SEARCH FOR THE STANDARD MODEL HIGGS BOSON IN THE DECAY CHANNEL H→ZZ→2l2q AT CMS

Matthew Kress

Adviser: Daniela Bortoletto
Ph.D. Defense
December 2, 2013

Outline

2l2q Team

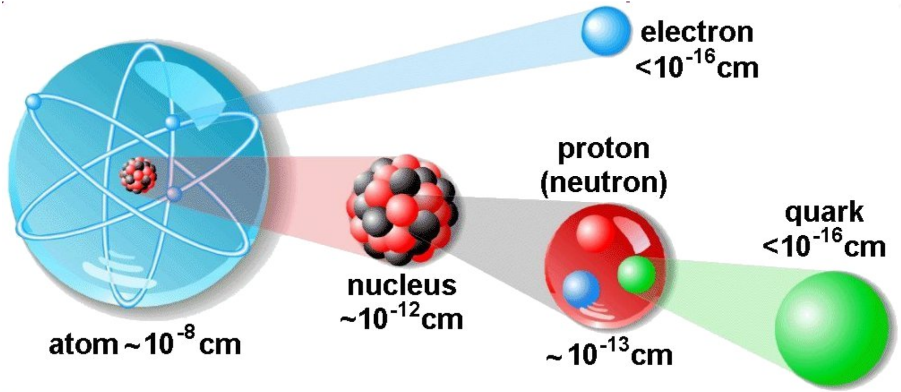
- *KIT:* M.U.Mozer
- *CERN:* A.Bonato, P.Lenzi, M.Mannelli
- *INFN-Padova:* K.Kanishchev
- *U. Basilicata & INFN-Napoli:* F. Fabozzi
- *CIEMAT:* O.González López, D.Domínguez Vázquez, P.García, J.Hernández, E. Navarro
- *Johns Hopkins U.:* A.Gritsan, S.Bolognesi, A.Withbeck
- *U & INFN-Napoli:* A.DeCosa
- *National Central U:* A.P.Singh, S.S.Yu, Y.J. Lu
- *Panjab:* L.K.Saini
- *Purdue:* D.Bortoletto, M.Kress, M.Vidal
- *Rochester:* R.Covarelli, P.Goldenzweig, R.Demina, A.Garcia-Bellido
- *UAM:* G.Codispoti, J.Fernández de Trocóniz
- *METU:* M.Yalvac
- *U & INFN-Firenze:* A.Tropiano, V.Gori, V.Ciulli, E.Gallo

Outline

1. Standard Model
2. Experiment
3. Event Selection
4. Signal Region Optimization
5. Statistical Analysis and Results
6. Conclusion

Sub-atomic World

Particle Physics is the study of the properties of the fundamental building blocks of the universe and the interactions between them.



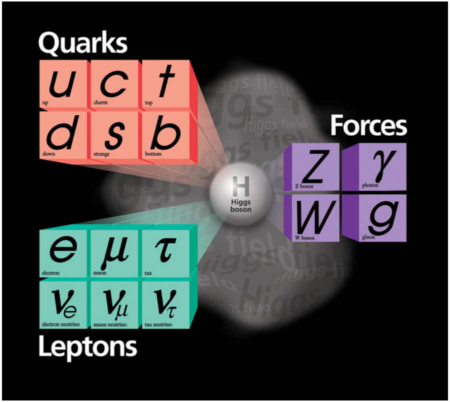
The Four Fundamental Forces

There are four fundamental forces that we know of.

Force	Boson	Charge	Mass
Gravitational	graviton(G)	0	?
Electromagnetic	photon(γ)	0	0
Weak	W boson($W^{\pm}$)	± 1	81 GeV
	Z boson(Z)	0	92GeV
Strong	gluon(g)	0	0

The Standard Model

The Standard Model is the compilation of over 100 years of scientific discoveries and is in excellent agreements with a wide range of experimental observations.



Particle Discoveries

Year	Discovery
1897	e discovery, by J.J. Thompson (cathode ray tube, UK)
1919	proton, Ernest Rutherford (UK)
1930	neutron, James Chadwick (UK)
1936	m, Carl D. Anderson at Caltech
1947	strange quark(K+=usbar, K-=subar)
1956	ν_e discovery (nuclear reactor)
1962	ν_μ discovery at BNL
1968	u and d quark (quark model)
1974	c quark (BNL, SLAC,J/y=ccbar)
1977	tau discovery (SLAC)
1977	b quark (Upsilon, FNAL)
1979	gluon (DESY)
1983	W and Z (CERN)
1995	top quark
2000	ν_τ discovery (Fermilab)
2012	Higgs (CERN)

- A simple solution for the nature of the electroweak symmetry breaking is introducing the Higgs field.
- The Higgs field is also the means by which elementary particles acquire mass.
- The Higgs boson is the smallest possible excitation of the Higgs field.
- The Higgs mass is a free parameter.

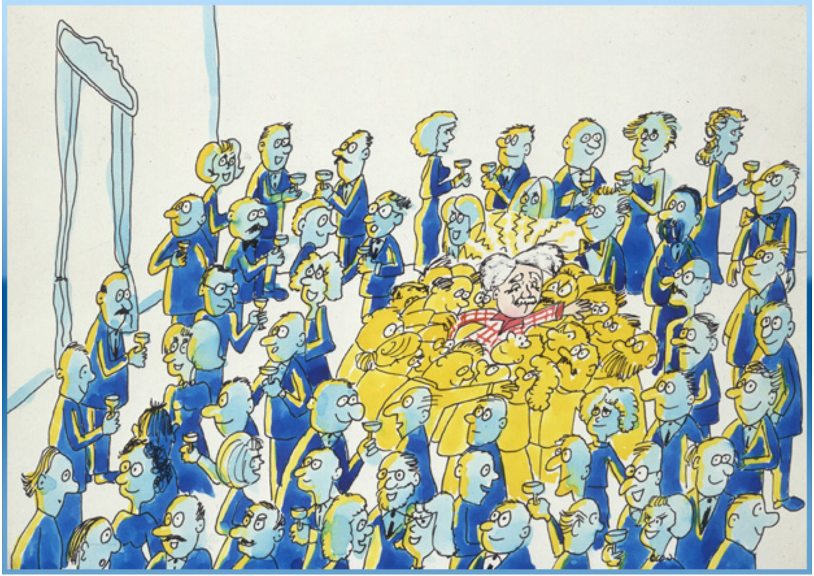
Higgs “Description”

Higgs “Description”

Higgs "Description"

Higgs "Description"

Higgs "Description"



Higgs “Description”

Higgs “Description”

Outline

Particle accelerators accelerate particles to high energies.

This allows us to:

- Look deeper into matter ($E \propto \dfrac{1}{size}$). "microscope"
- Discover new heavier particles ($E = mc^2$).
- Probe early conditions of the Universe ($E = kT$).

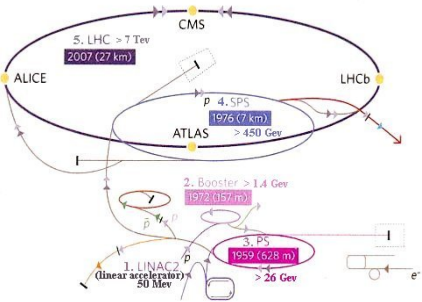
Before the particle accelerator

All this while being controlled in the laboratory.

The LHC Accelerator

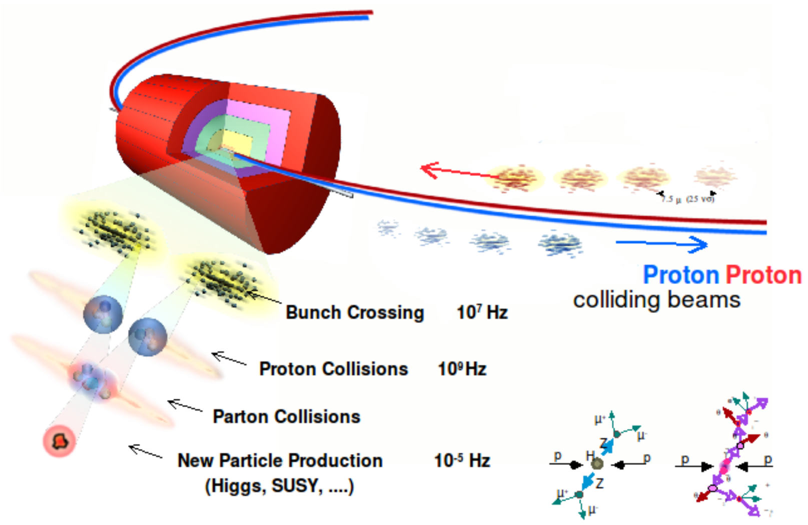
- Proton-proton collider
- Circumference: 26.7 km
- Tunnel: 100 meters underground
- dipoles operate at 8.3 T
- 1232 superconducting Niobium-Titanium magnets
- better vacuum and colder than inter-planetary space

Injection Scheme

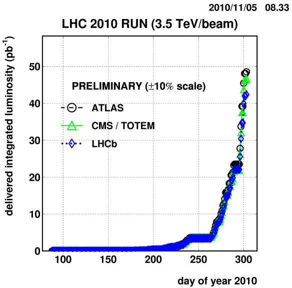


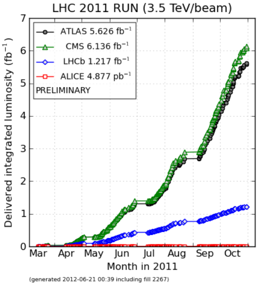
- Linac2 $\rightarrow$ 50 MeV
- Proton Synchrotron $\rightarrow$ 1.4 GeV
- Super Protron Synchrotron $\rightarrow$ 450 GeV
- LHC $\rightarrow$ 4.0 TeV

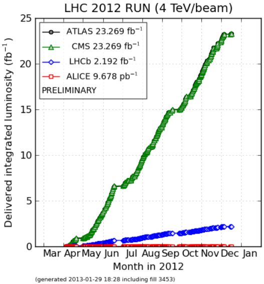
LHC Environment



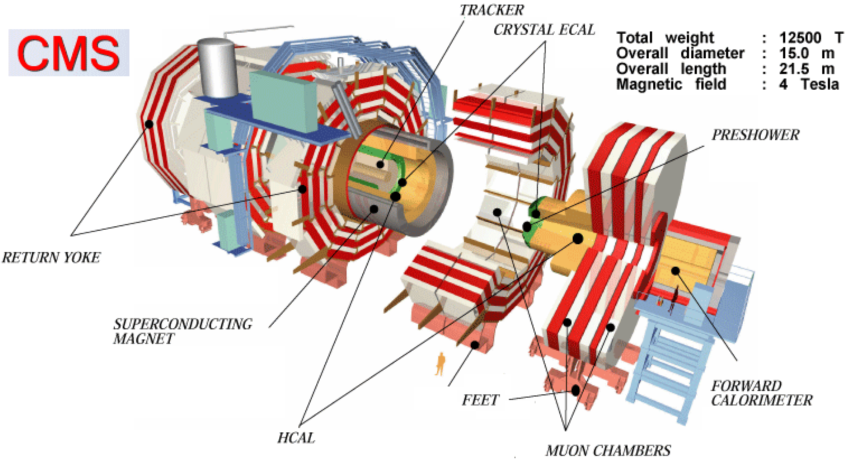
LHC Delivered Data



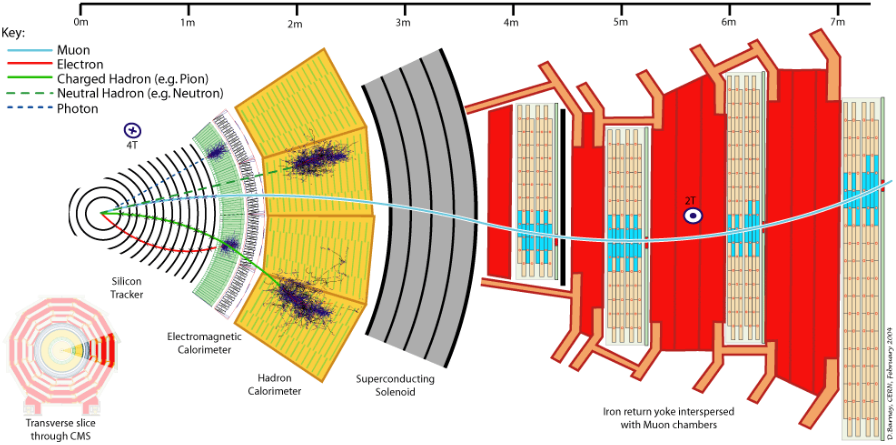




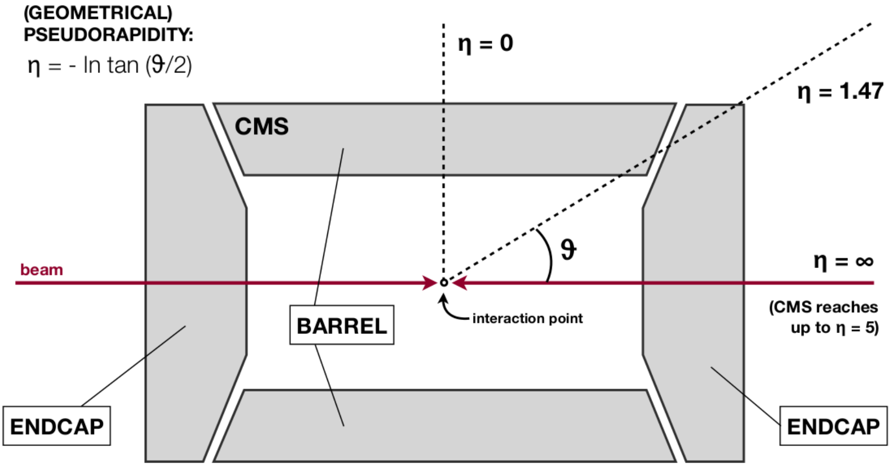
Compact Muon Solenoid



CMS Slice

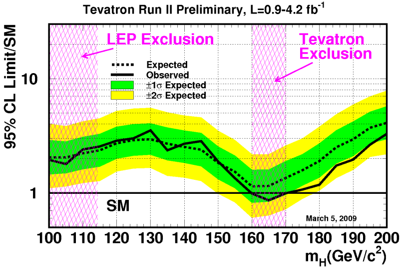

0m
1m
2m
3m
4m
5m
6m
7m
Key:
Muon
Electron
Charged Hadron (e.g. Pion)
Neutral Hadron (e.g. Neutron)
Photon
4T
2T
Silicon Tracker
Electromagnetic Calorimeter
Hadron Calorimeter
Superconducting Solenoid
Iron return yoke interspersed with Muon chambers
Transverse slice through CMS
D.Barney, CERN, February 2004

CMS Detector Glossary



The Missing Piece

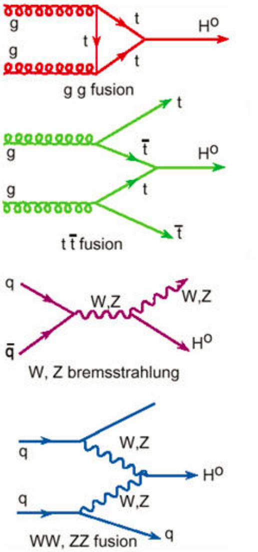
Before LHC



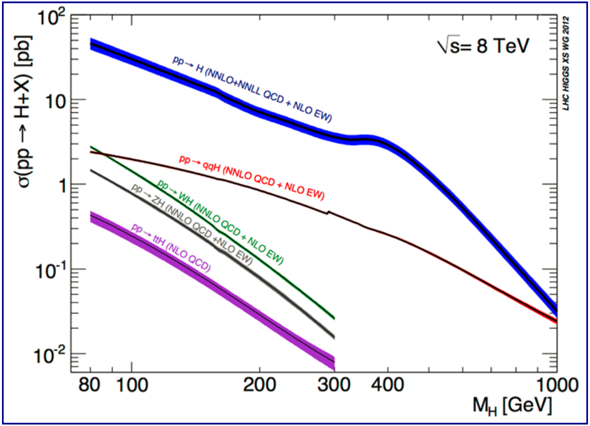
Experimental constraints:

- From direct searches at LEP and Tevatron.
- Indirect ones from LEP precision EWK measurements.

Higgs Production

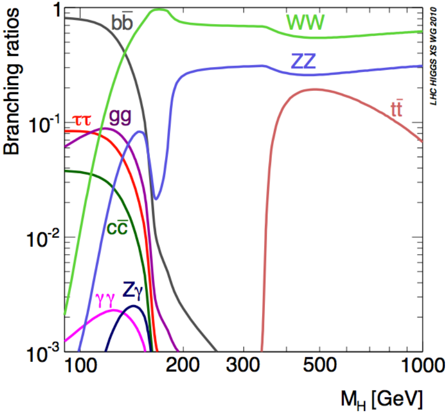


Gluon-gluon fusion ($gg \to H$) and vector-boson fusion ($qq \to qqH$) are dominant



Higgs Decay

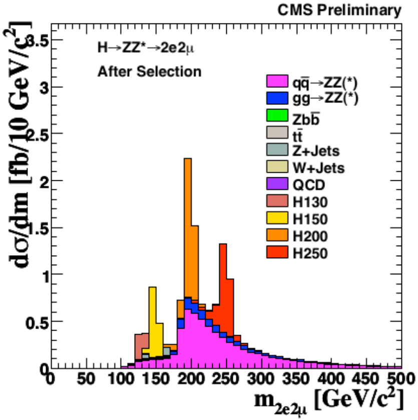
- Discovery strategy depends on the available decay channels.
- Decays with leptons provide clean signatures.

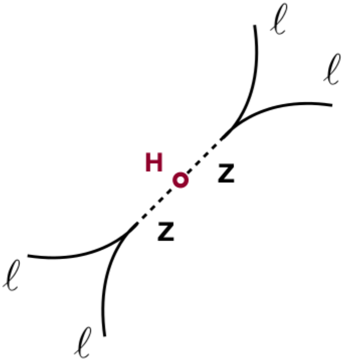


Main Discovery Channels

- $H \rightarrow \gamma\gamma$
- $H \rightarrow W^{+}W^{-}$
- $H \rightarrow ZZ$

- Decay to two Z bosons is the most promising discovery channel for $m_H > 180 GeV$.
- If both Z bosons decay to electrons or muons you get a very clean signature.

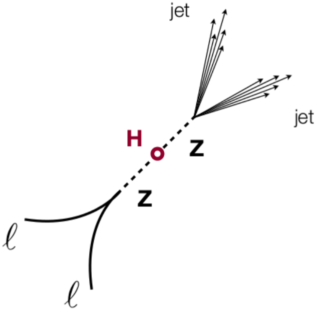




The $H \to ZZ \to llqq$ channel

Large Yields:

- BR($Z \to qq$) = 70%
- BR($ZZ \to 2l2q$) = 20 $\times$ $BR(ZZ \to 4l)$
- BR($ZZ \to 2l2q$) = 3.5 $\times$ $BR(ZZ \to 2l2\nu)$



Drawbacks

- Bad resolutions coming from jets.
- Large backgrounds coming from Z+jets

Full decay is reconstructed (closed kinematics)

Traditional Jet Approach: Caloremeter Jets

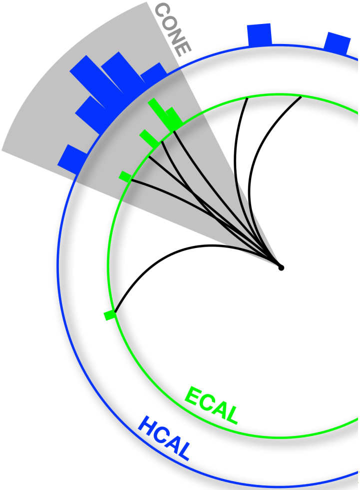
For CaloJets we assume most particles will reach the calorimeters close to each other. So all we have to do is cluster the energy deposits in the calorimeter.

Pros:

- straightforward
- fast and unaffected by event complexity

Cons:

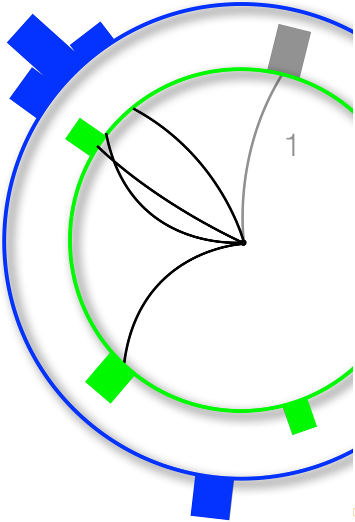
- response is p_T-dependent
- lose low p_T charged particles
- HCAL resolution



Particle Flow

The particle flow algorithm attempts to reconstruct all stable final state particles using all of the CMS sub-detectors.

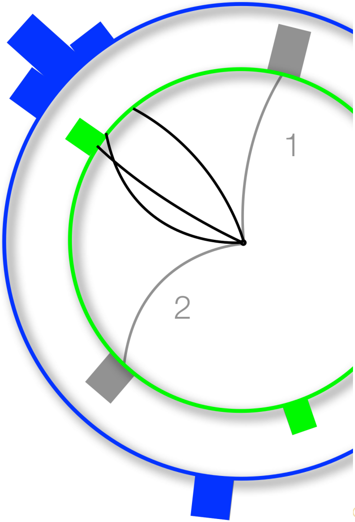
- 1. Get tracks linked to a single ECAL cluster if electron ID says OK create an electron.



Particle Flow

The particle flow algorithm attempts to reconstruct all stable final state particles using all of the CMS sub-detectors.

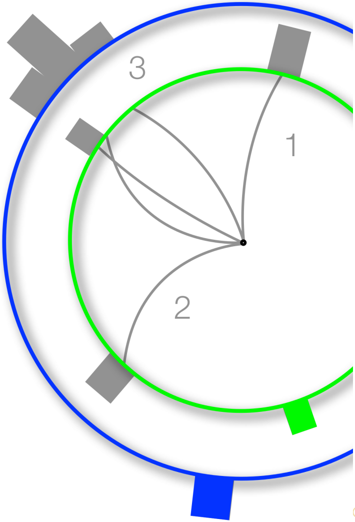
- 1. Get tracks linked to a single ECAL cluster if electron ID says OK create an electron.
- 2. If not then if there are compatible hits in muon chambers create a muon if not create a charged hadron.



Particle Flow

The particle flow algorithm attempts to reconstruct all stable final state particles using all of the CMS sub-detectors.

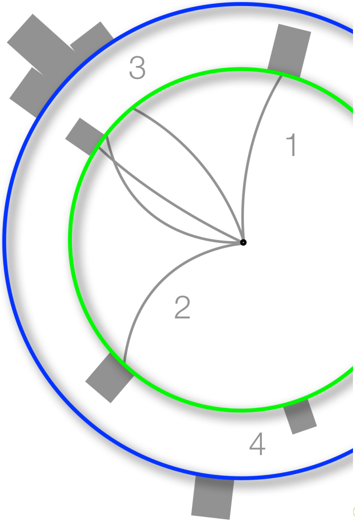
- 1. Get tracks linked to a single ECAL cluster if electron ID says OK create an electron.
- 2. If not then if there are compatible hits in muon chambers create a muon if not create a charged hadron.
- 3. For each HCAL cluster get all the linked tracks and all ECAL clusters linked to tracks. If the energy in the calorimeters is compatible with the tracks then create a charged hadron for each track. If the energy is greater than the tracks create photons or neutral hadrons equal to the missing calorimeter energy.



Particle Flow

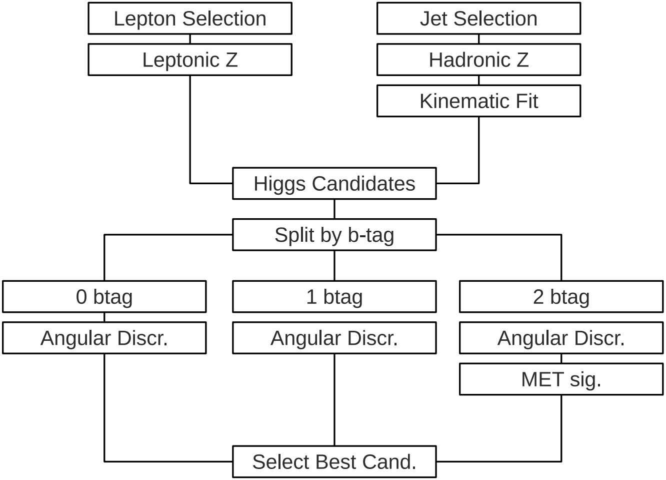
The particle flow algorithm attempts to reconstruct all stable final state particles using all of the CMS sub-detectors.

- 1. Get tracks linked to a single ECAL cluster if electron ID says OK create an electron.
- 2. If not then if there are compatible hits in muon chambers create a muon if not create a charged hadron.
- 3. For each HCAL cluster get all the linked tracks and all ECAL clusters linked to tracks. If the energy in the calorimeters is compatible with the tracks then create a charged hadron for each track. If the energy is greater than the tracks create photons or neutral hadrons equal to the missing calorimeter energy.
- 4. For the remaining ECAL (HCAL) clusters that are not linked to tracks create a photon (neutral hadron).



Outline

Analysis workflow



Samples

Data is Double Electron and Double Muon datasets.

Background

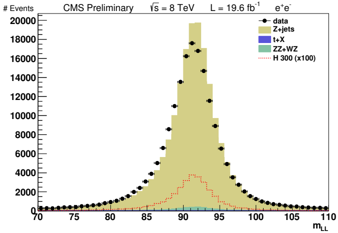
MC Sample	Generator	Percent of Final Region
Z+Jets	MADGRAPH	80%
tt	PYTHIA	7%
ZZ	PYTHIA	11%
WZ	PYTHIA	2%
WW	PYTHIA	<1%

The signal Monte Carlo samples are POWHEG.
m_H =120 to 1000 GeV

We are using the official prescriptions from the POGs for both Lepton ID and Isolation.

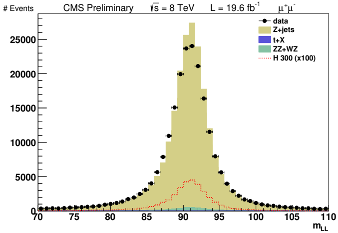
Electrons

- $p_T > 40/20$ GeV
- $|\eta| < 2.5$



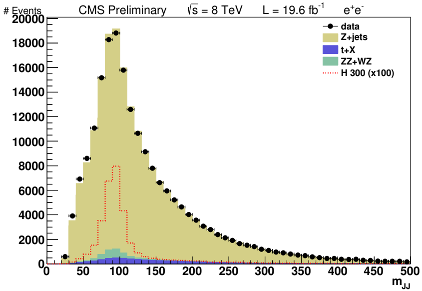
Muons

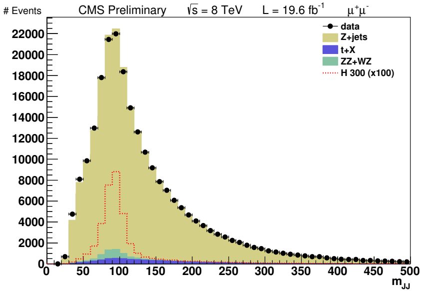
- $p_T > 40/20$ GeV
- $|\eta| < 2.4$



Jets

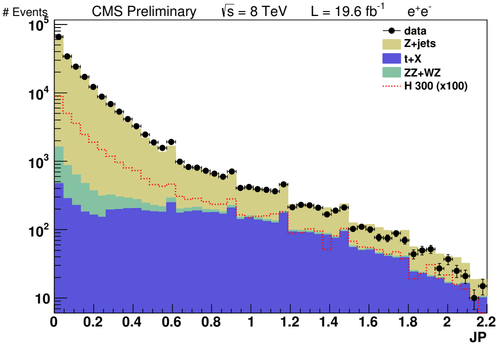
- AK5PF, L1FastJet+L2+L3
- p_T > 30 GeV
- $|\eta| < 2.4$





b-tagging

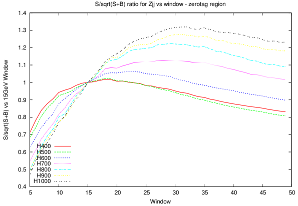
- Using **JP algorithm**
- We are looking for heavy quarks (b,c)
- There are "loose"(0.275) and "medium"(0.545) working points that are defined by the b-tag group that we use to classify our events.

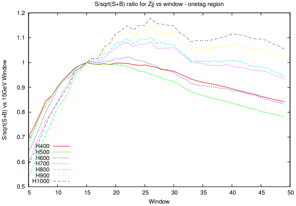


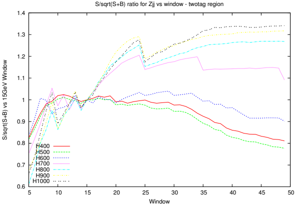
0 - tag	Both Jets < Loose
1 - tag	> Loose and < Medium
2 - tag	> Loose and > Medium

Window Cuts

Previous values of the windows around the m_{LL} and m_{JJ} peaks were done by eye. We optimized these windows to maximize $\frac{S}{\sqrt{S+B}}$ for a wide range of Higgs masses.

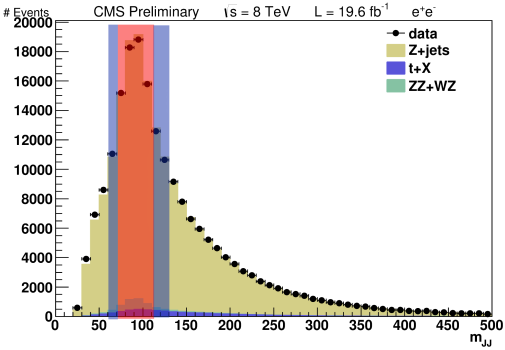






Signal and Sideband Regions
Study Regions

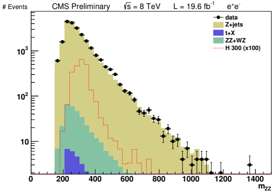
- Signal Region is $76 < m_{ll} < 106$ GeV and $71 < m_{jj} < 111$ GeV
- Also we are looking at the sidebands region, defined as $60 < m_{jj} < 71$ GeV || $111 < m_{jj} < 130$ GeV (i.e. outside signal window 75< mjj < 105 GeV).



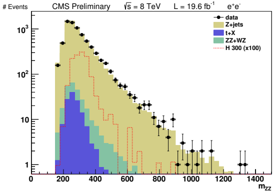
m_{ZZ} plots - Side Band Region

Electrons

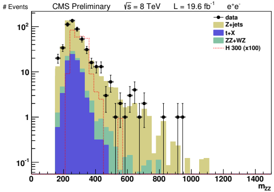
0-tag



1-tag

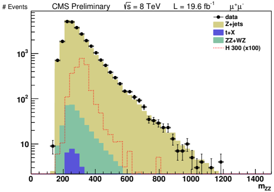


2-tag

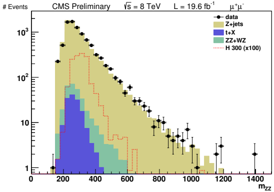


Muons

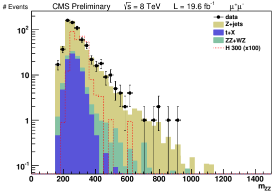
0-tag



1-tag

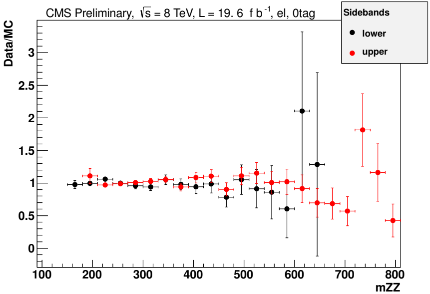


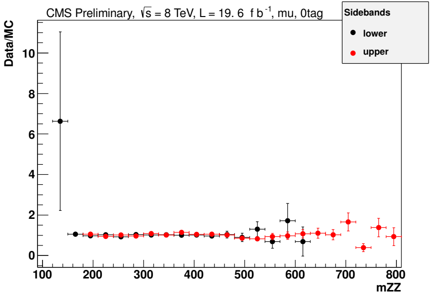
2-tag



Upper vs Lower Sideband

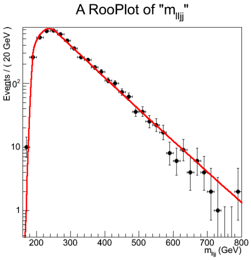
We also have very similar performance between the upper and lower sidebands in the m_{ZZ} distribution so we are able to use them together.



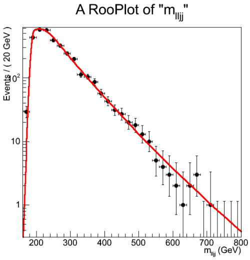


Background Estimation

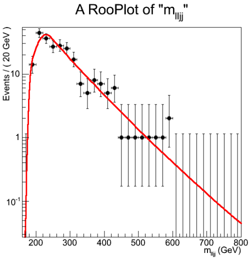
- We use the mjj sideband in data to get the normalization and a MC shape correction.
- The fit is to a Fermi*CrystallBall function



0-tag



1-tag

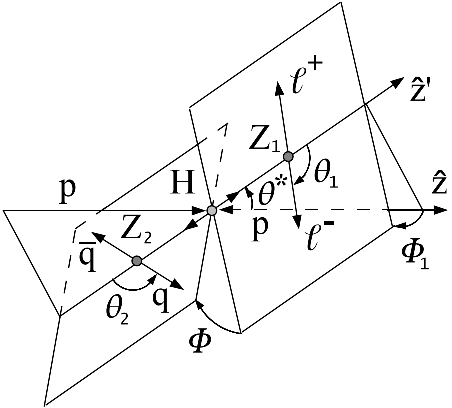


2-tag

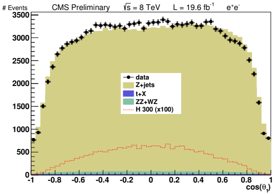
Outline

Helicity and Production Angles

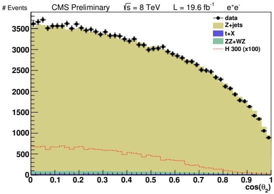
Final state kinematics completely determined by 5 angles.



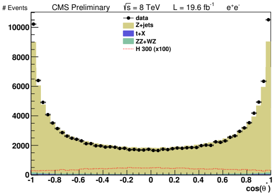
$cos(\theta^*), cos(\theta_1), cos(\theta_2), \Phi, \Phi_1$



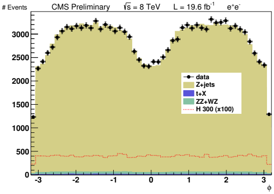
Events
CMS Preliminary
√s = 8 TeV
L = 19.6 fb⁻¹
e⁺e⁻
data
Z+jets
t+X
ZZ+WZ
H 300 (x100)
cos(θ1)



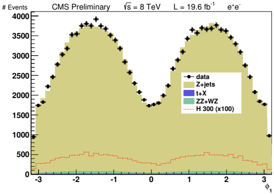
Events
CMS Preliminary
√s = 8 TeV
L = 19.6 fb⁻¹
e⁺e⁻
data
Z+jets
t+X
ZZ+WZ
H 300 (x100)
cos(θ2)



Events
CMS Preliminary
√s = 8 TeV
L = 19.6 fb⁻¹
e⁺e⁻
data
Z+jets
t+X
ZZ+WZ
H 300 (x100)
cos(θ*)

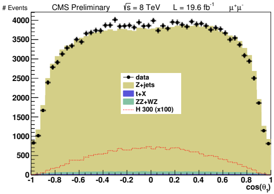


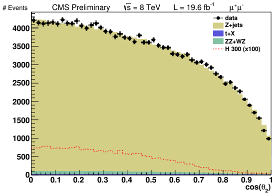
Events
CMS Preliminary
√s = 8 TeV
L = 19.6 fb⁻¹
e⁺e⁻
data
Z+jets
t+X
ZZ+WZ
H 300 (x100)
ϕ

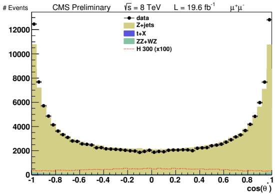


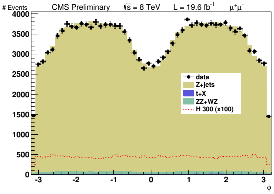
Events
CMS Preliminary
√s = 8 TeV
L = 19.6 fb⁻¹
e⁺e⁻
data
Z+jets
t+X
ZZ+WZ
H 300 (x100)
ϕ1

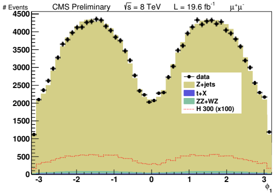
Muon Helicity and Production Angles





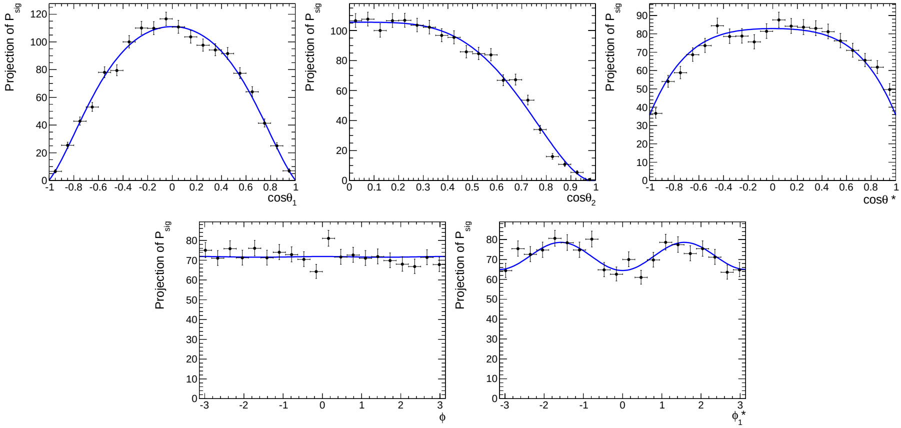




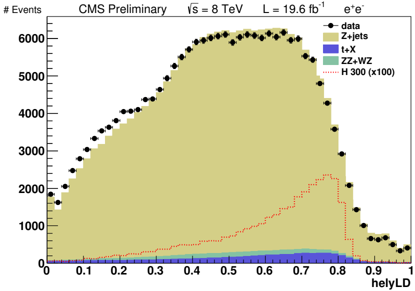


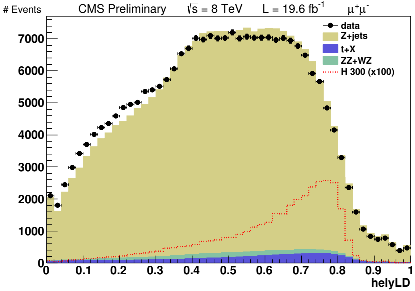
Signal Angular Distribution Fits

Example fits at for 500 GeV (475 - 550 GeV).



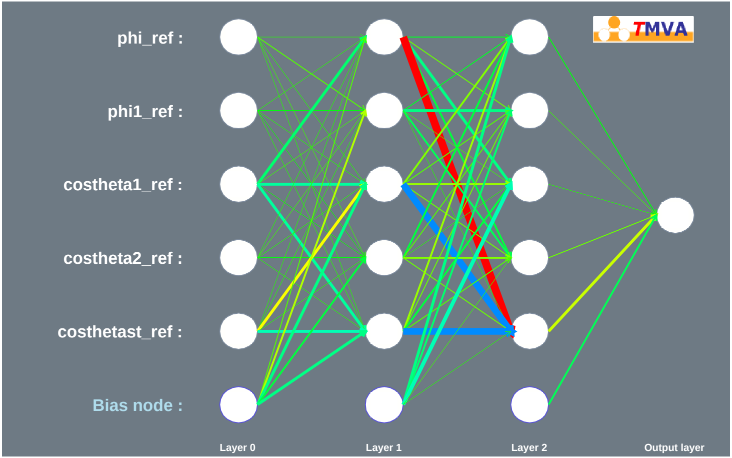
$$LD = \frac{P_{sig}}{P_{sig} + P_{bkg}}$$





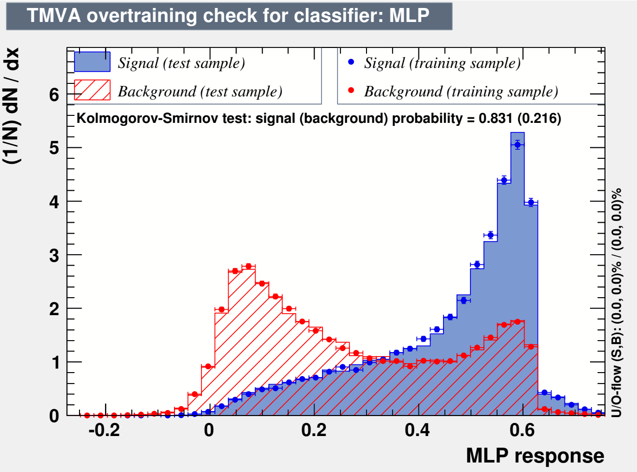
We cut on 0.5 for all of the b-tag regions.

Neural Network with TMVA Package

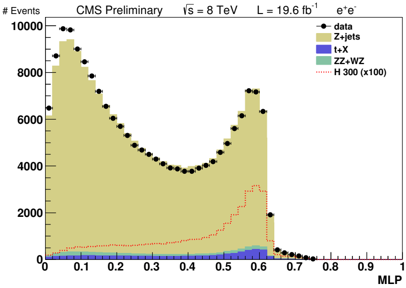


Neural Network Training and Testing

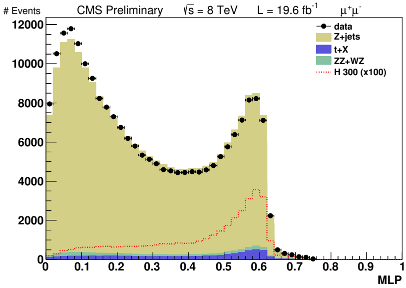
The trainings are done after preselection and additionally require at least one B-tagged Medium jet. (Trained on H 400 GeV)



MLP



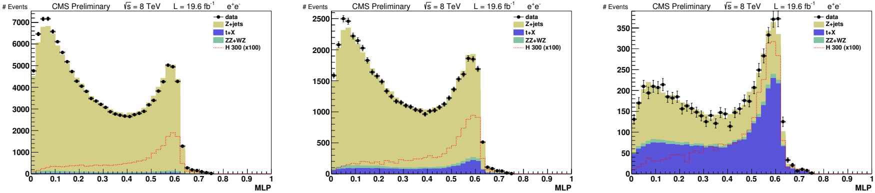
Events
CMS Preliminary
√s = 8 TeV
L = 19.6 fb-1
e+e-
data
Z+jets
t+X
ZZ+WZ
H 300 (x100)
MLP



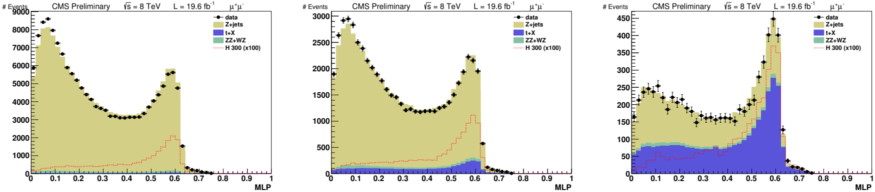
Events
CMS Preliminary
√s = 8 TeV
L = 19.6 fb-1
μ+μ-
data
Z+jets
t+X
ZZ+WZ
H 300 (x100)
MLP

MLP - In b-tag regions

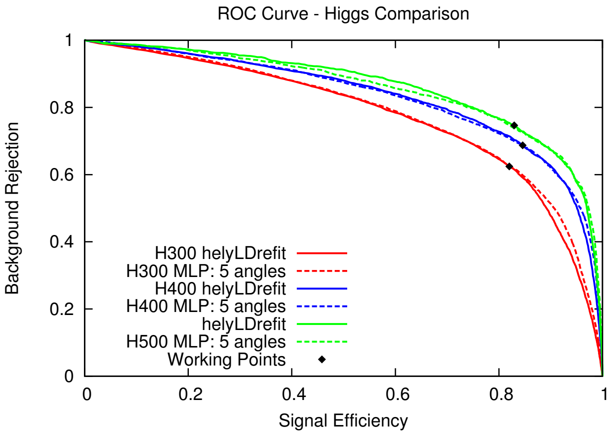
Electrons (zero,one,two)



Muons (zero,one,two)



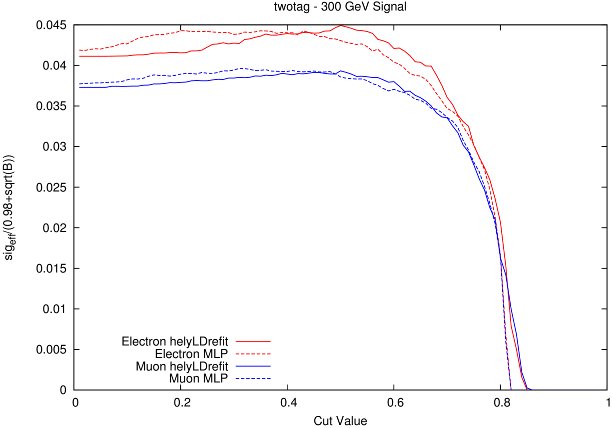
The working point is the background rejection point that we currently achieve in the two tag region in our analysis.



HelyLD Optimization

We do an optimization for the cut on the HelyLD maximizing the Punzi equation

$$\frac{sig_{eff}}{0.98+\sqrt{B}}.$$



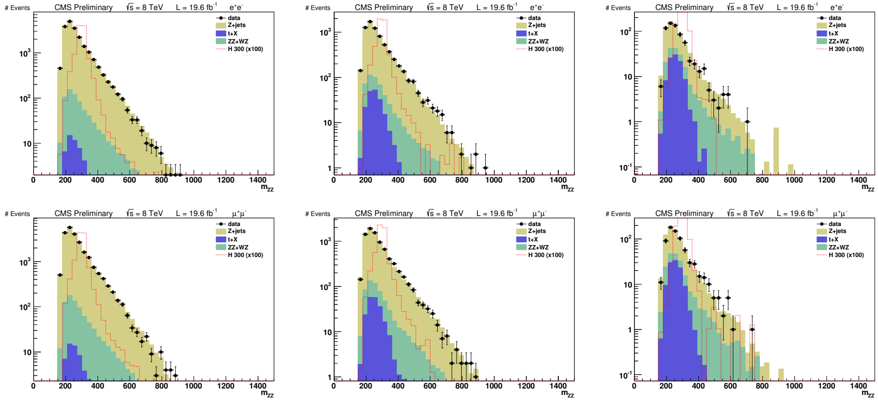
While this plot is for only one b-tag region and one mass we find 0.5 a very good number to cut on for all b-tag regions.

Outline

Selection Requirements

observable	0 b-tag	1 b-tag	2 b-tag
btag	none	JPL	JPL & JPM
helicity LD	> 0.5		
missing E_T significance	< 10		
m_{jj}	[71,111] GeV/c^2		
m_{ll}	[76,106] GeV/c^2		
$p_T(l^\pm)$	$>$ 40/20 GeV/c		
$p_T(jets)$	$>$ 30 GeV/c		
$\|\eta\|(l^\pm)$	$e^\pm < 2.5$, $\mu^\pm < 2.4$		
$\|\eta\|$(jets)	< 2.4		
lepton quailty jet quailty			

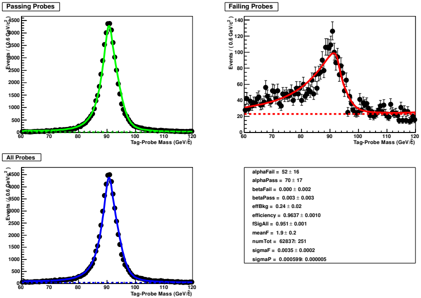
Final Region



Tag and Probe

- Method to use Z boson to calculate efficincy of Data and Monte Carlo.
- One lepton is “good” (tag) and the other is used for the calculations(probe).
- Compairing Monte Carlo to data gives us Scale Factors.

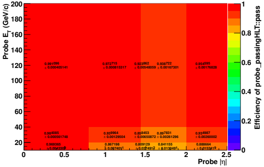
SuperCluster to GSF Electron



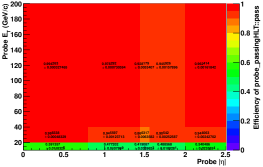
Triggers

- Muons
  - HLT_Mu17_Mu8
  - HLT_Mu17_TkMu8

Electron Leg 8 GeV WPLoose to HLT



Electron Leg 17 GeV WPLoose to HLT

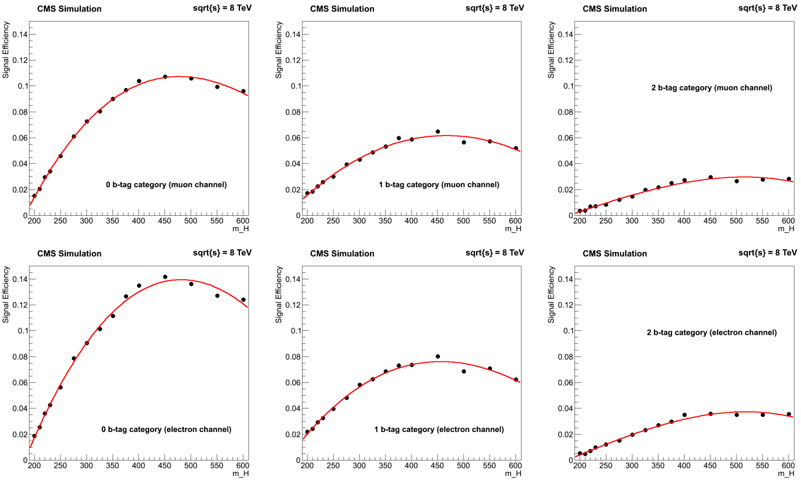


- Electrons
  - HLT_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL _Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL
- EMu (for backround estimation and analysis checks)
  - Mu8_Ele17_CaloIdT_CaloIsoVL
  - Mu17_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL

We are using the POG provided scale factors for electrons and computing the WP to HLT ourselves.

Efficiency Fit

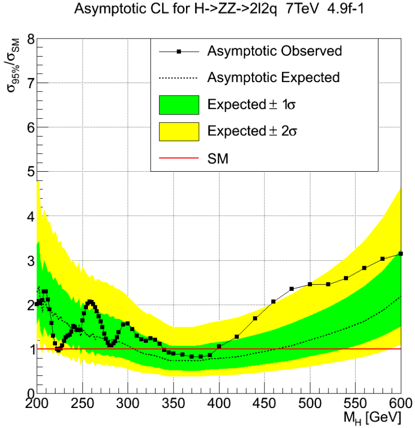
The signal efficiency as a function of the Higgs mass is fitted to a polinomial in order to be estimatated for those Higgs mass hypothesis where no Monte-Carlo sample is available



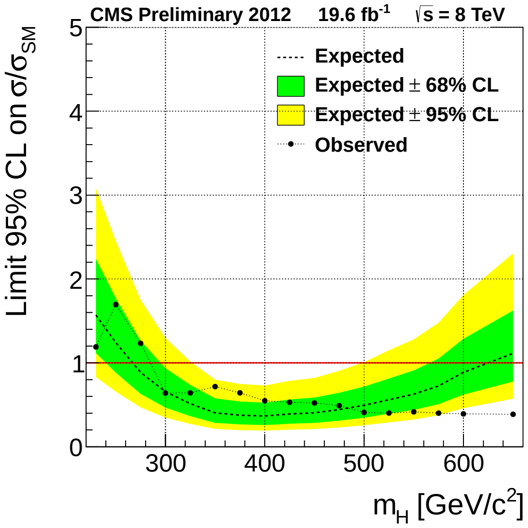
Systematics

source	0 b-tag	1 b-tag	2 b-tag	comment
muons reco		1.1%		tag-and-probe study
electrons reco		3.6%		tag-and-probe study
jet reco		1%–5%		JES-uncert., JER uncert. negligible; correlated between categ
pileup		1-2%		correlated between categ
b-tagging	1-4%	1-5%	5-8%	anti-correlated between categ.
MET	–	–	$< 1\%$	loose requirement
production mechanism (PDF)		1-4%		PDF4LHC, acceptance only
production mechanism (lineshape)		0-3%		only for $M_H > 400$
luminosity		4.4%		same for all analyses
Higgs cross-section (for R)		13–18%		detailed table from YR available

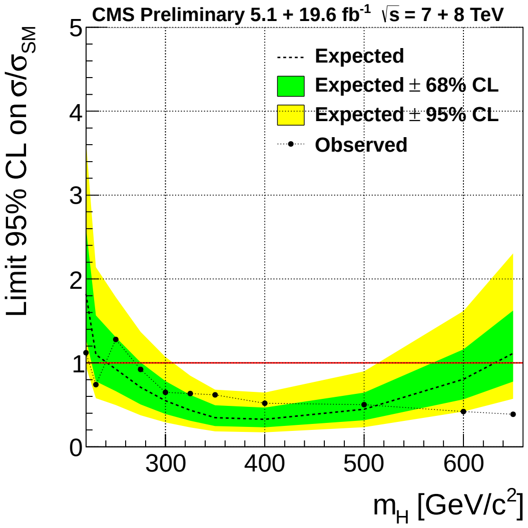
Limit on the expected 95% CL upper limit on the product of the Higgs boson production cross section and the branching fraction of H→ZZ (black dots.), for the recorded luminosity in 2011 of 4.9 fb^{-1} at 7TeV. Yellow and Green bands represent the 68% and 95% ranges of expectation.



2012 Preliminary Results

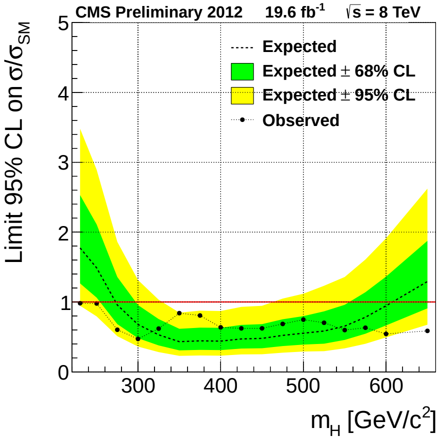


Combined Results



MVA Limit

Using the MVA cutting at Higgs 400 GeV equivalent working point (as shown before) is only slightly worse results than the HelyLD results.



Outline

Conclusion

- No evidence of a Higgs boson signal is found.
- At $95\%_{CL}$ a Standard Model Higgs boson is excluded in the range of 300 Gev to 650 GeV.
- Similar Results are obtained using both the HelyLD and the MLP classifier using the kinematic information of the events.

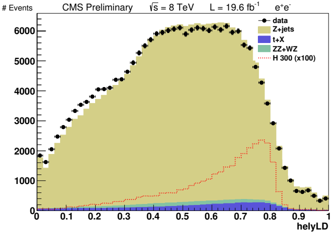
Thank You

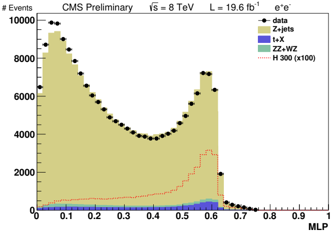
Backup

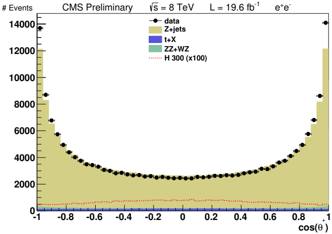
BACKUP

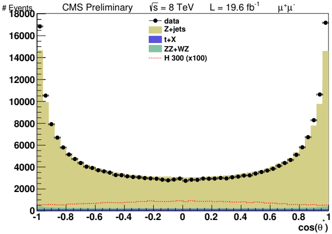
MLP and HelyLD Data/MC Discrepancy

Both the HelyLD and the MLP don't agree in the background like regions between data and MC. This is because $cos(\theta^*)$ is not well modeled in MC.



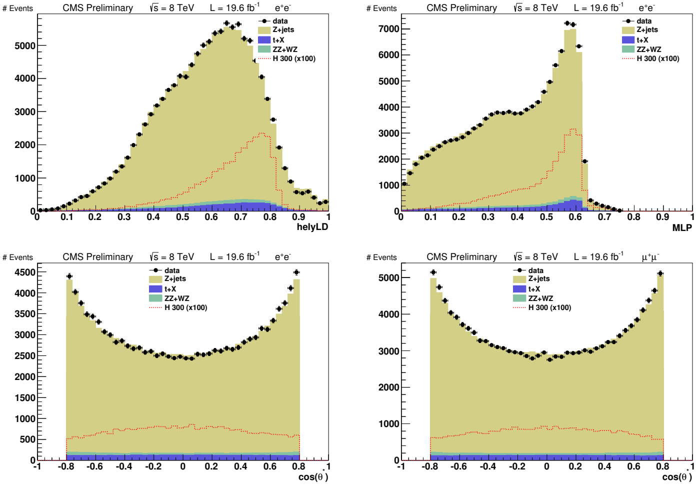






Solution

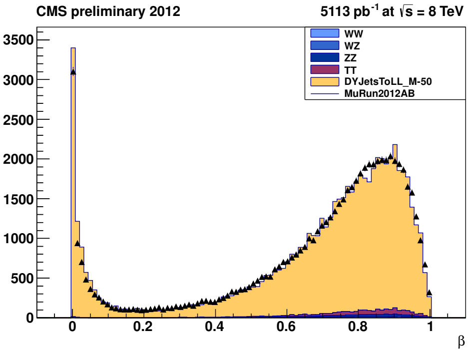
This is a MC issue that can be fixed by cutting on $|cos(\theta^*)| > 0.8$ or by refitting MC to data.



Pile-up Rejection

β is the sum of transverse momenta of all charged particles in the jet coming from the primary vertex, normalized to the total sum of transverse momenta of all charged particles in the jet.

- Using β variable to remove candidates with PU-like jets
- Cutting on $\beta \geq 0.2$



$\frac{Signal}{\sqrt{Signal + Background}}$

400GeV Training

	Electron			Muon		
Sample	zero	one	two	zero	one	two
GG600	0.215	0.224	0.309	0.232	0.210	0.345
GG700	0.086	0.089	0.121	0.088	0.083	0.134
GG800	0.035	0.036	0.052	0.036	0.033	0.057
GG900	0.016	0.016	0.023	0.016	0.015	0.025
GG1000	0.009	0.009	0.012	0.009	0.008	0.013

800GeV Training

	Electron			Muon		
Sample	zero	one	two	zero	one	two
GG600	0.204	0.223	0.297	0.225	0.198	0.344
GG700	0.084	0.091	0.118	0.088	0.082	0.140
GG800	0.035	0.039	0.053	0.038	0.034	0.061
GG900	0.017	0.018	0.023	0.017	0.016	0.027

Difference in $\frac{Signal}{\sqrt{Signal + Background}}$ 400 vs 800

	Electron			Muon		
Sample	zero	one	two	zero	one	two
GG600	4.93%	0.18%	4.17%	2.98%	5.49%	0.33%
GG700	1.72%	-2.75%	2.58%	0.36%	1.57%	-3.91%
GG800	-1.71%	-7.89%	-2.51%	-4.42%	-2.47%	-7.02%
GG900	-2.87%	-8.16%	0.30%	-5.16%	-4.40%	-8.41%
GG1000	-2.62%	-8.38%	-1.75%	-4.26%	-2.13%	-10.91%