Search for a Standard Model-like Higgs boson decaying into $W^+W^- \longrightarrow l\nu q\bar{q}'$ final state in high mass

Bibhuti Parida

for the CMS Collaboration

Tata Institute of Fundamental Research, Mumbai

New Perspectives 2015

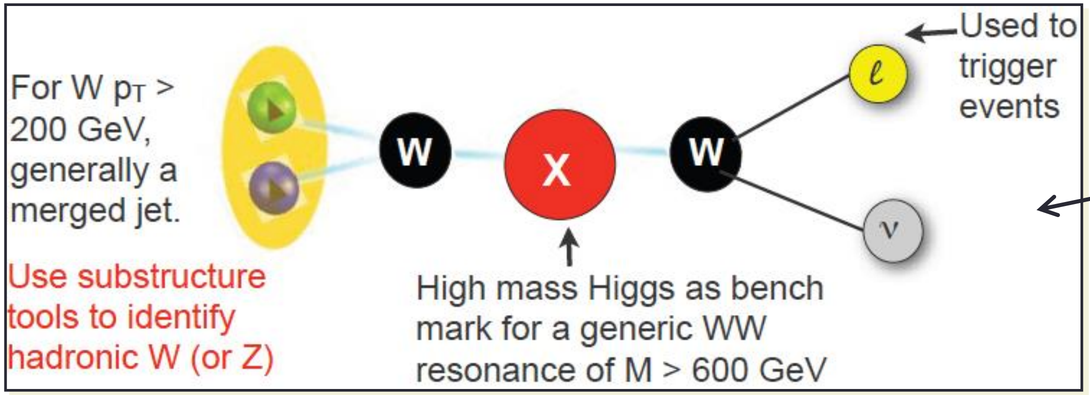
Fermi National Accelerator Laboratory

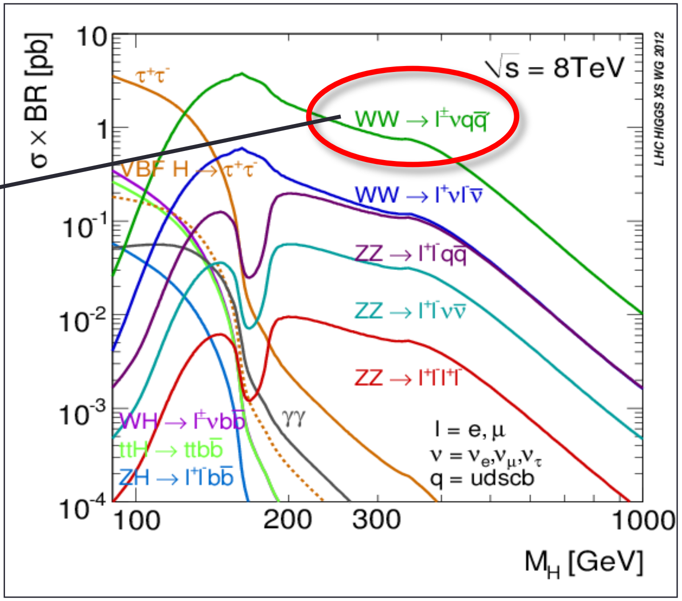
Outline

- Introduction
- Search for a high mass Higgs in $W^+W^- \longrightarrow l\nu qq'$ final state
  - Analysis feature
  - W-tagging
  - Definition of exclusive jet bins
  - Data-MC comparison
- Background extraction
- Statistical interpretation
- Conclusions

Introduction

- Standard Model-like Higgs boson is discovered at ~ 125 GeV
- Still it is important to keep searching in the high mass regime
- We search for a SM-like Higgs Boson decaying into WW $\rightarrow$ lvqq' in the high mass region (600- 1000 GeV)
- The search is performed in the semi-leptonic final state in exclusive jet bins
- This analysis is also the benchmark for future analysis of W^+W^- scattering





Production mechanisms
- Gluon gluon fusion (ggH)
- Vector boson fusion (VBF)

Analysis feature

$W \rightarrow l\nu$
High pT lepton + high MET

$W \rightarrow qq'$
High boost, jets merged together to form a "fat jet"

Signal:
ggH + qqH,
(m_H= 600-1000 GeV)

Backgrounds:
W+ jets (dominant)
ttbar, single top ,
WW/WZ/ZZ

Signature
WW $\rightarrow$ {l + MET} + merged jet

Used **jet substructure techniques** to identify a single jet containing decay products of hadronic W

Kinematics: boosted leptonic W back-to-back to a merged jet
Discriminating observables: pruned jet mass (m_J) and three-body mass ($m_{l\nu J}$)
Unbinned shape limits using $m_{l\nu J}$ distributions

Data:
Full 2012 , $\sqrt{s} = 8$ TeV
Ele + Muon Sample,
19.3 fb^{-1}

Background estimation
Data-driven Background extraction using m_J sideband for W+ jets and top enriched control regions

Object definition and event selection

Muon ID + Selection:

Muon Tight ID

p_T> 24 GeV, $|\eta| < 2.1$

Electron ID + Selection:

Electron MVA ID tight

p_T > 27 GeV, $|\eta| < 2.5$

- MC lepton efficiencies corrected by Tag &Probe on Z

Jets

Merged jet: CA8 with CHS (Loose Jet ID)
- AK7 JEC applied on the CA8

VBF jets: Standard AK5
(Loose Jet ID + loose pileup jet ID)

b-veto: CSV with a medium working point

MET

Type -1 corrected PF MET
+
Φ modulation correction

Leptons: p_T μ (e) > 30 (35) GeV; veto any 2nd μ or e with p_T>10 (20) GeV

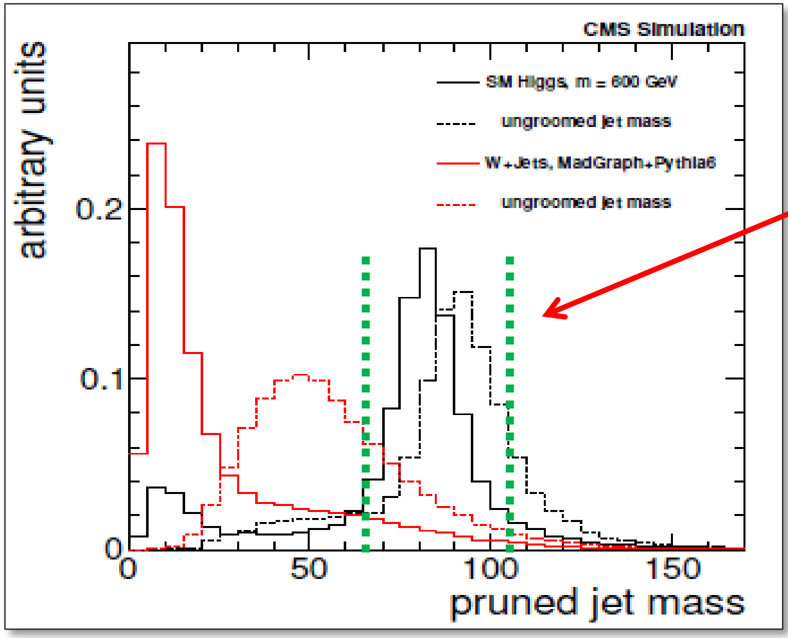
Missing transverse energy: MET μ (e) > 50 (70) GeV

Boosted regime: p_T(Jet) > 200 GeV and p_T(lν) > 200 GeV

Topological back-to-back angular cuts:

$|\Delta\Phi_{W\text{-lep, CA8}}| > 2.0$, $|\Delta\Phi_{MET, CA8}| > 2.0$ and $|\Delta R_{lep, CA8}| > \pi/2$

Selection on the merged jet (W- tagging)



Grooming techniques: remove soft radiation and pileup from the jet (e.g., **pruning**)

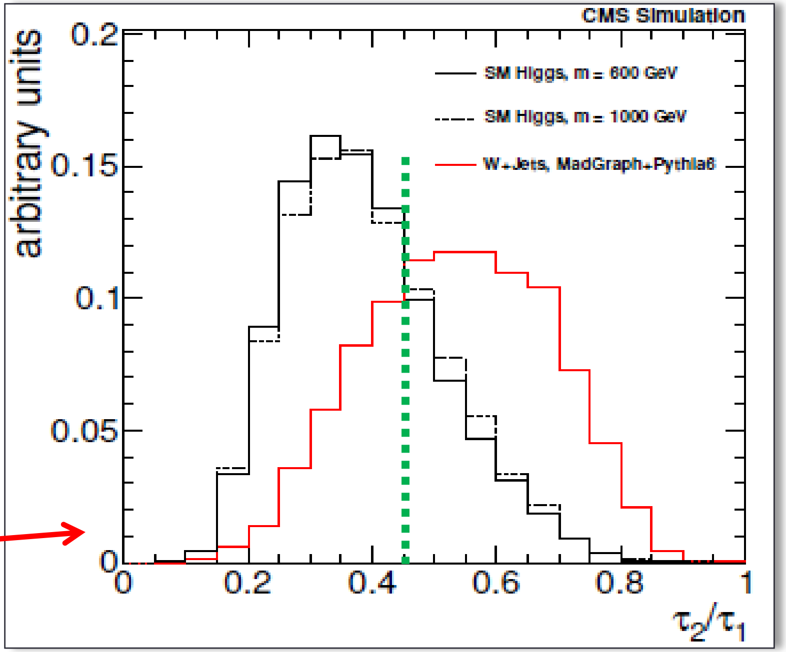
Pruned jet mass: [65-105] GeV → **signal region**
[40,65] and [105,130]GeV → **sideband**

N Subjettiness :

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min\left(\Delta R_{1,k}, \Delta R_{2,k}, \ldots\ldots, \Delta R_{N,k}\right)$$

- τ_N tends to be zero when the jet is consistent with N subjets

- **N-subjettiness ; $\tau_2/\tau_1 < 0.5$**



Categorization of jet bins

- Select the highest p_T CA8 jet as the hadronic W candidate
- Collect the AK5 jets at a distance $\Delta R > 0.8$ from the hadronic W candidate (with p_T > 30 GeV)

if

0 or 1 additional AK5 jets in the event

(0-1 jet bin category)

- Electron channel
- Muon channel

2 or more additional AK5 jets in the event

(2 jet bin category)

(The two AK5 jets with max. pT in the collection are taken as tag jets)

VBF selection

B-veto + top mass veto

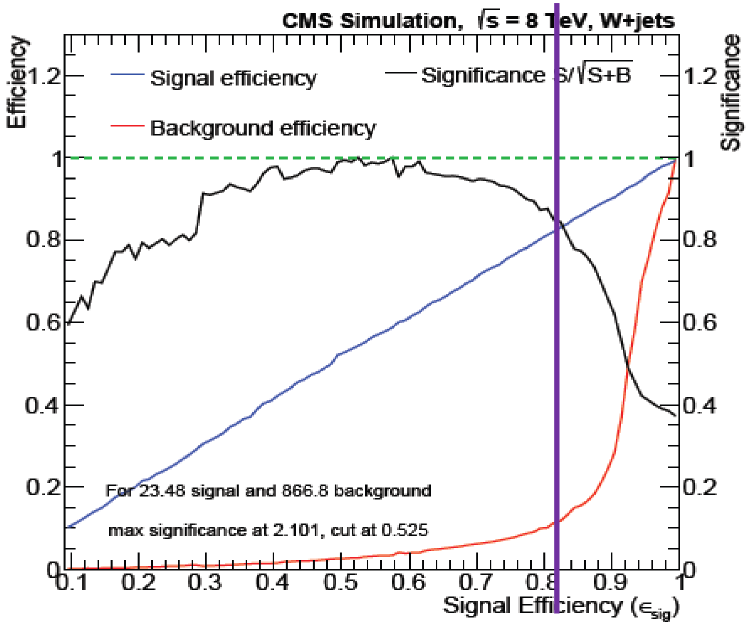
- Electron +Muon (lack of statistics)

2-jet bin: additional selection

VBF jets: Two powerful observables to discriminate signal against background

$\Delta\eta_{jj}$ and M_{jj}

VBF Selection and optimization:



- In the TMVA optimization **VBF = signal**
- ggH is a part of the background together with ttbar, W+Jets, VV and Single-Top in order to enhance the presence of VBF in the final selected sample

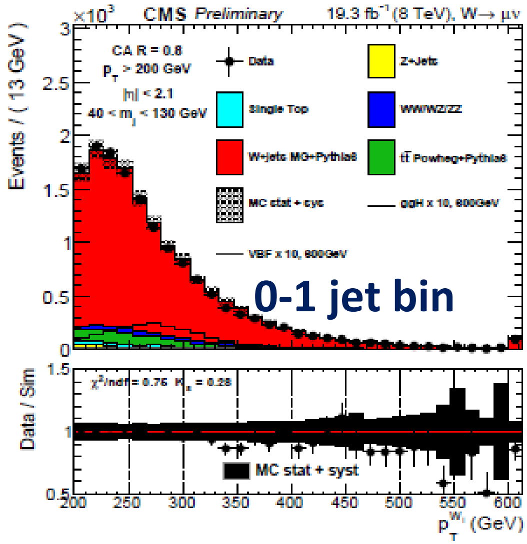
Working point:

$\Delta\eta_{jj}$ > 3.0 && M_{jj} > 250 GeV

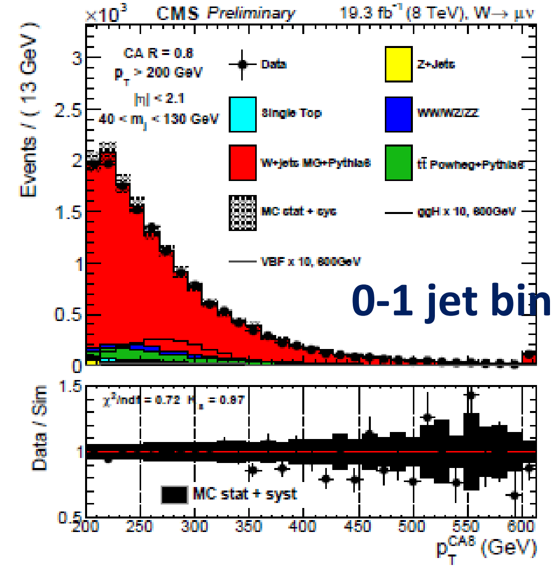
(signal efficiency: ~ 80%)

Data-MC comparison

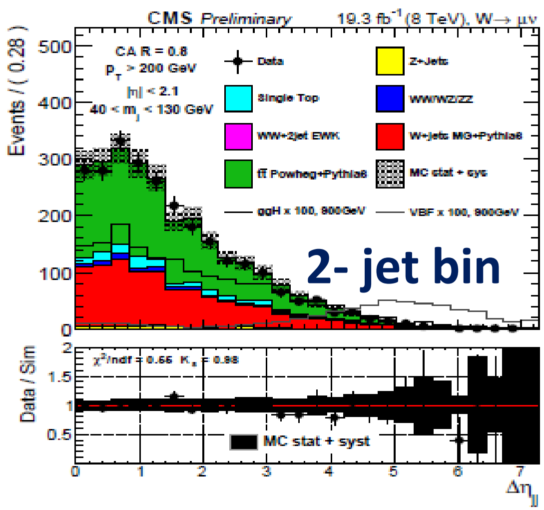
Leptonic $p_T(W)$



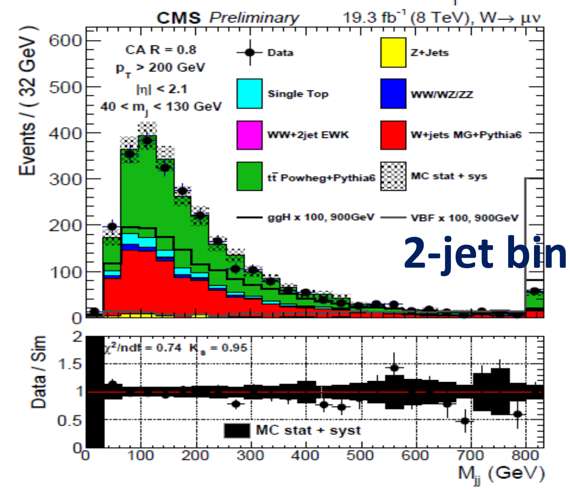
Hadronic $p_T(W)$



$\Delta\eta_{jj}$



M_{jj}



Background extraction

- The limit is calculated from the unbinned shapes in m_{lvJ}

Signal Shapes:

Fits from MC, reweighted line shape

Background Shapes

- **ttbar, Single top, WW/WZ (non-dominant)**
  - Fits from MC, normalization corrected by data-to-MC scale factors
  - ttbar normalization taken from a top-enriched control region
- **W+ jets (dominant)**
  - **Data driven technique:**
    - W+jets normalization estimated from sideband fit of m_j
    - W+jets shape in $\mathbf{m_{lvJ}}$ is taken from the m_J sideband region and then extrapolated into the signal region
    - Extrapolation function ($\mathbf{m_{lvJ}}$ – dependent) is taken from MC

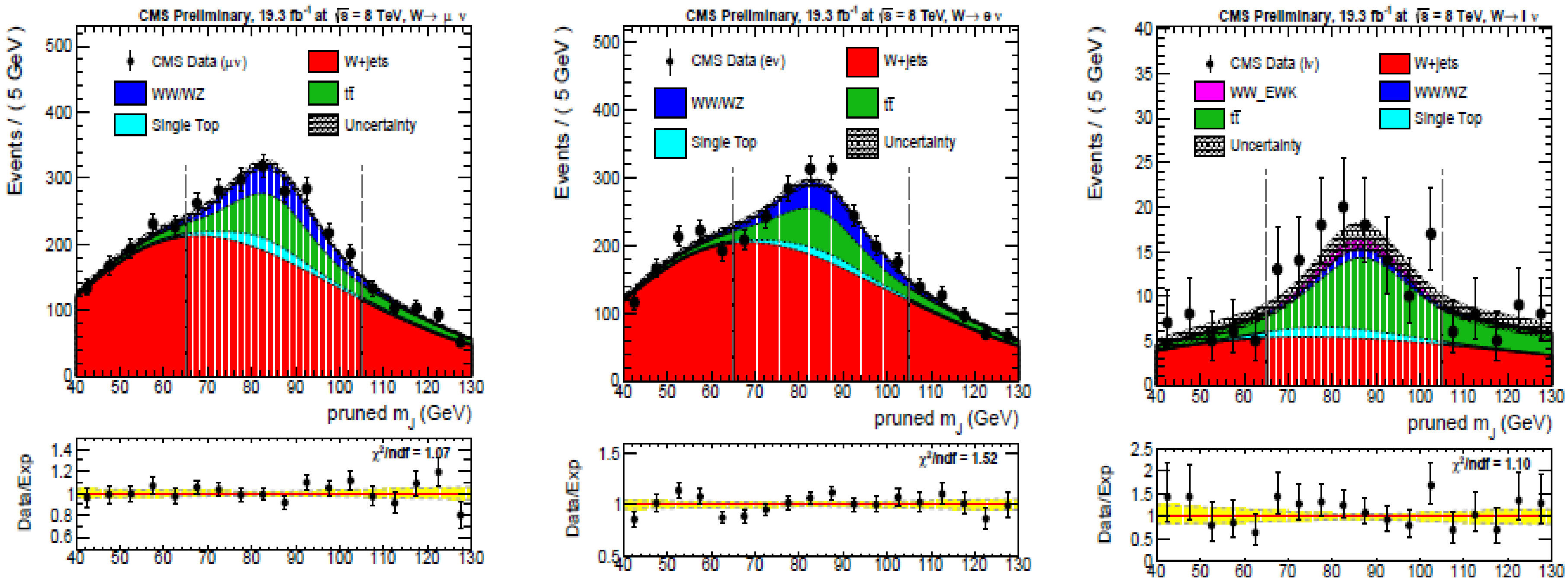
$$F_{\mathrm{data,SR}}(m_{lvj}) = \alpha_{\mathrm{MC}}(m_{lvj}) \times F_{\mathrm{data,LSB}}(m_{lvj})$$

where

$$\alpha_{\mathrm{MC}}(m_{lvj}) = \frac{F_{\mathrm{MC,SR}}(m_{lvj})}{F_{\mathrm{MC,LSB}}(m_{lvj})}$$

and SR= Signal region: m_J = [65-105] GeV
LSB= Low sideband region: m_J = [40-65] GeV & m_J =[105-130] GeV

W + jet Bkg: Normalization procedure (m_J sideband fit)



Ttbar, VV, Single top:
Normalization and shape in Mj → fixed from MC

→

Fit on data in the sideband → extraction of W+jets normalization

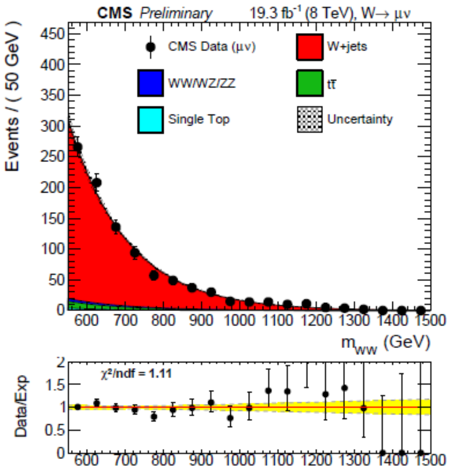
→

Extrapolation of W+jets normalization to the signal region

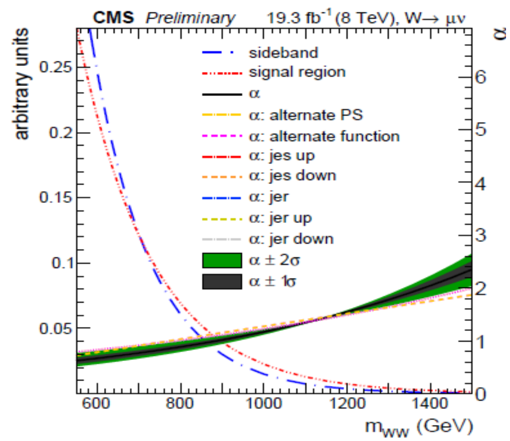
Background estimation

$$F_{\mathrm{data,SR}}(m_{\ell\nu j}) = \alpha_{\mathrm{MC}}(m_{\ell\nu j}) \times F_{\mathrm{data,LSB}}(m_{\ell\nu j})$$

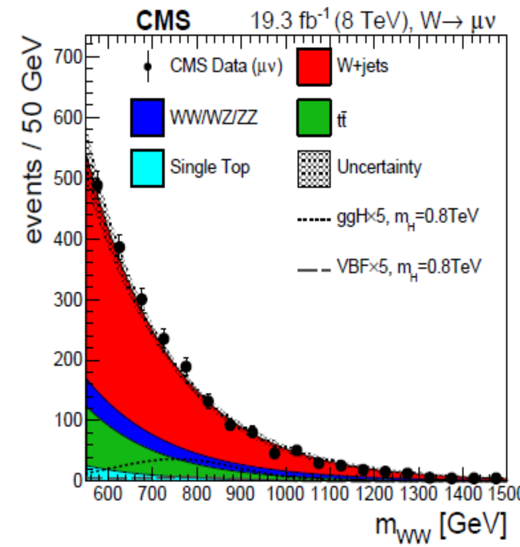
$$\alpha_{\mathrm{MC}}(m_{\ell\nu j}) = \frac{F_{\mathrm{MC,SR}}(m_{\ell\nu j})}{F_{\mathrm{MC,LSB}}(m_{\ell\nu j})}$$

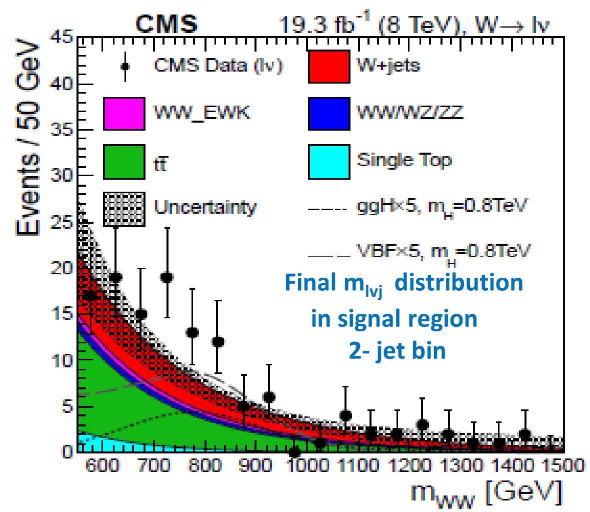


X



=





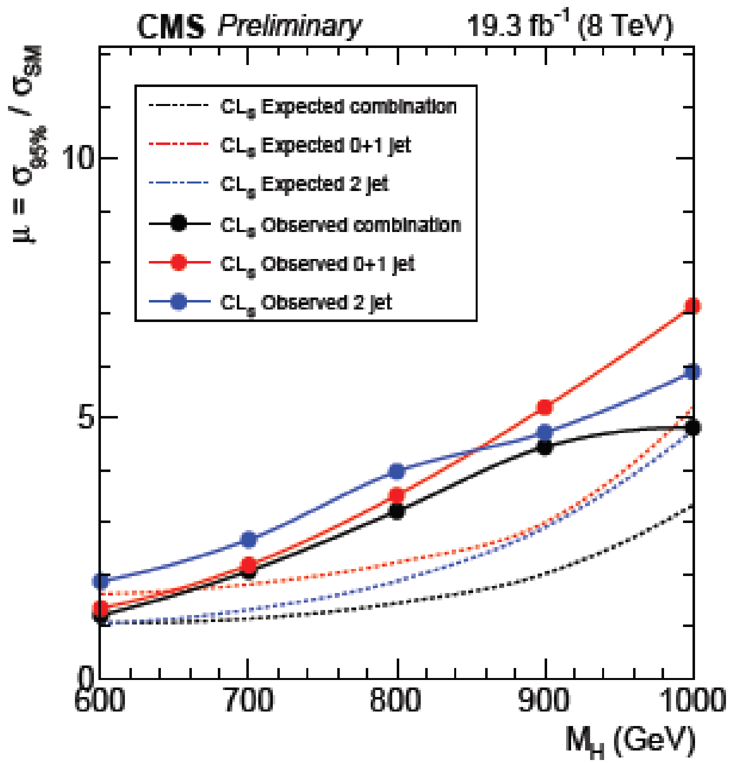
Uncertainties on W+ jets shape/normalization from:

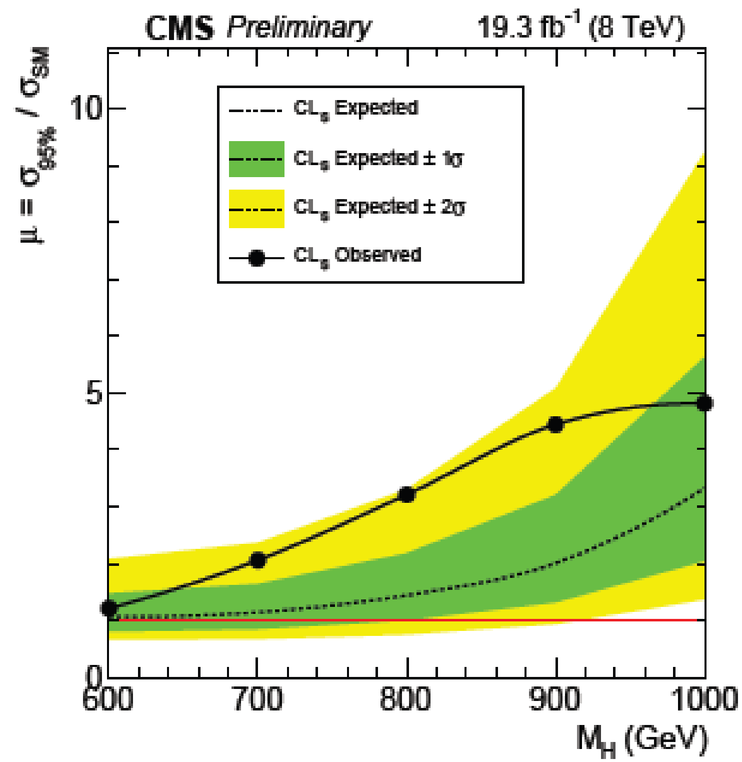
W + jet sideband fit

α function fit shape uncertainty

Shape uncertainty from alternate parton shower and alternate fitting function

Statistical interpretation





- No significant excess observed in the investigated mass region
- An expected sensitivity to exclude at 1.1 (3.3) times the SM cross section for a Higgs boson mass of 600 (1000) GeV

Conclusions

- A search is presented for a heavy Higgs boson (m_H= 600-1000 GeV) in the semi-leptonic WW final state in exclusive jet bins
  - Novel jet substructure techniques are employed in the search
  - Limits are presented for the SM Higgs case for 0+1 and 2-jet bin case
  - No evidence for a SM-like Higgs boson is found in the investigated region.

Thank you

Data and MC Samples

- We use Prompt-Reco dataset
- Pat-Tuples frozen from spring 2012 , processed in CMSSW_5_3_X

Monte Carlo:

Signal : SM Higgs, qqH & ggH at 8TeV-powheg-pythia
(Private production) [m_H =600-1000GeV]

Background:

W+ Jets: (Boosted W+1 jet sample, Herwig++ and madgraph+Pythia6)
exclusive MG+Pythia6 samples (Exclusive W+1jet, W+2Jets , W+3Jets etc.)

ttbar: Powheg+ pythia6 (also mc@NLO: to assign sys uncertainty)

Single top: Powheg+ pythia6

VV (WW;WZ;ZZ): Pythia6

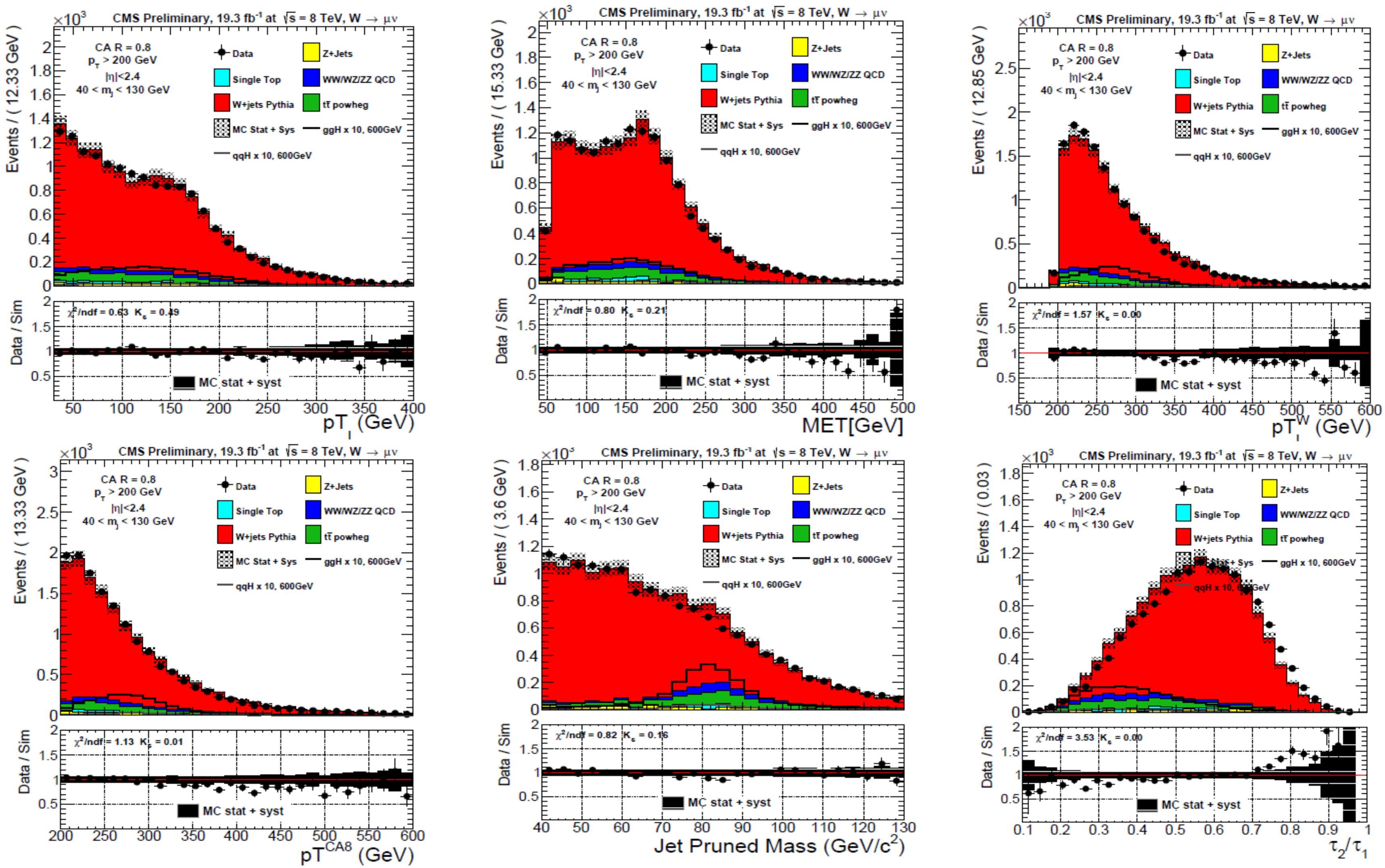
WW EWK: phantom + pythia6 (private production)

DY + Jets: Madgraph+ pythia6

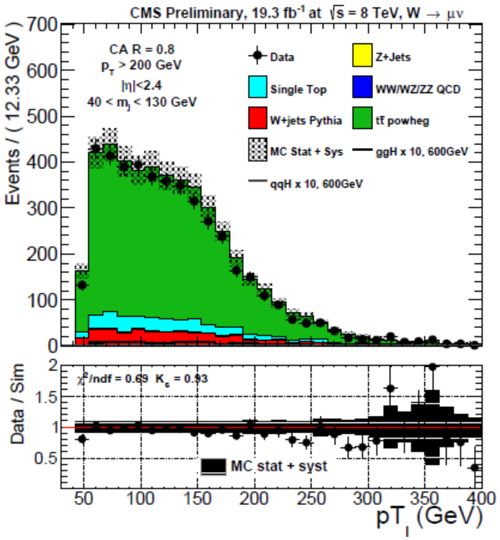
Data:

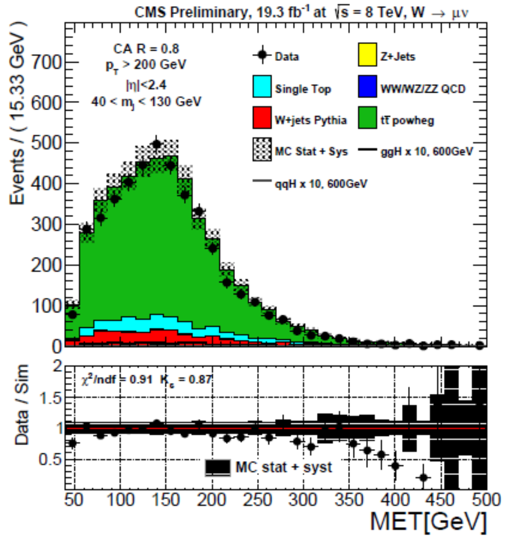
Full 2012 Ele + Muon,
$\sqrt{s} = 8$ TeV Sample,
19.3 fb^{-1}

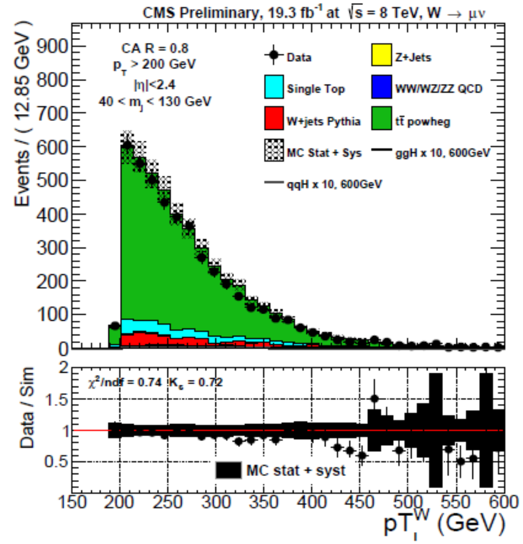
Data-MC comparison at pre-selection level (0+1 jet bin muon channel)

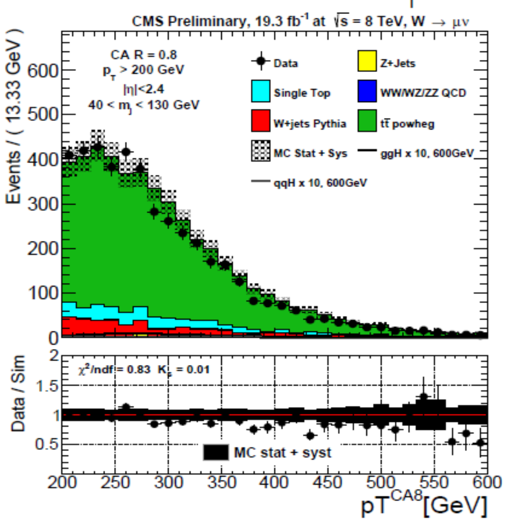


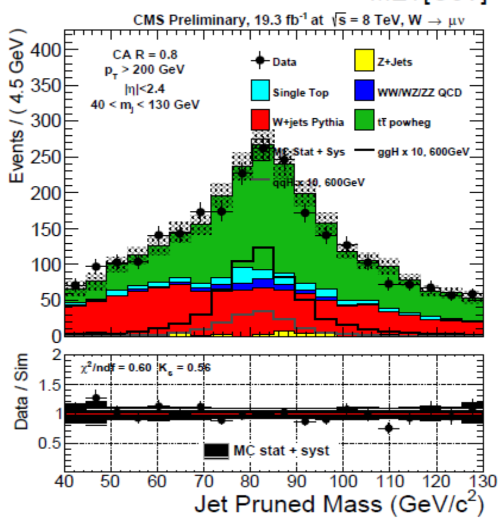
Data-MC comparison at pre-selection level (2- jet bin muon channel)

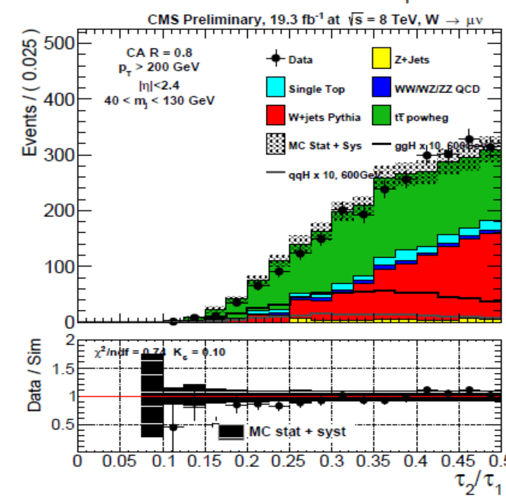




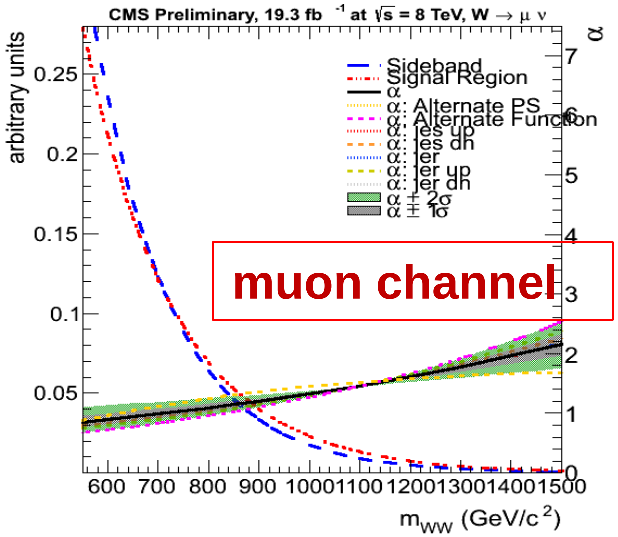




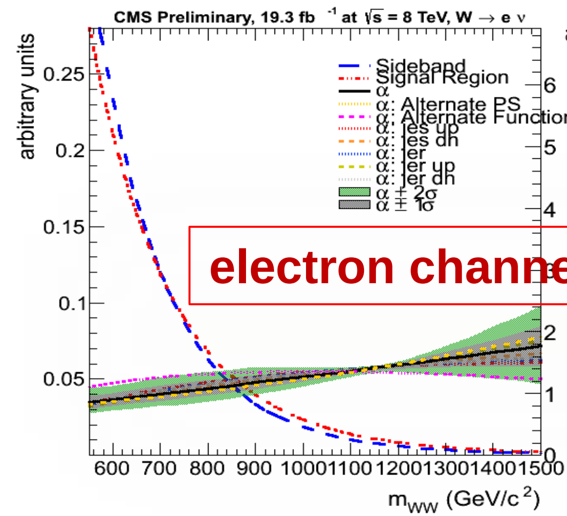


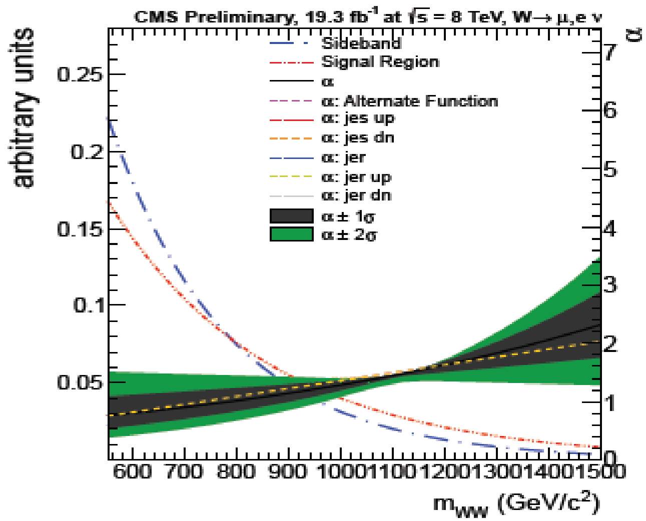


extrapolation function α_{MC} (m_{lvJ})



0-1 jet bin case





$$\alpha_{MC}(m_{lvj}) = \frac{F_{MC,SR}(m_{lvj})}{F_{MC,LSB}(m_{lvj})}$$

Systematics

Syst. uncertainty	W+jets	$t\bar{t}$	single t	VV	W+jets	$t\bar{t}$	single t	VV	WW_{ewk}
Luminosity	-	2.6%	2.6%	2.6%	-	-	2.6%	2.6%	2.6%
Bkg. Cross Sec.	-	-	30%	20%	-	-	30%	20%	20%
Trigger Eff.	-	1%	1%	1%	-	1%	1%	1%	1%
Lepton Eff.	-	2%	2%	2%	-	2%	2%	2%	2%
B-Tagging	-	1.7%	3.3%	0.6%	-	1.5%	3%	0.5%	0.7%
W-Tagging	-	-	-	9.3%	-	-	-	9.3%	9.3%
Top Normalization	-	6.5%	-	-	-	26.5%	-	-	-
W+jet Normalization	5-8%	-	-	-	22%	-	-	-	-
Lepton Scale	-	0.4%	-	1%	-	0.5%	-	1.5%	1%
Lepton Res.	-	-	-	-	-	-	-	-	0.8%
Jet Scale (JES)	2.7%	4%	4.1%	3%	2.1%	4.1%	7.1%	7.5%	4.6%
Jet Res. (JER)	1%	0.4%	0.9%	0.7%	1.9%	3.1%	8.3 %	4.3%	6.3%

List of systematic uncertainties on background normalisation: left part of the table refers to 0+1 jet bin, right to 2-jet bin category.

Background Systematics

Syst. uncertainty	ggH	VBF	ggH	VBF
Luminosity	2.6%	2.6%	2.6%	2.6%
PDF gg	-	9.1%*	-	9.1%*
PDF qq	-	5%*	-	5%*
ggH0In	26%	-	-	-
ggH2In	6%	-	19%	-
Int ggH	10%	-	10%	-
Int vbfH	-	10%	-	10%
Trigger eff.	1%	1%	1%	1%
Lepton eff.	2%	2%	2%	2%
B-Tagging	0.5%*	0.2%*	0.5%*	0.2%*
W-Tagging	9.3%	9.3%	9.3%	9.3%
Lepton Scale	2.1%*	1.5%*	3.5%*	1.8%*
Jet Scale (JES)	3.9%*	4.4%*	5.0%*	4.5%*
Jet Res (JER)	2.5%*	3.5%*	8.0%*	10.6%*

List of systematic uncertainties on signal (ggH and VBF) normalisation: left part of the table refers to 0+1 jet bin, right to 2-jet bin category.(* stands for mass dependent systematics)

Signal Systematics