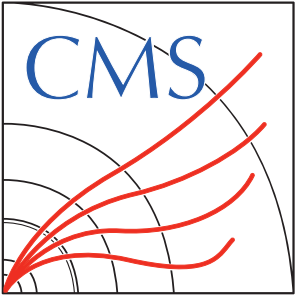


Search for the Standard Model Higgs Boson Produced in Vector Boson Fusion and Decaying to Bottom Quarks

Gurpreet Singh*, Burin Asavapibhop, Narumon Suwonjandee

Chulalongkorn University, Thailand

Siam Physics Congress - 2015

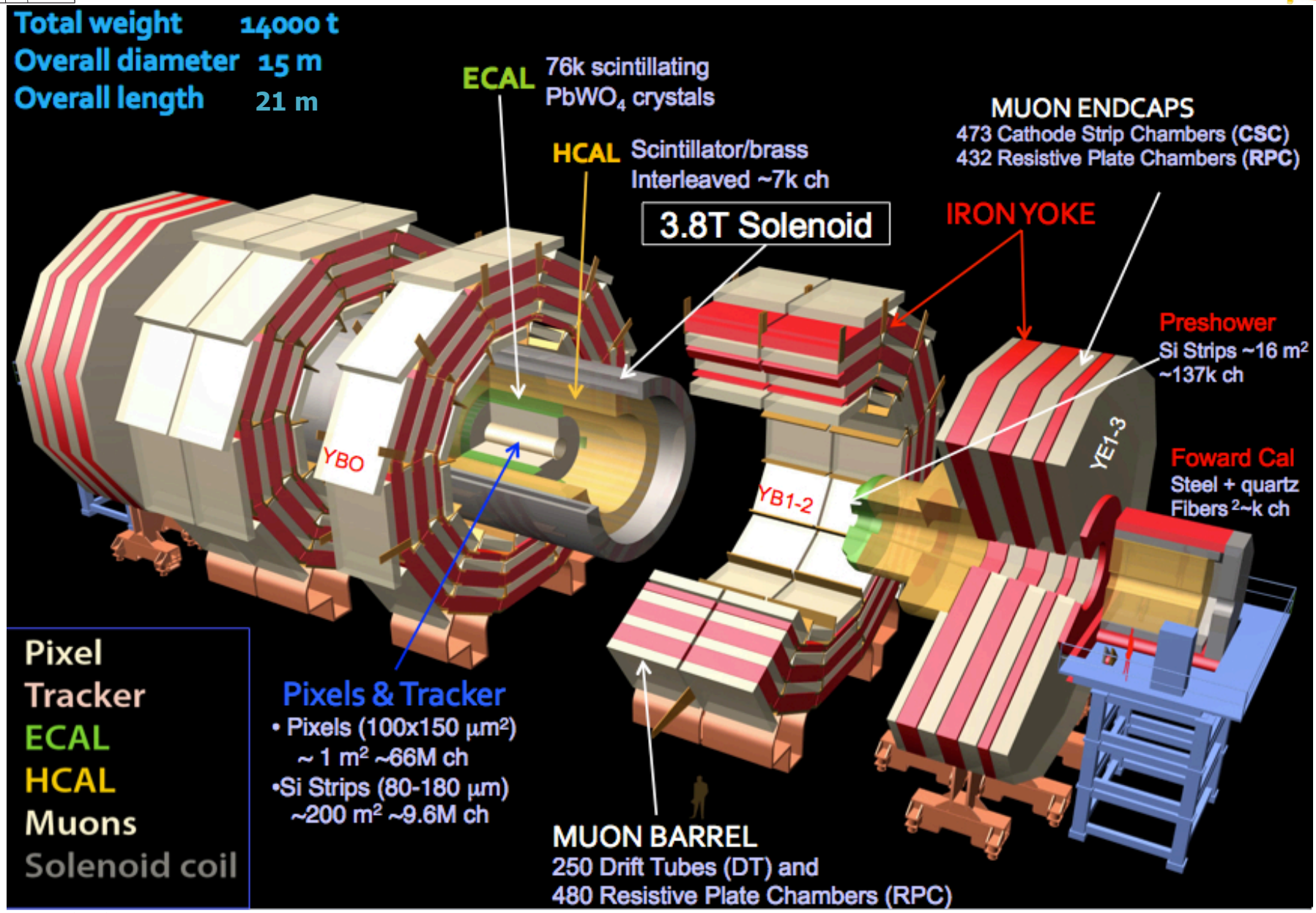
May 20, 2015

*Gurpreet.Singh@cern.ch

Prologue

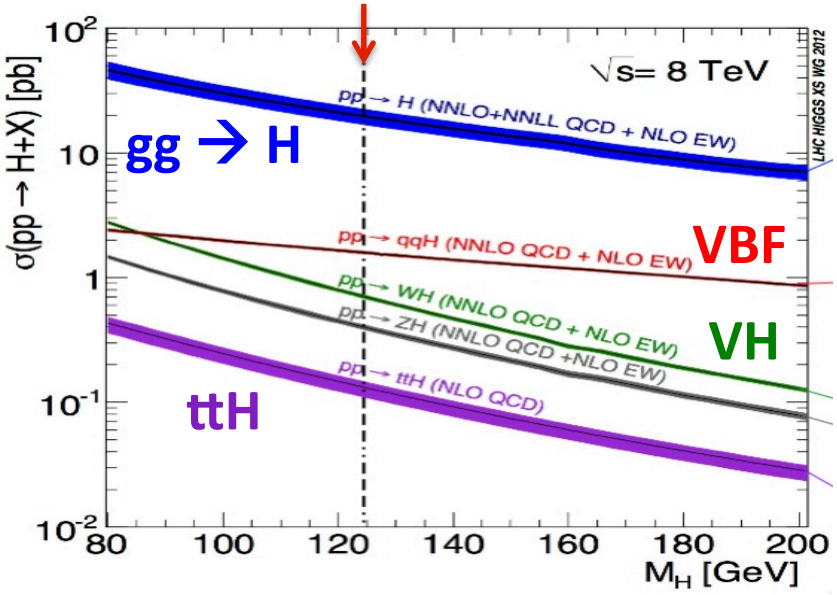
- Compact Muon Solenoid (CMS) Experiment
- Standard Model Higgs Boson Production at Large Hadron Collider (LHC)
- Vector Boson Fusion - Hbb search in a Nutshell
- Event Selection Strategy
- Pre-selection
- Multivariate Analysis
- Results
- Summary

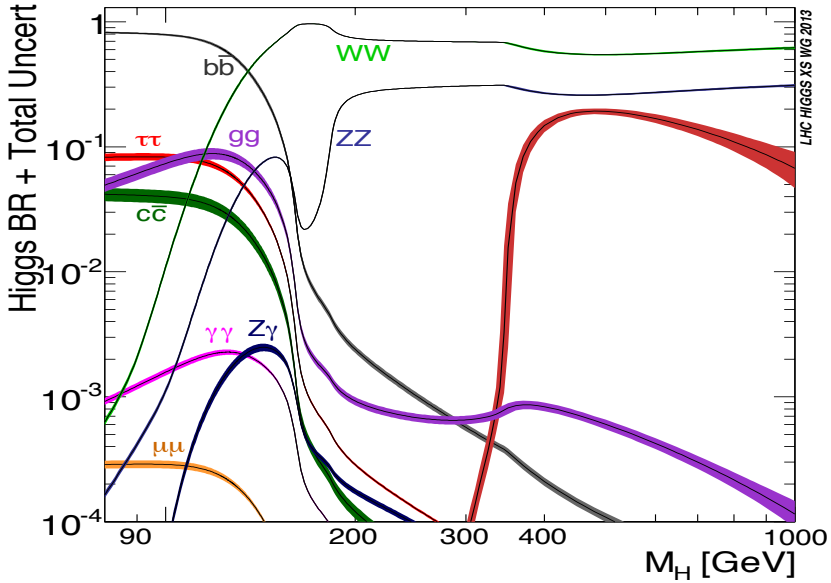
The CMS Detector

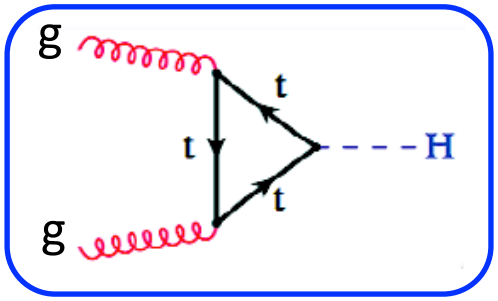


Standard Model Higgs Boson Production at LHC

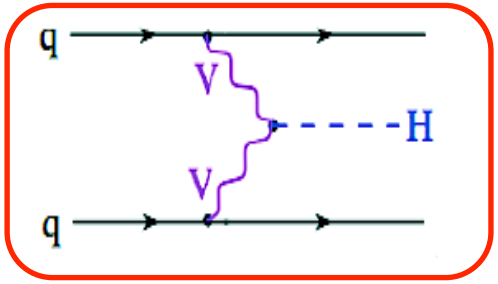
We are here ~125 GeV



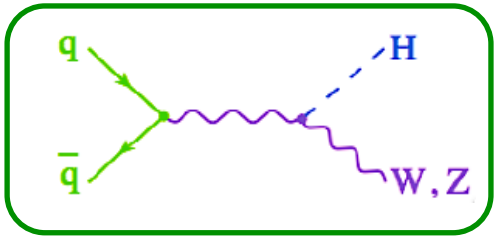




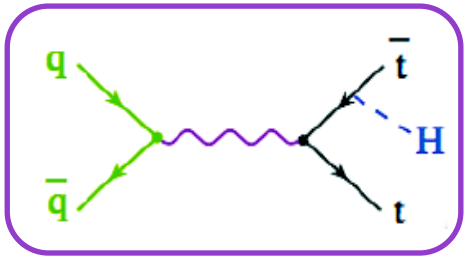
Gluon – gluon fusion



Vector – boson fusion (VBF)



Associated production with Vector bosons

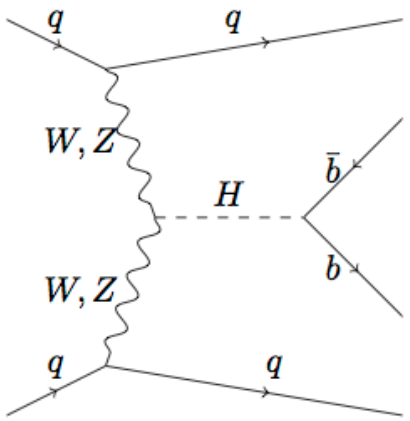


Associated production with top quark

VBF Hbb in a Nutshell

Overview

- Important to establish the nature of Higgs boson
- Anomaly in the Higgs boson coupling - hint for New Physics.
- At m_H=125, H→bb has ~60% branching fraction
  - QCD background pp→bb for **ggF** is > 10^7 times larger.
- Improve the sensitivity of H→bb searches (VBF, VH, ttH)
- Experimental challenges are trigger selection, overwhelming QCD background
- Significantly large cross section



Proton-proton collision data = 19 fb^{-1}
Center-of-mass energy ($\sqrt{s}$) = 8 TeV

Signal Topology

- 4 energetic hadronic jets
- Two quark-jets with high mass (m_{qq}) in forward/backward direction; large pseudo-rapidity (η) difference; low hadronic activity between them
- Two b-jets in central region of detector
- No QCD color exchange between b-jets and VBF jets
- ggF → H with associated 2 QCD jets – considered as signal

Backgrounds

QCD multijet production
Z+jets
ttbar
single-top
W+jets

Event Selection Strategy - I

Search Strategy

- Topological trigger on the signal main properties
- Multivariate methods to exploit significant differences between signal and QCD
- Perform a search for a resonance structure ("bump hunt") on the m_{bb} spectrum

Event Interpretation

- At least one good primary vertex and 4 reconstructed PF jets
- Order jets according to:
  - Transverse momentum (p_T)
  - 4 leading jets in p_T define the searched signal final state
  - By Combined Secondary Vertex (CSV) BTag value
    - b quarks jets [highest CSV]; qq jets [lowest CSV]
  - By η value
    - b quarks jets in central region; qq jets in forward – backward region

pre-selection

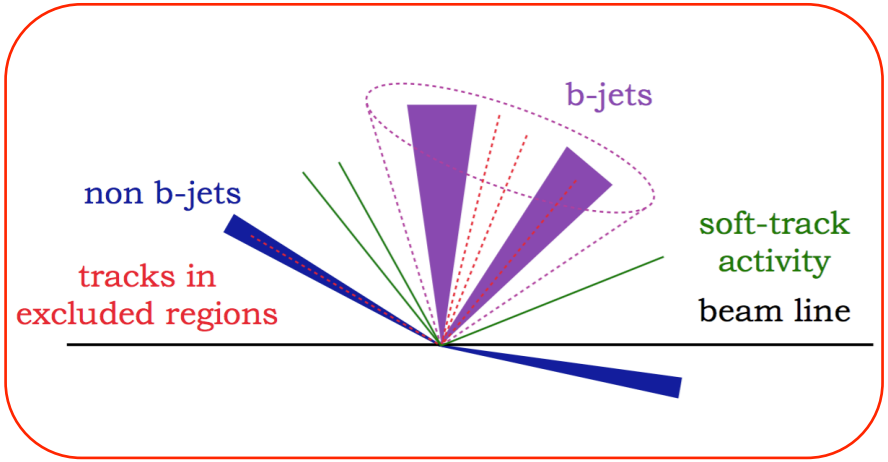
- $p_{T,1} > 85$ GeV
- $p_{T,2} > 70$ GeV
- $p_{T,3} > 60$ GeV
- $p_{T,4} > 40$ GeV
- $m_{q\bar{q}} > 300$ GeV
- $\Delta\eta_{q\bar{q}} > 2.5$
- $\Delta\phi_{b\bar{b}} < 2.0$

Event Selection Strategy - II

- CSV b-tagger: Get 2 most b-tagged jets
  - Identification of jets that likely originate from the hadronization of b-quarks
    - To reject fake jets from detector noise,
    - Jets heavily contaminated with pileup energy

- Quark-Gluon Jet discriminator: Scrutinize 2 least b-tagged jets
  - The non b-jets in signal originate from quarks while in QCD they originate mainly from gluons
  - Exploits the differences in the showering and fragmentation of gluons and quarks

- Additional hadronic activity
  - Quantify the hadronic activity not contained in jets
  - Absence of color flow between the VBF qq jets and the b-jets in the signal
  - Effective against QCD background



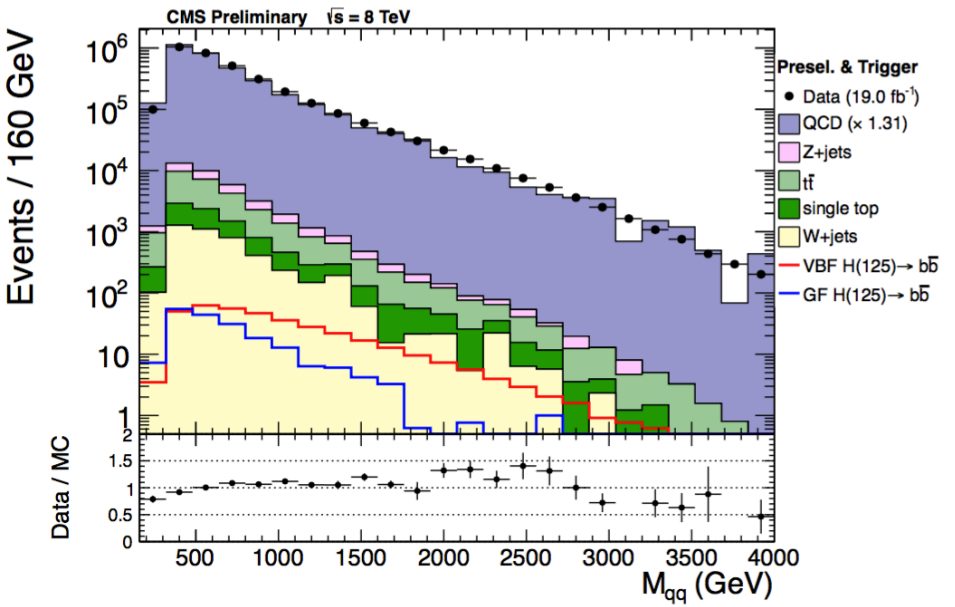
Event Selection Strategy - III

- A jet-by-jet correction factor (CF) is derived by combining various jet properties in a multivariate regression analysis

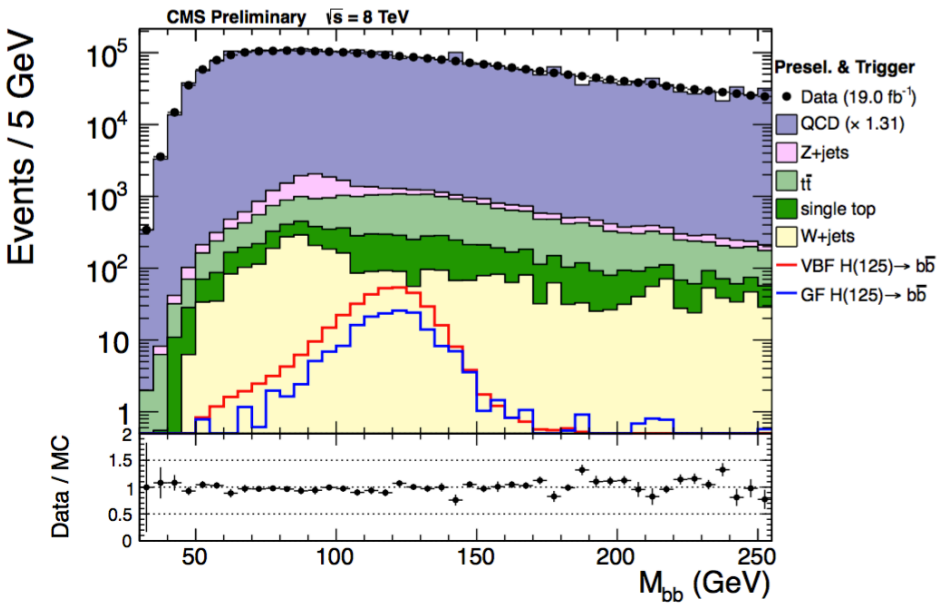
- CF is used to define a corrected m_{bb} mass of the two most b-tagged jets
  - To improve the resolution of the b-jets invariant mass spectrum

Pre-selection Results

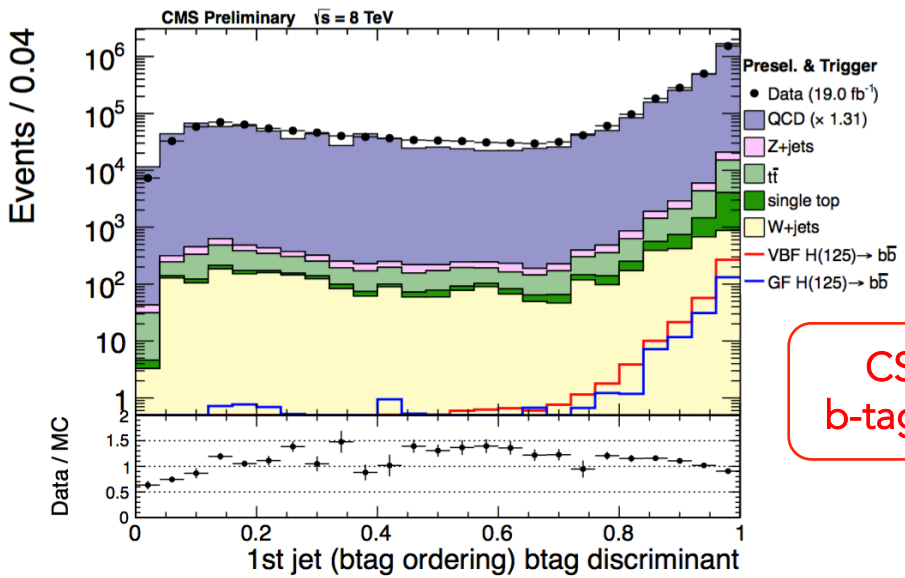
M_{qq}



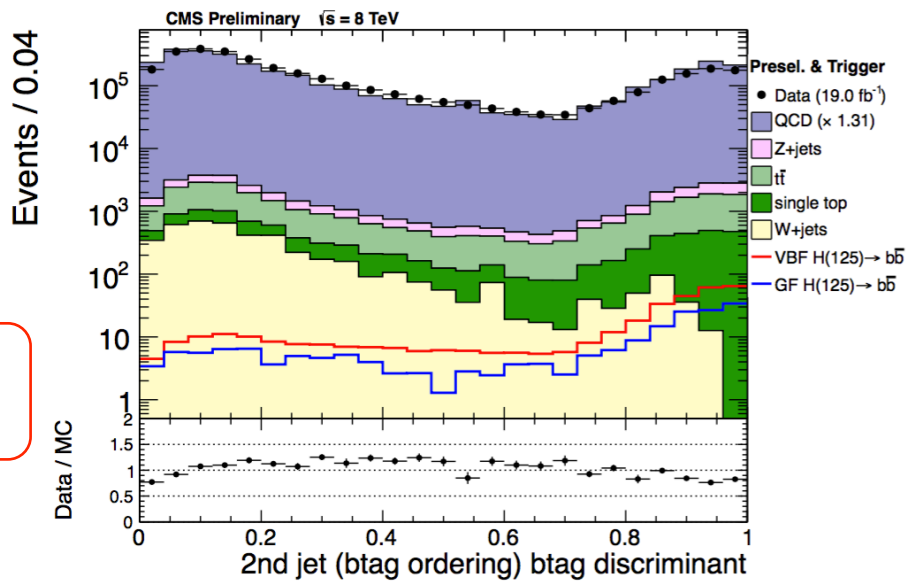
M_{bb}

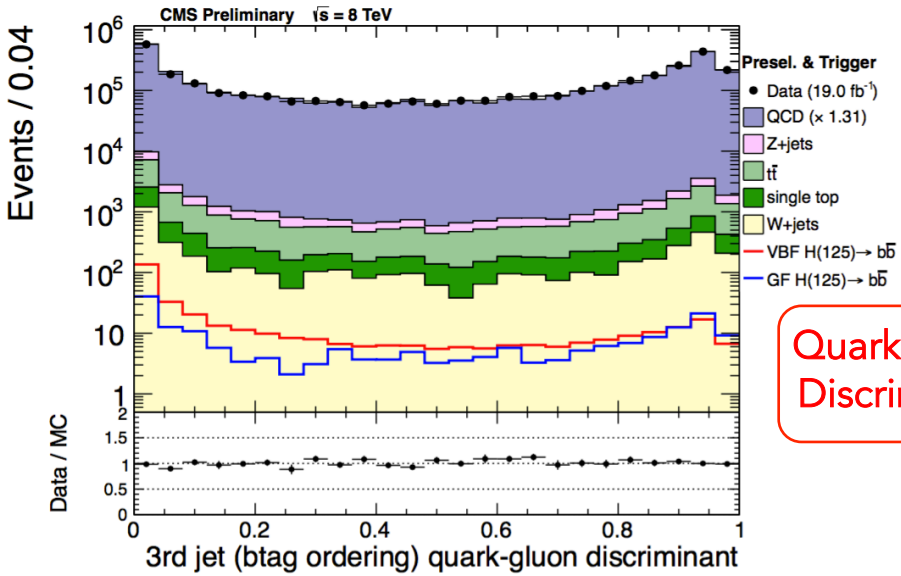


Pre-selection Results

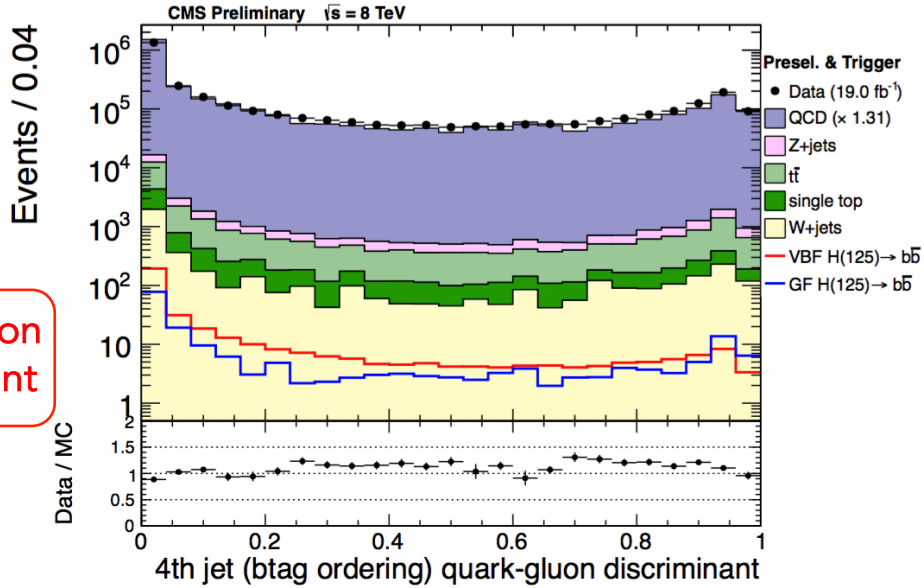


CSV b-tagger



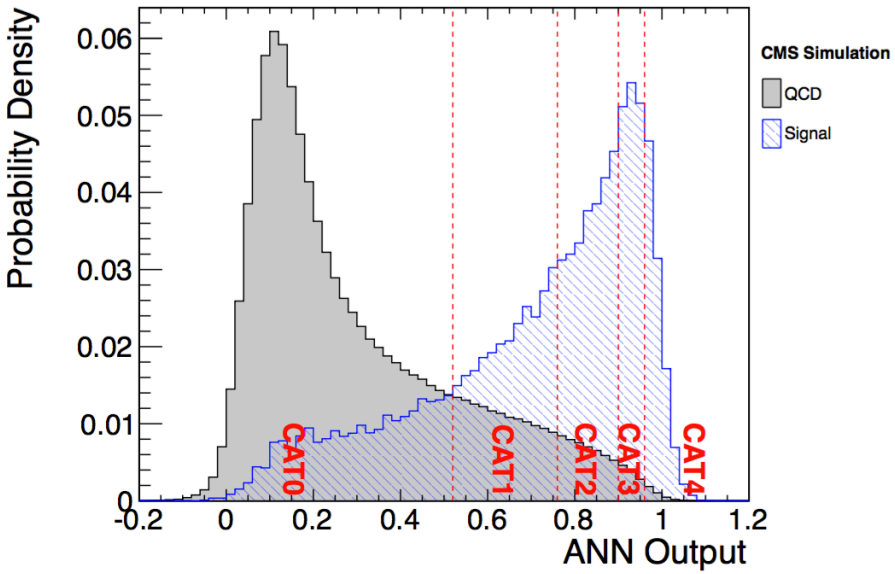


Quark-Gluon Discriminant



Multivariate Analysis

- Artificial Neural Network (ANN) implemented in TMVA
- Preselected events are further characterized by the response of ANN
  - Trained to separate signal events from background ones
- Two most b-tagged jets: CSV b-tagging values are used
- Two less b-tagged jets: Output of a quark-gluon jet discriminator applied

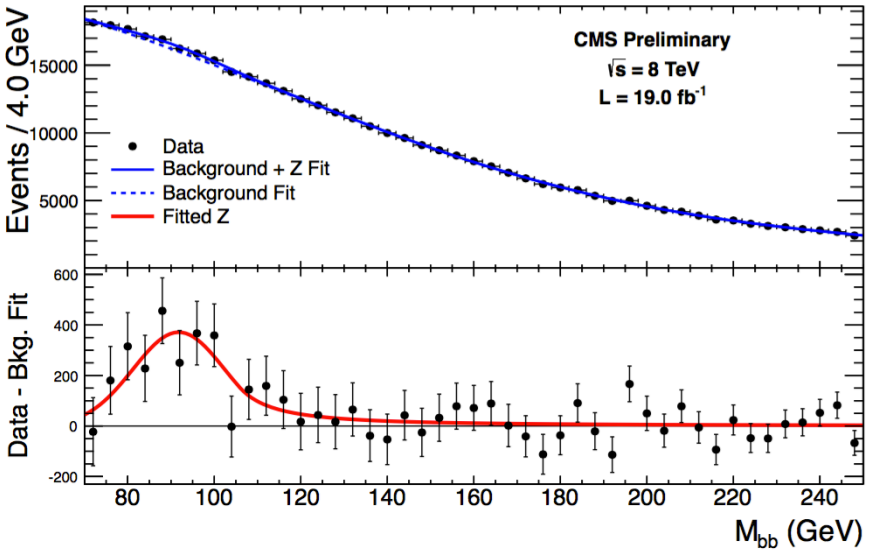


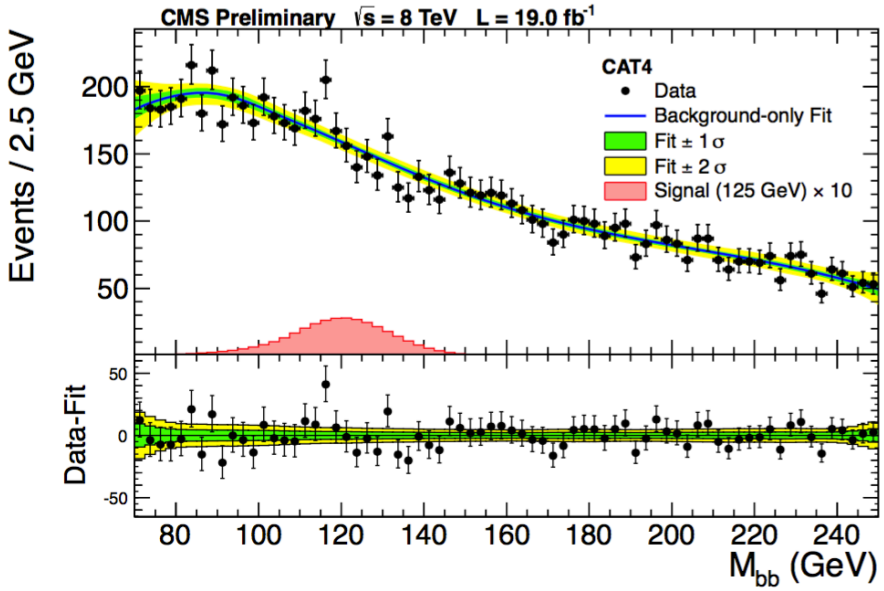
Cat. 0	Cat. 1	Cat. 2	Cat. 3	Cat. 4
$ANN < 0.52$	$0.52 \leq ANN < 0.76$	$0.76 \leq ANN < 0.90$	$0.90 \leq ANN < 0.96$	$ANN \geq 0.96$

Invariant Mass Distribution of bb

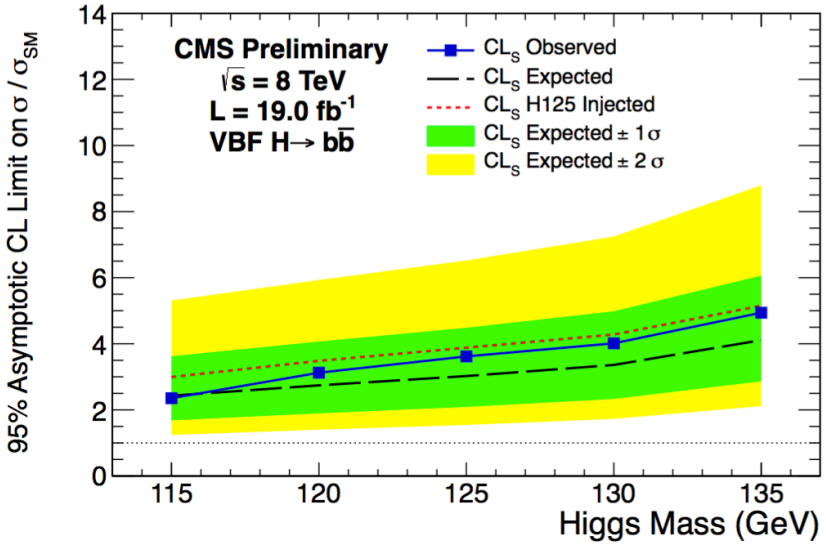
- Final M_{bb} distribution is analyzed in each ANN category separately
- M_{bb} spectrum is fitted with a background template with three parts:
  - QCD background - Normalization and shape of Bernstein polynomial
  - Z+jets/W+jets
  - top (tt +single-top)

  } Normalization and shape from Monte Carlo
- Fits of the Z → bb peak performed, for each category, in M_{bb} distribution by using the same techniques employed to search for the VBF Higgs boson signal.
  - To validate the methodology used to search Higgs boson bb signal.
  - Z → bb signal strength (μ) = 0.99 ± 0.12

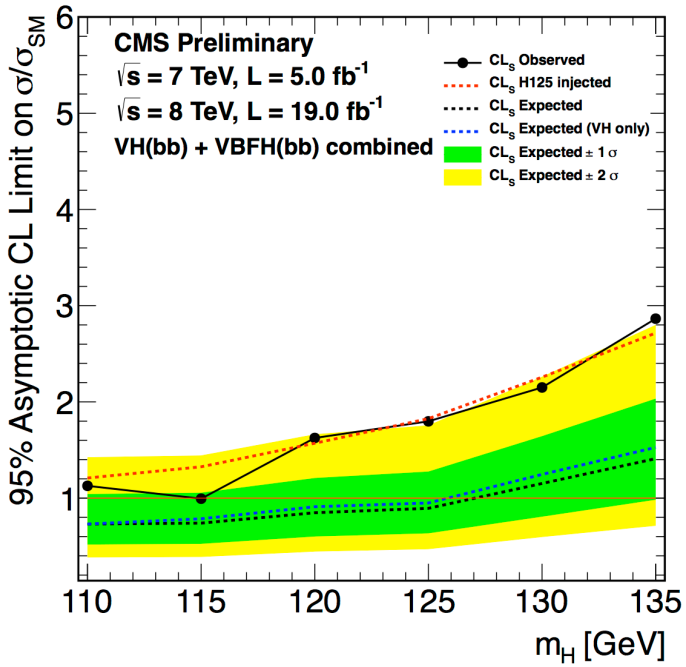




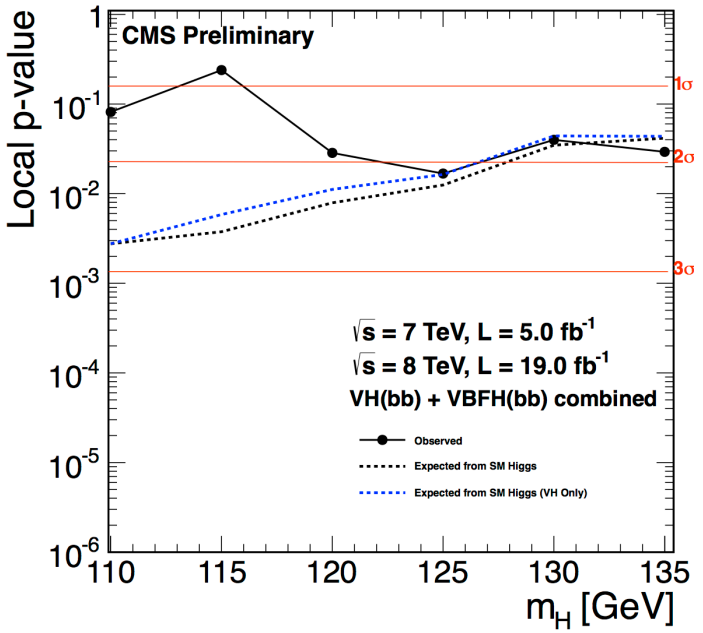
Results



VBF



VBF + VH

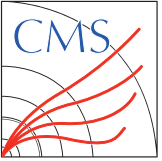


Conclusions

- Search for the standard model Higgs boson produced in vector boson fusion and decaying to bottom quarks is presented
  - pp collisions data corresponds to $\sqrt{s}$ = 8 TeV and integrated luminosity of 19 fb^{-1}

- At a Higgs boson mass of 125 GeV
  - Expected upper limit is 3.0 times the standard model prediction
  - Observed limit is 3.6 times the standard model prediction
    - Excess corresponds to a signal strength (μ) = 0.7 ± 1.4

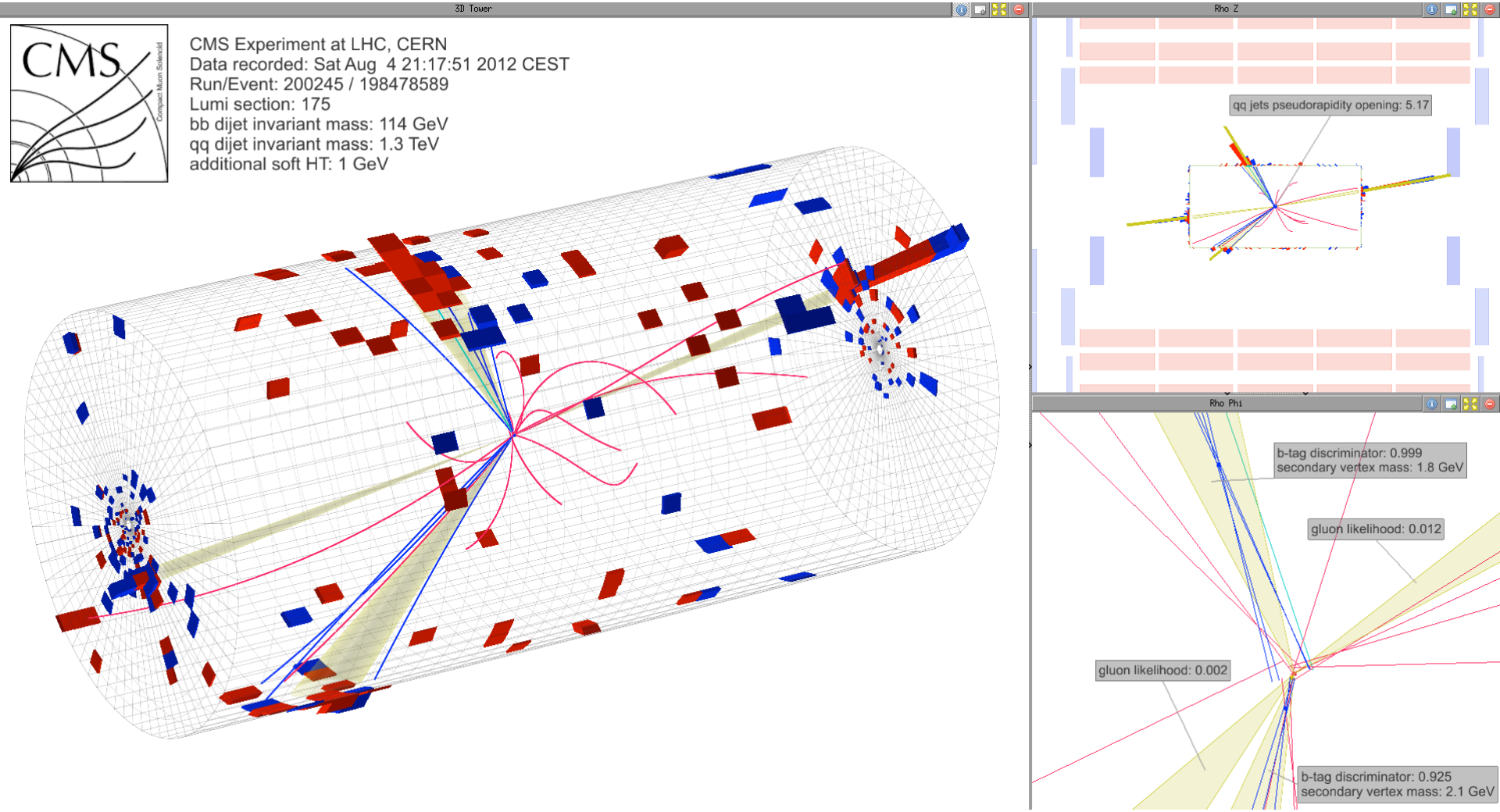
- To improve the sensitivity of H → bb search, VBF channel combined with VH
  - An excess of 2.2 standard deviations is reported at m_H = 125 GeV
    - Signal strength (μ) = 0.97 ± 0.48



Thank You

Back Up

Event Display



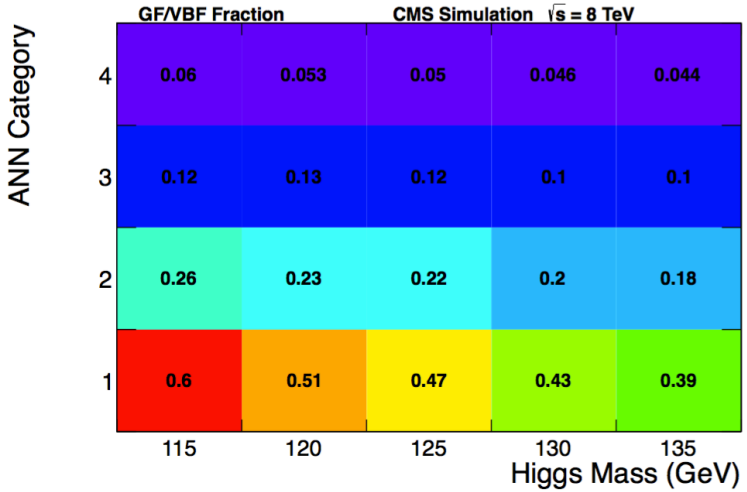
Miscellaneous

Systematic Uncertainties

Source	Uncertainty
Background fit	depending on the statistics of each category
Z+jets cross section	±20%
top cross section	±20%
Signal and Z peak position (JES)	±1.5%
Signal and Z resolution	±10%
Luminosity	±4.4%
Trigger efficiency	±5 − 8%
Signal acceptance due to JES	±10%
Signal acceptance due to JER	±2%
VBF cross section	±3%
VBF Monte Carlo acceptance	±10%
PDF	±5%
VBF ANN shape due to b-tag	±2%
VBF ANN shape due to quark-gluon discriminator	±2%
VBF ANN shape due to UE modeling	−8 − +2%
GF cross section	±15%
GF Monte Carlo acceptance	±50%
GF ANN shape	±50%

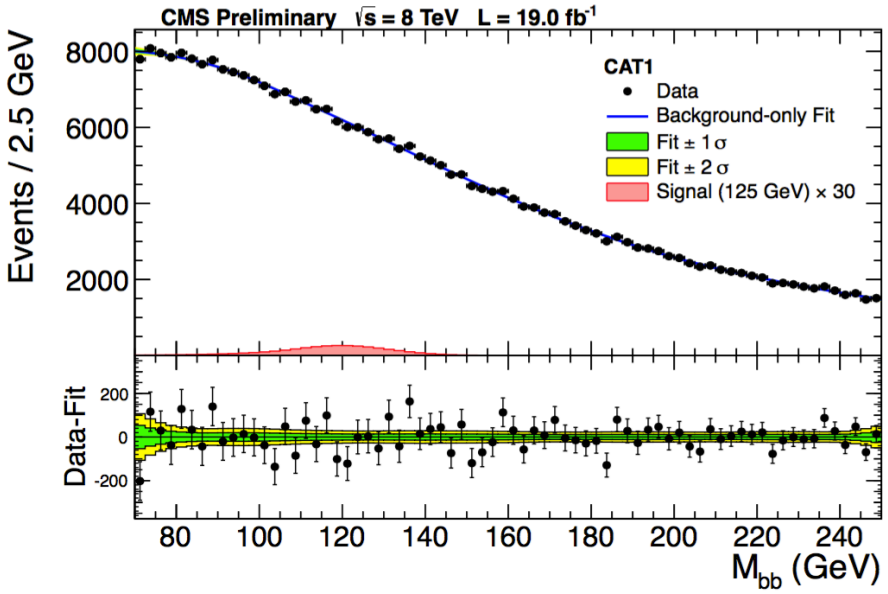
MVA Inputs

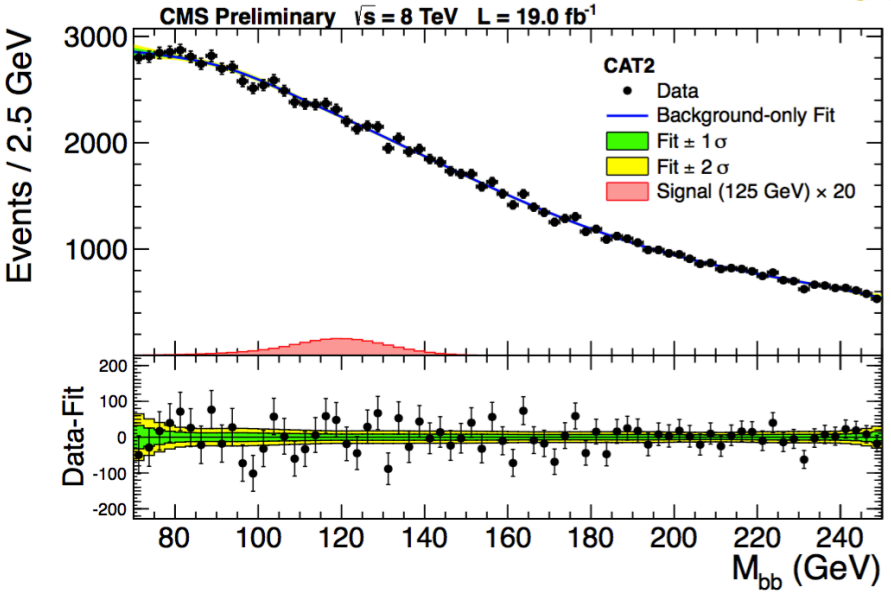
1. $\Delta\eta_{qq}$ the pseudorapidity separation between the b-tag sorted qq jets.
2. $\delta\Delta\eta_{qq}$ the pseudorapidity separation difference between the b-tag sorted and the η sorted qq jets. This difference is expected to be mostly zero for preselected signal events where the non-b-tagged (VBF) jet pair is often the most forward-backward.
3. m_{qq} the invariant mass of the b-tag sorted qq jet pair.
4. η_{qq}^{boost} the average pseudorapidity of the b-tag sorted qq jet pair system.
5. $\mathbf{CSV_0}$ the CSV b-tagging output for the most b-tagged jet.
6. $\mathbf{CSV_1}$ the CSV b-tagging output for the second most b-tagged jet.
7. $\mathbf{QGL_2}$ the quark/gluon likelihood discriminator output for the third b-tagged jet.
8. $\mathbf{QGL_3}$ the quark/gluon likelihood discriminator output for the least b-tagged jet.
9. η_2 the pseudorapidity of the third b-tagged jet.
10. H_T^{soft} the scalar p_T sum of the additional "soft" Track-Jets with $p_T > 1$ GeV.
11. $\cos\theta$ the cosine of the polar angle of the vector $\vec{p}_{q_1} \times \vec{p}_{q_2}$ in the Higgs boson rest frame (the frame where the momenta of the two most b-tagged jets are back-to-back), where q_1 and q_2 are the least b-tagged jet pair. The angle θ is essentially the angle between the qq and bb planes.
12. $\cos\alpha$ the polar angle of the vector $\vec{p}_{q_1} + \vec{p}_{q_2}$ in the Higgs boson rest frame.

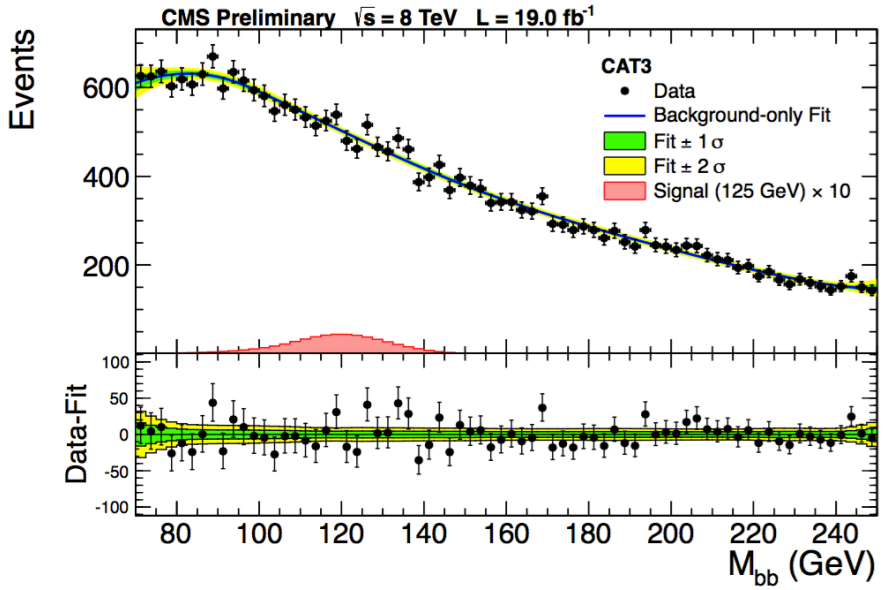


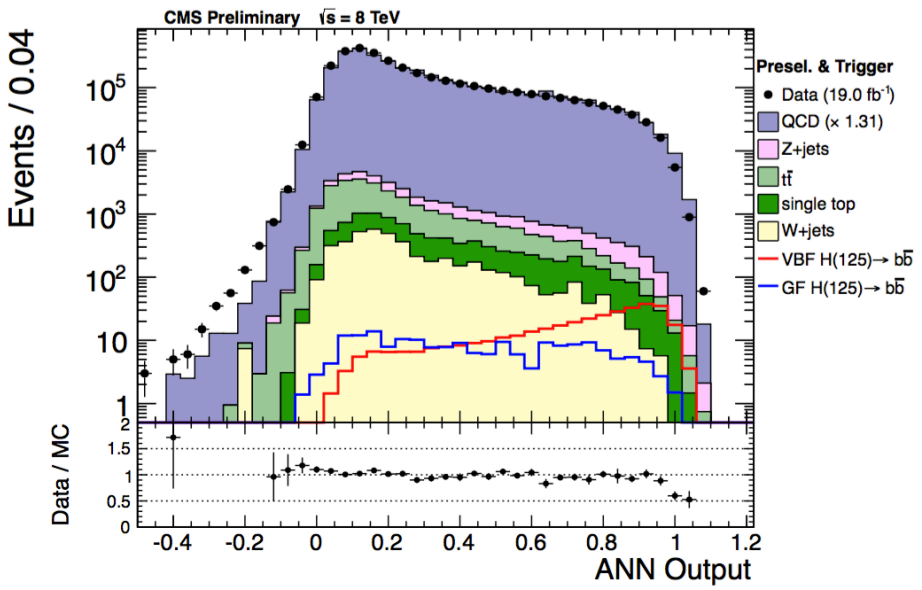
Sample/ANN range	< 0.52	$0.52 - 0.76$	$0.76 - 0.90$	$0.90 - 0.96$	> 0.96
QCD	1.9e+6	3.2e+5	1.1e+5	2.7e+4	8.7e+3
$Z + jets$	5531	1222	531	124	54
$t\bar{t}$	12730	1032	190	33	15
t	1839	383	128	25	10
$\bar{t}$	895	226	73	15	7
$W + jets$	2033	226	50	4	<1
VBF $M_H(125)$	66	79	84	49	33
GF $M_H(125)$	94	37	18	6	2

Results


CMS Preliminary √s = 8 TeV L = 19.0 fb-1
CAT1
Data
Background-only Fit
Fit ± 1 σ
Fit ± 2 σ
Signal (125 GeV) × 30
Events / 2.5 GeV
Data-Fit
Mbb (GeV)


CMS Preliminary √s = 8 TeV L = 19.0 fb-1
CAT2
Data
Background-only Fit
Fit ± 1 σ
Fit ± 2 σ
Signal (125 GeV) × 20
Events / 2.5 GeV
Data-Fit
Mbb (GeV)


CMS Preliminary √s = 8 TeV L = 19.0 fb-1
CAT3
Data
Background-only Fit
Fit ± 1 σ
Fit ± 2 σ
Signal (125 GeV) × 10
Events
Data-Fit
Mbb (GeV)


CMS Preliminary √s = 8 TeV
Presel. & Trigger
Data (19.0 fb-1)
QCD (× 1.31)
Z+jets
tt̄
single top
W+jets
VBF H(125)→ bb̄
GF H(125)→ bb̄
Events / 0.04
Data / MC
ANN Output

Transverse Slice through CMS

- 85 % - 90 % efficiency for collecting LHC delivered data
- High efficiency and resolution in object (e, μ, tau etc.) reconstruction
- The CMS detector provides good tracking and particle ID all around the interaction point ($0 < \phi < 2\pi$, $|\eta| < 3$)