VBF Analysis in H->γγ Channel at ATLAS

Jin Wang

On behalf of ATLAS Collaboration

IHEP, China

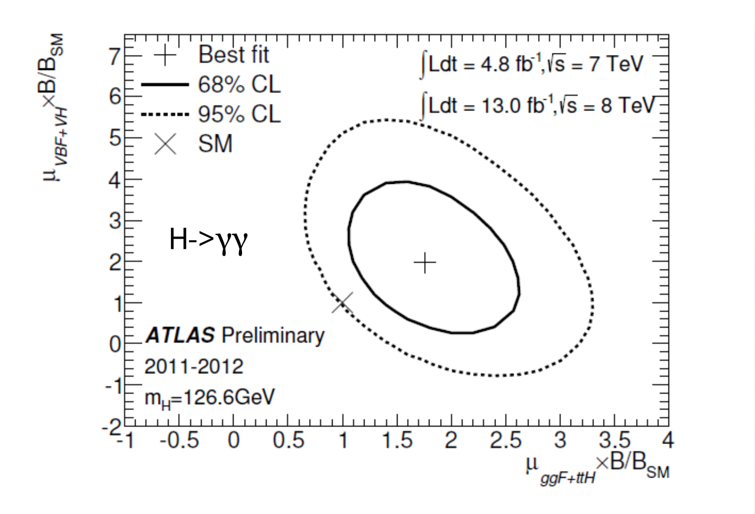
2013-08-14 International Symposium on Higgs Physics

Outline

- Introduction
- Samples and Event Selection
- Multivariate Analysis (MVA) Training and Validation
- Categorization and Results
- Summary

Introduction

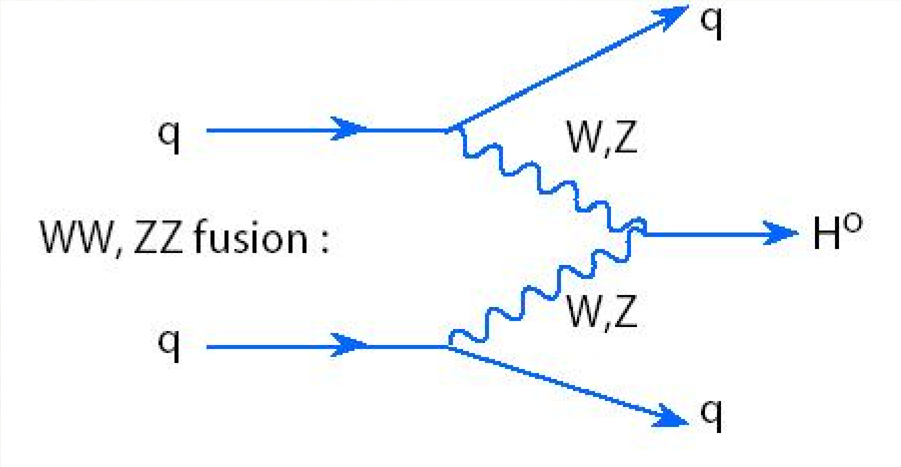
Introduction



The best fit value (+) of $\mu_{ggF+ttH} \times B/B_{SM}$ and $\mu_{VBF+VH} \times B/B_{SM}$,
from public results in December 2012.
B is the branching ratio for *H->γγ*.

- Higgs-like boson properties: its couplings to fermions and bosons are of particular relevance
- Gluon fusion and Vector Boson Fusion (VBF) are two main mechanisms how Higgs is produced
  - VBF production is of primary importance in determining Higgs coupling to gauge bosons
- Higgs->γγ is the main channel with sensitivity to VBF production
- VBF is also important in Spin-*CP* studies

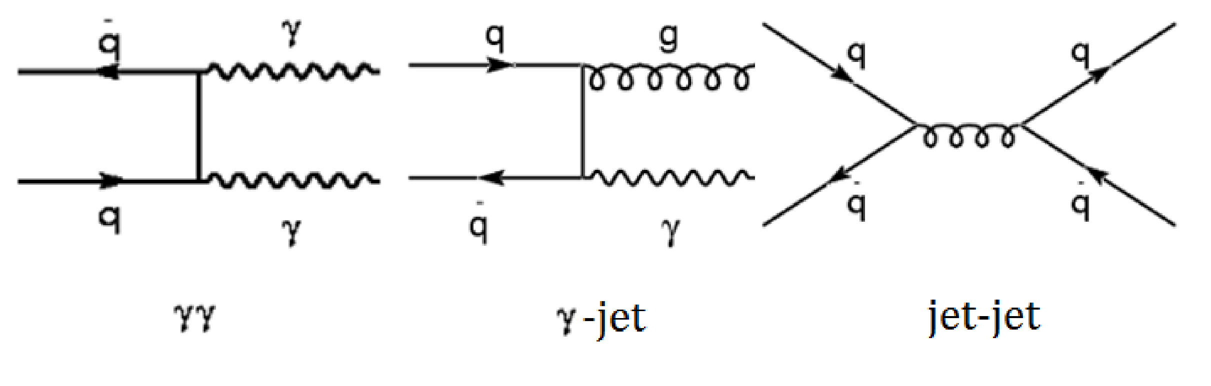
Vector Boson Fusion Higgs Production



- Two tagging jets in the forward regions with high transverse momentum and large rapidity separation
  - large invariant mass of the two leading jets
  - large jets pseudorapidity separation
  - forward jets eta distribution
- Decay products of the vector bosons lying in the central-rapidity region between jets.
  - photons/jets angular correlations
- A peak in the diphoton invariant mass distribution

Backgrounds

- Non-resonant processes
  - γγ pairs produced in association with at least two jets
  - direct single γ events with at least three jets (with one jet faking a photon)
  - multi-jet events where two jets are misidentified as photons



- Resonant backgrounds
  - Higgs bosons production via gluon fusion and in association with at least two jets

Discriminating Variables

Variable	Description
m_{jj}	the invariant mass of the two leading jets
$\Delta\eta_{\mathrm{jj}}$	pseudorapidity separation between the leading two jets
η_{jet1}	pseudorapidity of the leading jet
η_{jet2}	pseudorapidity of the subleading jet
p_{Tt}	diphoton p_{T} projected perpendicular to the diphoton thrust axis
$\Delta\phi_{\gamma\gamma,\mathrm{jj}}$	azimuthal angle difference between the diphoton and dijet systems
Δy_{jj}	rapidity separation between the leading two jets
$p_{\mathrm{T}\gamma\gamma\mathrm{jj}}$	p_{T} of the $\gamma\gamma\mathrm{jj}$ system
$\Delta p_{\mathrm{T}\gamma\gamma}$	p_{T} difference between the the two photons
η^*	the diphoton system pseudorapidity in the frame of the tagging jet pseudorapidity, define as $\eta^* = \eta_{\gamma\gamma} - \frac{\eta_{j1}+\eta_{j2}}{2}$
$\min(\Delta \mathrm{R}_{j\gamma})$	minimum $\Delta\mathrm{R}$ between either leading/subleading jet and leading/subleading photon

Strategies in VBF Analysis

- To increase the significance of VBF signal events
  - Strategy 1: sequential cuts on discriminating variables
    - easy understanding of different steps of the VBF separation
    - not able to capture the details of the correlations between the different variables
    - do not analyze the full-shape of variables
    - complex in optimization when dealing with many variables
  - Strategy 2: Multivariate Analysis (MVA)
    - use the full event topology
    - exploits the correlations of jets and photons kinematic quantities
    - “black-box”: less transparent how the variables are combined

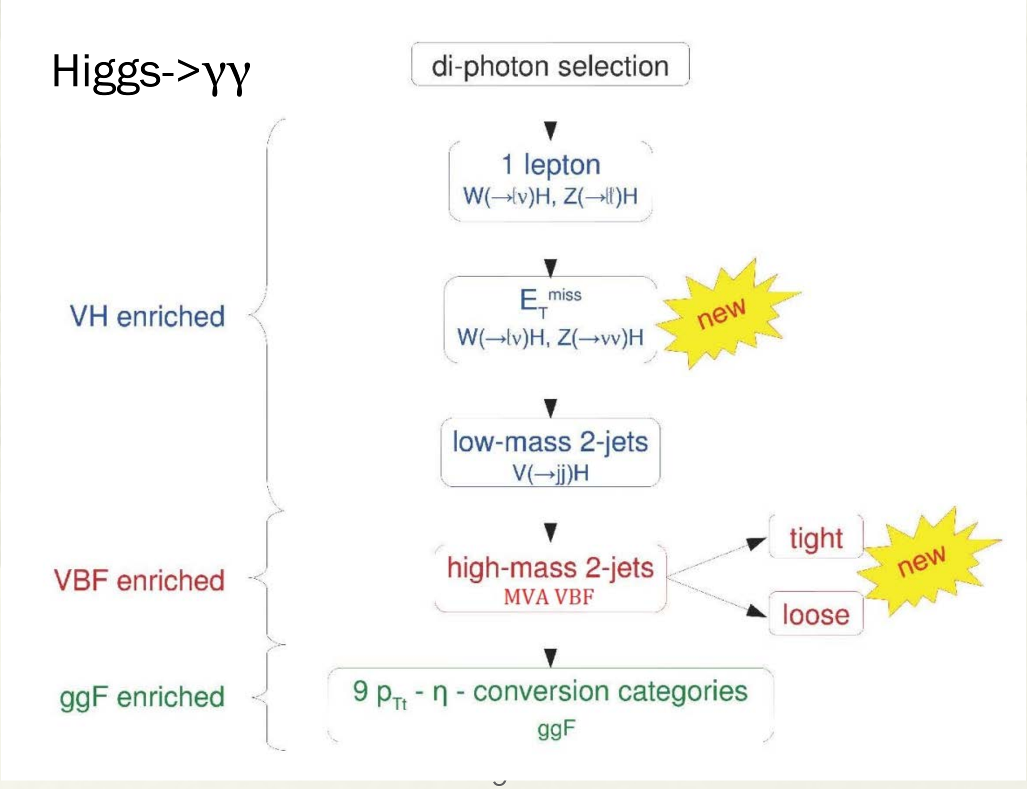
VBF in Higgs->γγ Analysis

In previous results of December 2012, the cut-based analysis is used:

- $m_{jj}>400GeV$, $\Delta\eta_{jj}>2.8$, $\Delta\Phi_{\gamma\gamma,jj}>2.6$

Presently focus on MVA VBF analysis since March 2013:

- Moriond Note: ATLAS-CONF-2013-012, Paper: arXiv:1307.1427



Samples and Event Selection

Data and MC Samples

- Data samples
  - 2012 ATLAS data collected at a center-of-mass $\sqrt{s} = 8$ TeV
    - diphoton trigger with 35 GeV and 25 GeV on leading and subleading photons
    - total integral luminosity: $20.7 \pm 0.7\ fb^{-1}$

- MC Samples
  - full simulated POWEHEG MC for gluon fusion and VBF signal
  - other signal processes are produced with PYTHIA
  - irreducible $\gamma\gamma$ backgrounds are produced with Sherpa MC

- Control regions from data
  - for the modeling of reducible backgrounds
    - one or two photons fail the isolation criteria (RevISO)

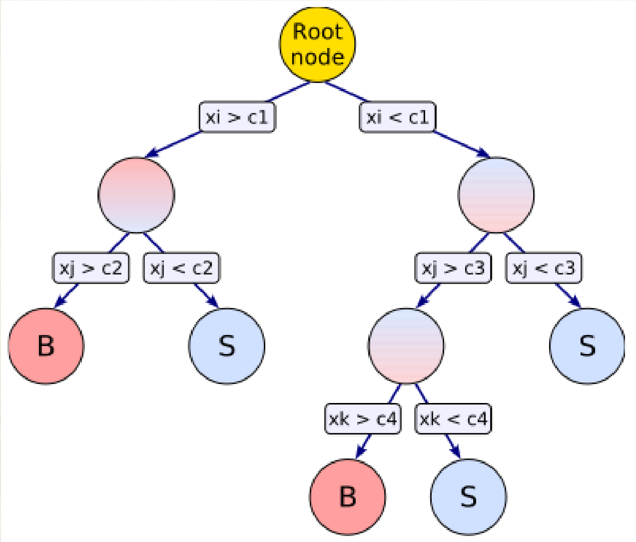
Event Selection

- Inclusive event selection
  - leading photon p_T>40 GeV, subleading photon p_T >30 GeV
  - require photon track and calorimetric isolation
  - invariant mass of two photons: 100GeV<$m_{\gamma\gamma}$<160GeV
  - see ATLAS-CONF-2013-012 for more detailed selections
- Two jet event selection
  - fail the VH category selection
    - where requiring the presence of a lepton or missing energy or a di-jet system with invariant mass consistent with gauge boson
  - at least two jets in the event
    - p_T>25GeV for $|\eta|$<2.4, p_T>30GeV for 2.4<$|\eta|$<4.5
- Pre-selection for MVA analysis
  - $\Delta\eta_{jj}$>2, $|\eta^*|$<5
  - not use $\Delta\Phi_{\gamma\gamma,jj}$ shape information for events with $\Delta\Phi_{\gamma\gamma,jj}$>2.94
    - avoid divergences in the predicted gluon fusion cross section
    - in MVA training, set $\Delta\Phi_{\gamma\gamma,jj}$=2.94 for such events

MVA Training and Validation

MVA Method and Samples

- MVA Method: Boosted Decision Trees

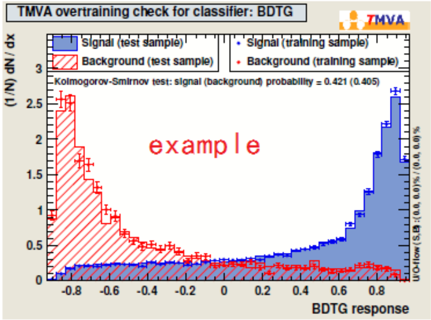


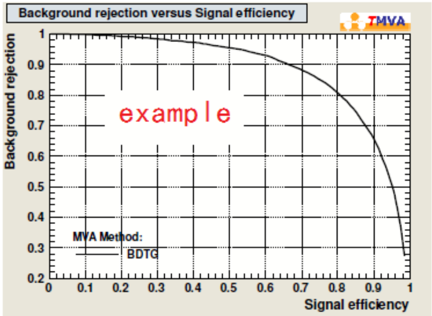
- MVA Training and Test Samples
  - Signal: VBF MC events
    - Wegithed to event yields of 20.7fb-1
  - Background: Sherpa for γγ and RevISO for rj+jj
    - weighted according to the background composition measured in data
  - Separated into two parts with event entry

MVA Variables and Training

- 8 Variables are used in BDT training:
  - p_{Tt}, m_{jj}, $\Delta\eta_{jj}$, $\Delta\Phi_{\gamma\gamma,jj}$, η_{jet1}, η_{jet2}, η^*, $\Delta R(\gamma\gamma jj)_{min}$
    - little to no correlation to $m_{\gamma\gamma}$
    - high signal/background separating power
    - low susceptibility to systematic uncertainties

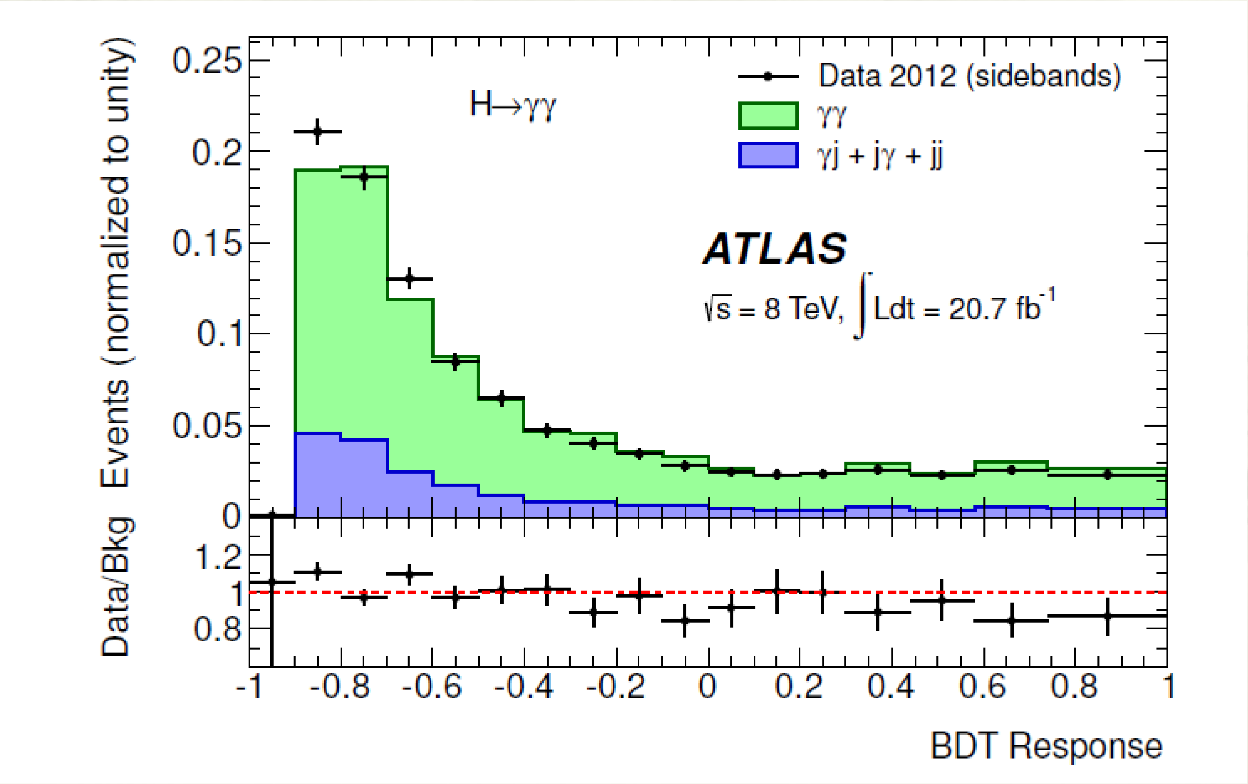
- MVA training options are optimized:
  - maximum integral of the background rejection v.s signal efficiency curve (ROC curve)
  - overtraining K-S test to be greater than 0.1





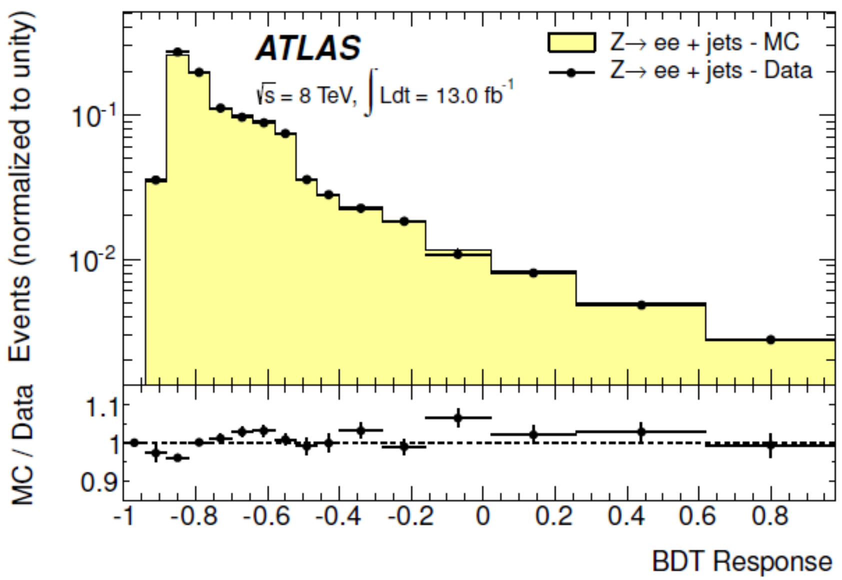
Backgrounds BDT Response Validation

- The response of the VBF BDT to the data excluding the signal region ($m_{\gamma\gamma} \in$ [120-130] GeV) and to the background model used in training



BDT Validation

- Validate that BDT gives reliable and reproducible answer that does not depend on the specific details of the data to MC difference



BDT response for $Z \to ee$+jet events in data and SHERPA MC and their ratio. Events are required to have $84 < m_{ee} < 96$ GeV and at least two jets. Electrons are taken as photons for the BDT input variables.

Categorization and Results

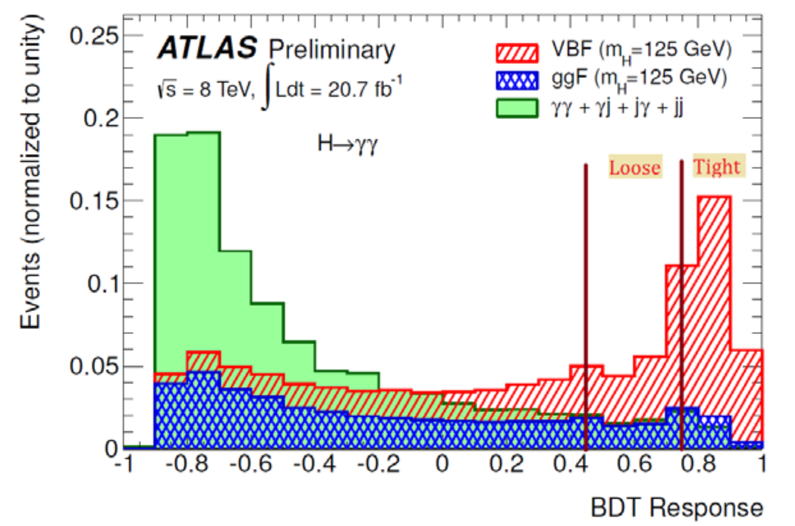
VBF Category Optimization

- VBF categories are selected by cutting on event BDT response
  - optimized by maximizing the VBF signal significance against ggF and non-resonant backgrounds

$$\sigma_{VBF} = \sqrt{2 \times ((N_{VBF} + N_{ggH} + N_{Background}) \times ln\left(1 + \frac{N_{VBF}}{N_{ggH} + N_{Background}}\right) - N_{VBF})}$$

  - use sherpa+RevISO backgrounds in optimization
  - two categories are selected: loose/tight VBF category

Loose VBF	Tight VBF
0.44<BDT<0.74	BDT≥0.74



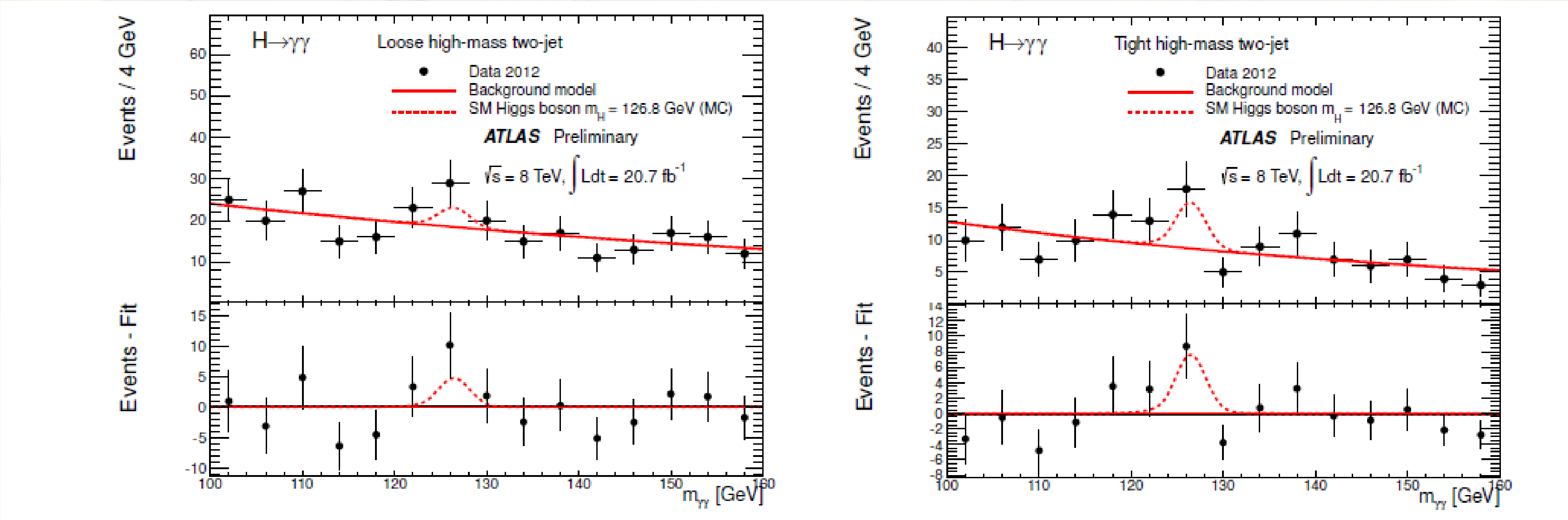
Event Yields

Number of events in the data (N_D) and expected number of SM Higgs signal events (N_S) for m_H = 126.5 GeV from the $H \to \gamma\gamma$ analysis, for each category in the mass range 100-160 GeV at $\sqrt{s}$ = 8 TeV. The fractions of expected signal events from the $gg \to H$, VBF, WH, ZH, ttH processes are detailed.

$\sqrt{s}$	8 TeV						
Category	N_D	N_S	$gg \to H$ [%]	VBF [%]	WH [%]	ZH [%]	ttH [%]
Unconv. central, low p_{Tt}	10900	51.8	93.7	4.0	1.4	0.8	0.2
Unconv. central, high p_{Tt}	553	7.9	79.3	12.6	4.1	2.5	1.4
Unconv. rest, low p_{Tt}	41236	107.9	93.2	4.0	1.6	1.0	0.1
Unconv. rest, high p_{Tt}	2558	16.0	78.1	13.3	4.7	2.8	1.1
Conv. central, low p_{Tt}	7109	33.1	93.6	4.0	1.3	0.9	0.2
Conv. central, high p_{Tt}	363	5.1	78.9	12.6	4.3	2.7	1.5
Conv. rest, low p_{Tt}	38156	97.8	93.2	4.1	1.6	1.0	0.1
Conv. rest, high p_{Tt}	2360	14.4	77.7	13.0	5.2	3.0	1.1
Conv. transition	14864	40.1	90.7	5.5	2.2	1.3	0.2
Loose high-mass two-jet	276	5.3	45.0	54.1	0.5	0.3	0.1
Tight high-mass two-jet	136	8.1	23.8	76.0	0.1	0.1	0.0
Low-mass two-jet	210	3.3	48.1	3.0	29.7	17.2	1.9
E_T^{miss} significance	49	1.3	4.1	0.5	35.7	47.6	12.1
One-lepton	123	2.9	2.2	0.6	63.2	15.4	18.6
All categories (inclusive)	118893	395.0	88.0	7.3	2.7	1.5	0.5

Background Fit in Two VBF Categories

- Background-only fits to the diphoton invariant mass spectra for categories.
- The bottom inset displays the residual of the data with respect to the background fit.
- The Higgs boson expectation for a mass hypothesis of 126.8 GeV corresponding to the SM cross section is also shown.

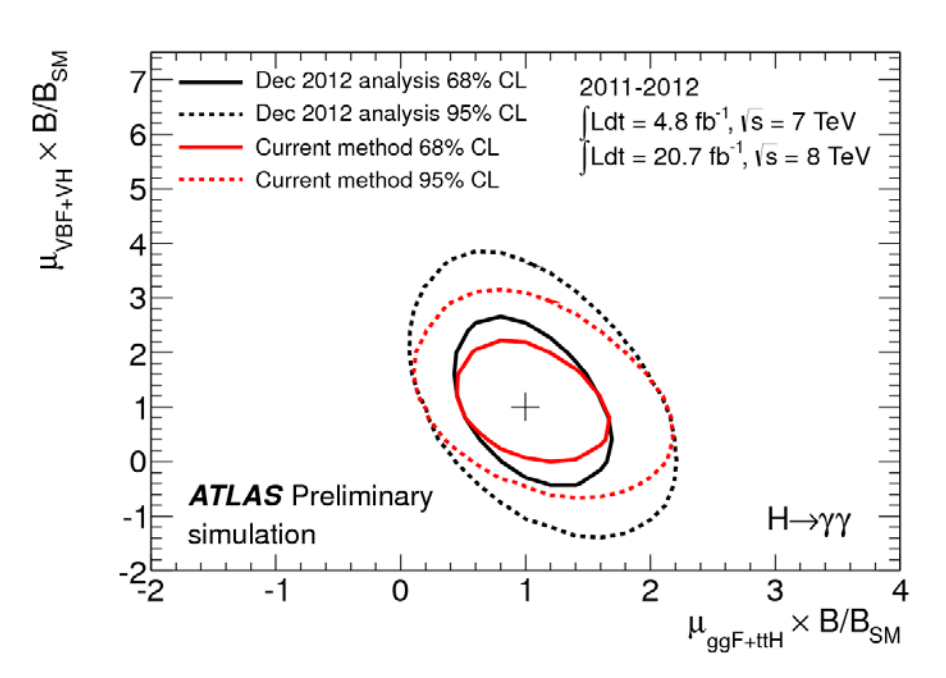


Systematics

- The uncertainties associated with the MC prediction of the SM Higgs yields in the two categories
  - Detector modelling systematics
    - jet energy scale and resolution
    - jet vertex fraction selection
  - Theoretical systematics
    - higher order perturbative correction for *gg->H*
    - modelling of η^*
    - modelling of $\Delta\Phi_{\gamma\gamma,jj}$
    - underlying event uncertainty
- The uncertainties related to the choice of the functional form used to model the $m_{\gamma\gamma}$ background spectra
  - Spurious signal studies

VBF Signal Strength Measurement in H->γγ

- Signal strength $\mu_{\mathrm{VBF}} = N_{obs}(\mathrm{VBF})/N_{exp}(\mathrm{VBF})$
- Moriond MVA VBF analysis: $$\mu_{\mathrm{VBF}} \times B/B_{\mathrm{SM}} = 1.7^{+0.8}_{-0.8}(\mathrm{stat})^{+0.5}_{-0.4}(\mathrm{syst})$$
- Cross check of refined cut-based analysis shows consistency.



Compare to December results, improvements on expected uncertainty are:

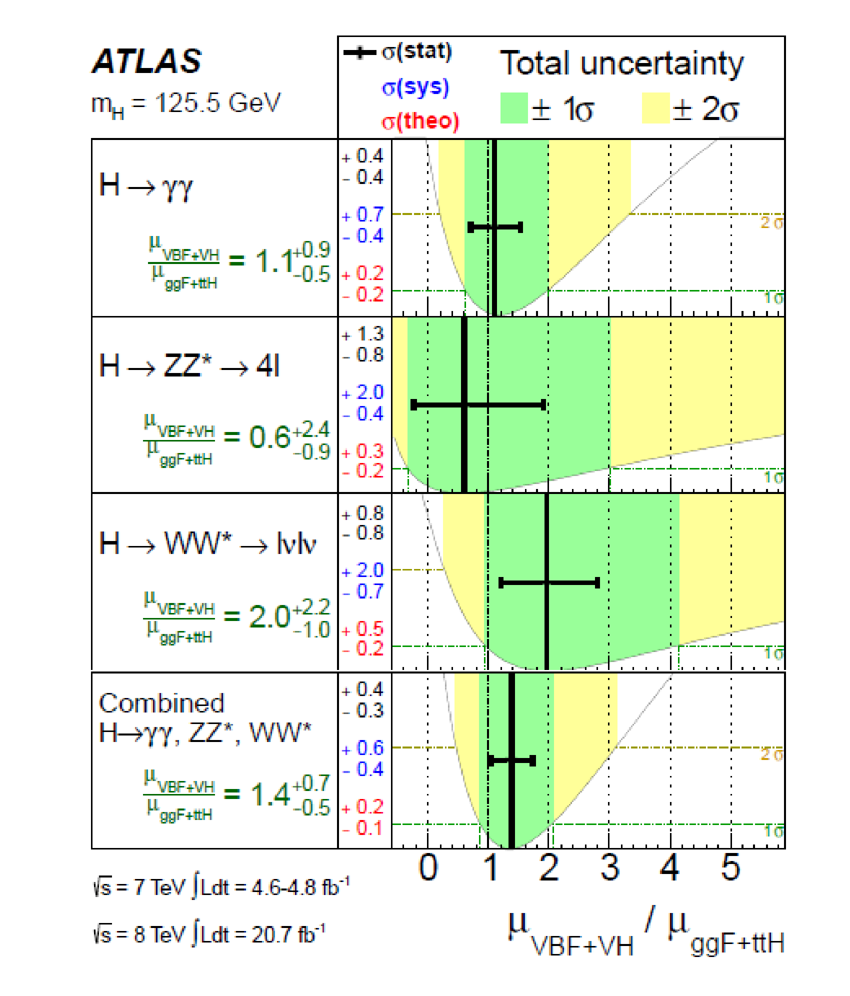
μ_{VBF}	μ_{VH}
32%	27%

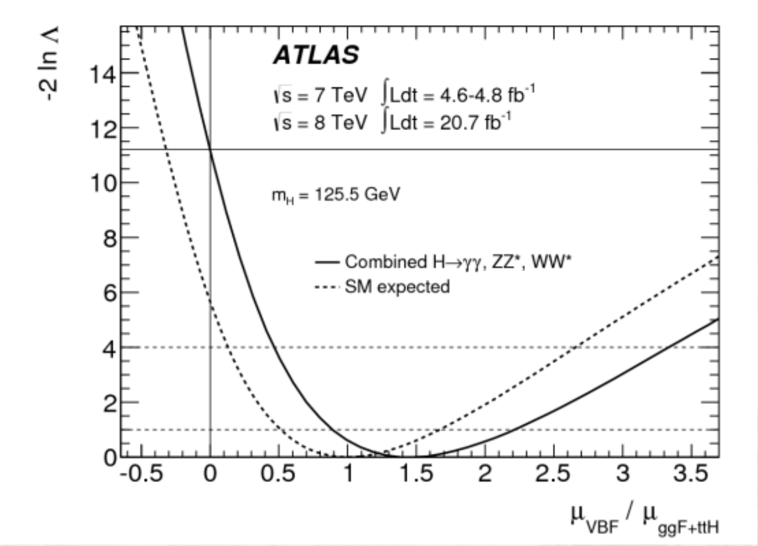
Note that only 13 fb^{-1} 2012 data is used in December results

Combined Coupling Measurement

(arXiv:1307.1427)

Measurements of the $\mu_{VBF}/\mu_{ggF+ttH}$ ratios for diboson final states and their combination





Likelihood curve for the ratio $\mu_{VBF}/\mu_{ggF+ttH}$ for the combination

- The result provides evidence at the 3.3σ level for VBF production
  - main sensitivity comes from H->γγ channel

Summary

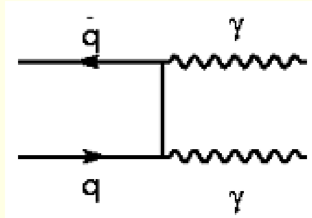
Summary

- The VBF analysis in Higgs->γγ channel is presented
- A multivariate BDT approach is implemented to enhance VBF signal significance
  - used 8 variables in Moriond MVA VBF analysis
  - two VBF categories are selected based on event MVA response
- An improvement of 32% on expected μ_{VBF} is reached comparing to previous public results
- Combination result gives 3.3σ evidence for VBF production
- Future possible improvements:
  - update with larger statistics of training samples
  - improve the variable modeling in background training samples by reweighting
  - explore more potential variables in VBF MVA analysis
- Spin-*CP* analysis in VBF is another interesting direction especially with larger statistics in future
  - VBF is rich of discriminating variables for spin analysis
  - VBF is an important component in *CP*-mixing studies.

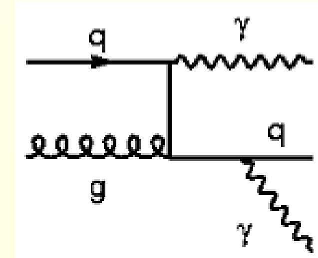
Back Up

backgrounds

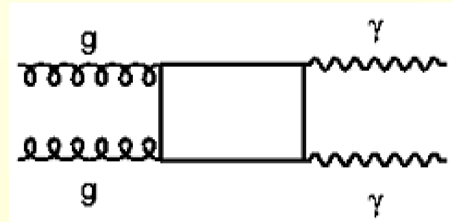
- Born(α_{QED}^2)



- Bremsstrahlung (correction to Born)

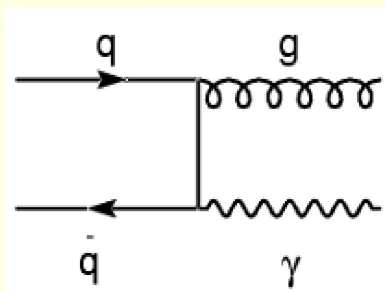


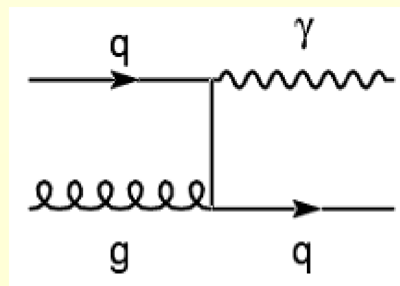
- Box gg→γγ ($\alpha_s^2\alpha_{QED}^2$)



- γ+jet

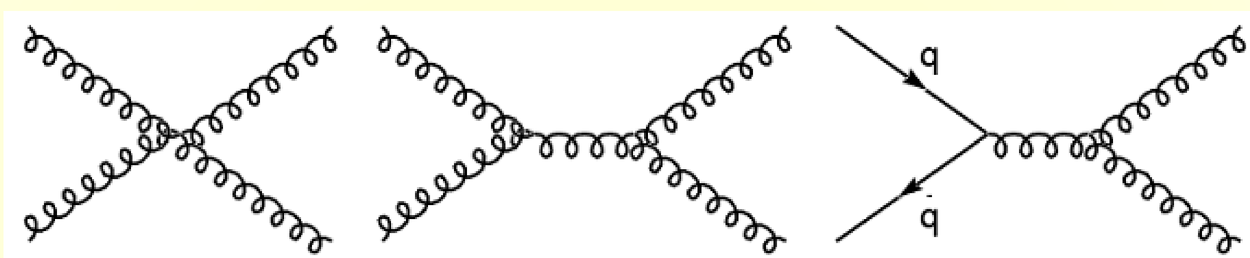
Partons→hadrons (« jets »)



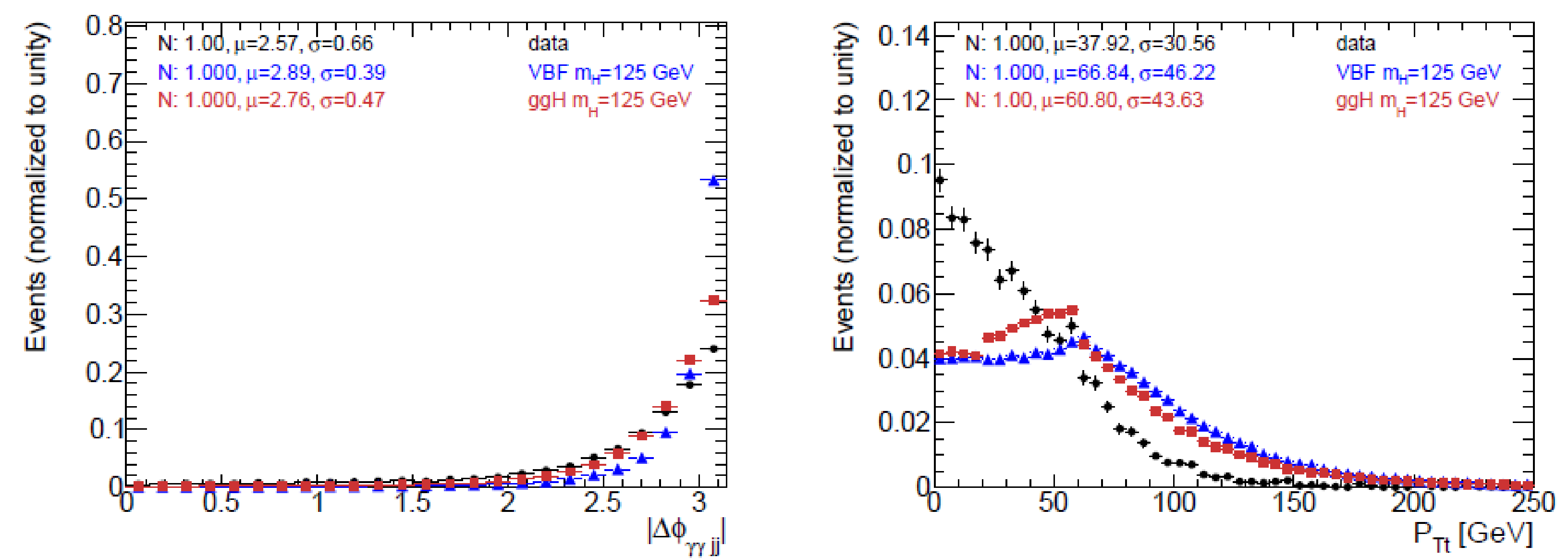


- jet+jet

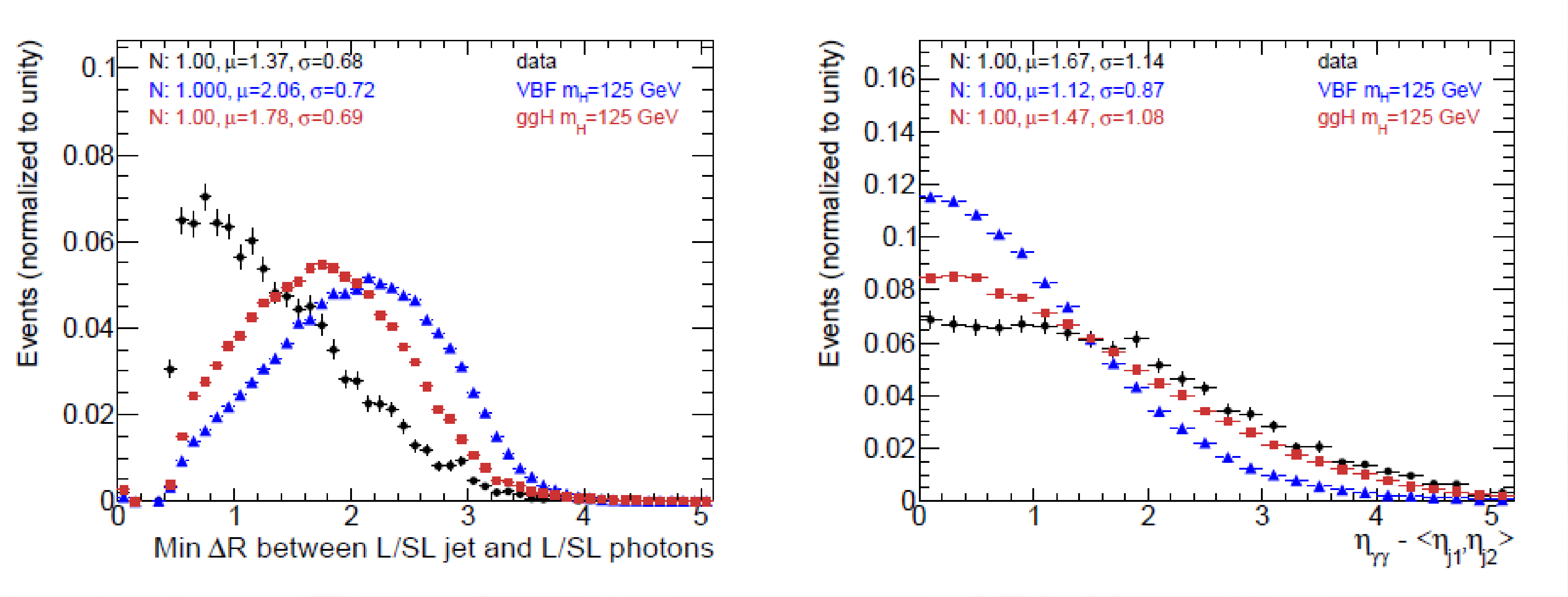
58 % jet gluon



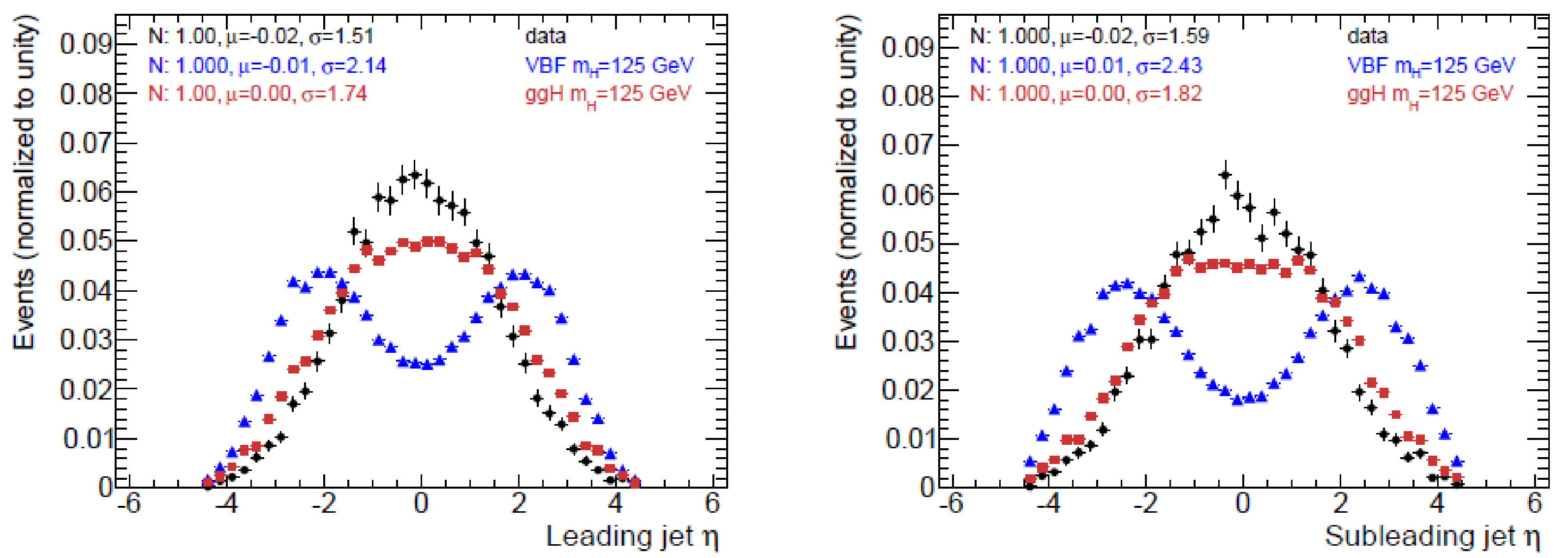
Discriminating Variables



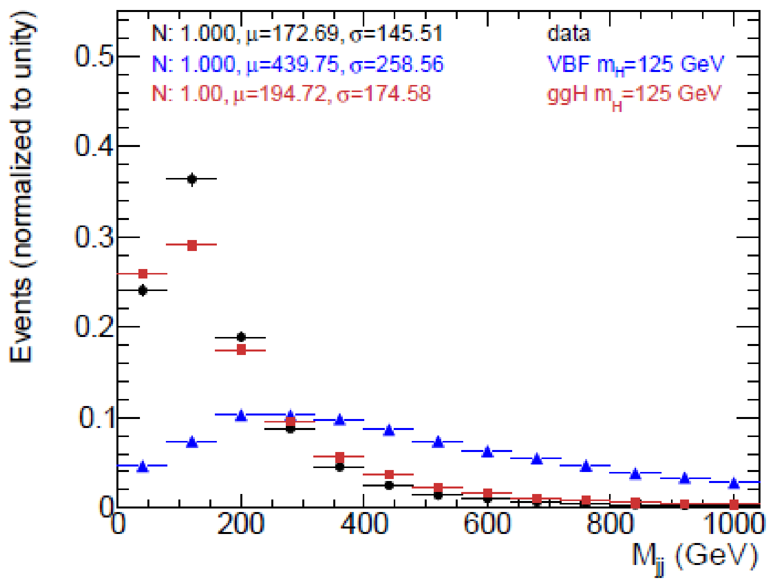
Discriminating Variables

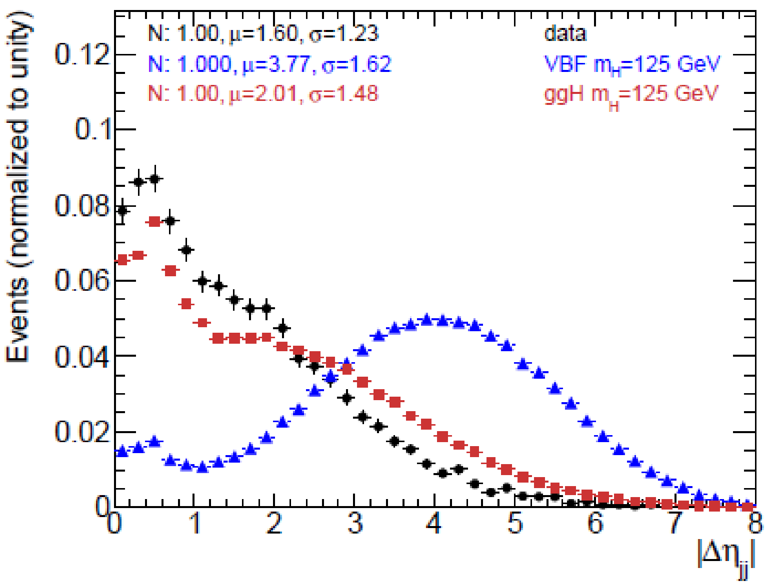


Discriminating Variables

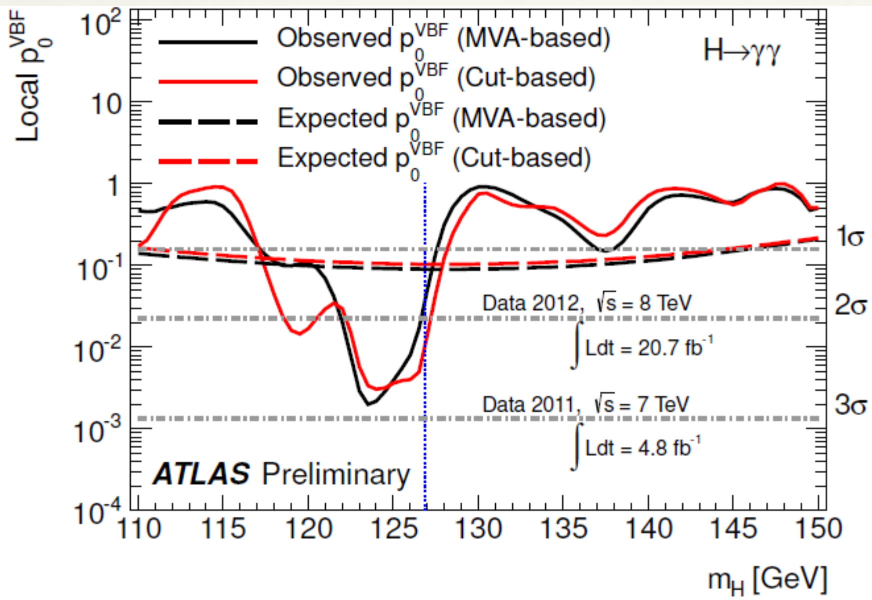


Discriminating Variables





Results from refined cut-based analysis



The observed local p_0^{VBF} value for VBF production as a function of m_H for the combination of $\sqrt{s}$ = 7 TeV and $\sqrt{s}$ = 8 TeV data (solid line) and the corresponding expected local p_0^{VBF} values for the SM Higgs boson signal plus background hypothesis (dashed line). In parallel to the baseline MVA analysis, the cut based selection was re-optimised: two categories are defined, with more or less stringent cuts. The "tight high-mass two-jets" category is defined by $\Delta\eta_{jj} > 2.4$, $m_{jj} >$ 520 GeV, $\Delta\phi_{\gamma\gamma;jj} > 2.6$, $\eta^* < 2.4$ and $\Delta R^{\gamma j}_{\min} > 2$. Events failing those cuts but passing the ones used in Ref. [6] enter into the "loose high-mass two-jets" category. Both the baseline MVA analysis and the reoptimised cut based analysis are shown.