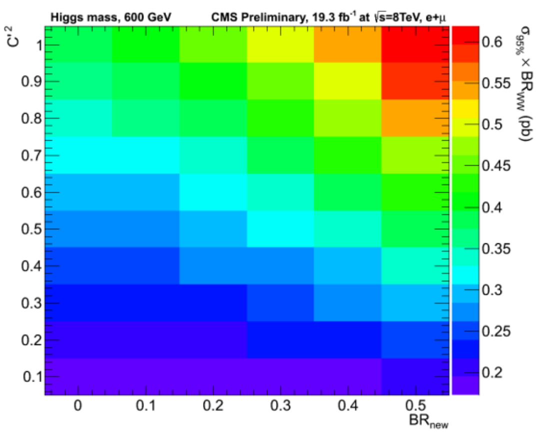
mH=800

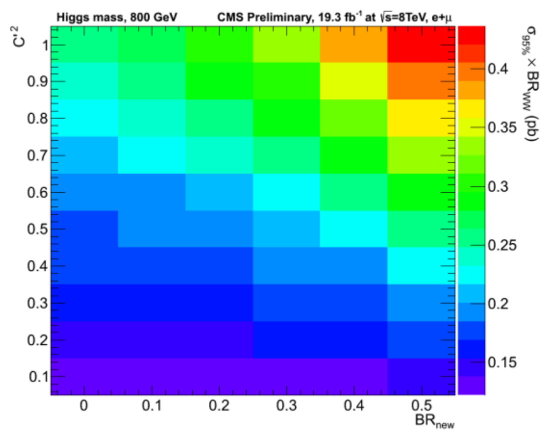


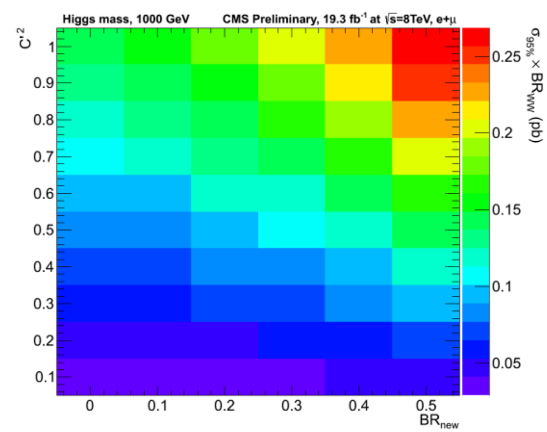
mH=1000



Obs







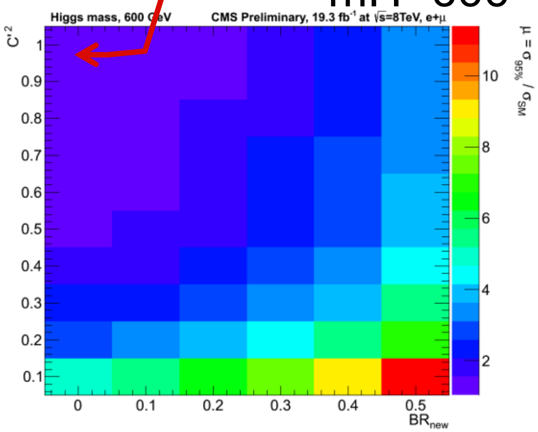
BSM: μ

BR_{new} vs. C'^2

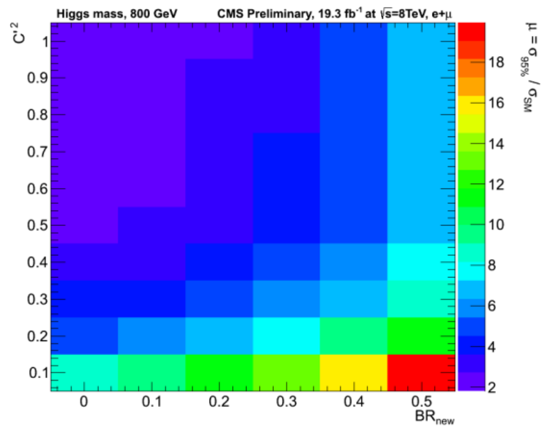
SM case

Exp

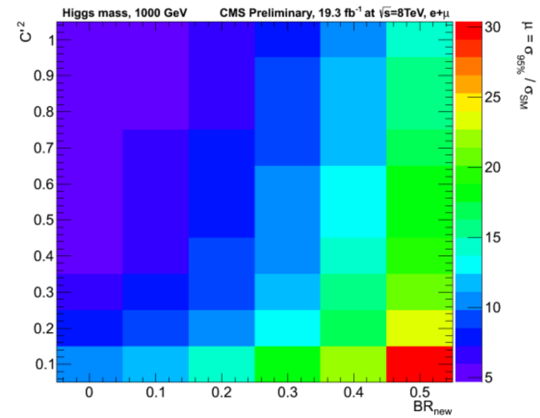
mH=600



mH=800



mH=1000



Obs