



# Search for the Standard Model Higgs boson in $H \rightarrow WW \rightarrow \ell \nu q \bar{q}$

$\ell$ nujj analysis team

Using 8 TeV 2012 data, and combined result of 7+8 TeV

**5.1 fb<sup>-1</sup> top-up approval, July 3, 2012**

# Data & MC samples and trigger



dataset name
/SingleElectron/Run2012A-PromptReco-v1/AOD
/SingleElectron/Run2012B-PromptReco-v1/AOD
/SingleMu/Run2012A-PromptReco-v1/AOD
/SingleMu/Run2012B-PromptReco-v1/AOD

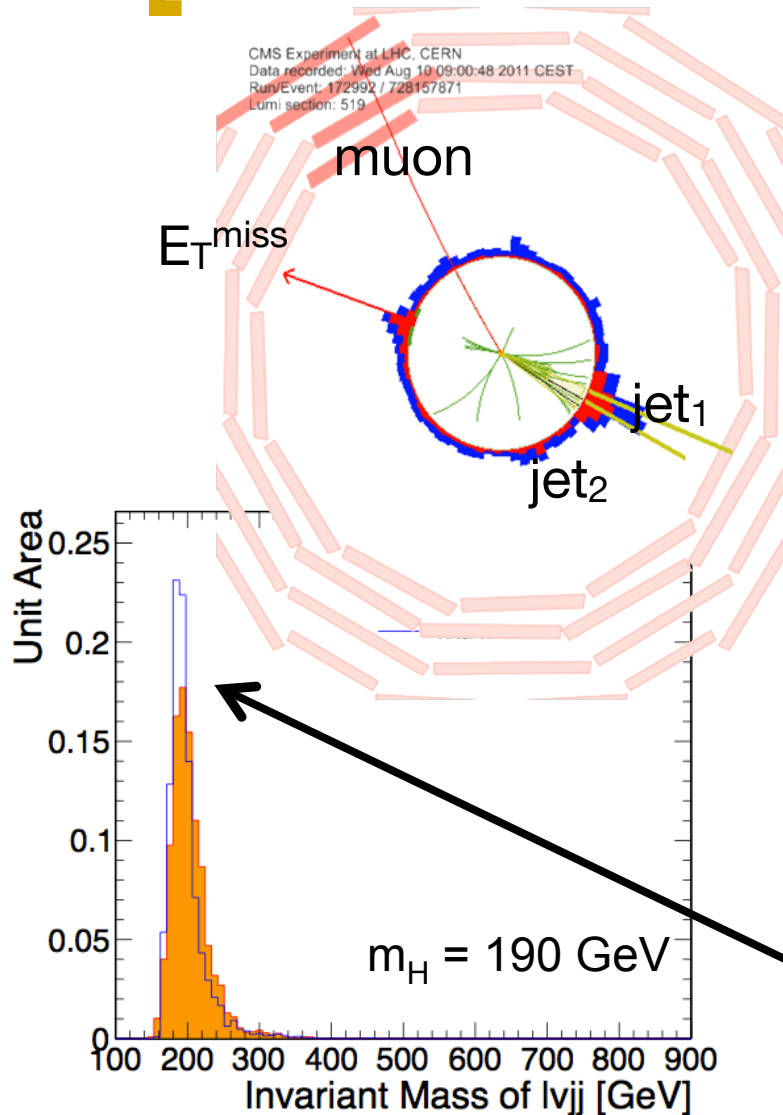
Certified “golden” JSON  
(**5.1 fb<sup>-1</sup>**) used

electron trigger paths	run range
HLT_Ele27_WP80_v*	190456- ICHEP cutoff

muon trigger paths	run range
HLT_IsoMu24_v*, HLT_IsoMu24_eta2p1_v*	190456- ICHEP cutoff

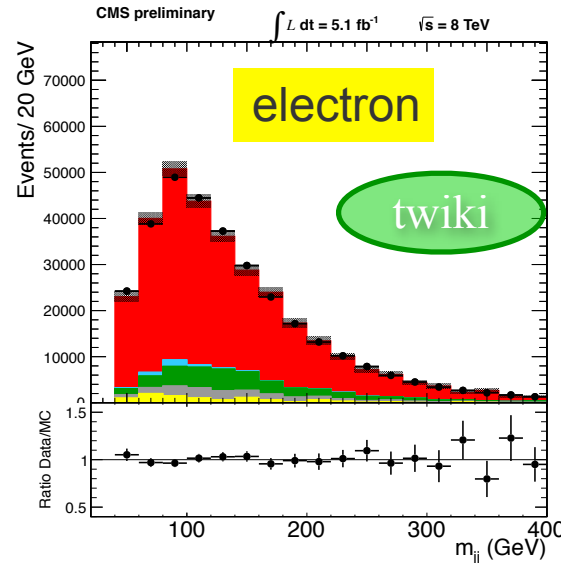
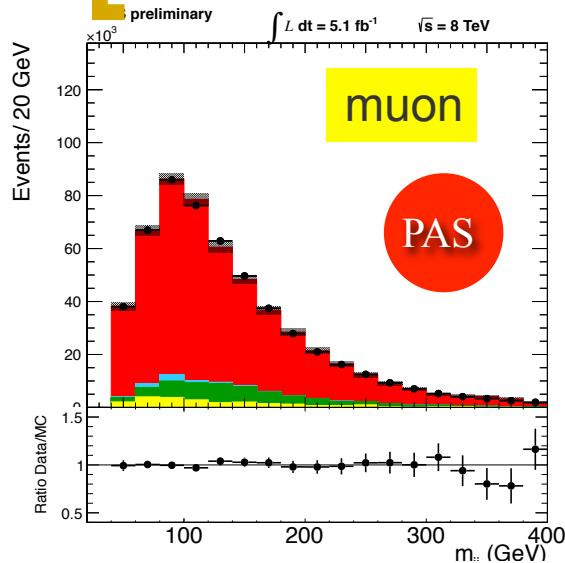
11 signal mass points at 8 TeV: **180, 190, 200, 250, 300, 350, 400, 450, 500, 550, 600 GeV**. Same granularity as last year, and same granularity as other high mass analyses.

# Event (pre-) selection

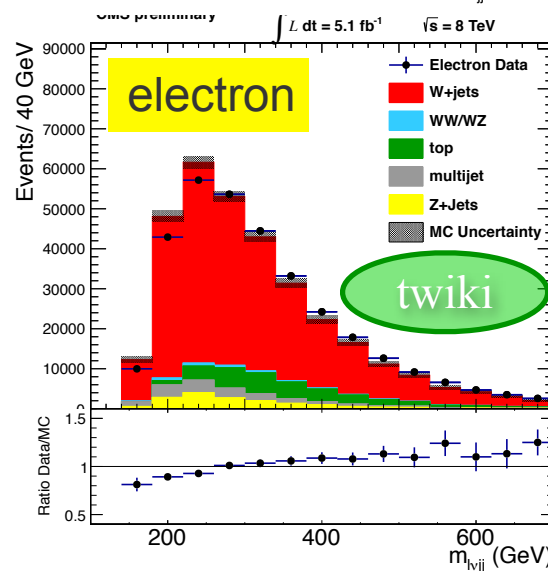
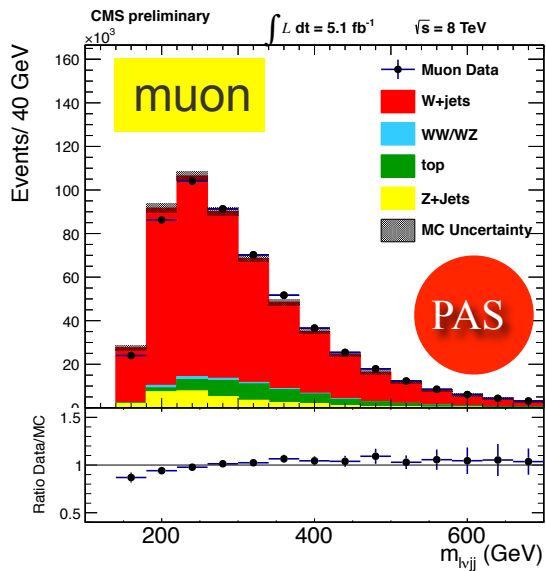


- ◆ One isolated, high- $p_T$  lepton
  - $p_T > 35$  (25) GeV for electrons (muons)
- ◆ High  $E_T^{\text{miss}}$  from 1 neutrino
  - $E_T^{\text{miss}} > 30$  (25) GeV for electrons (muons)
  - $m_T(\text{lepton} + E_T^{\text{miss}}) > 30$  GeV
- ◆ Two high  $p_T$  jets with  $m_{jj} \sim 80$  GeV
  - Anti-Kt 0.5 particle flow jets
  - $p_T > 30$  GeV,  $|\eta| < 2.4$
  - $\Delta R(\text{jet-lepton}) > 0.3$
  - $N_{\text{extra-jets}} = 0, 1$
- ◆ WW inv. mass  $\rightarrow$  reconstruct Higgs signal
  - neutrino  $p_z$  from  $m_W$  constraint
  - We do a kinematic fit on lepton,  $E_T^{\text{miss}}$ , hadronic W to improve Higgs mass resolution and to remove the correlation between  $m_{WW}$  from  $m_{jj}$ .

# Data/MC comparison @pre-sel level: $m_{jj}$ , $m_{lvjj}$



Corrected MC error calculation that takes into account the pileup reweighting efficiency (~30%)

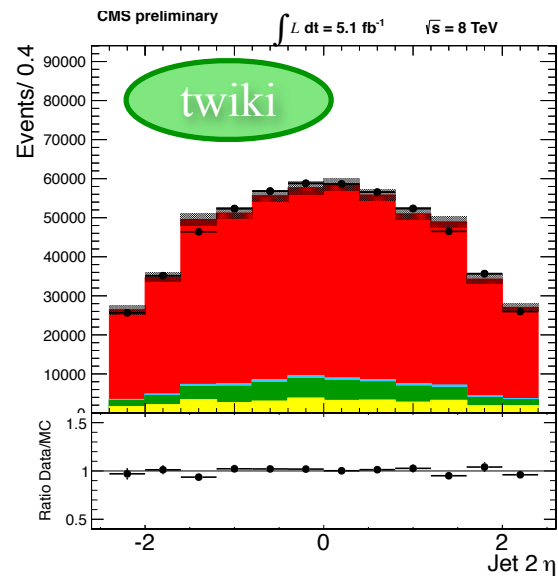
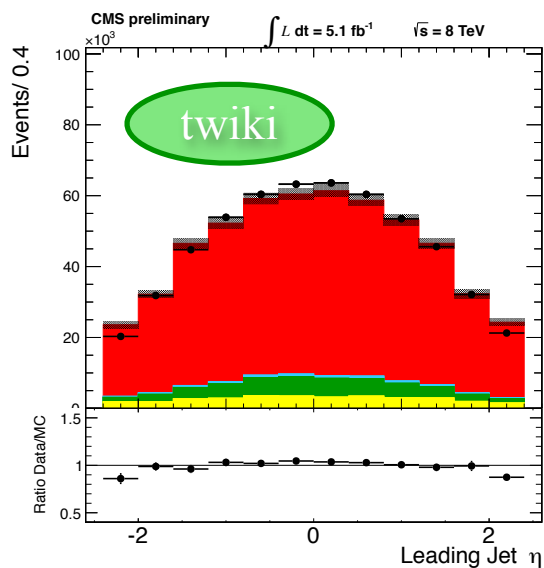
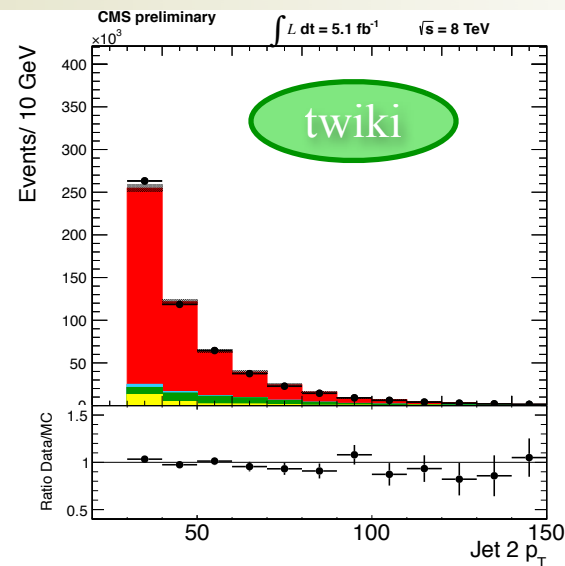
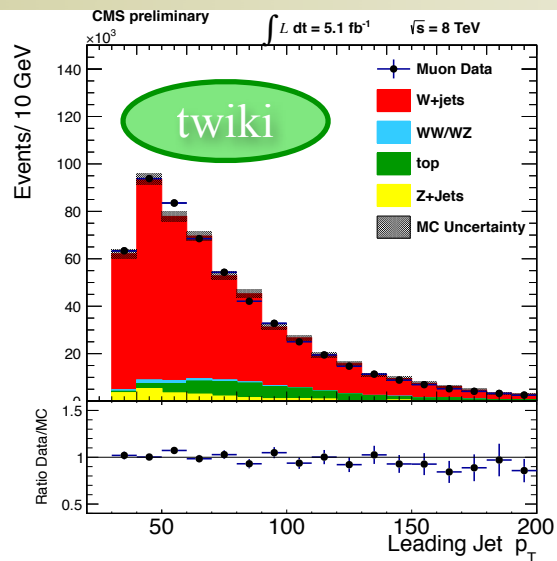


Disagreement seen here is understood. The LO MadGraph simulation doesn't get the  $m_{lvjj}$  spectrum right. Hence, we use data-driven technique for W+jets shape.

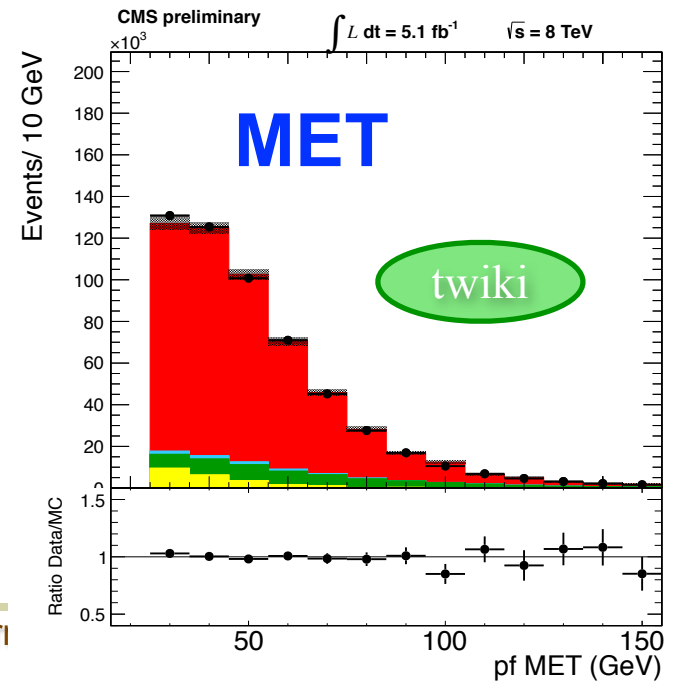
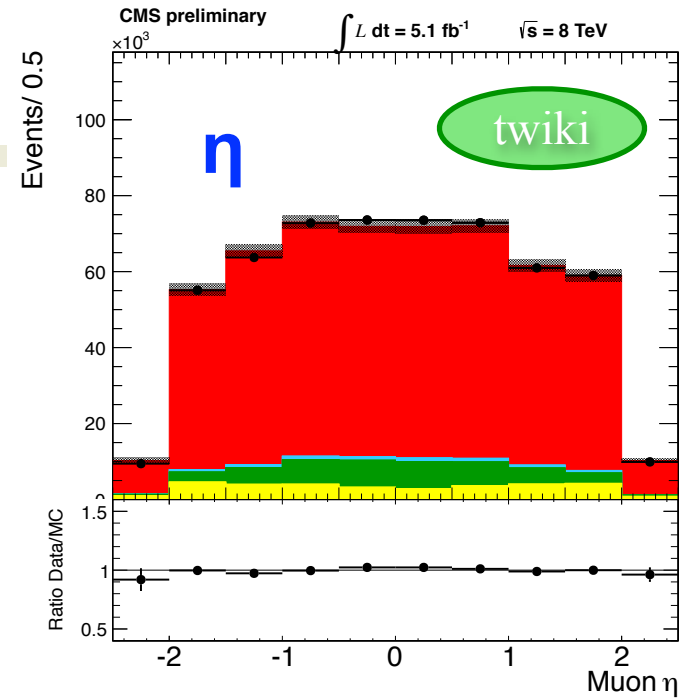
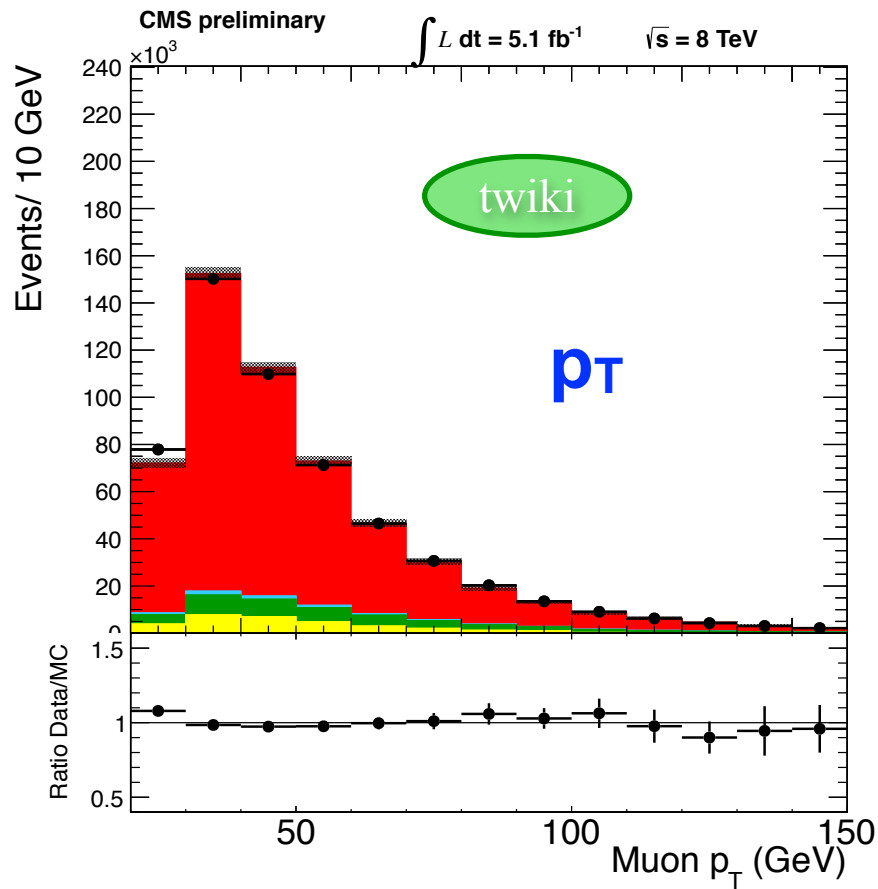
# Comparison plots: jets



muon data

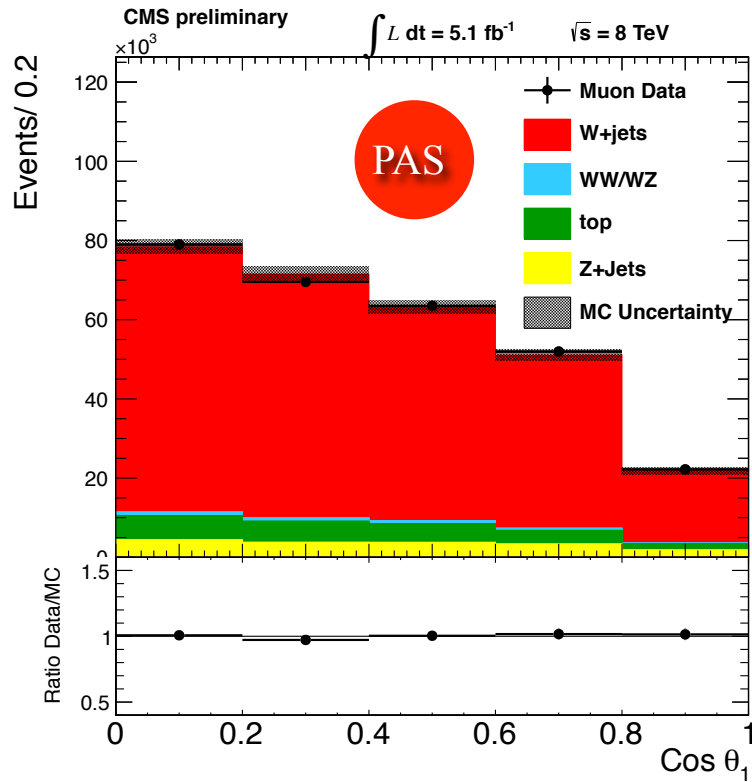


# Comparison plots: leptons



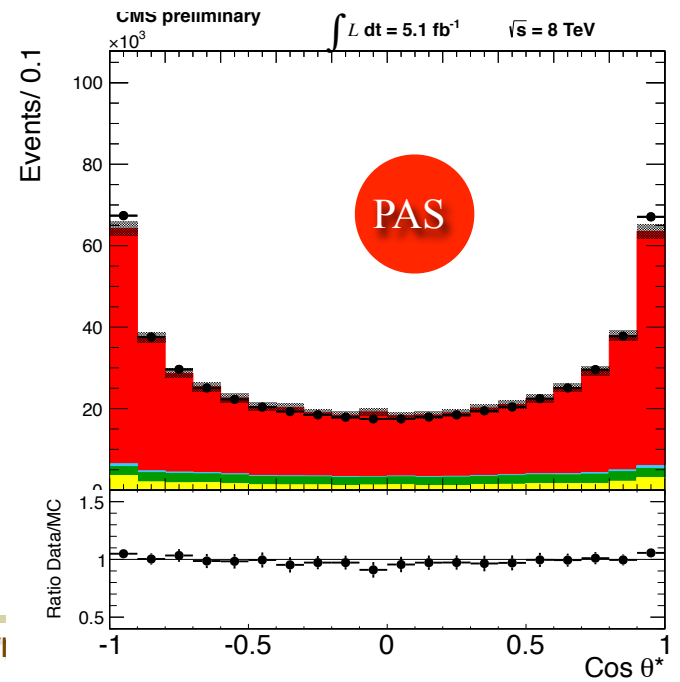
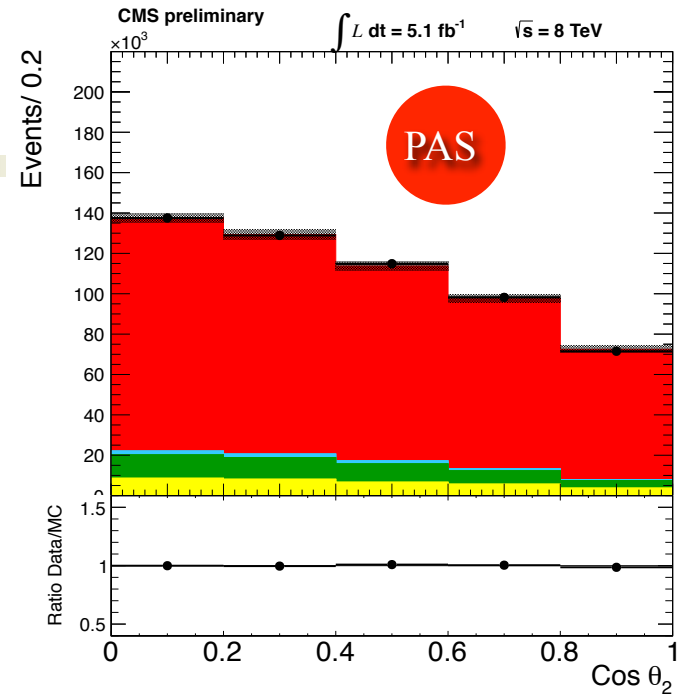
# Likelihood input variables (I)

muon data



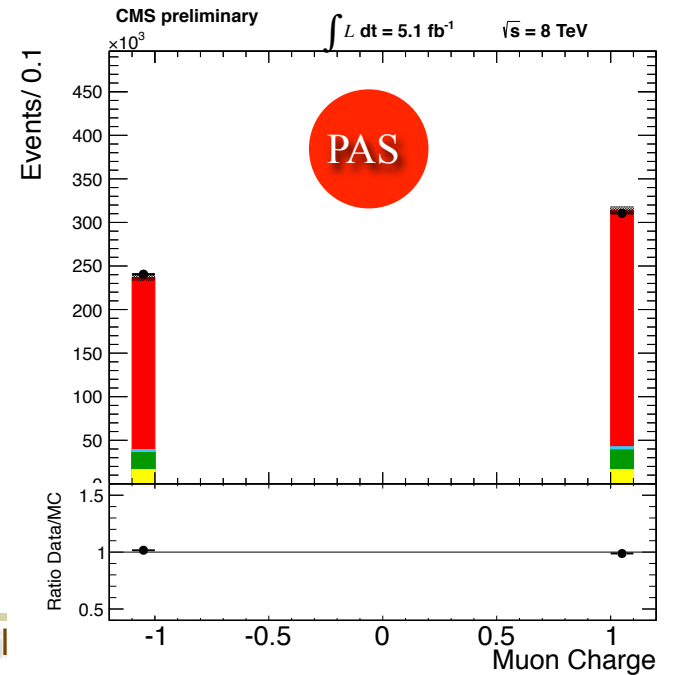
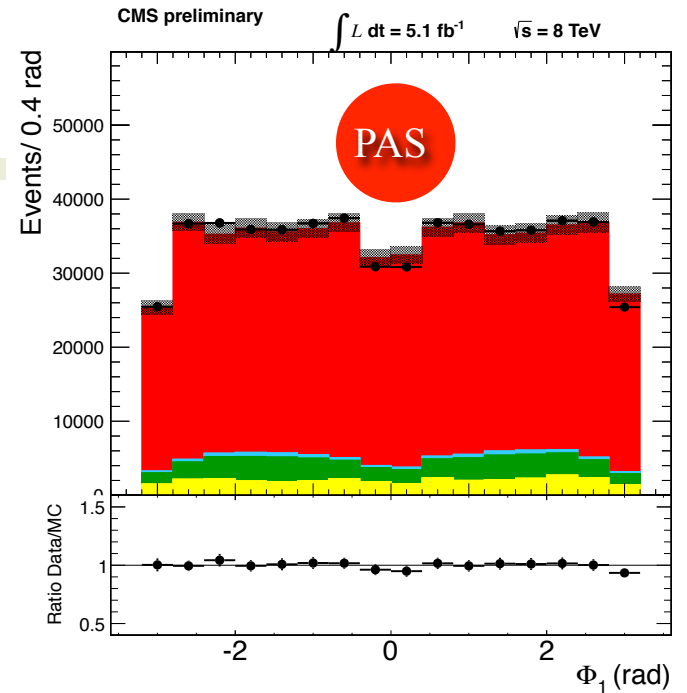
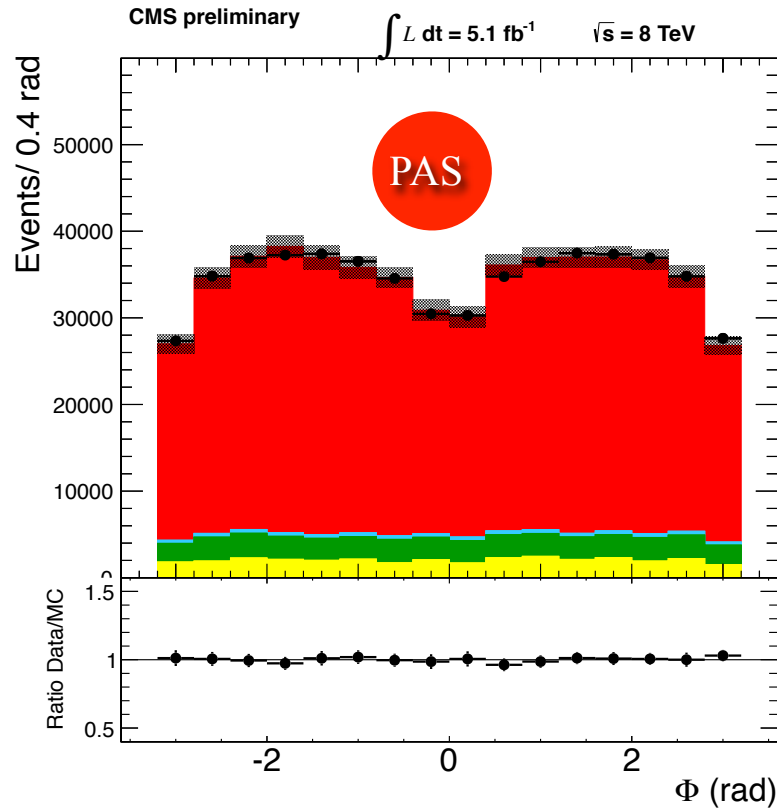
Reasonable modeling of kinematics in simulation.

Kalanand Mishra, Feri



# Likelihood input variables (II)

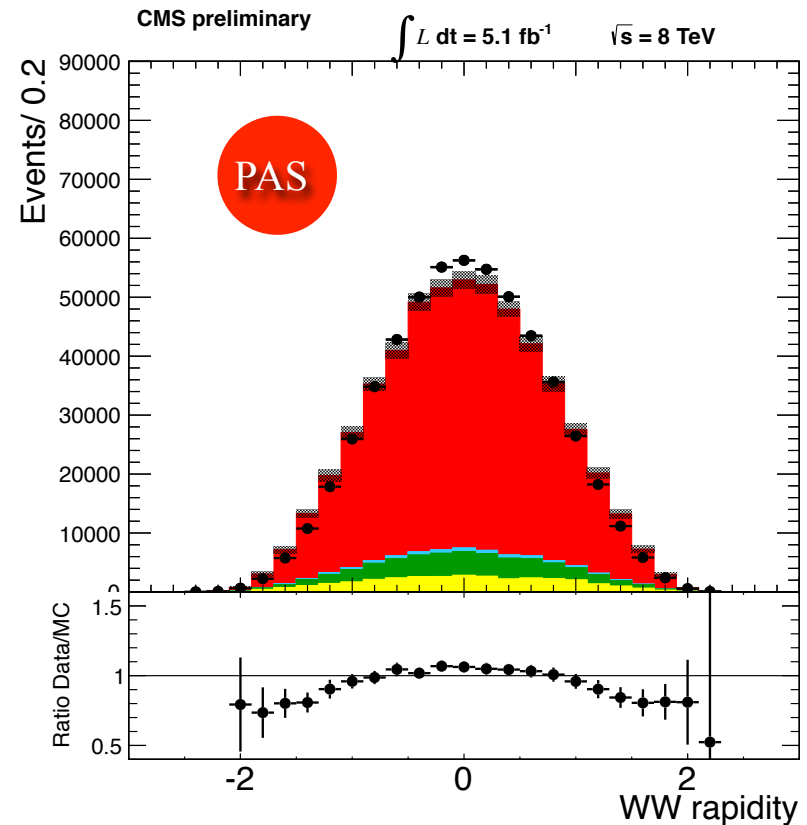
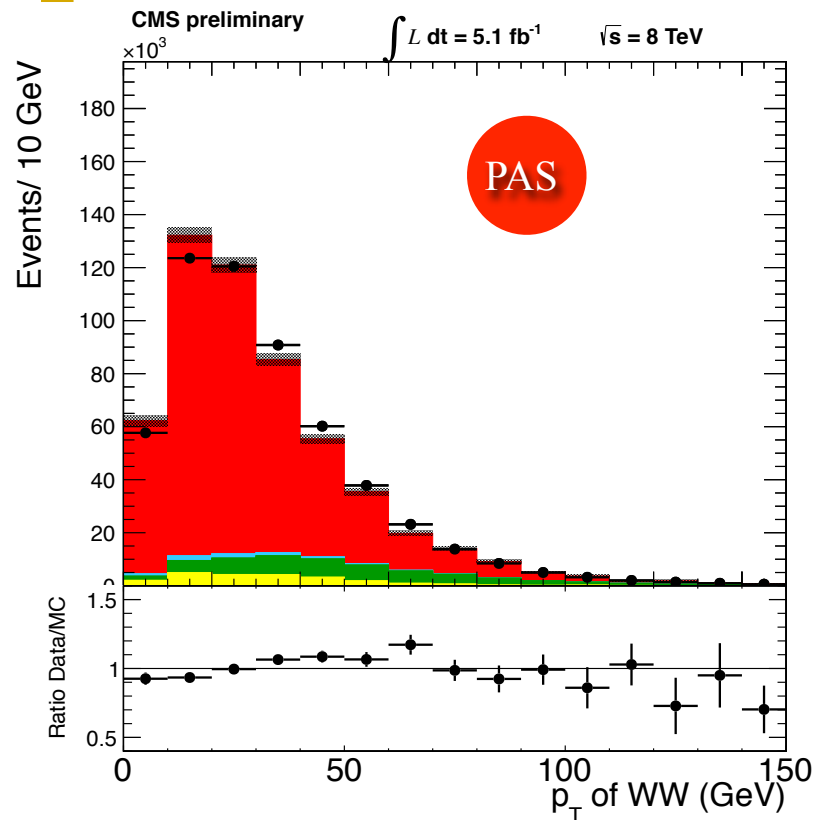
muon data



**Reasonable modeling of kinematics  
in simulation.**



# Likelihood input variables (III)

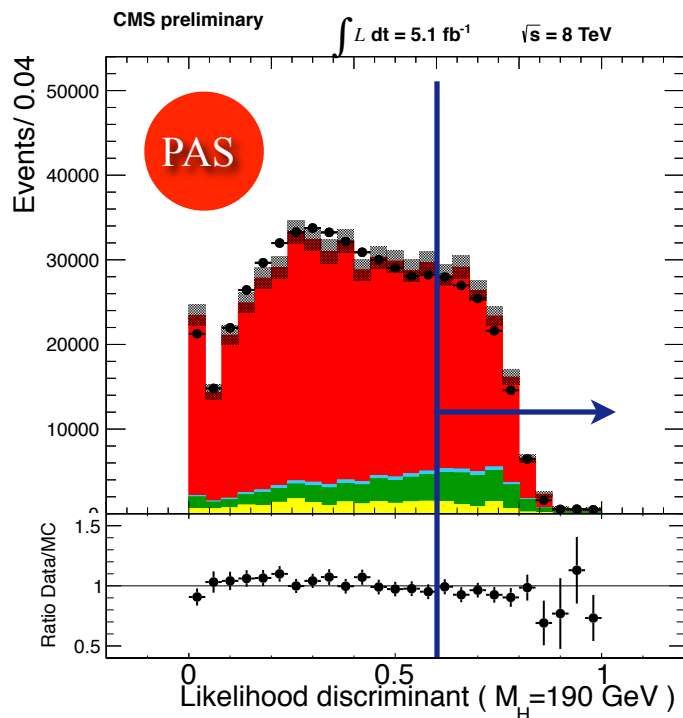


**Agreement at the same level as in last year's data.**

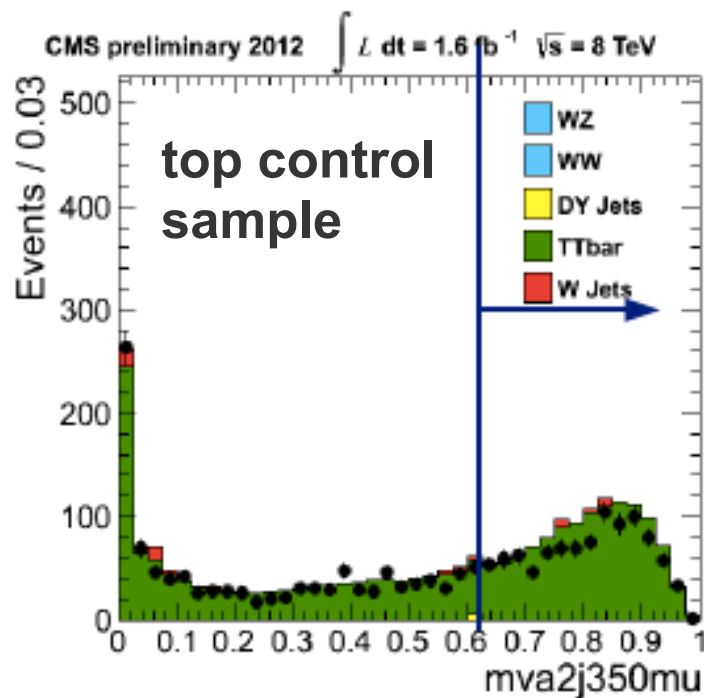


# Example of likelihood output

- We look at data/MC comparison of the likelihood for the analysis final state, dominated by **W+jets** (left), and for a **tt-enriched** final state (right)
- Residual differences between tt-data and tt-MC taken as **systematic error on MVA-signal efficiency**



H190 - mu - 2j

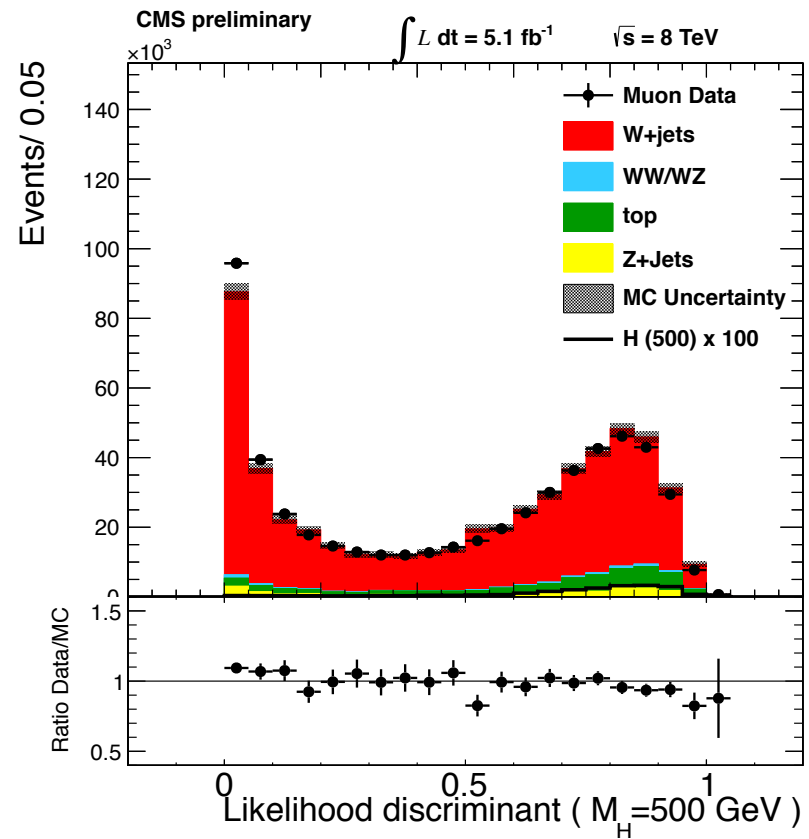
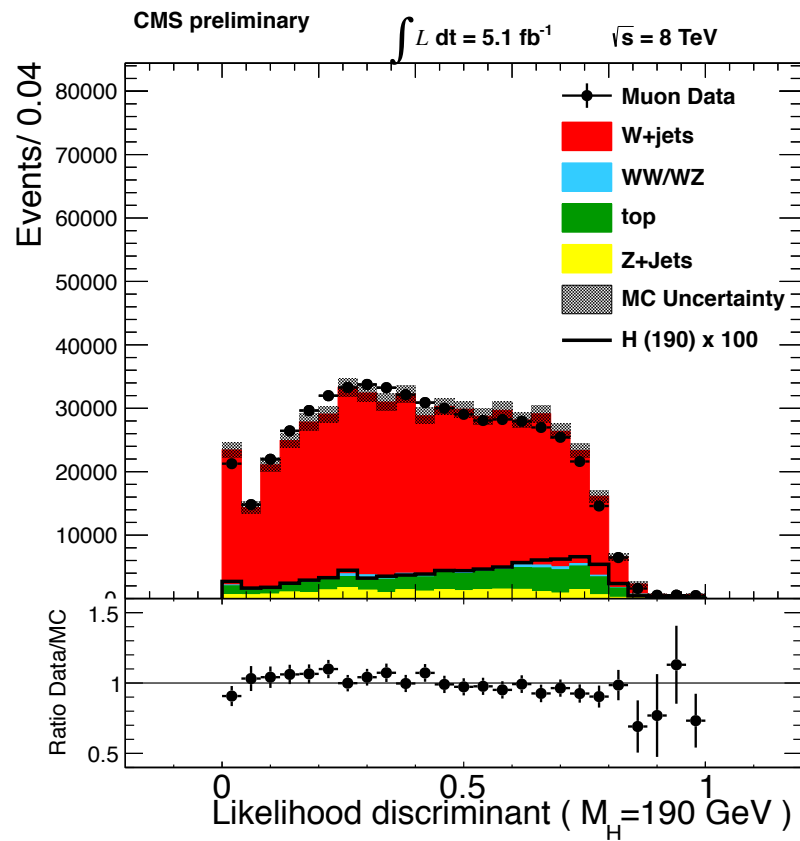


H350 - mu - 2j

# Plots requested by ARC (Vitaliano)



Show signal shape in the MVA distribution

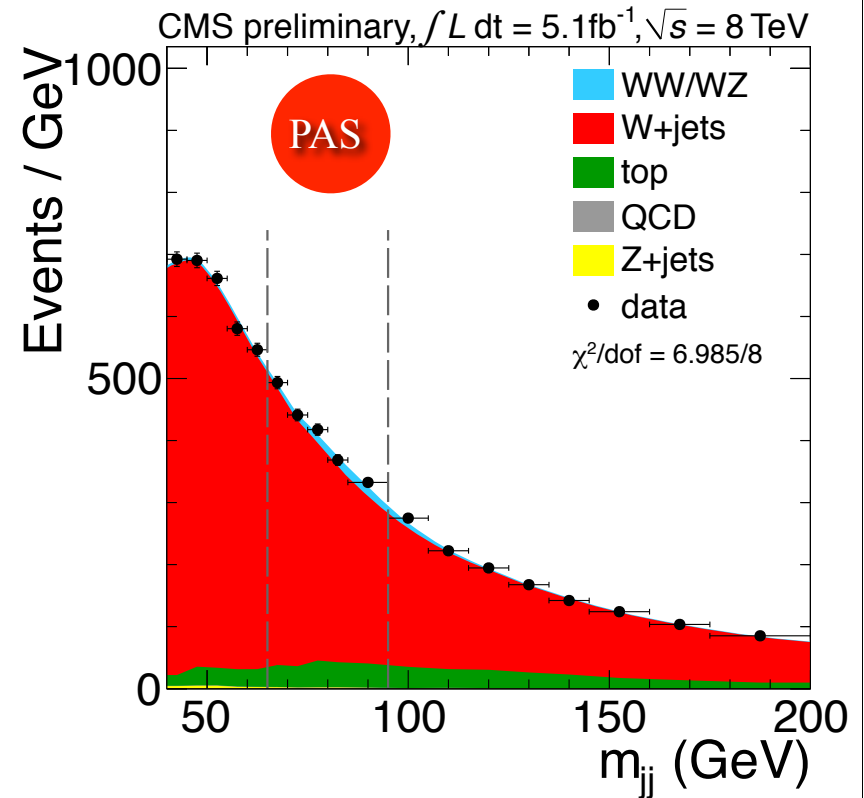


Should we replace bottom two plots in the PAS Fig 2 by these ?

# The background composition

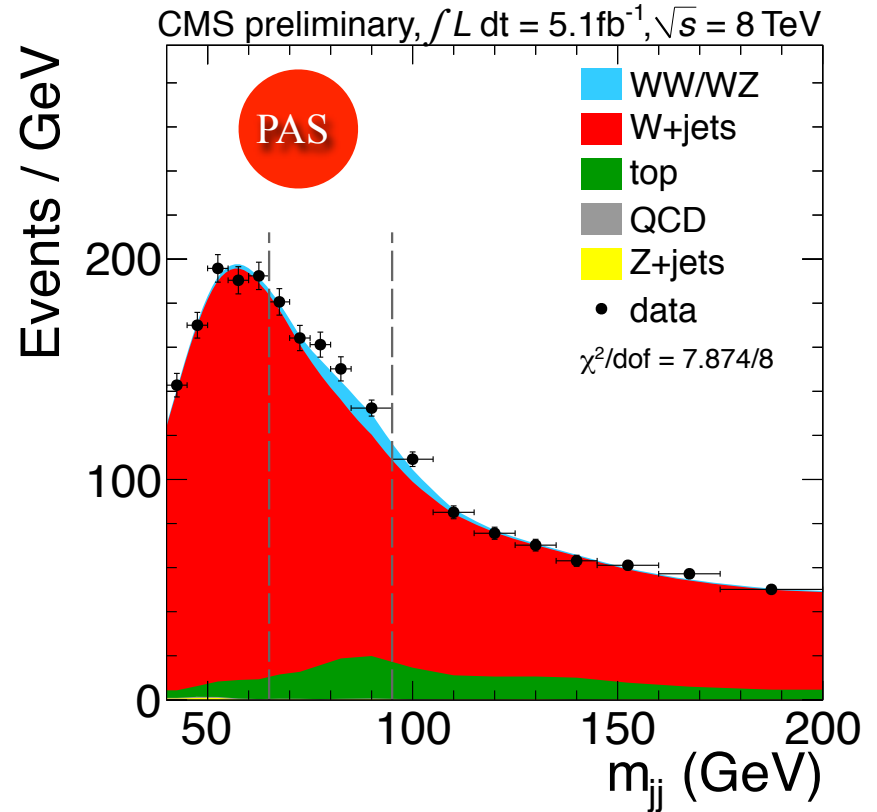
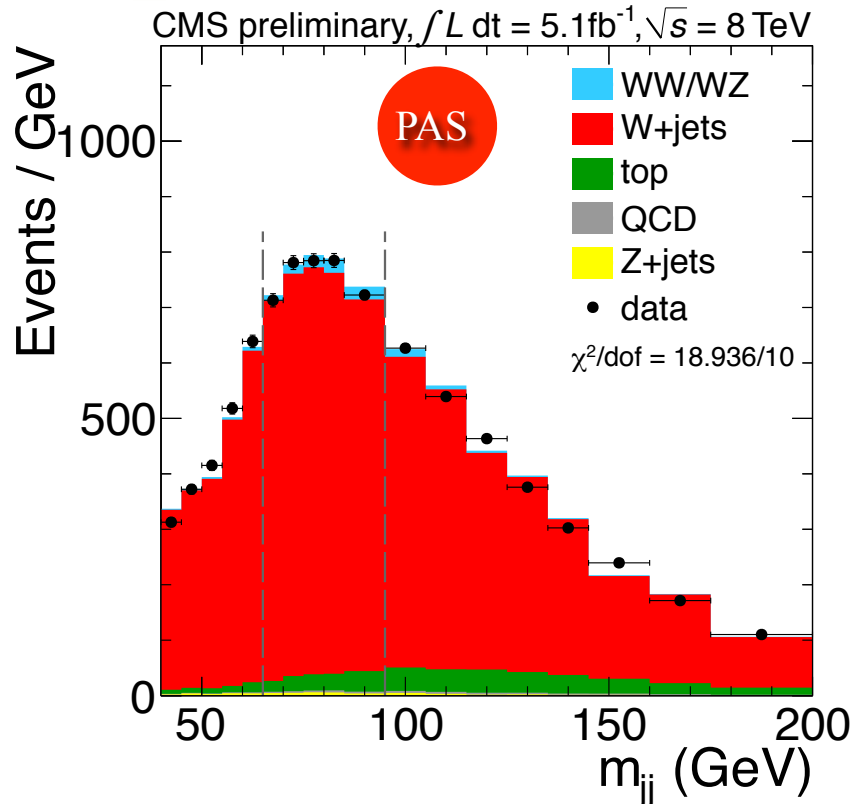
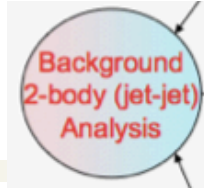


- We exclude the region from 65 – 95 GeV from the fit.
- Good agreement between the fitted composition and data.
- The normalizations are extrapolated into the signal region and passed to the next stage, i.e., 4-body analysis.

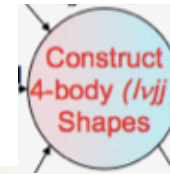


$m_H = 300 \text{ GeV}$

# Additional $m_{jj}$ examples

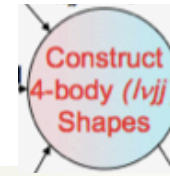


## Analysis of $m_{WW}$ distribution

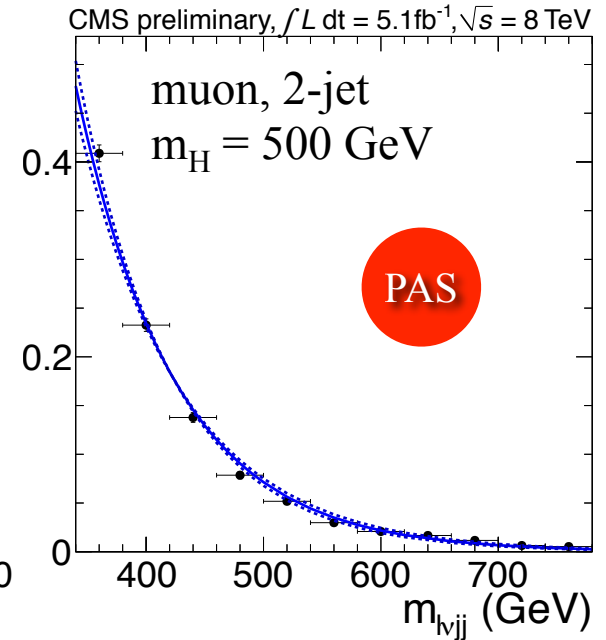
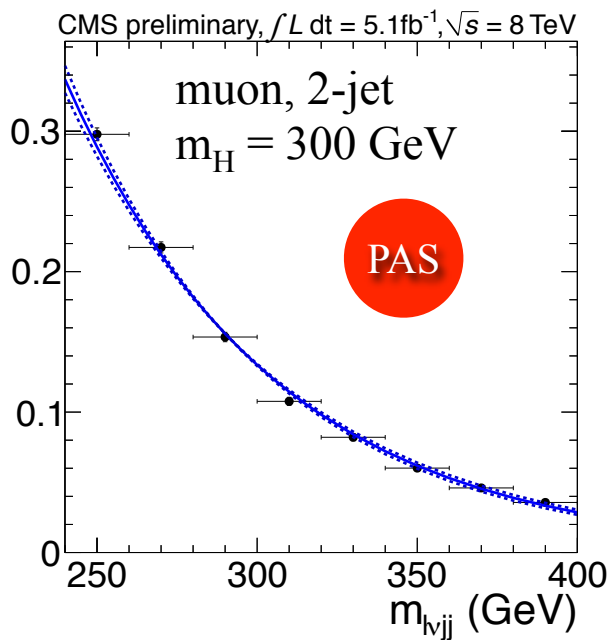
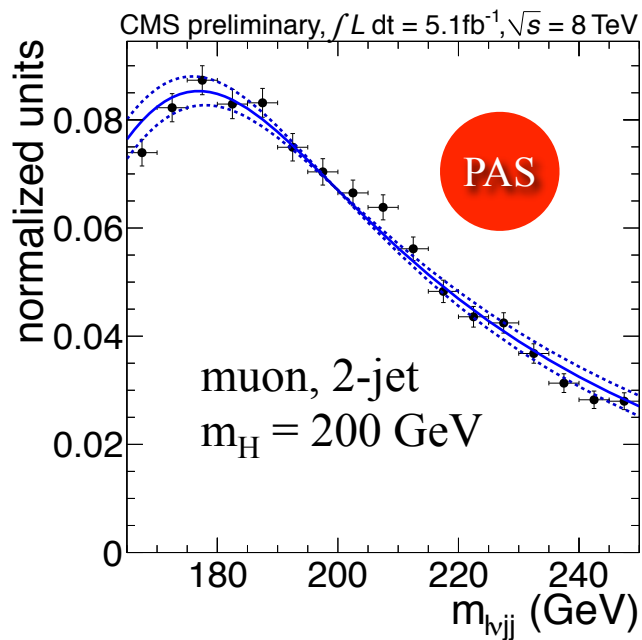


- ◆ The fit to the  $m_{jj}$  spectrum determines the relative normalization of the backgrounds.
- ◆ Then we plot the  $m_{WW}$  spectrum.
- ◆ The background components are stacked up and compared with the data with the additional selection ( $65 < m_{jj} < 95$ ) GeV.
- ◆ Shapes of the minor backgrounds are taken from MC.
  - Again, QCD is taken from the data-driven sample.
  - The  $W$ +jets shape is constructed from the  $m_{jj}$  sidebands (see next slide).

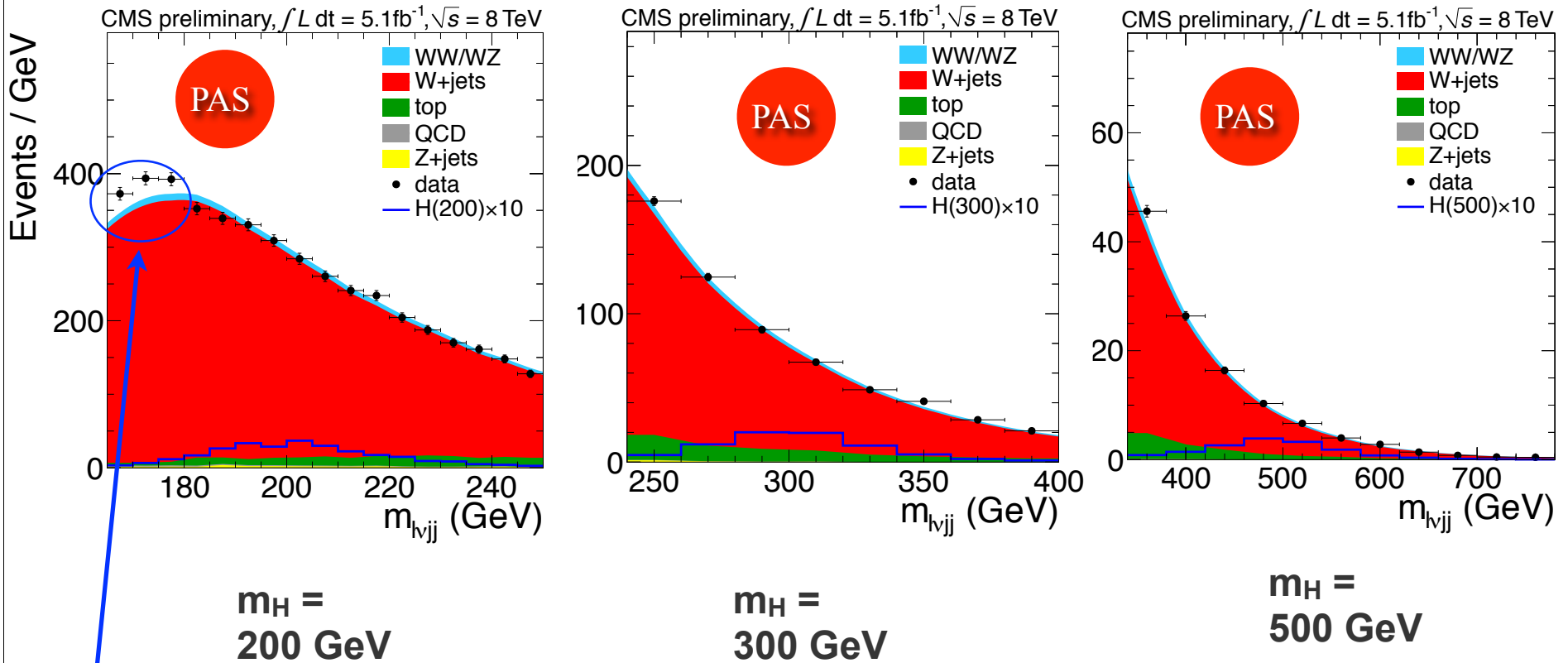
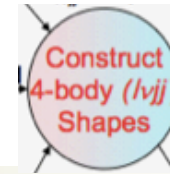
# W+jets shape: from data sidebands



- The W+jets shape obtained from data sidebands is smoothed using an exponential function.
  - Statistical uncertainty of the smoothing is combined with the uncertainty due to  $\alpha$  and used as a systematic error.
  - The dotted lines are the total shape systematic envelope.



# The $m_{WW}$ spectrum



Difference is covered by systematic uncertainty in W+jets shape. See previous slide.



# Systematic uncertainties



Source	uncertainty
Higgs line shape	0 – 30%
Signal cross-section	15 – 20%
Signal efficiency x acceptance	10%
Luminosity	4.4%
Jet energy scale, resolution and MET	< 1%
Theory (PDFs)	1 – 2%
Lepton trigger efficiency	1%
Lepton selection efficiency	2%

## Background systematics

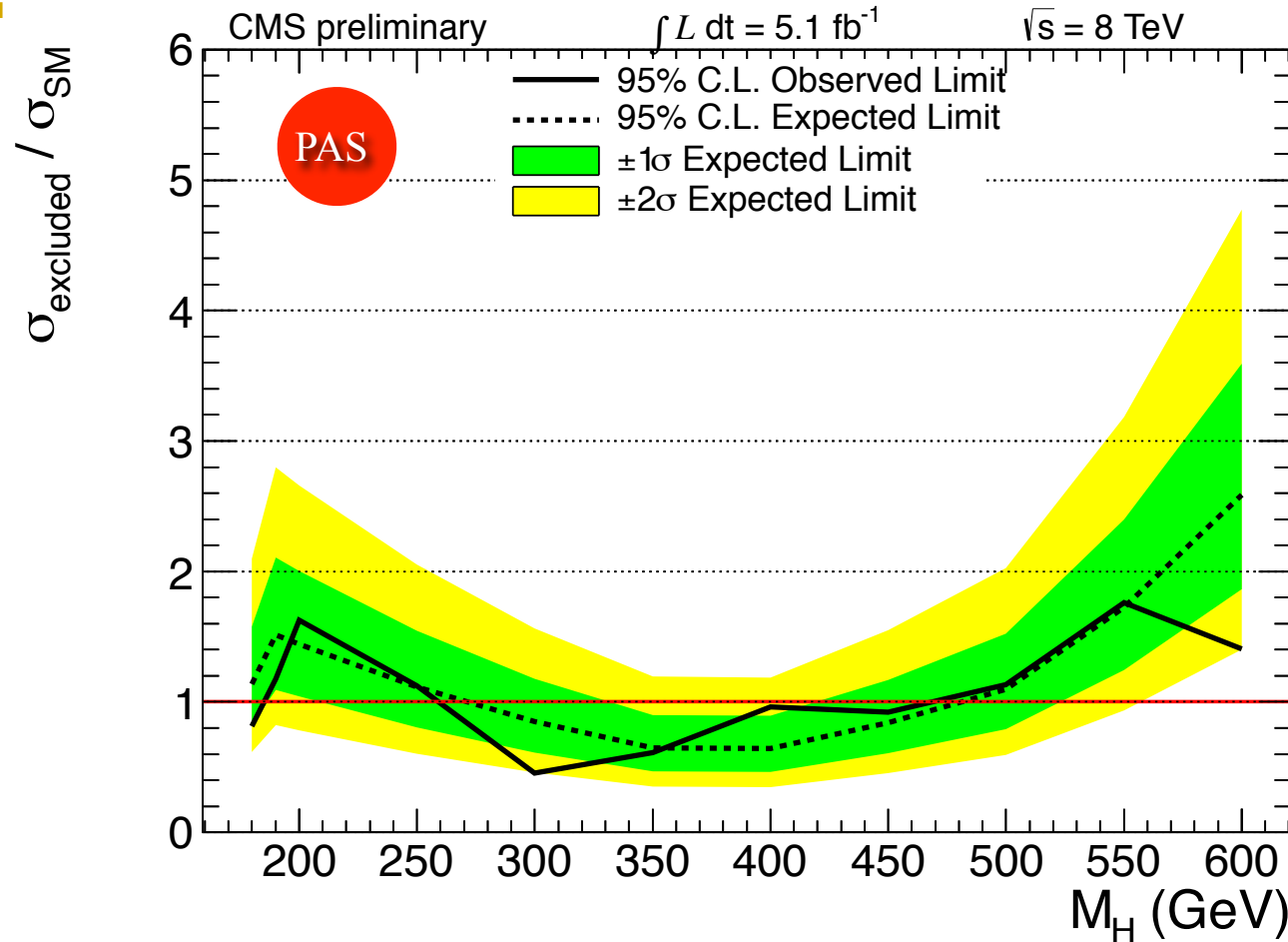
normalization uncertainty	$\leq 2\%$
W+jets fit uncertainty	shape

already described previously

- ◆ Since background is  $\sim 100x$  signal, the background systematics is dominant.
- ◆ Signal efficiency x acceptance syst. is evaluated, using a pure  $t\bar{t}b\bar{b}$  control sample, as the difference between data and MC.

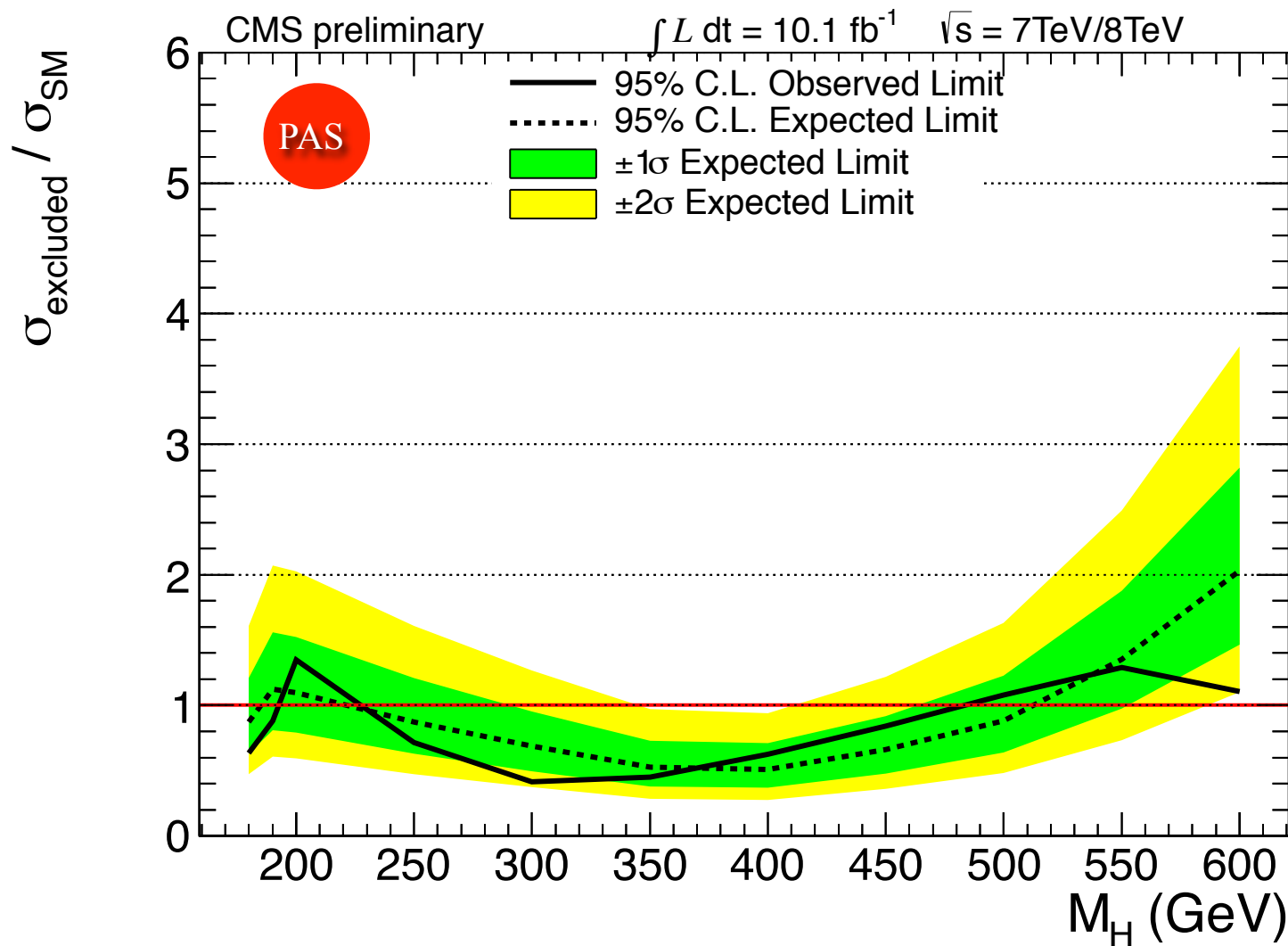
Added jet veto uncertainty which is in the range 4–28% (depending on  $m_H$  and nJets)

# The 8 TeV limit



- ◆ About 30% improvement over 2011 sensitivity
- higher Higgs xsection at 8 TeV, improved trigger

# Combination of 7 TeV and 8 TeV results



## Summary



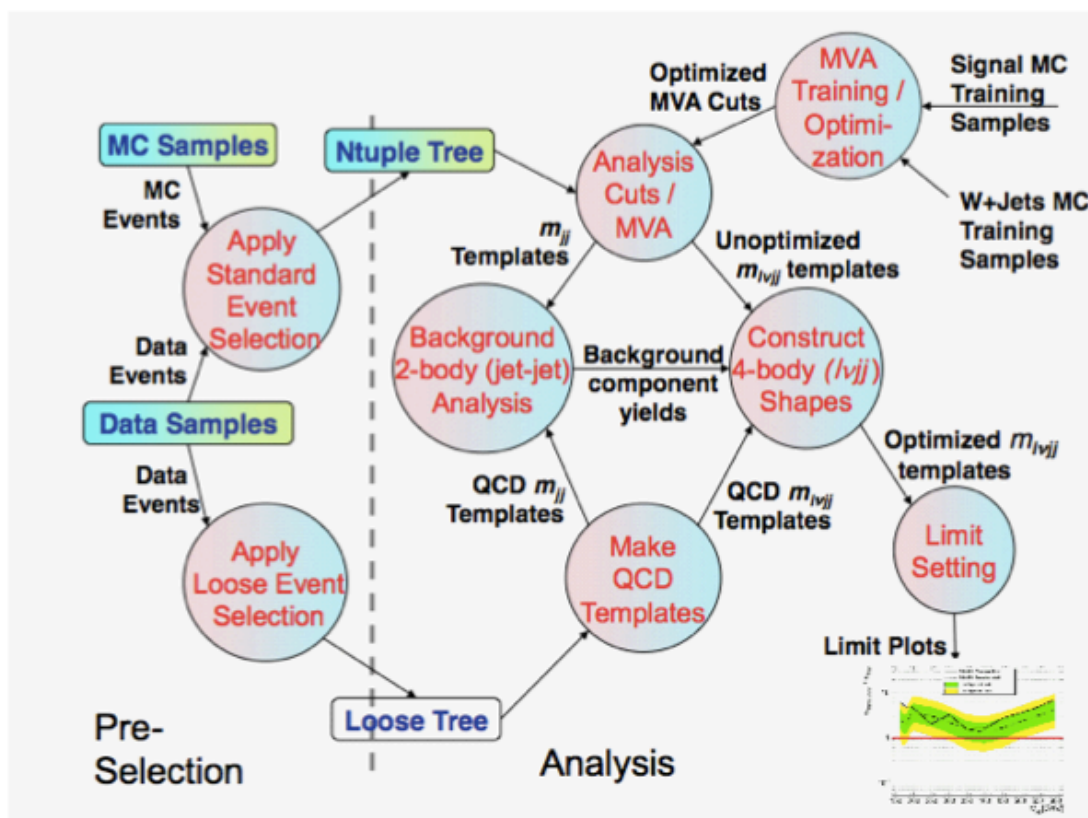
- We have topped-up the  $H \rightarrow WW \rightarrow \ell v jj$  analysis using  $5.1 \text{ fb}^{-1}$  data available for ICHEP
  - In 8 TeV data, we exclude the Standard Model Higgs boson in the mass range 260 – 460 GeV at 95% confidence level.
  - Combining 7 TeV and 8 TeV results, we exclude Higgs boson in the mass range **230 – 480 GeV** at 95% confidence level.

backup slides



# Fitting and limit setting

- The analysis has two major parts
  - Fit to the  $m_{jj}$  spectrum to determine the backgrounds
  - limit setting from  $m_{WW}$  spectrum



# 2012 physics objects



- Detailed summary in Sarah Boutle talk [\*]:

[\*]<https://indico.cern.ch/getFile.py/access?contribId=45&sessionId=1&resId=0&materialId=slides&confId=188820>

## *Short summary*

- Use of PFBReco objects
  - PF isolation for electrons and muons
  - MVA Id for electrons (WP70) and cut based for muons (tight WP)
- AK5PFJets CHS Jets with JEC
  - Cut based JetId - loose WP
- PFMET with type 1 correction



# Jet and MET



- Use PF2PAT jets (with CHS)
- PF MET with type 1 correction

- Jet energy corrections (standard):

The minimum correction levels to be applied on any CMS analysis using Monte Carlo and Data are:

Monte Carlo	L1(Pile Up)+L2(Relative)+L3(Absolute)
Data	L1(Pile Up)+L2(Relative)+L3(Absolute)+L2L3Residuals

- Jet ID: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetID>:

PF Jet ID	Loose (Recommended)	Medium	Tight
Neutral Hadron Fraction	<0.99	<0.95	< 0.90
Neutral EM Fraction	<0.99	<0.95	< 0.90
Number of Constituents	>1	>1	> 1
And for $\eta < 2.4$ , $\eta > -2.4$ in addition apply			
Charged Hadron Fraction	>0	>0	>0
Charged Multiplicity	>0	>0	>0
Charged EM Fraction	<0.99	<0.99	<0.99

- Will include the information for PU jet ID in our PATtuples: <https://twiki.cern.ch/twiki/bin/view/CMS/PileupJetID>
- Will also include information for computing mvaMET: <https://twiki.cern.ch/twiki/bin/view/CMS/MVAMet>

# Muon



- **Muon ID:** Use POG recommendations: <https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideMuonId>

Plain-text description	Technical description
The candidate is reconstructed as a Global Muon	<code>recoMu.isGlobalMuon()</code>
Particle-Flow muon id	<code>recoMu.isPFMuon()</code>
$\chi^2/\text{ndof}$ of the global-muon track fit < 10	<code>recoMu.globalTrack()-&gt;normalizedChi2() &lt; 10.</code>
At least one muon chamber hit included in the global-muon track fit	<code>recoMu.globalTrack()-&gt;hitPattern().numberOfValidMuonHits() &gt; 0</code>
Muon segments in at least two muon stations This implies that the muon is also an arbitrated tracker muon, see <a href="#">SWGuideTrackerMuons</a>	<code>recoMu.numberOfMatchedStations() &gt; 1</code>
Its tracker track has transverse impact parameter $d_{xy} < 2$ mm w.r.t. the primary vertex	<code>fabs(recoMu.innerTrack()-&gt;dxy(vertex-&gt;position())) &lt; 0.2</code> Or <code>dB() &lt; 0.2 on pat::Muon [1]</code>
The longitudinal distance of the tracker track wrt. the primary vertex is $d_z < 5$ mm	<code>fabs(recoMu.innerTrack()-&gt;dz(vertex-&gt;position())) &lt; 0.5</code>
Number of pixel hits > 0	<code>recoMu.innerTrack()-&gt;hitPattern().numberOfValidPixelHits() &gt; 0</code>
Cut on number of tracker layers with hits > 5	<code>track()-&gt;hitPattern().trackerLayersWithMeasurement() &gt; 5</code>

Use tighter cut of 0.02 here

- **Isolation:** Use POG recommendations: <https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideMuonId>

Algorithm	Type	Expression	PU correction	Cone size ( $\Delta R$ )	Tight cut	Loose cut
Subdetector based	Tracker relative	$(\sum p_T(\text{TRK})) / p_T$	none	0.3	0.05	0.10
PF based	Combined relative	$(\sum E_T(\text{chHad from PV}) + \sum E_T(\text{neutHad}) + \sum E_T(\text{photons})) / p_T$	Reference correction using DeltaB corrections	0.4	0.12	0.20

# Loose muon for veto and jet cleaning



- **Muon ID:** Use POG recommendations: <https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideMuonId>

Plain-text description	Technical description	Comments
Particle-Flow muon id	<code>recoMu.isPFMuon()</code>	Can be complemented by muon quality cuts similar to those used in the Tight Muon selection.
Is Global OR Tracker Muon	<code>recoMu.isGlobalMuon()    recoMu.isTrackerMuon()</code>	Avoid using muons which are only Standalone Muons

- **Isolation:** Use POG recommendations: <https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideMuonId>

Algorithm	Type	Expression	PU correction	Cone size ( $\Delta R$ )	Tight cut	Loose cut
Subdetector based	Tracker relative	$(\sum p_T(\text{TRK})) / p_T$	none	0.3	0.05	0.10
PF based	Combined relative	$(\sum E_T(\text{chHad from PV}) + \sum E_T(\text{neutHad}) + \sum E_T(\text{photons})) / p_T$	Reference correction using DeltaB corrections	0.4	0.12	0.20

Not used in first round of 2012 data processing

# Electrons



- **Electron ID:** Use POG recommendations: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/Electrons4Higgs>
- Retain possibility to do MVA ID
- But start with cut-based ID:

## Barrel Cuts

	Veto	Loose	Medium	Tight
dEtaIn	0.007	0.007	0.004	0.004
dPhiIn	0.8	0.15	0.06	0.03
sigmaEtaEta	0.01	0.01	0.01	0.01
H/E	0.15	0.12	0.12	0.12
d0 (vtx)	0.04	0.02	0.02	0.02
dZ (vtx)	0.2	0.2	0.1	0.1
fabs(1/E - 1/p)	N/A	0.05	0.05	0.05
PF isolation / pT	0.15	0.15	0.15	0.10
Conversion rejection: vertex fit probability	N/A	1e-6	1e-6	1e-6
Conversion rejection: missing hits	N/A	1	1	0

## Endcap Cuts

pT > 20 (pT < 20)	Veto	Loose	Medium	Tight
dEtaIn	0.01	0.009	0.007	0.005
dPhiIn	0.7	0.10	0.06	0.02
sigmaEtaEta	0.03	0.03	0.03	0.03
H/E	N/A	0.10	0.10	0.10
d0 (vtx)	0.04	0.02	0.02	0.02
dZ (vtx)	0.2	0.2	0.1	0.1
fabs(1/E - 1/p)	N/A	0.05	0.05	0.05
PF isolation / pT	0.15	0.15(0.10)	0.15(0.10)	0.10(0.07)
Conversion rejection: vertex fit probability	N/A	1e-6	1e-6	1e-6
Conversion rejection: missing hits	N/A	1	1	0

- **Isolation:** Use PF isolation (cone 0.3) with EA rho correction

# Loose electron for veto and jet cleaning



- **Electron ID:** Use POG recommendations: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/Electrons4Higgs>
- Retain possibility to do MVA ID
- But start with cut-based ID:

## Barrel Cuts

	Veto	Loose	Medium	Tight
dEtaIn	0.007	0.007	0.004	0.004
dPhiIn	0.8	0.15	0.06	0.03
sigmaEtaEta	0.01	0.01	0.01	0.01
H/E	0.15	0.12	0.12	0.12
d0 (vtx)	0.04	0.02	0.02	0.02
dZ (vtx)	0.2	0.2	0.1	0.1
fabs(1/E - 1/p)	N/A	0.05	0.05	0.05
PF isolation / pT	0.15	0.15	0.15	0.10
Conversion rejection: vertex fit probability	N/A	1e-6	1e-6	1e-6
Conversion rejection: missing hits	N/A	1	1	0

## Endcap Cuts

pT > 20 (pT < 20)	Veto	Loose	Medium	Tight
dEtaIn	0.01	0.009	0.007	0.005
dPhiIn	0.7	0.10	0.06	0.02
sigmaEtaEta	0.03	0.03	0.03	0.03
H/E	N/A	0.10	0.10	0.10
d0 (vtx)	0.04	0.02	0.02	0.02
dZ (vtx)	0.2	0.2	0.1	0.1
fabs(1/E - 1/p)	N/A	0.05	0.05	0.05
PF isolation / pT	0.15	0.15(0.10)	0.15(0.10)	0.10(0.07)
Conversion rejection: vertex fit probability	N/A	1e-6	1e-6	1e-6
Conversion rejection: missing hits	N/A	1	1	0

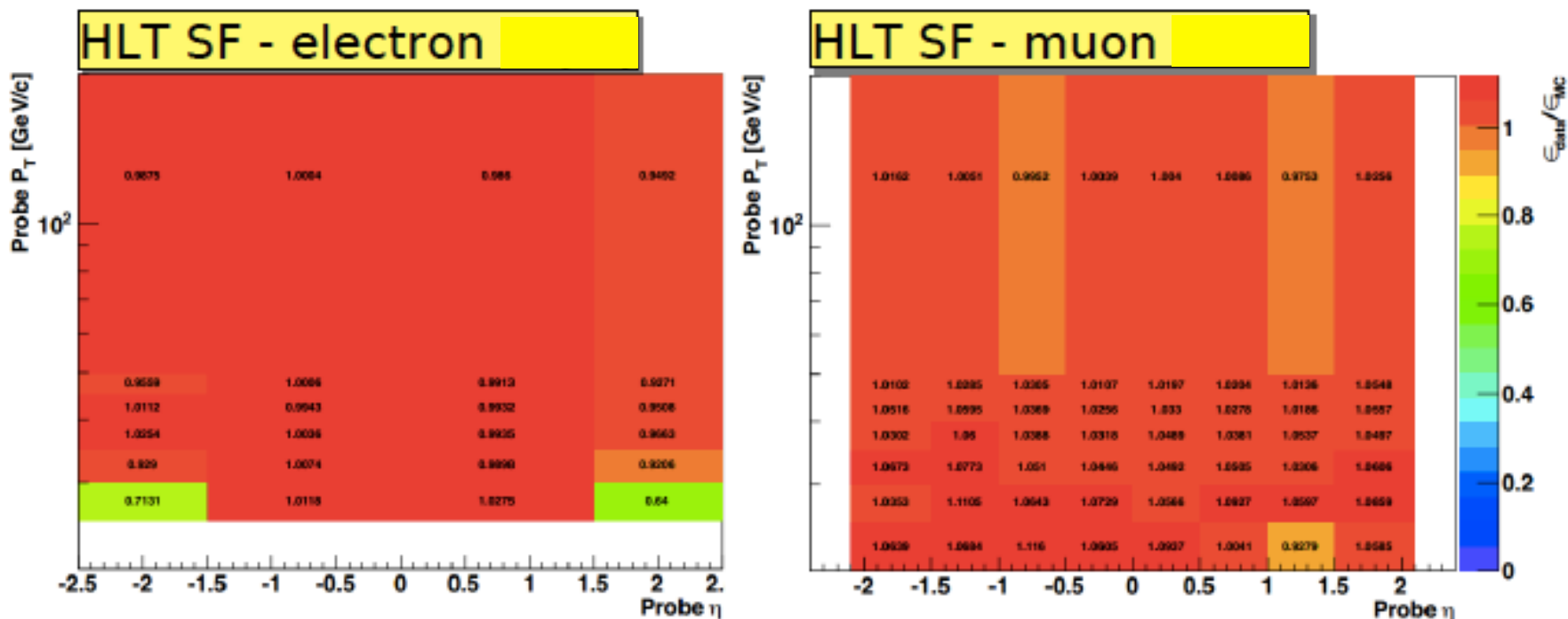
- **Isolation:** Use PF isolation (cone 0.3) with EA rho correction



# Efficiency corrections

- MC lepton efficiencies corrected with **“Tag and Probe”** on Z events
- **HLT efficiency and reconstruction/identification** scale factors
- Detailed description in Gordon Kaussen talk [\*]

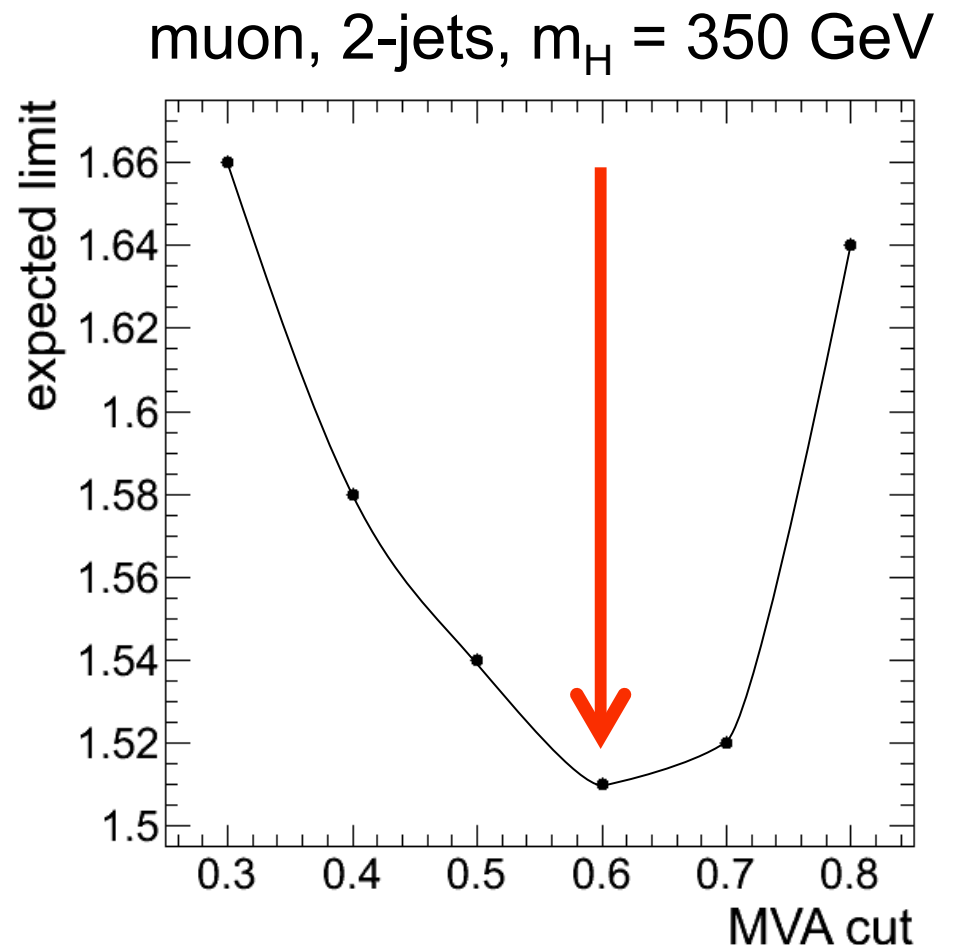
[\*]<https://indico.cern.ch/getFile.py/access?contribId=23&sessionId=10&resId=0&materialId=slides&confId=193619>



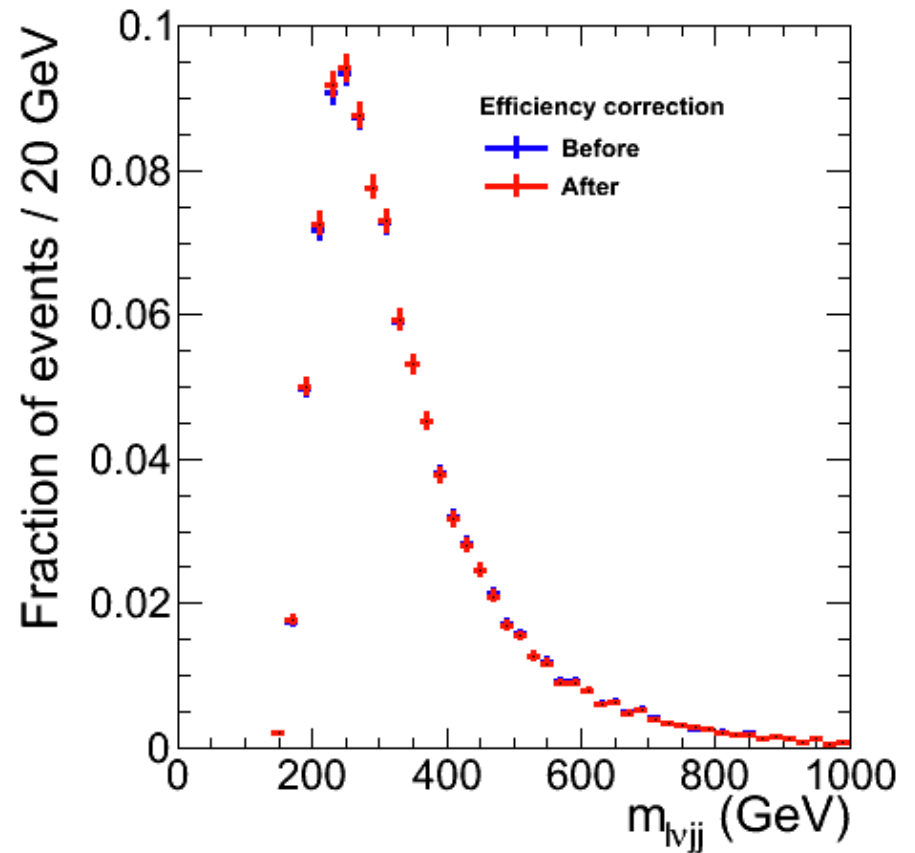
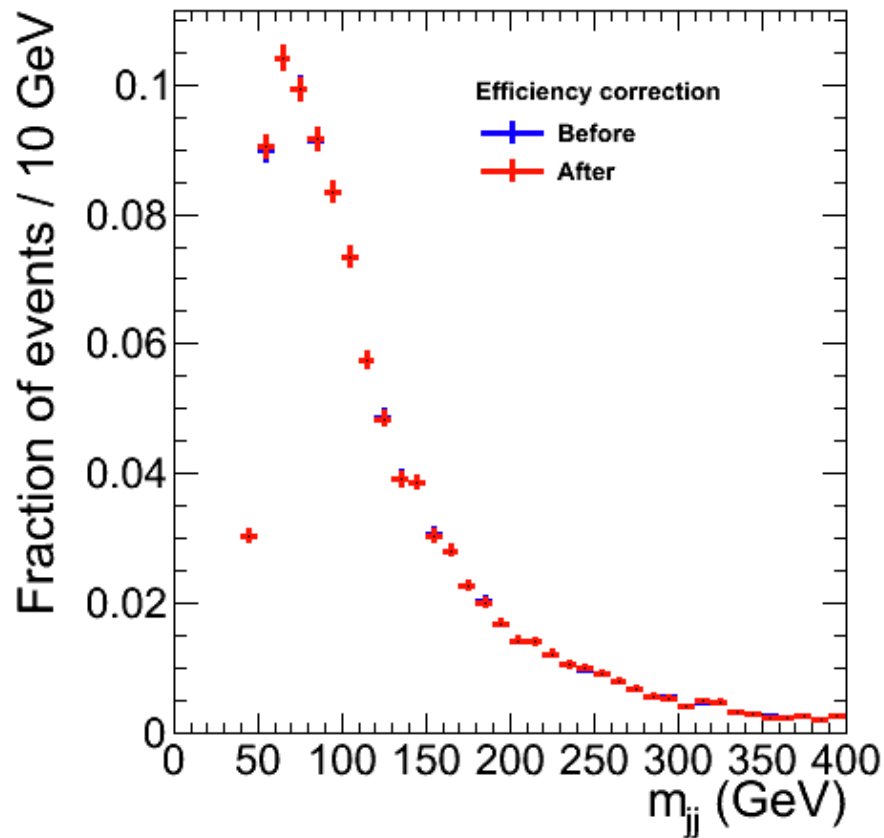
Apply similar scale factor for lepton reconstruction and ID efficiencies.

# MVA cut optimization

- We select the MVA cut value based on running the full asymptotic limit setting machinery and using the expected limit.
- Once the optimal cut is selected for each of the 48 analysis points they can be combined using the standard Higgs combination package.



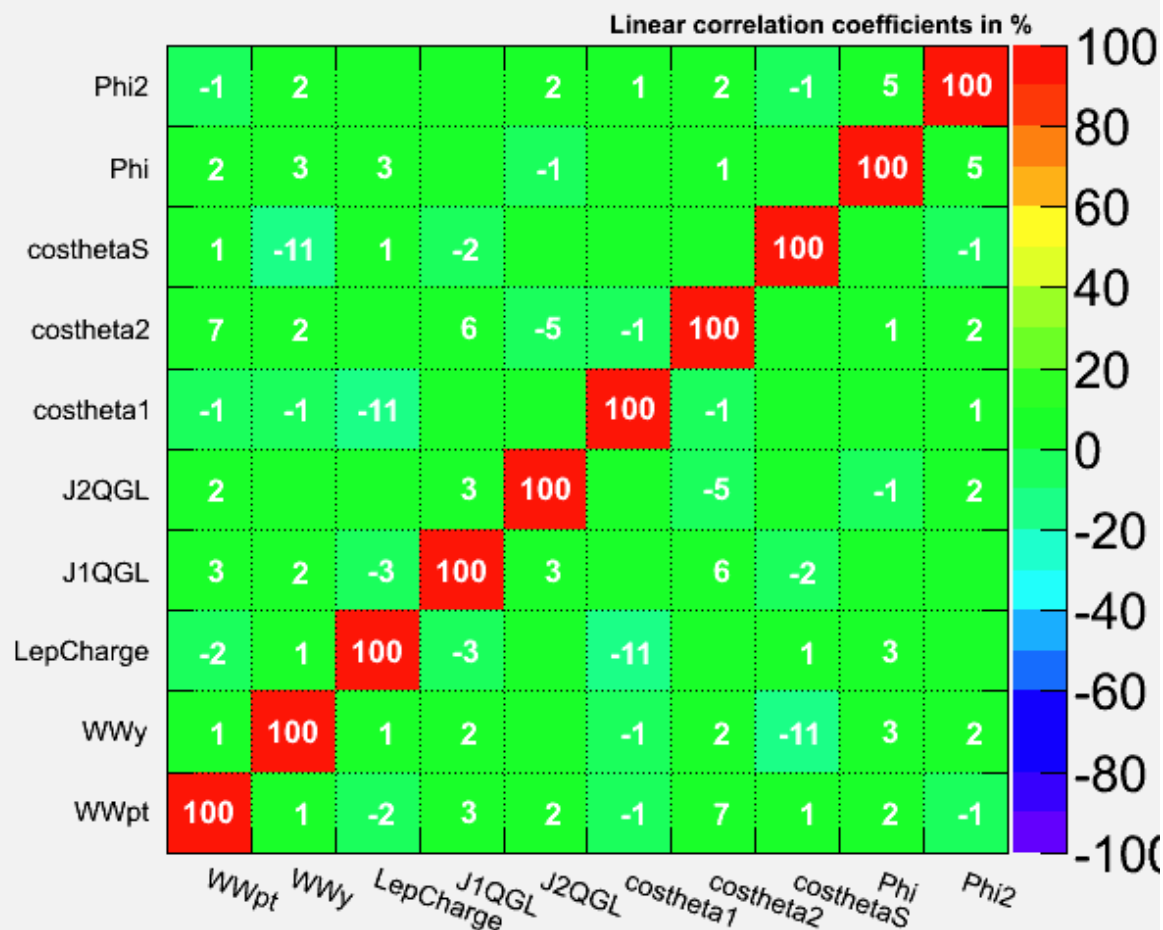
# Trigger effect on key distributions



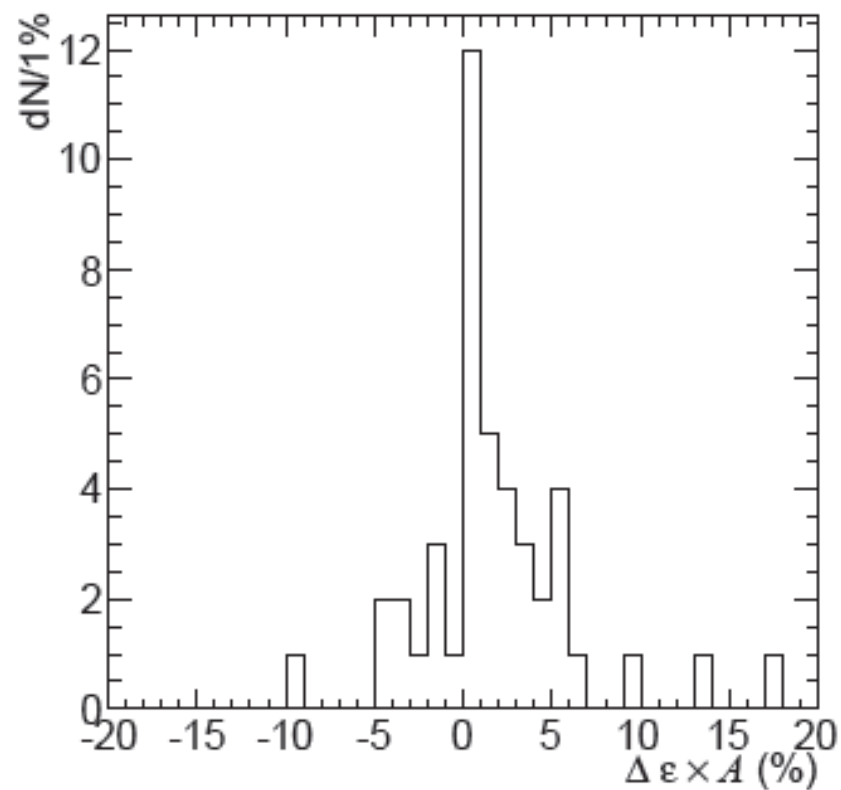
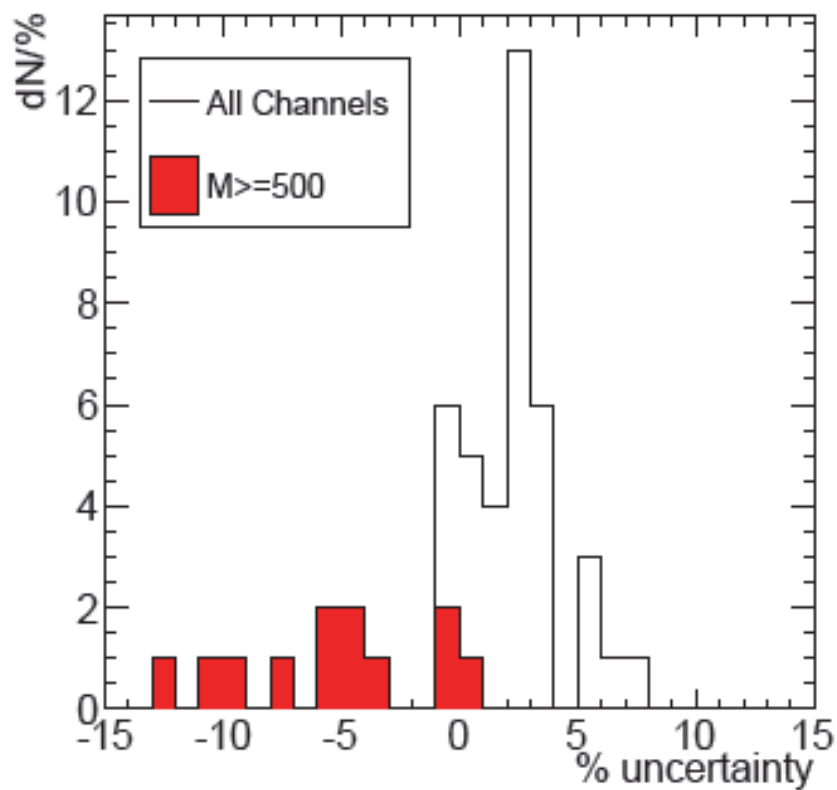


# MVA correlations

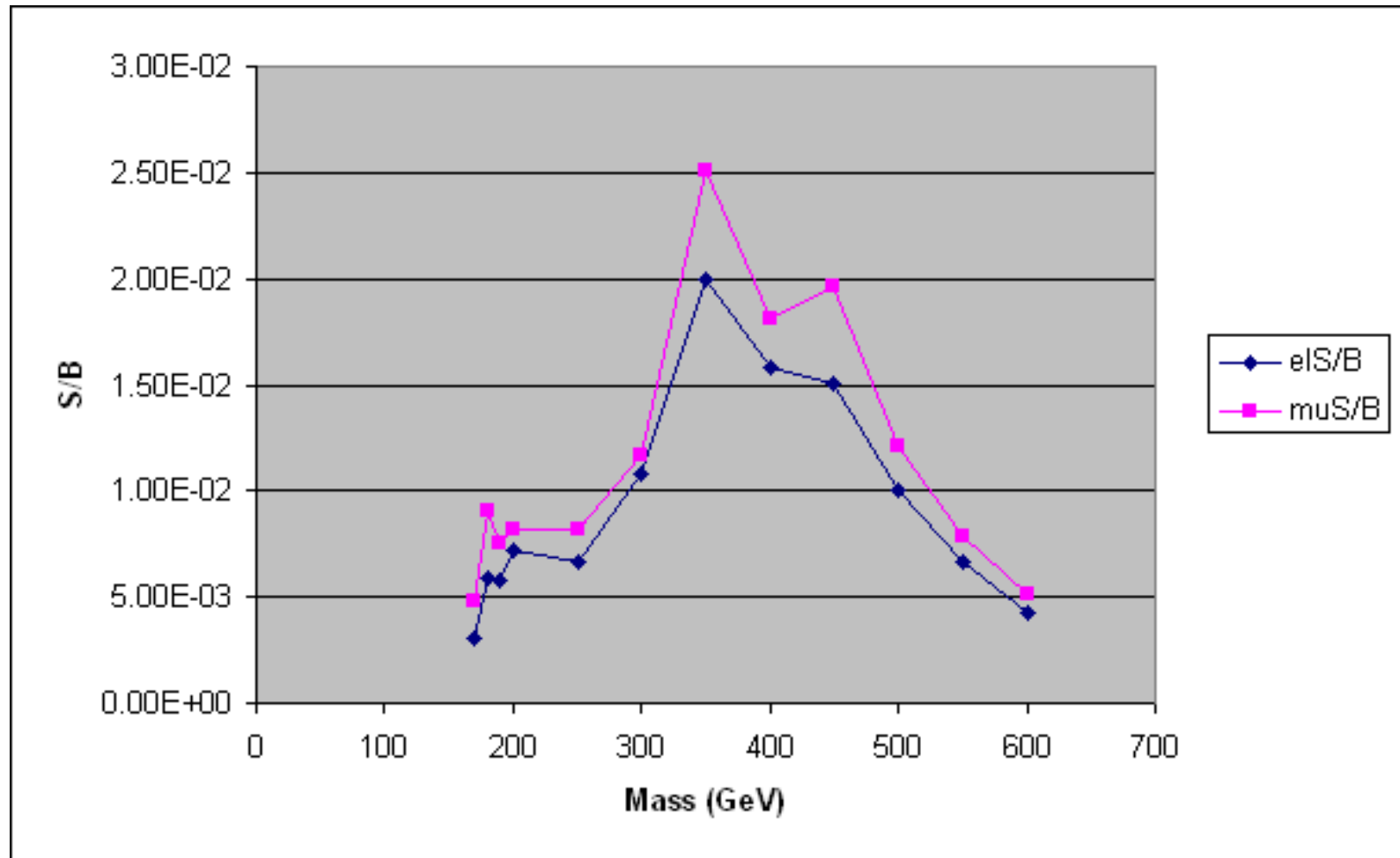
Correlation Matrix (signal)



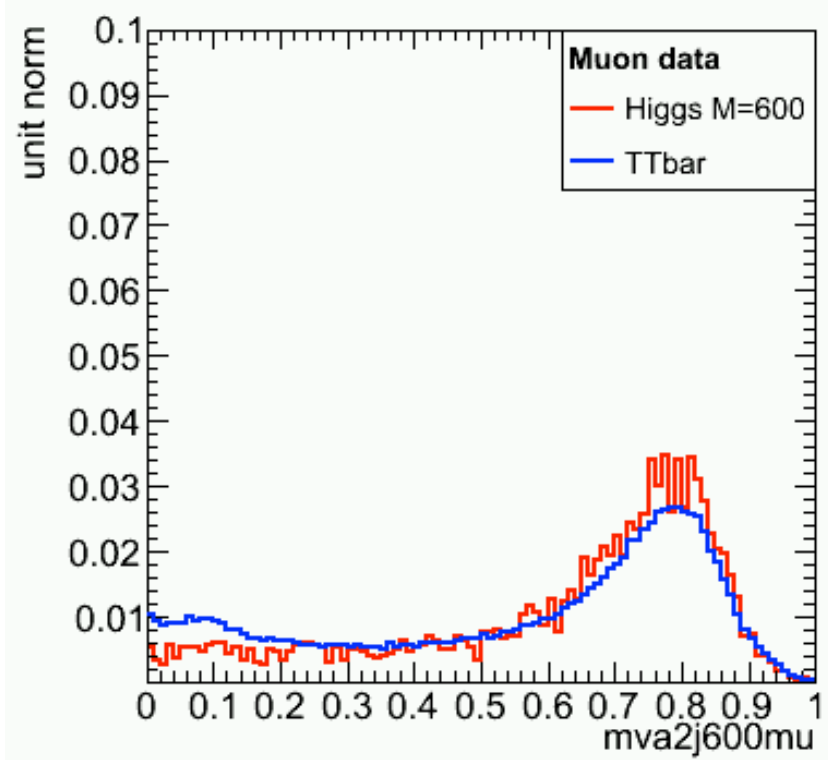
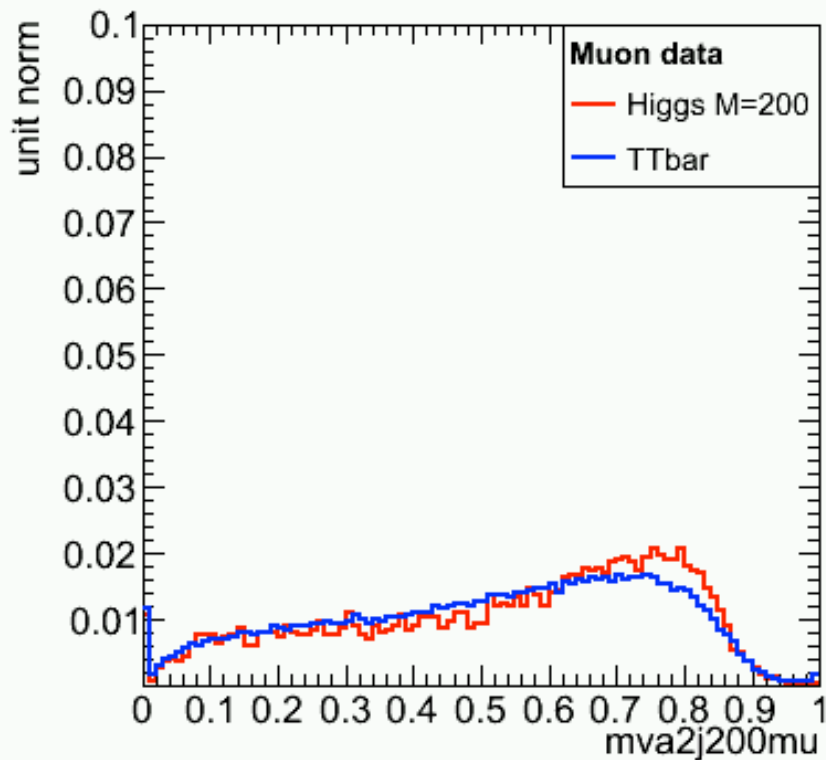
# Likelihood selection efficiency



# S/B comparison between muon and electron



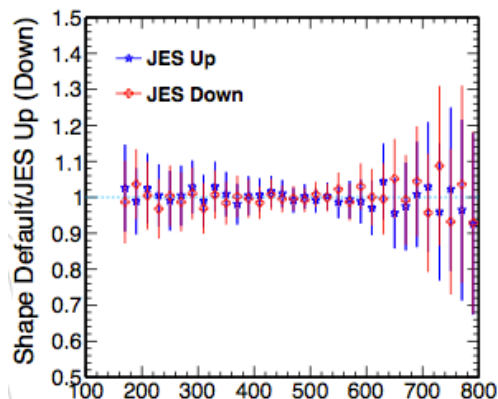
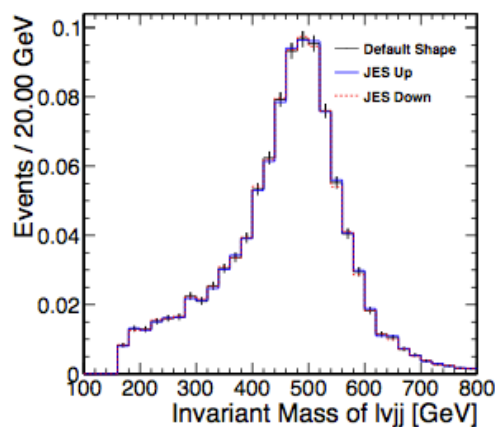
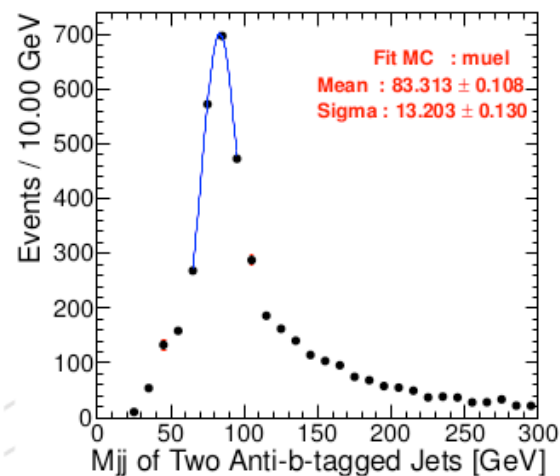
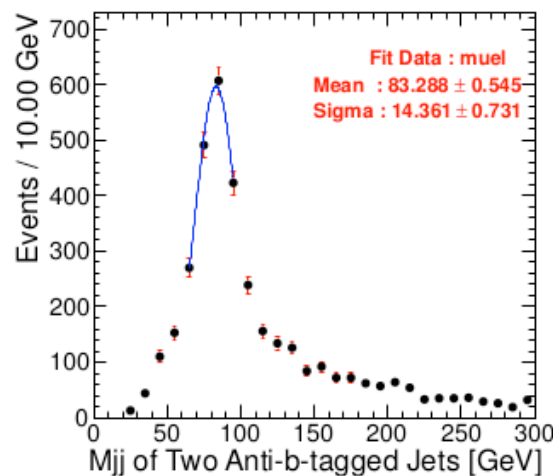
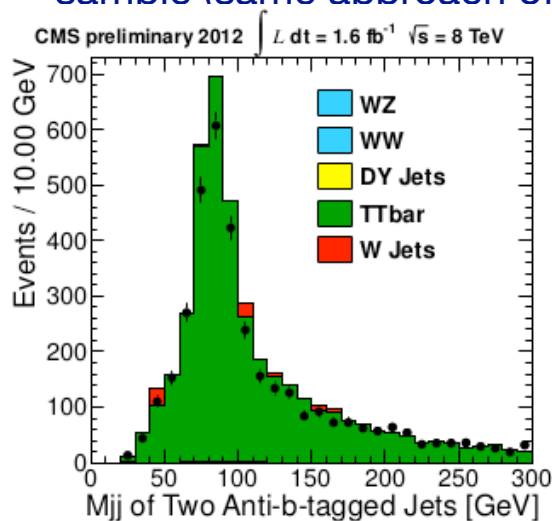
# Comparison of MVA output for tt and Higgs



# JES/JER

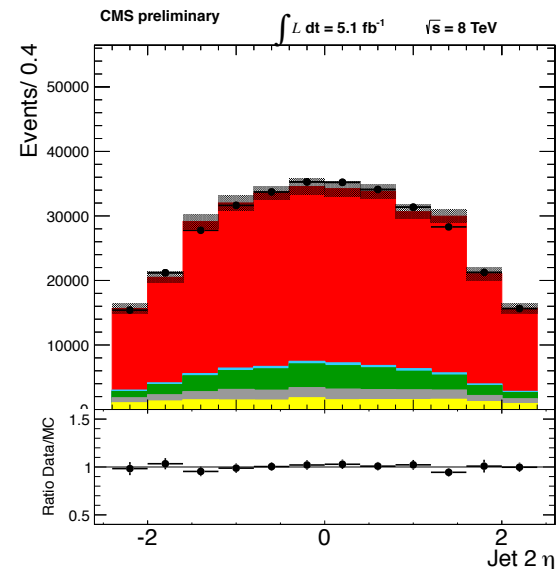
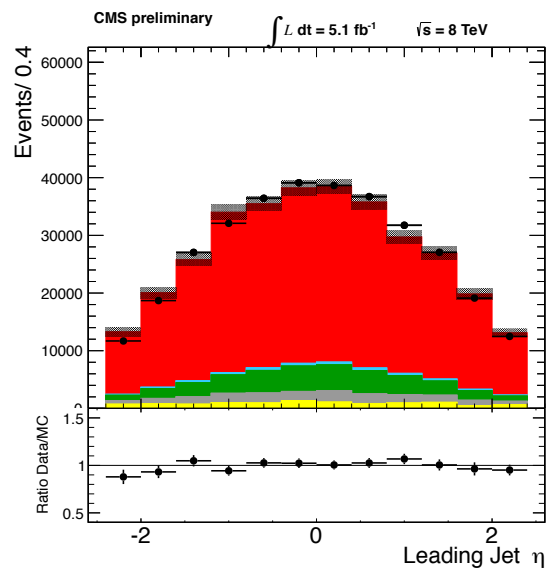
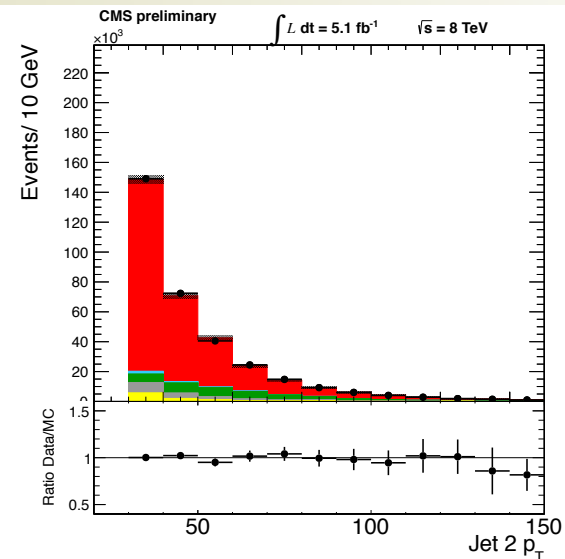
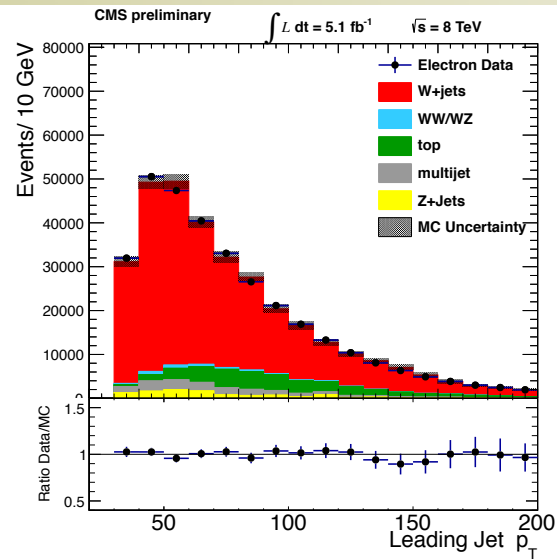


- The uncertainty on JES is evaluated from a fit to a had. W mass in a top-enriched sample (same approach of **EWK-11-017** and **SMP-12-015**)

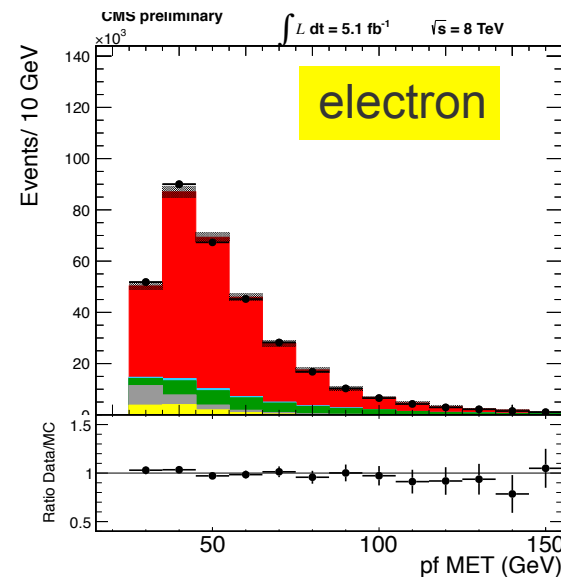
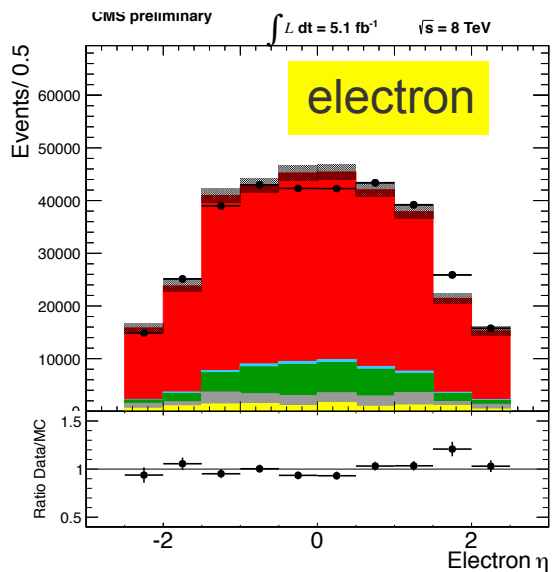
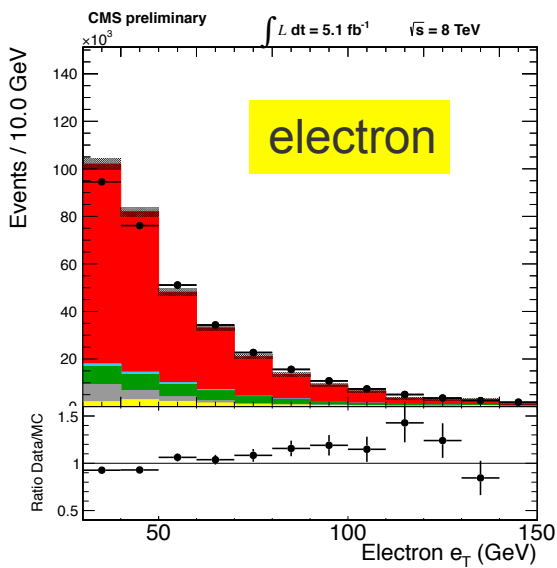
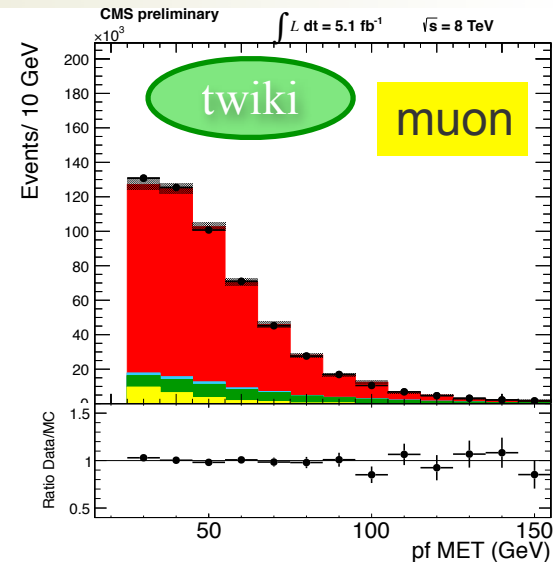
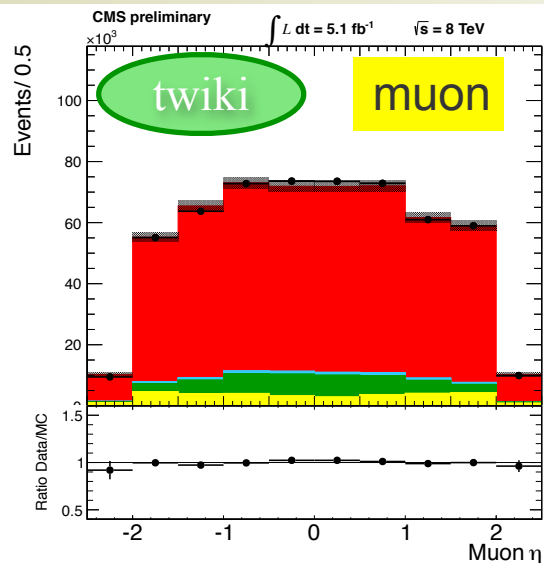
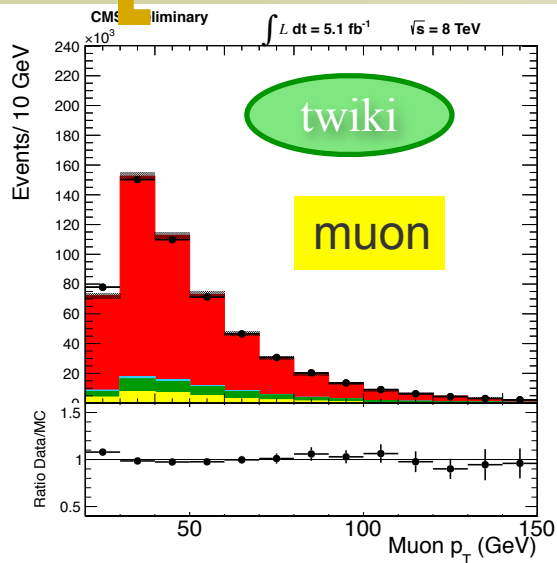


- no impact on the Higgs mass is observed
- overall effect is negligible

# Comparison plots: jets (electron data)



# Comparison plots: leptons



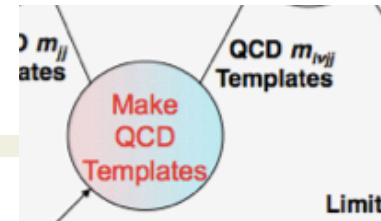
## Use optimized selection to improve sensitivity



- To **improve the limit** on SM Higgs cross-section and **reduce the systematics** on background evaluation:
  - ✓ drastically **reduce** the number of **background** events
  - ✓ **preserve** a reasonable **efficiency** on the **signal**
- Use event kinematics to build a simple **likelihood** discriminator:
  - a different likelihood is built for each operating point
    - 11 mass points  $\times$  2 flavors (e/ $\mu$ )  $\times$  2 jet bins (2j/3j) = **44 different likelihoods**
- Mostly sensitive at high mass, but have extended the analysis down to  $m_H = 180$  GeV

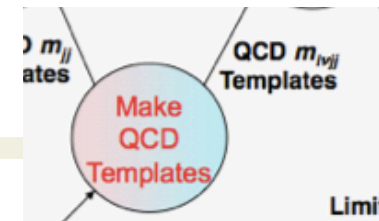


# QCD data driven



- To get the **QCD shape from data**: define a  $\sim$  pure QCD sample applying all standard cuts but
  - invert isolation cut:  $I_{\text{comb}}/p_T > 0.3$
  - for electrons, remove ID requirement
  - relax  $E_T^{\text{miss}}$  cut from  $>30$  to  $>20$  GeV
- Take the **signal W shape** from MC
- To get the **QCD fraction in data**
  - fit the leptonic  $W M_T$  distribution in data with the above described shapes
  - account for different acceptances between  $E_T^{\text{miss}} > 20$  and  $E_T^{\text{miss}} > 30$  GeV

# QCD fraction estimates



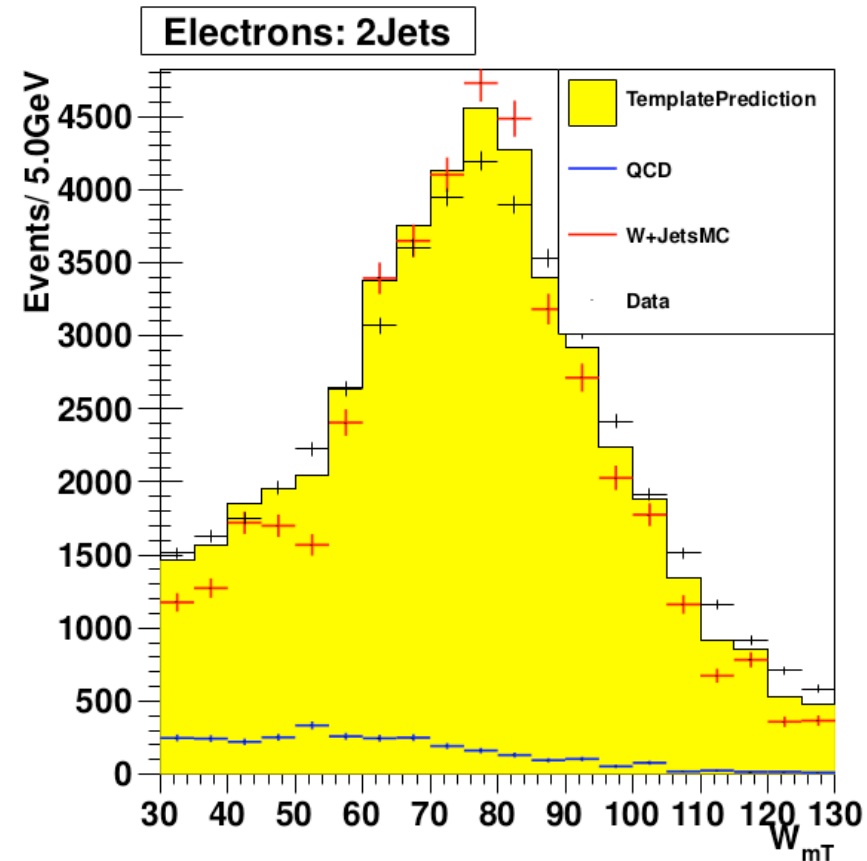
## QCD fractions in data

	2 jets	3 jets
electron	$6.4 \pm 0.7\%$	$2.1 \pm 0.7\%$
muon	$0.2 \pm 0.4\%$	$0.0 \pm 0.4\%$

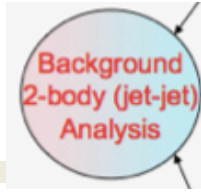
statistical fit errors only!

## Uncertainty on QCD normalization

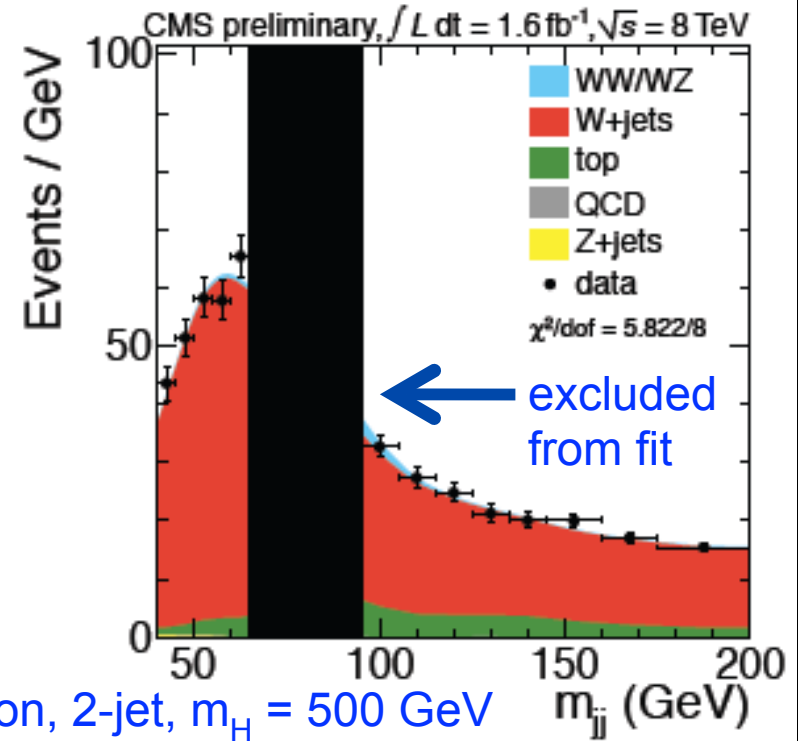
	2 jets	3 jets
electron	50%	50%
muon	100%	100%



# Normalization from $m_{jj}$ fit



- ◆ Determine backgrnd composition in a 1D, unbinned, max likelihood fit to the di-jet invariant mass spectrum.
- ◆ Background shapes are taken from:
  - MC for all minor backgrounds (not for W+jets and QCD)
  - data-driven approach for QCD
  - analytic description/ MC for W+jets



Process	Shape	External constraint on normalization
W+jets	data/MC	Unconstrained
Diboson	MC	Constrained: (NLO) $89.4 \text{ pb} \pm 10\%$ [48]
$t\bar{t}$	MC	Constrained: (approx. NNLO) $225 \text{ pb} \pm 7\%$ [49]
single top	MC	Constrained: (approx. NNLO) $113 \text{ pb} \pm 5\%$ [50–52]
Drell-Yan+jets	MC	Constrained: (NNLO, $m_{\ell\ell} > 50 \text{ GeV}$ ) $3504 \text{ pb} \pm 4.3\%$ [53]
Multi-jet	data	Constrained: $\cancel{E}_T$ fit in data $\pm 50\%$ (100%) for electrons (muons)

# W+jets: the dominant background



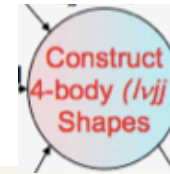
- For  $m_H \leq 180$  GeV, use MC shape template, because statistics are plentiful.
- For higher masses MC statistics are much lower so we take an analytic approach.
- The analytic shapes are chosen based on MC but the functional parameters are allowed to vary within constraints in the  $m_{jj}$  fit to the data.

$$\mathcal{F}_{W+jets} = \text{erf}(m_{jj}; m_0, \sigma) \times \left[ (m_{jj})^{-\alpha - \beta \ln(m_{jj}/\sqrt{s})} \right]$$

$$\mathcal{F}_{W+jets \text{ low mass, 2 jets}} = \text{erf}(m_{jj}; m_0, \sigma) \times (m_{jj})^{-\alpha} \times \exp(m_{jj}\tau)$$

$$\mathcal{F}_{W+jets \text{ low mass, 3 jets}} = (m_{jj})^{-\alpha - \beta \ln(m_{jj}/\sqrt{s})} \times \exp(m_{jj}\tau)$$

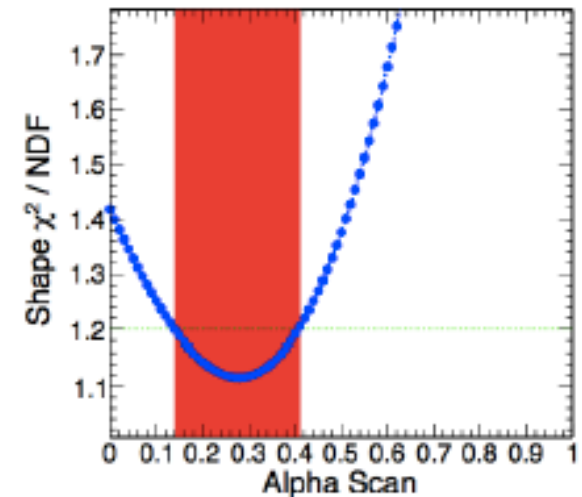
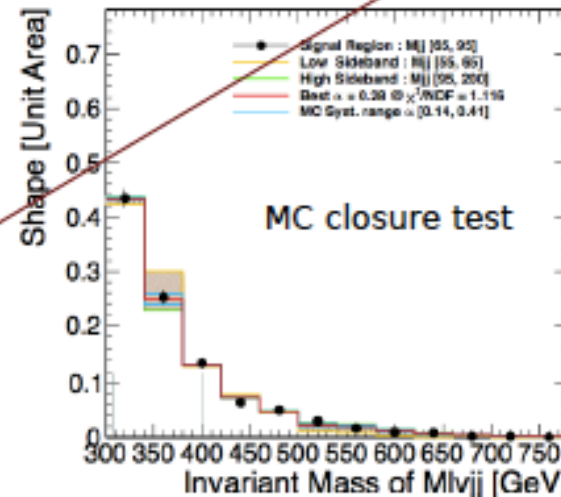
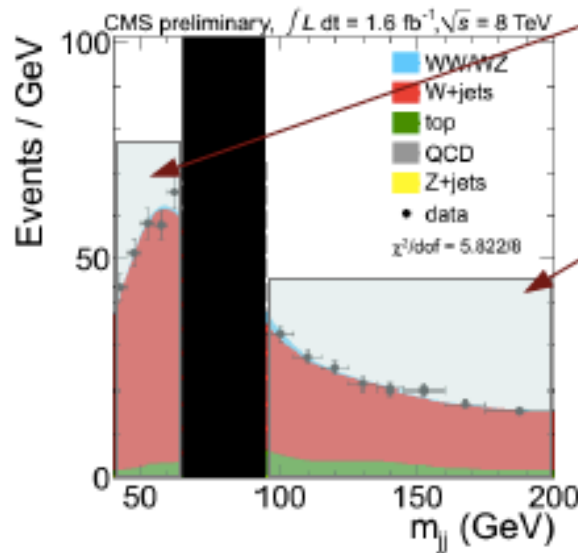
# W+jets shape for $m_{WW}$ : data driven



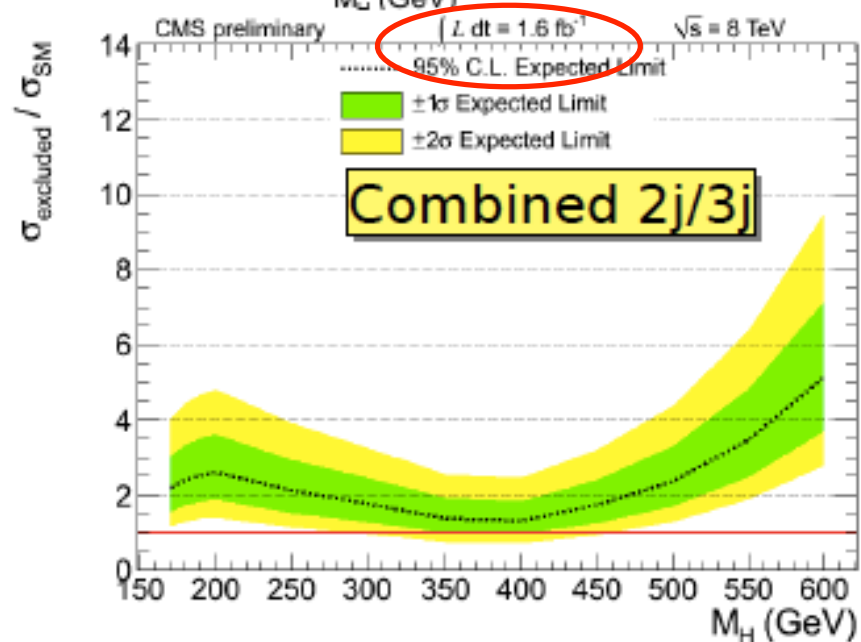
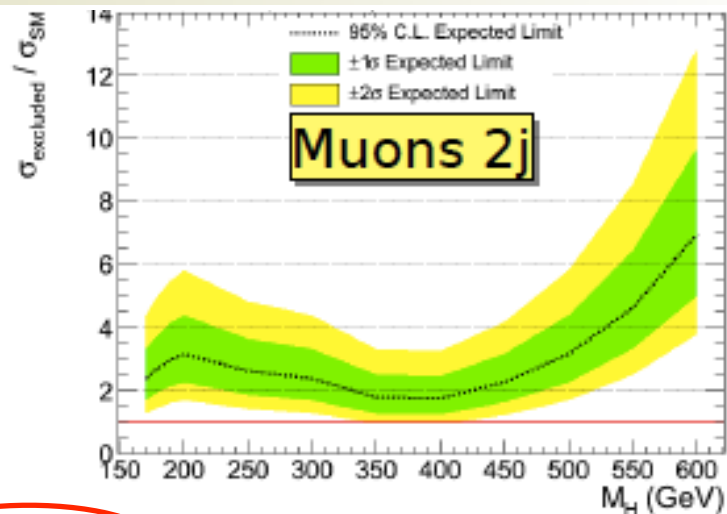
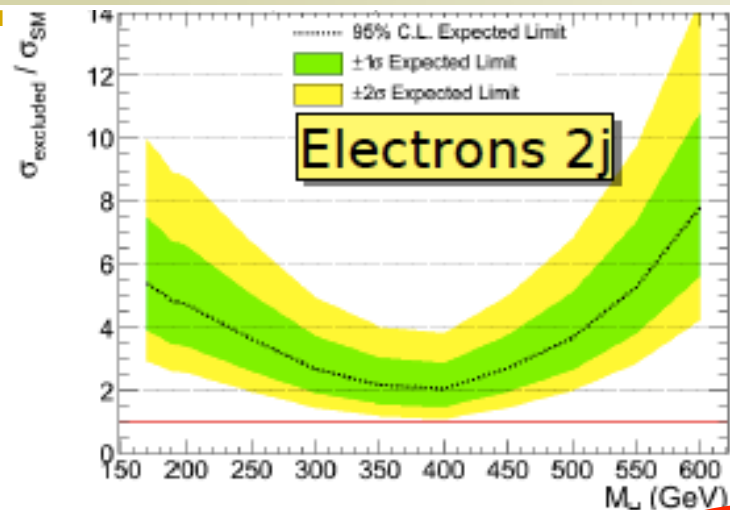
- Following the sideband method devised for **HIG-12-003**
- For **each of the 48 working points**, 3 regions are defined
  - $m_{ij} \in [65,95]$  GeV/c<sup>2</sup> (signal region)
  - $m_{ij} \in [55,65]$  GeV/c<sup>2</sup> (lower sideband, SBL)
  - $m_{ij} \in [95,115]$  or  $[95,200]$  GeV/c<sup>2</sup> for  $m_{\nu} <$  or  $\geq 250$  GeV/c<sup>2</sup> (higher sideband, SBH)
- The shape is extracted, in each bin  $i$ , as

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

the best  $\alpha$  value, as well as its uncertainty, are obtained on MC

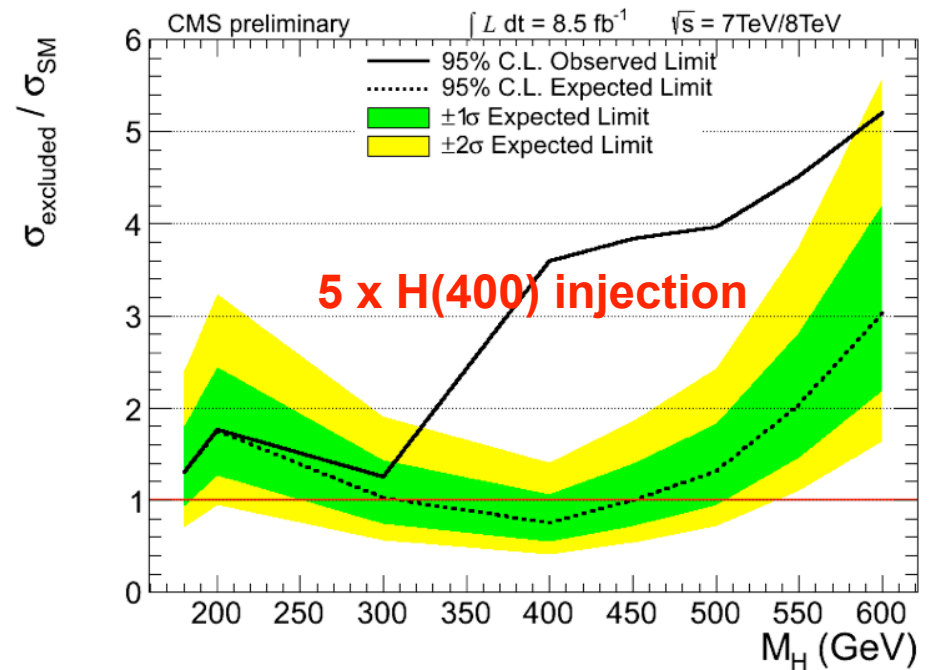
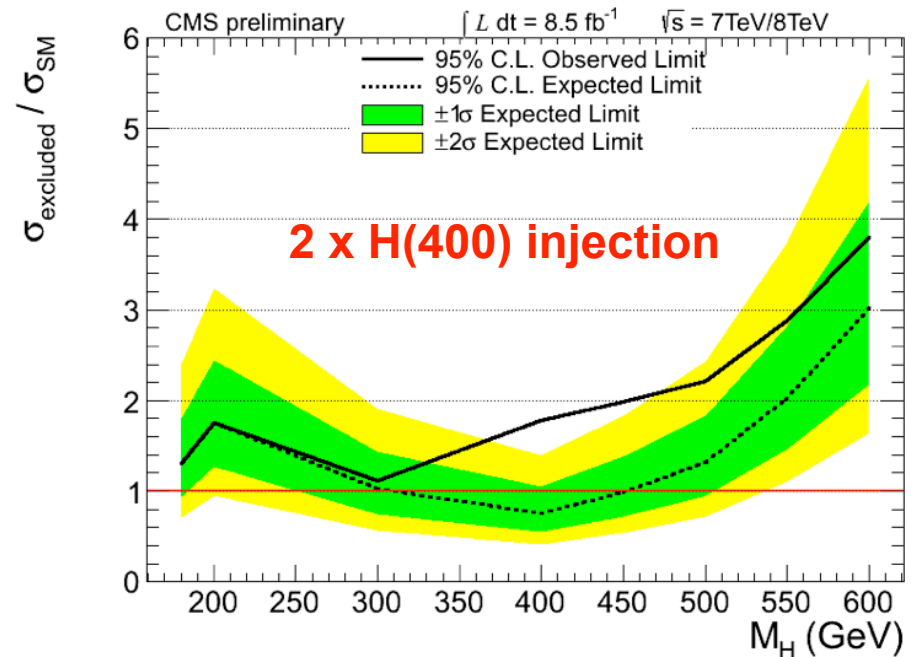
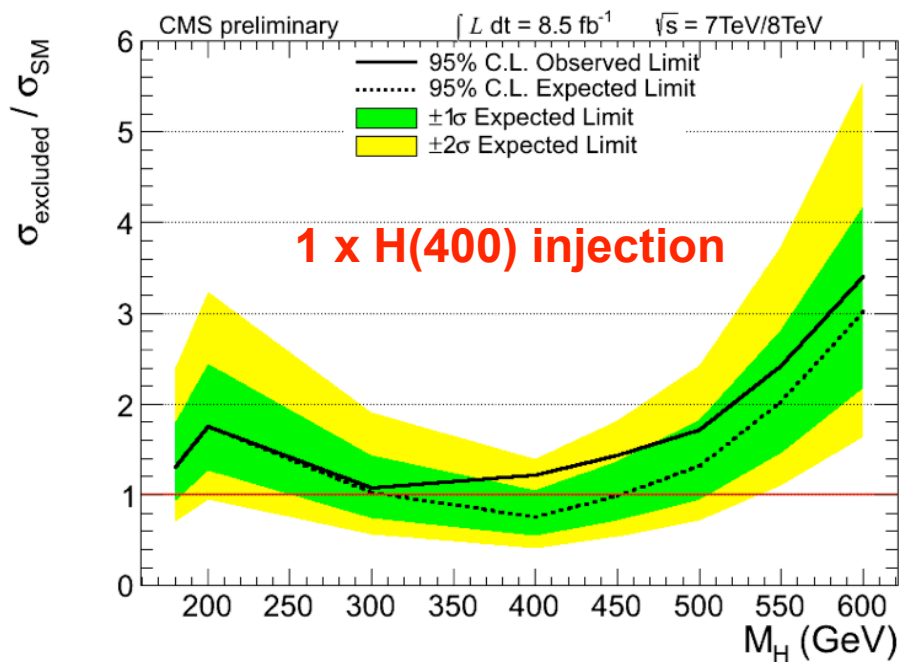


# Limit using $1.6 \text{ fb}^{-1}$ 8 TeV data

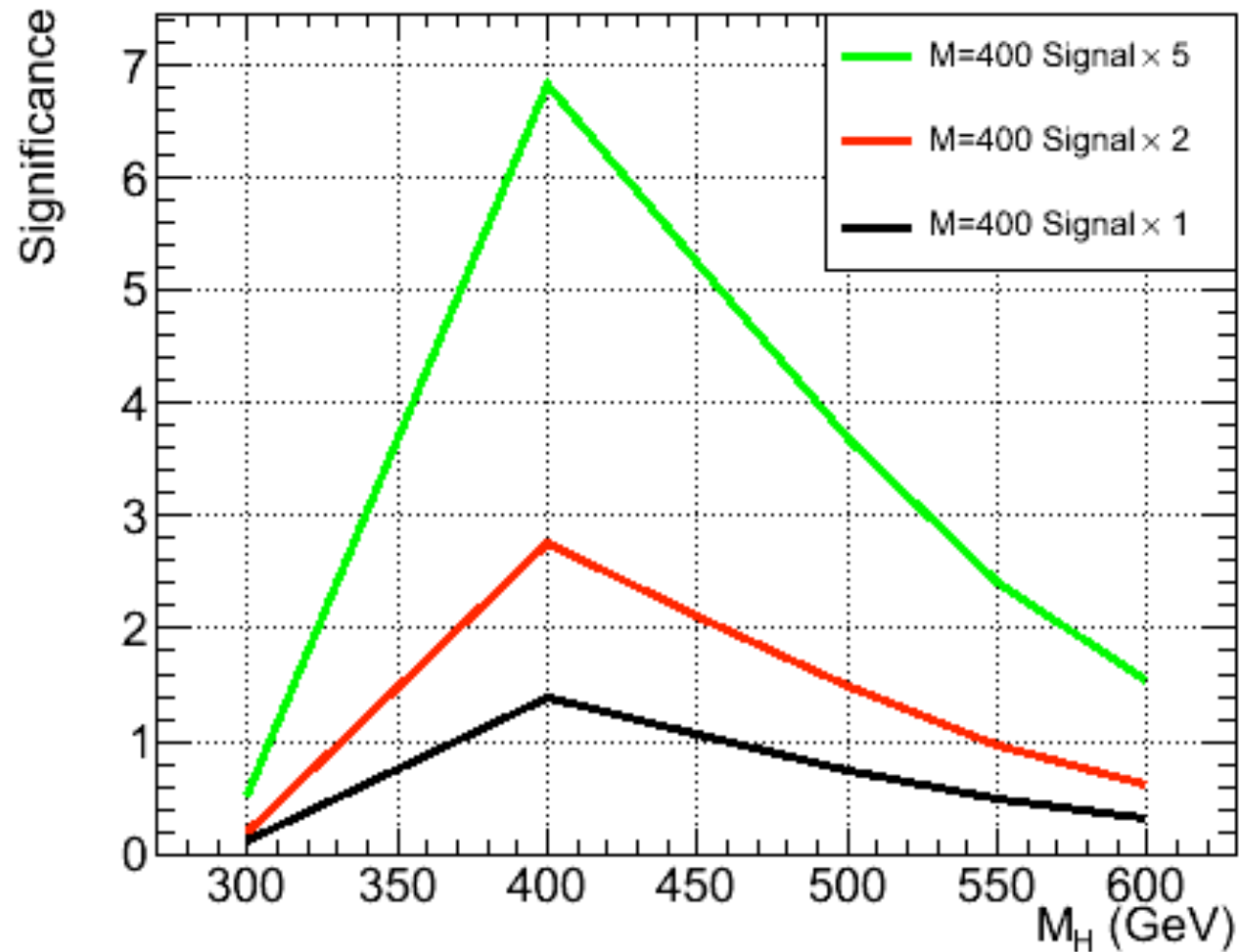


# Signal injection test

Inject Higgs signal and study its effect on observed limit (the expected limit doesn't change because background estimation comes from sideband data)



# Signal injection test: significance



If there was a signal we would have found it !