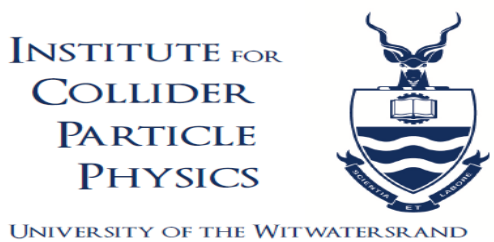
# Testing the search for new resonances in the di-photon channel with topological requirements on the production of the Standard Model Higgs boson at the LHC

Joshua Choma

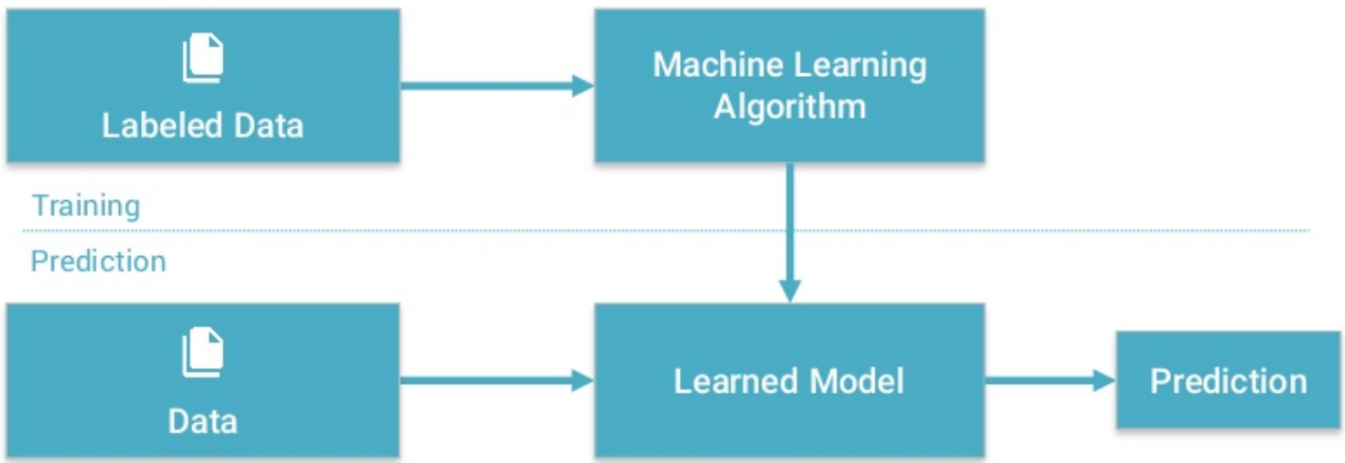
University of the Witwatersrand

January 30, 2020



# Machine Learning

- Machine learning is a subset of artificial intelligence that gives computers the ability to learn without being explicitly programmed
- Algorithms that can learn from and make predictions on data
  - Classification
  - Clustering
  - Regression
- Machine Learning plays important role from analysis to event reconstruction and identification in HEP

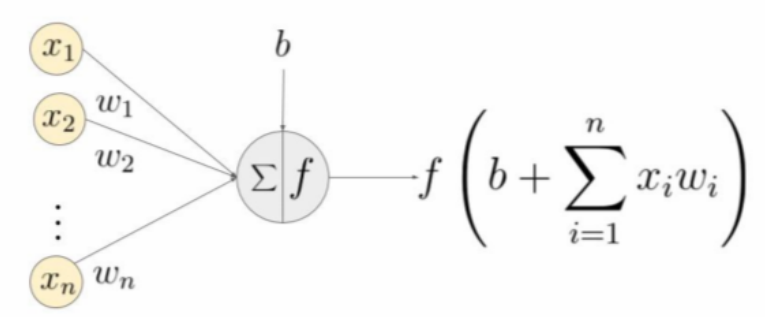


# Types of Learning

- Supervised: Learning with a labelled training set
- Unsupervised: Discover patterns in unlabelled data
- Semi-Supervised: Semi-supervised learning is describes a task of learning a functional form from data that has both labelled and unlabelled instances.

# Introduction to ANN

- Basic building block of artificial neural network

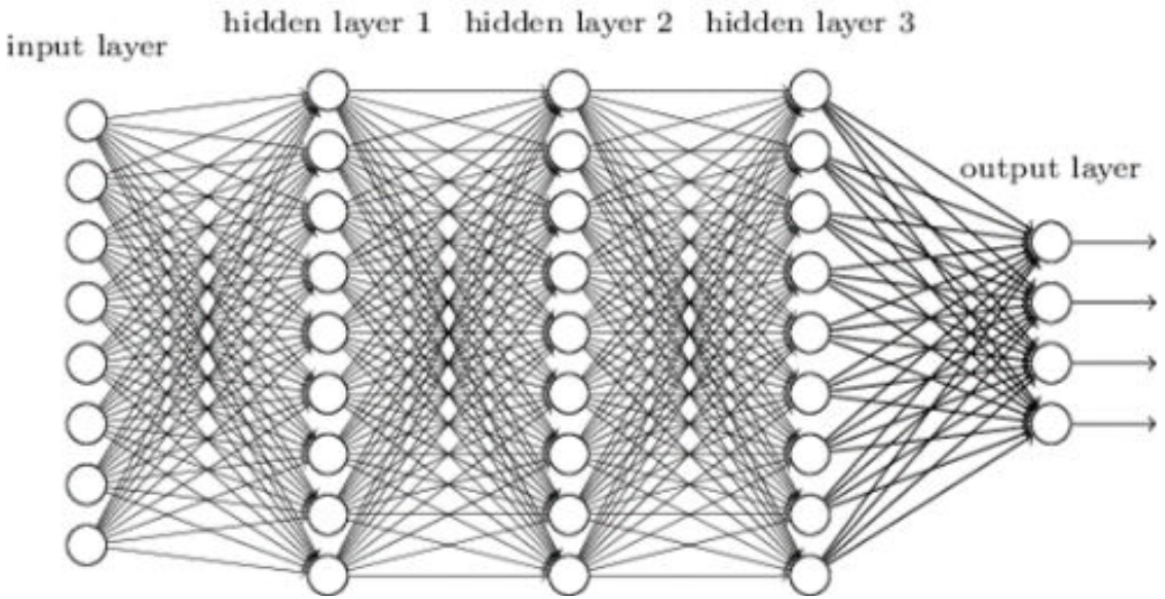



- Activation functions: *sigmoid, ReLu and tanh*

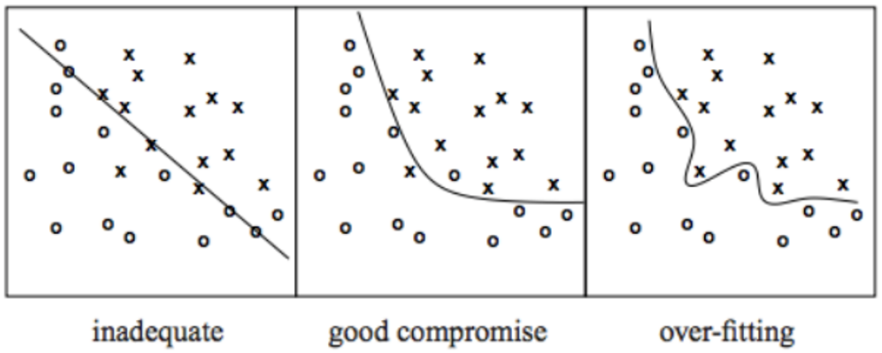
$$f(x) = \frac{1}{1 + e^{-x}} \tag{1}$$

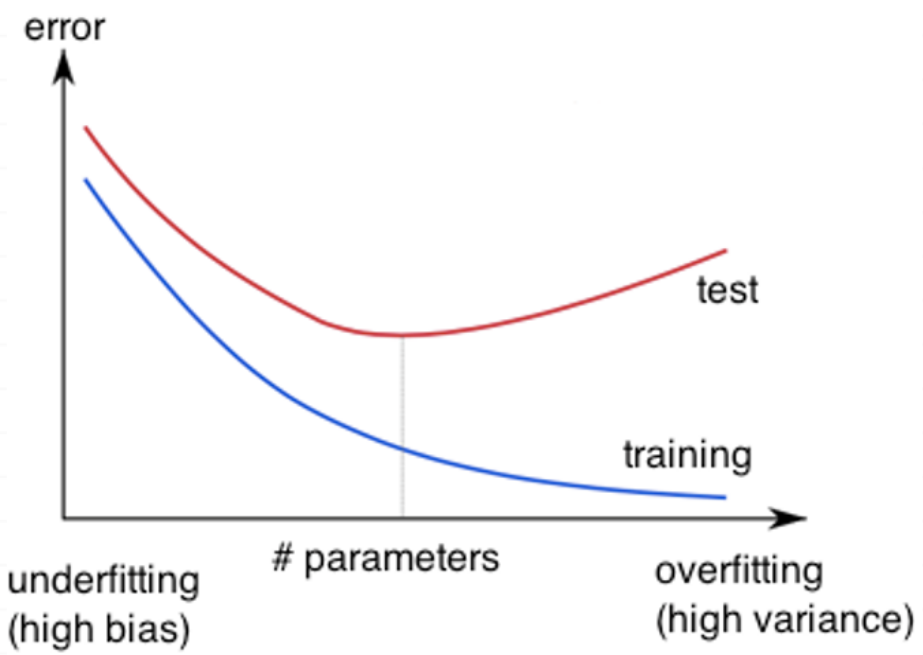
# Introduction to Deep Learning

- It uses many-layered deep neural networks (DNN) to learn levels of representation and abstraction that makes sense of data.



# Overfitting



# Objectives

- Search for new resonances beyond the Standard Model (SM) using Machine Learning techniques.
- Combine weak-supervision and full-supervision in conjunction with Deep Neural Network algorithms.
- Classifying events as Signal or Background:
  - Signal (Monte Carlo, period 15-17): ggH,WH,ZH, VBFH, ttH
  - Background (Monte Carlo, period 15-17): $\gamma\gamma$, $\gamma$jet, V$\gamma$, V$\gamma\gamma$.
- Region of focus:
  - Background Side Band:
    - 115 GeV $< m_{\gamma\gamma} <$ 120 GeV || 130 GeV $< m_{\gamma\gamma} <$ 135 GeV
    - Sample 1
  - Signal Mass Window
    - 120 GeV $< m_{\gamma\gamma} <$ 130
    - Sample 2 (including background in the region)

# Data: Background

| | Entries | Weights |
|---|---|---|
| $\gamma\gamma$ | $9.930x10^6$ | $1.137x10^5$ |
| $\gamma$jet | $1.362x10^5$ | $4.739x10^4$ |
| V$\gamma$ | 3907 | 979.6 |
| V$\gamma\gamma$ | 16168 | 199.3 |

# Data: Signal

| | Entries | Weights |
|---|---|---|
| WH | $3.045x10^5$ | 83.68 |
| ZH | $2.970x10^5$ | 47.26 |
| ggH | $3.515x10^6$ | 3328 |
| VBFH | $1.436x10^6$ | 268.5 |
| ttH | $1.648x10^6$ | 37.1 |

# Model Architecture

- **Fully connected layers**
  - 7 hidden layers
  - 250, 170, 150 (last 5 layers) nodes
  - ReLu activation function
- **Output [0,1]**
  - Sigmoid
- **Hyper-parameters:**
  - Number of epochs: 150
  - Batch size: 500
  - Optimizer: Adam
  - Learning rate: $8.52x10^{-4}$
  - Decay = $3.005x10^{-2}$
- Framework: Keras with TensorFlow backend
- The network architecture is kept constant

# VBF125

# m_yy

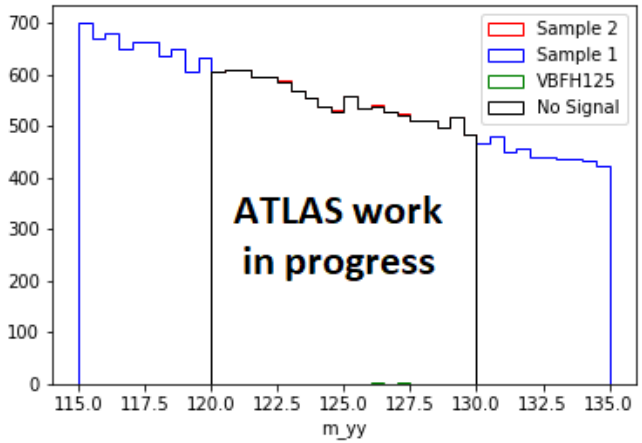



Figure: 10 VBF125 events

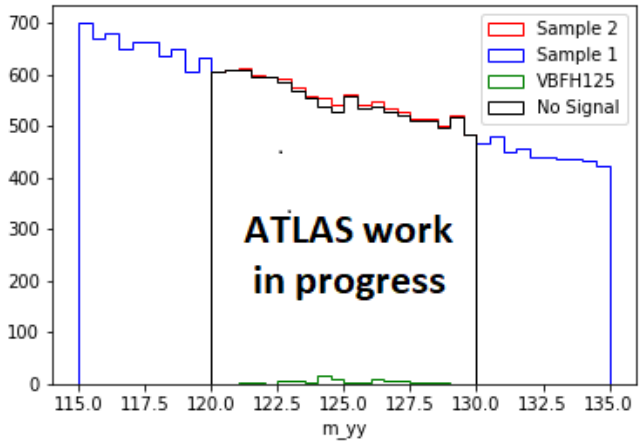



Figure: 100 VBF125 events

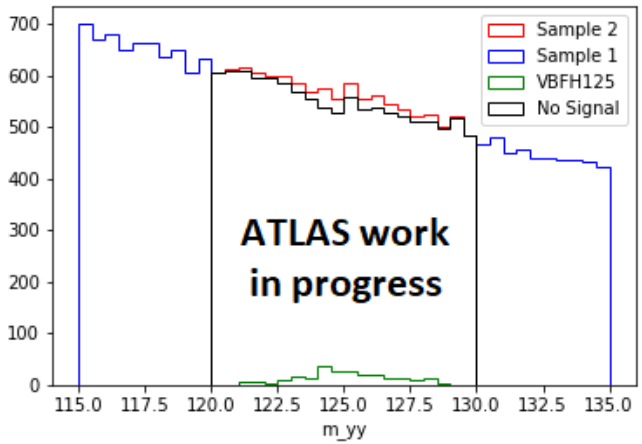



Figure: 250 VBF125 events

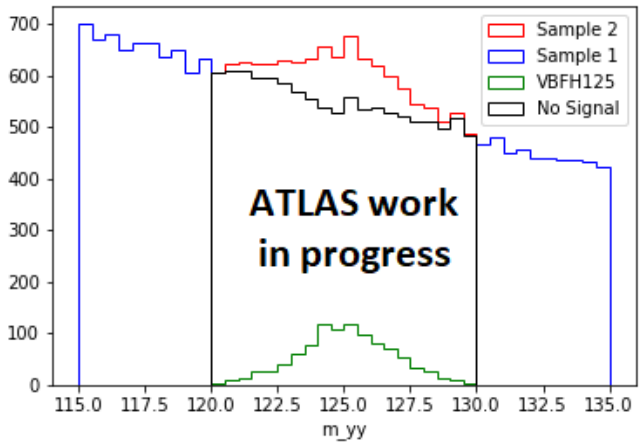



Figure: 1000 VBF125 events

# Response Distributions

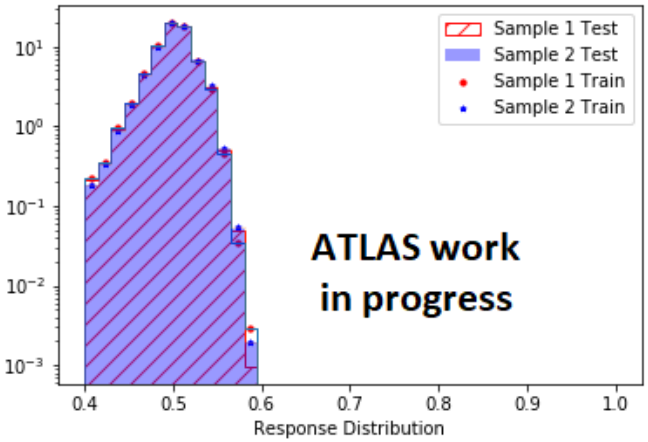



Figure: 10 VBF125 events

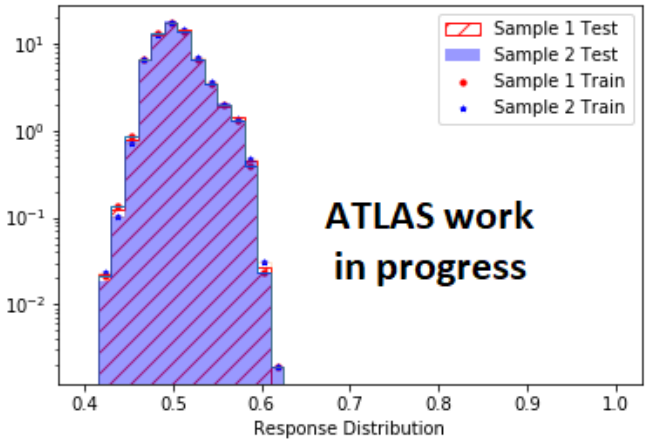



Figure: 100 VBF125 events

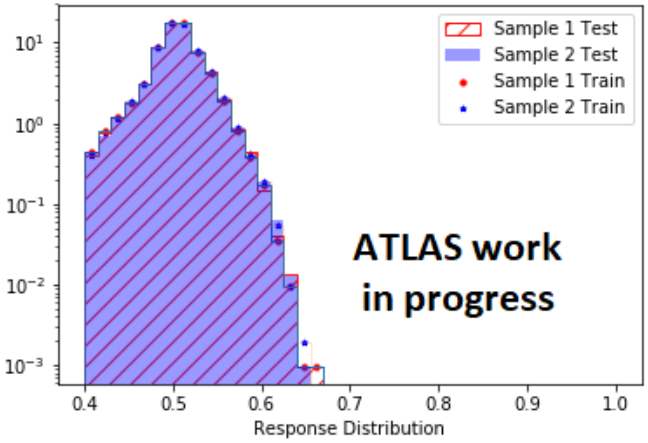



Figure: 250 VBF125 events

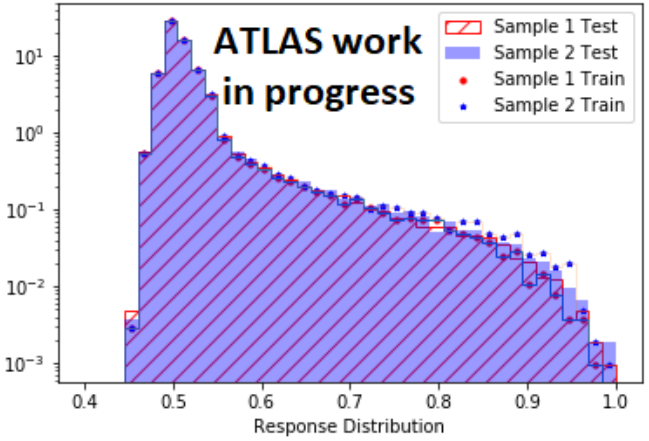



Figure: 1000 VBF125 events

# Response Distribution: m_yy cuts

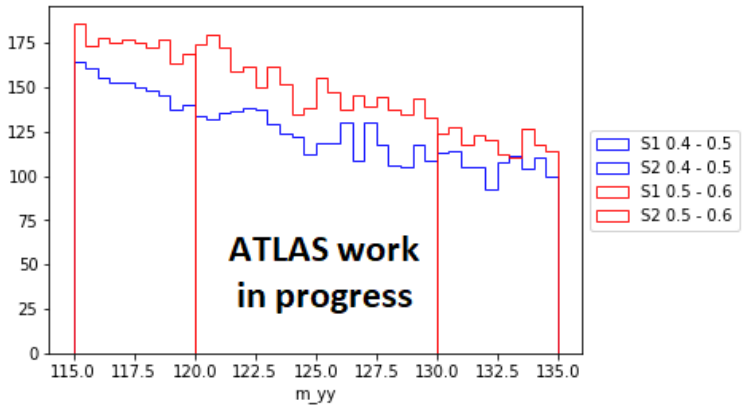



Figure: 10 VBF125 events

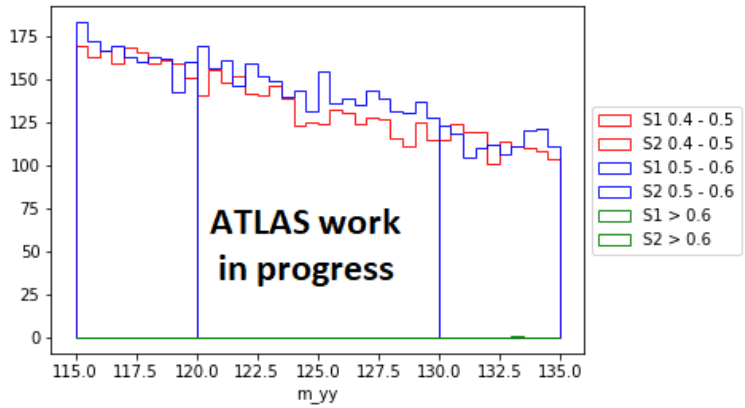



Figure: 100 VBF125 events

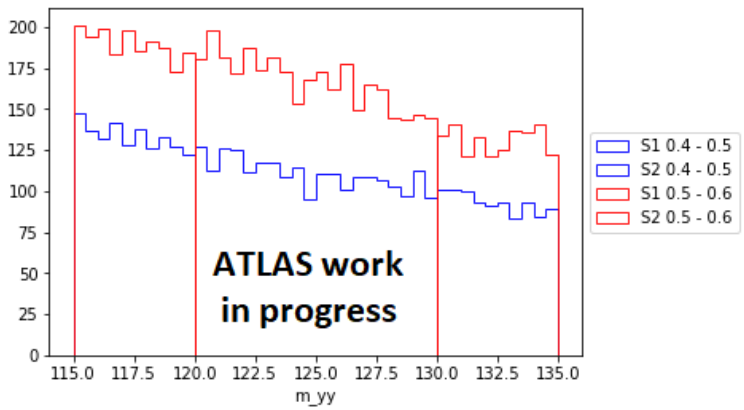



Figure: 250 VBF125 events

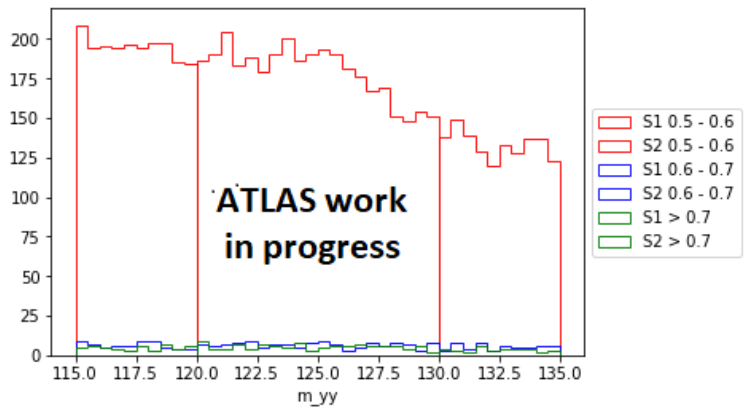



Figure: 1000 VBF125 events

# ttH125

# m_yy

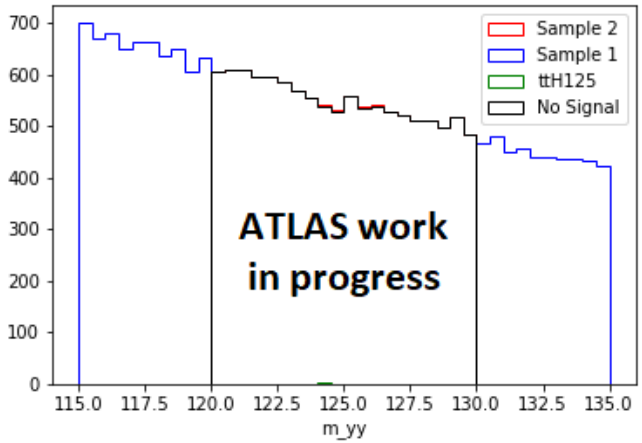



Figure: 10 ttH125 events

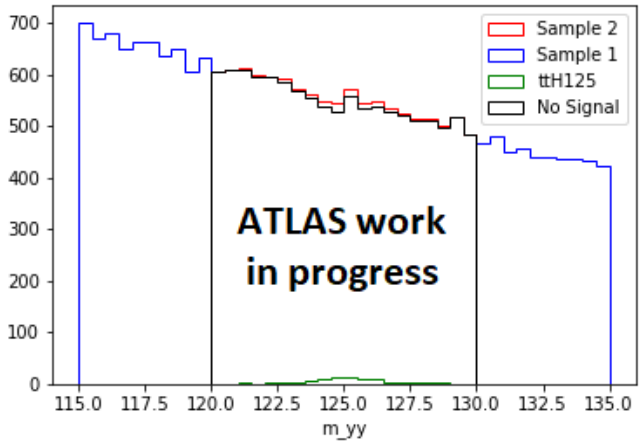



Figure: 100 ttH125 events

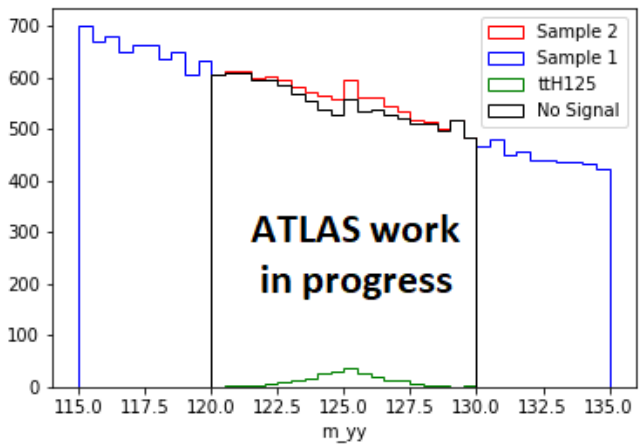



Figure: 250 ttH125 events

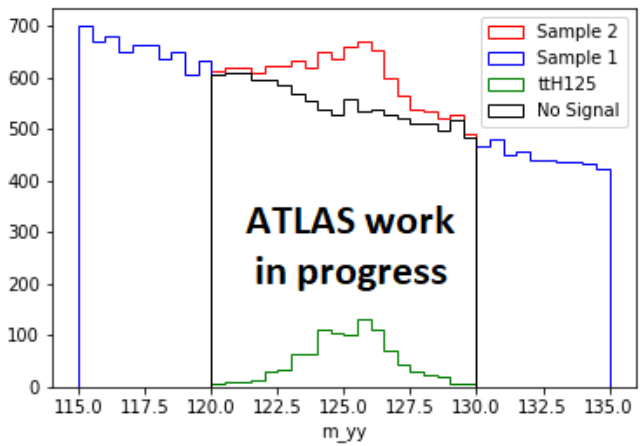



Figure: 1000 ttH125 events

# Response Distributions

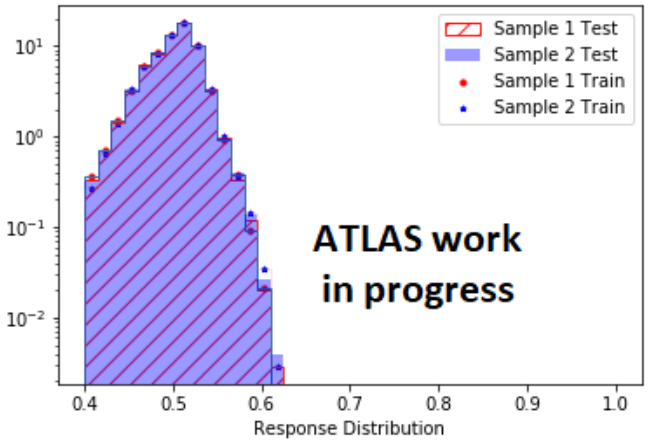



Figure: 10 ttH125 events

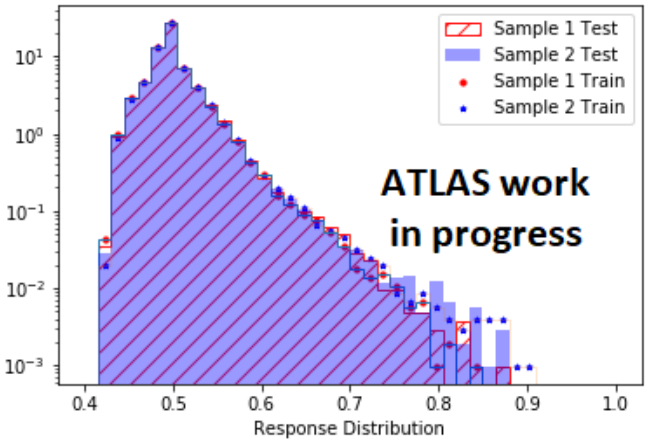



Figure: 100 ttH125 events

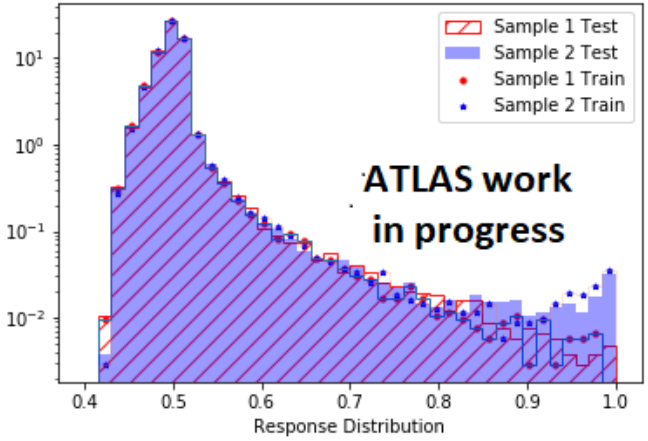



Figure: 250 ttH125 events

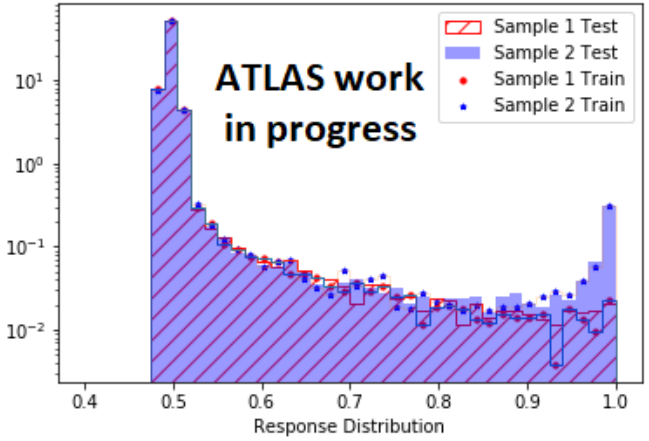



Figure: 1000 ttH125 events

# Response Distribution: m_yy cuts

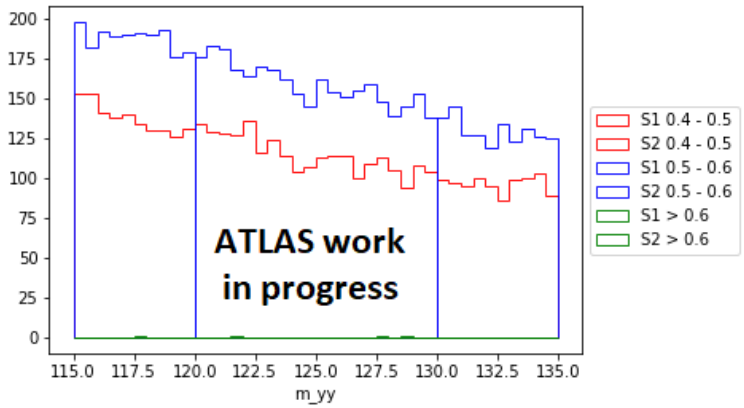



Figure: 10 ttH125 events

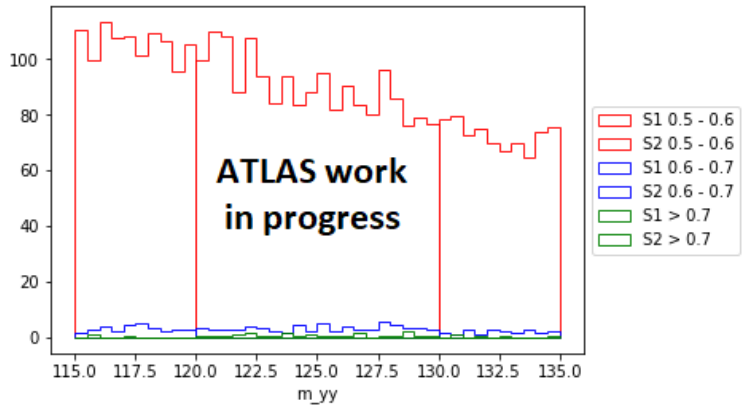



Figure: 100 ttH125 events

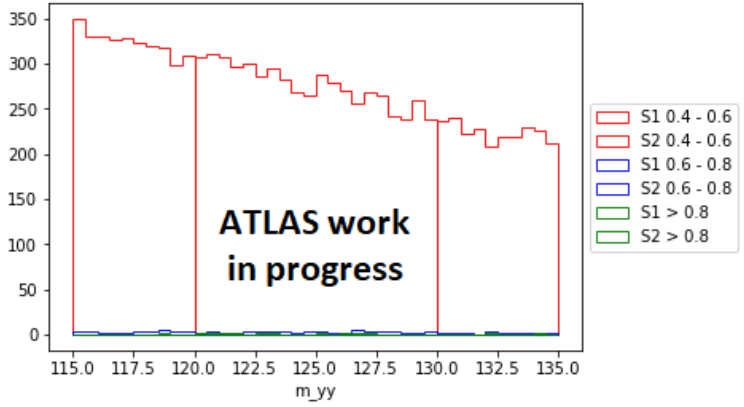



Figure: 250 ttH125 events

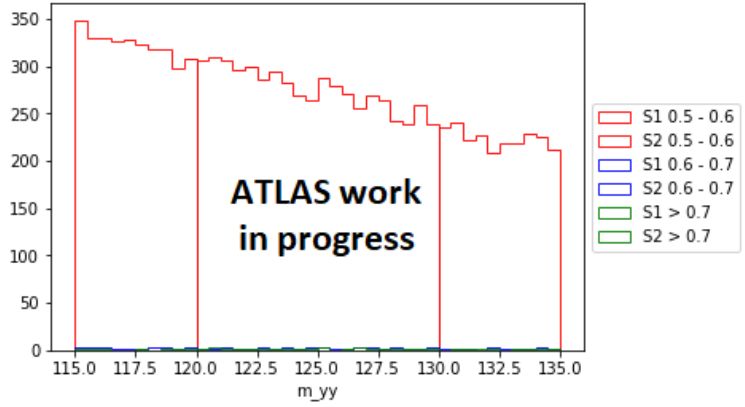



Figure: 1000 ttH125 events

# ggH125

# m_yy

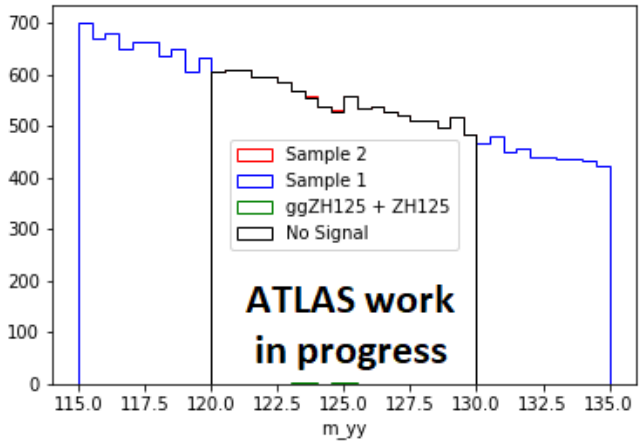



Figure: 10 ggH125 events

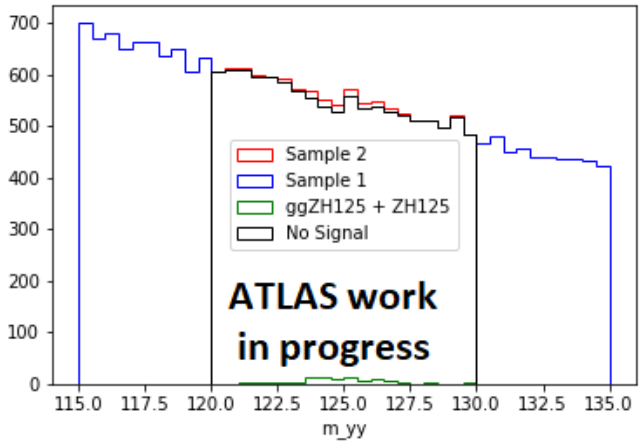



Figure: 100 ggH125 events

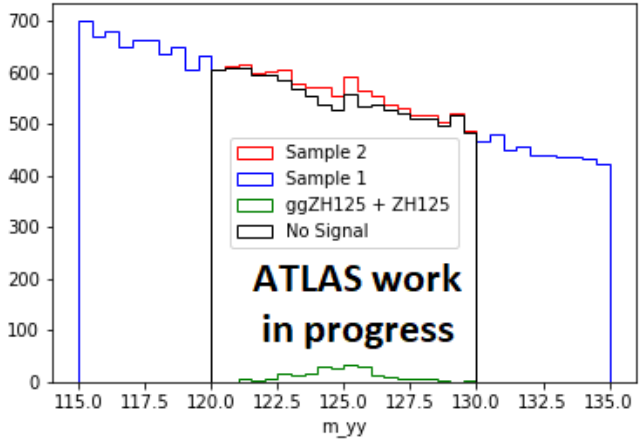



Figure: 250 ggH125 events

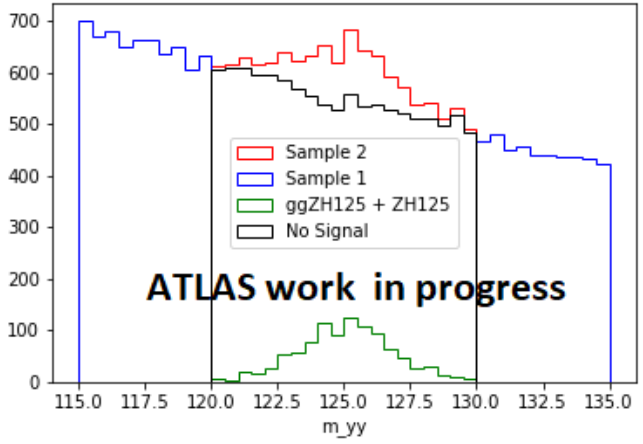



Figure: 1000 ggH125 events

# Response Distributions

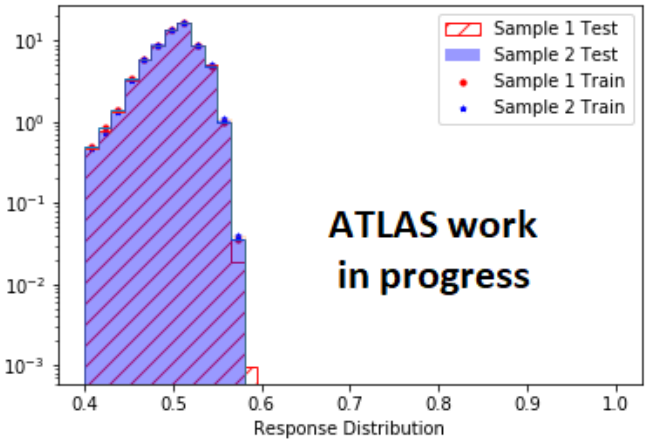



Figure: 10 ggH125 events

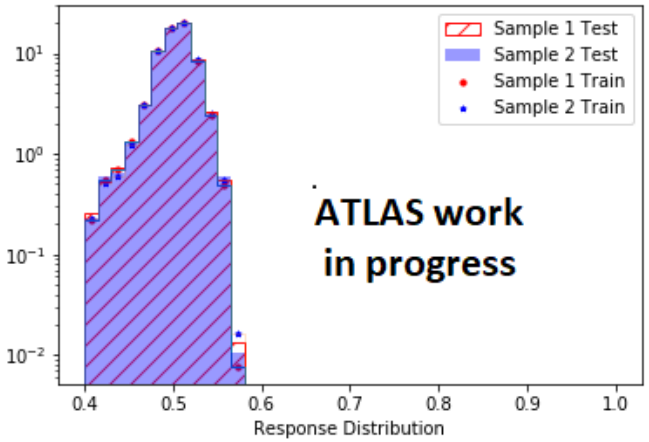



Figure: 100 ggH125 events

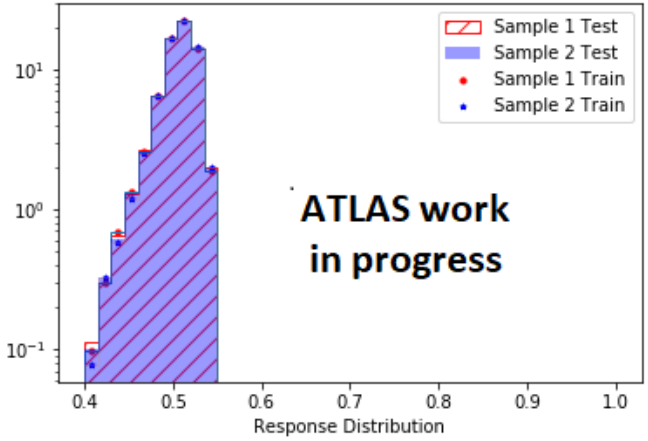



Figure: 250 ggH125 events

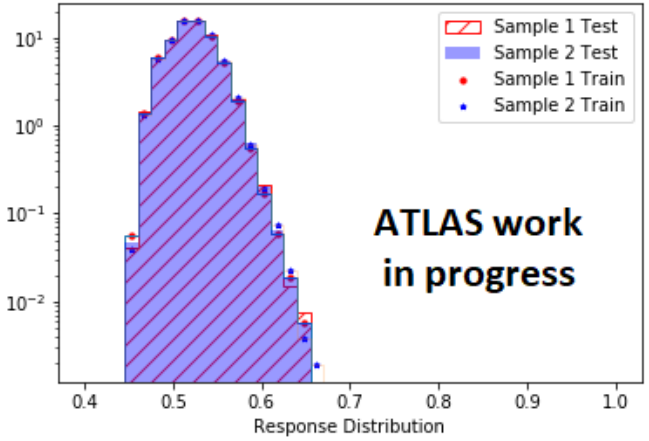



Figure: 1000 ggH125 events

# Response Distribution: m_yy cuts

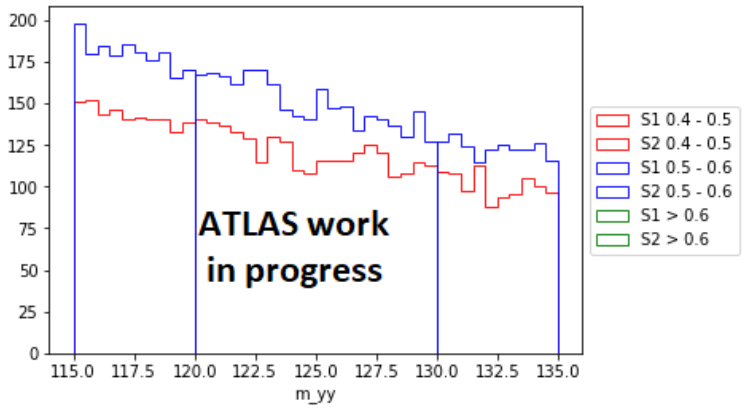



Figure: 10 ggH125 events

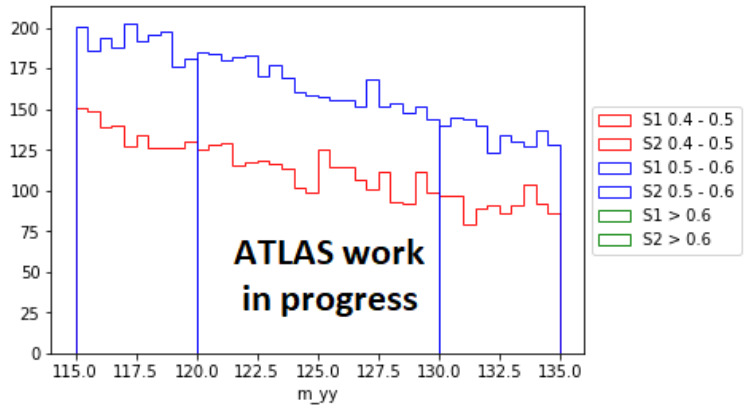



Figure: 100 ggH125 events

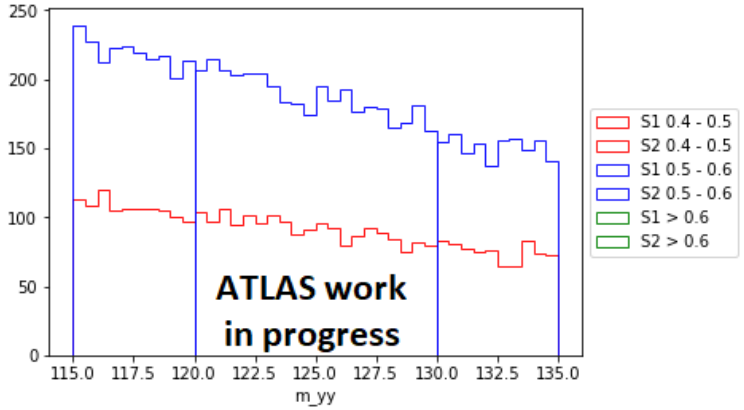



Figure: 250 ggH125 events

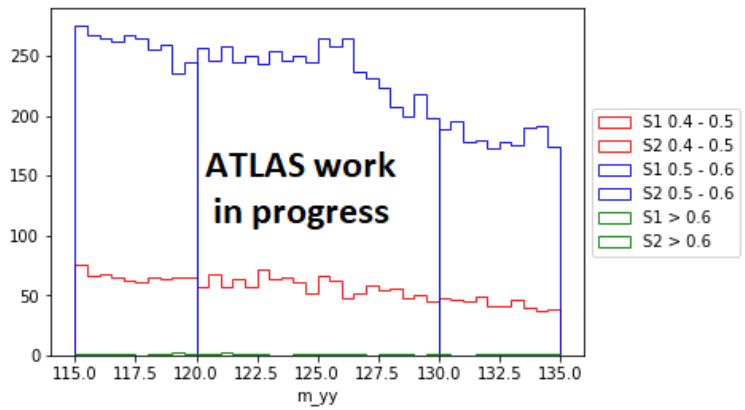



Figure: 1000 ggH125 events

# ggZH125

# m_yy

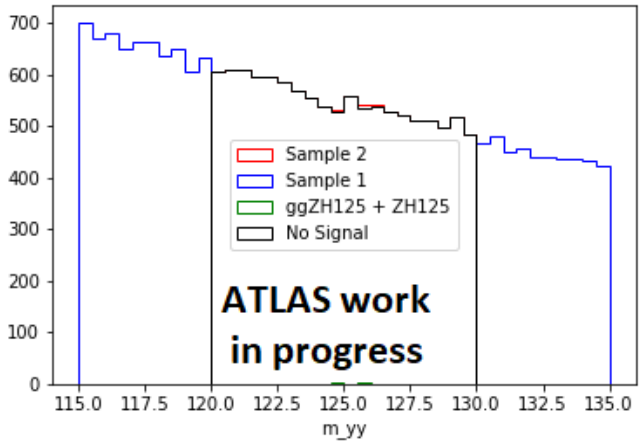



Figure: 10 ggZH125 events

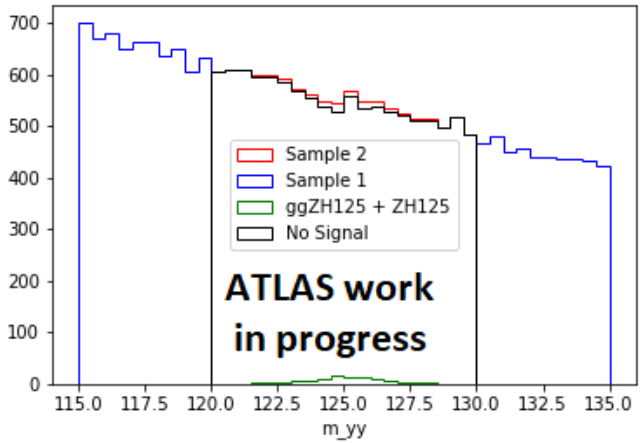



Figure: 100 ggZH125 events

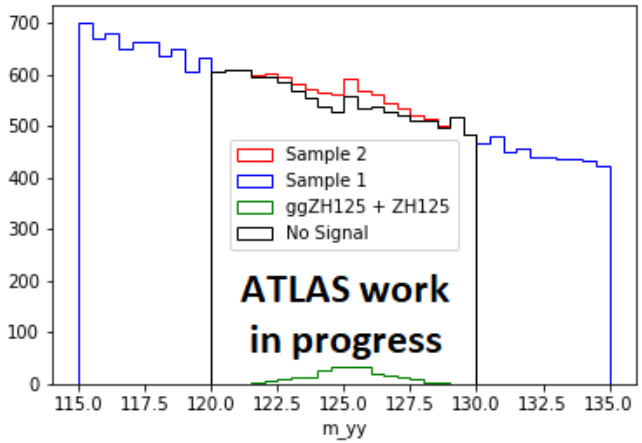



Figure: 250 ggZH125 events

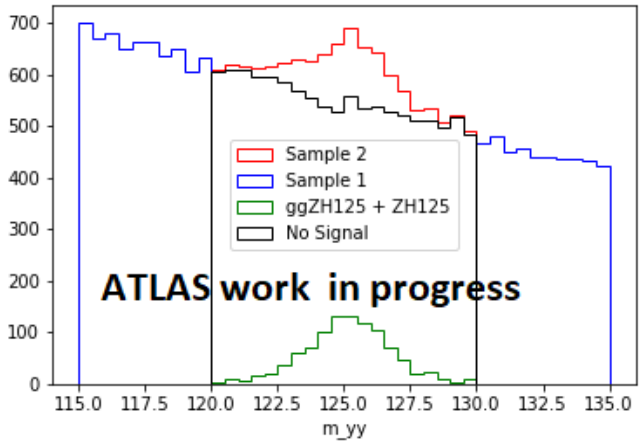



Figure: 1000 ggZH125 events

# Response Distributions

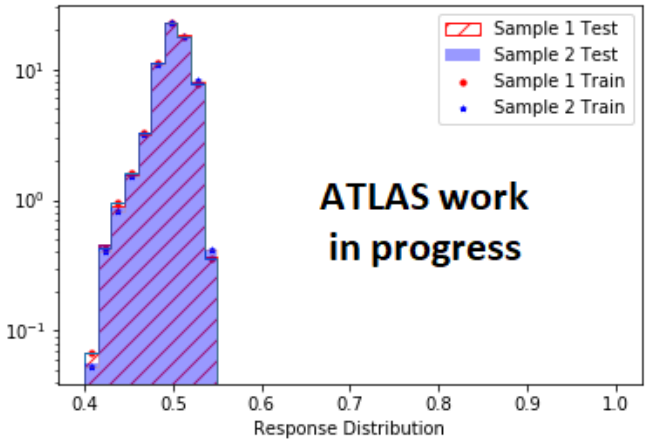



Figure: 10 ggZH125 events

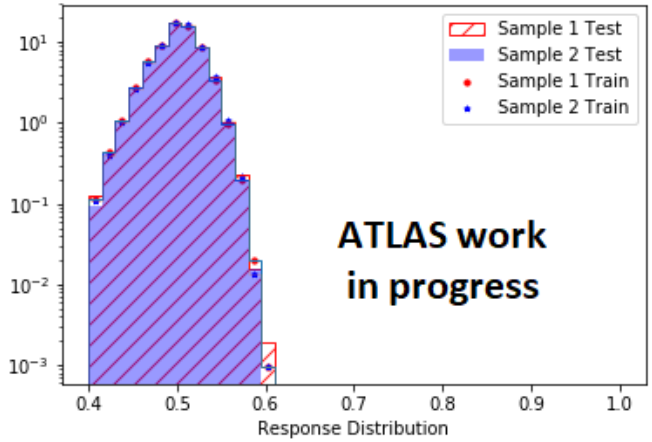



Figure: 100 ggZH125 events

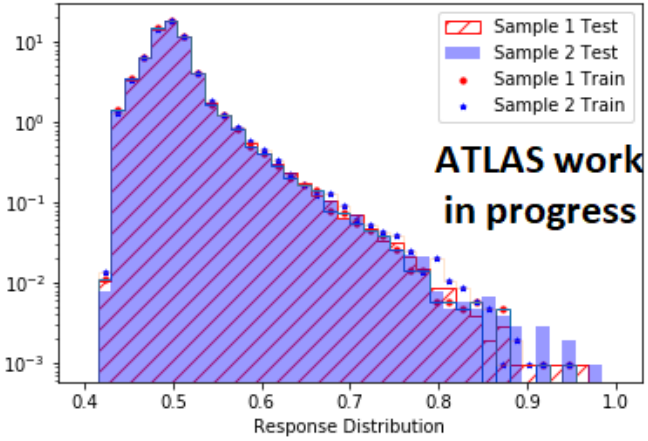



Figure: 250 ggZH125 events

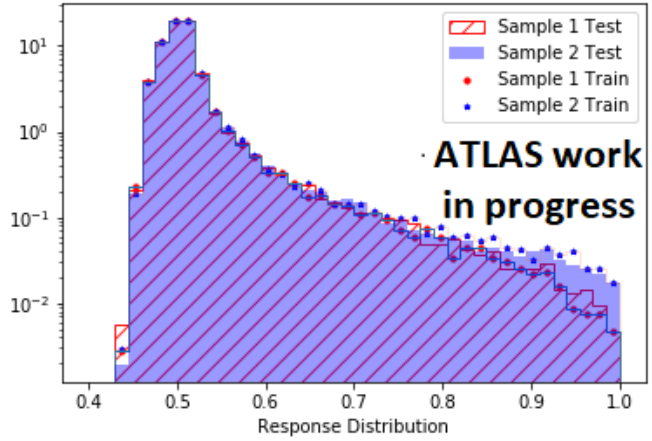



Figure: 1000 ggZH125 events

# Response Distribution: m_yy cuts

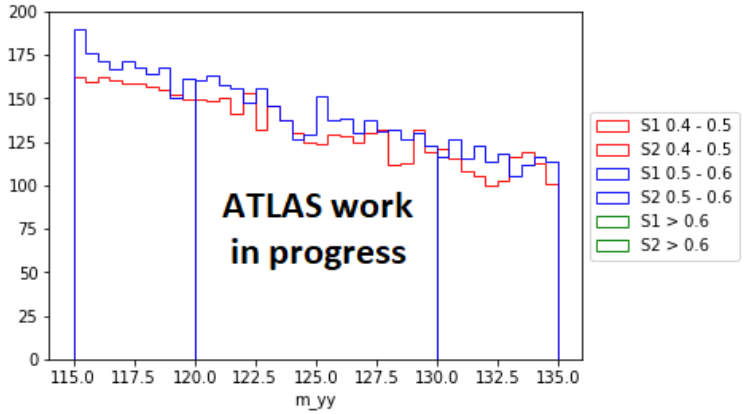



Figure: 10 ggZH125 events

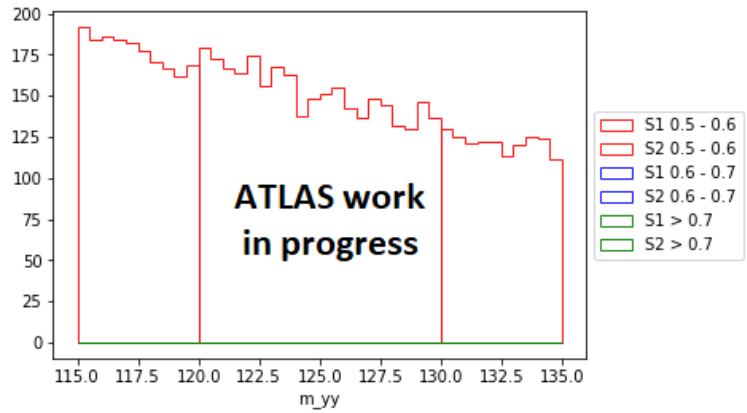



Figure: 100 ggZH125 events

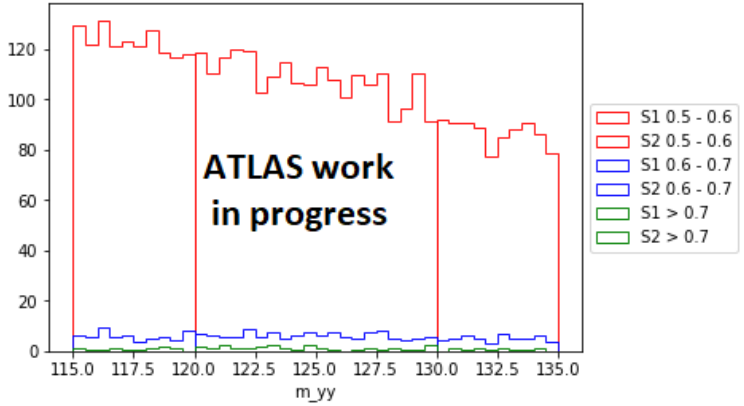



Figure: 250 ggZH125 events

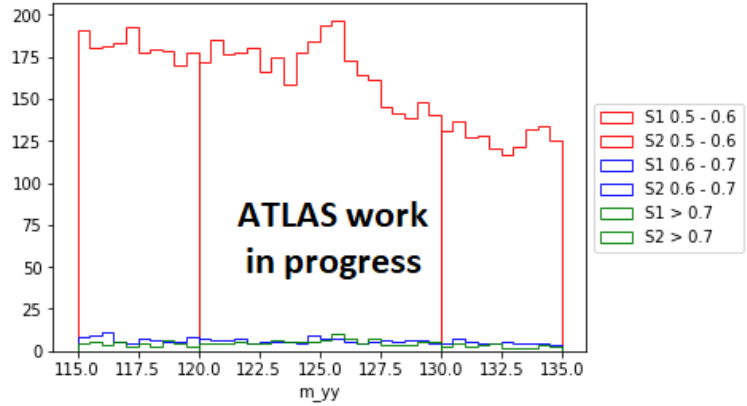



Figure: 1000 ggZH125 events

# WH125

# m_yy

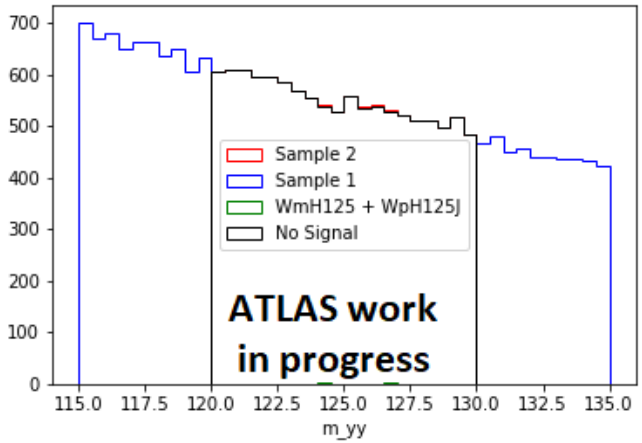



Figure: 10 WH125 events

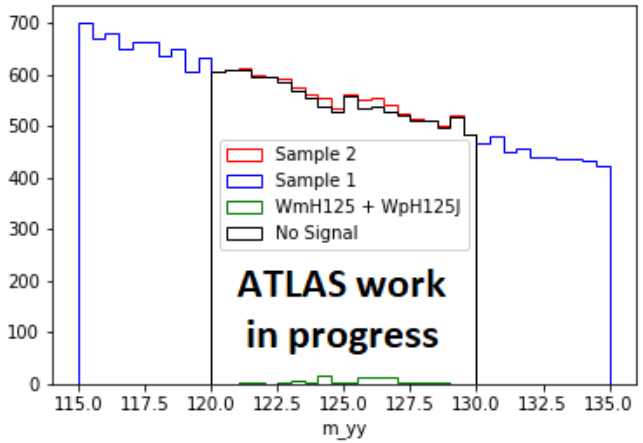



Figure: 100 WH125 events

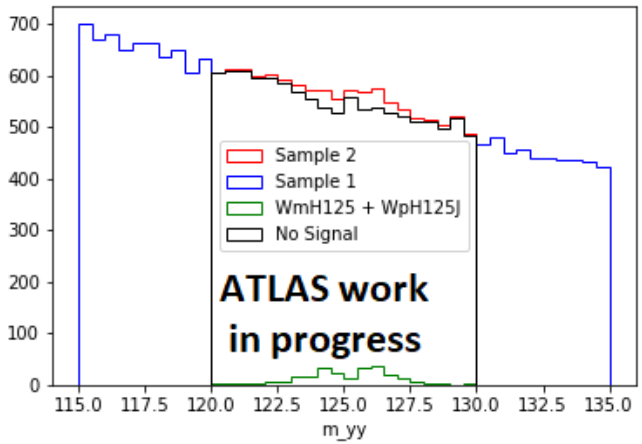



Figure: 250 WH125 events

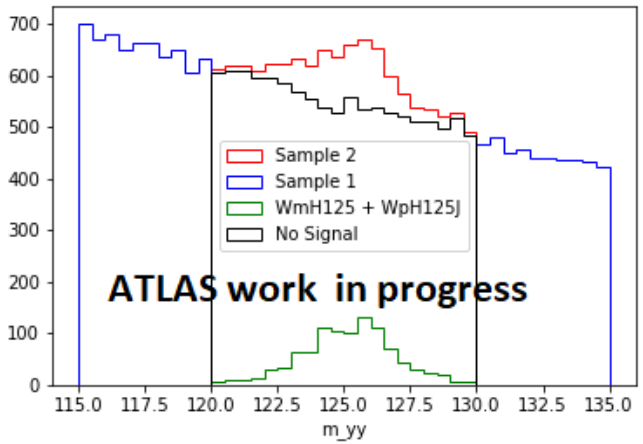



Figure: 1000 WH125 events

# Response Distributions

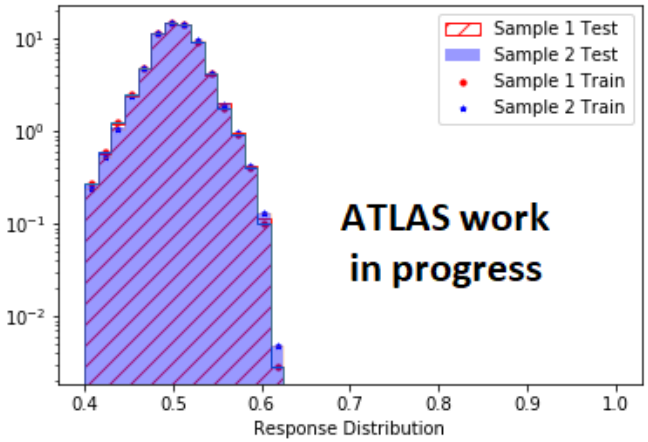



Figure: 10 WH125 events

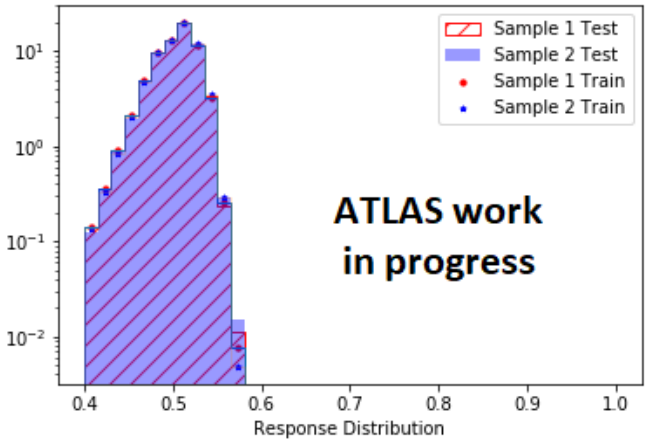



Figure: 100 WH125 events

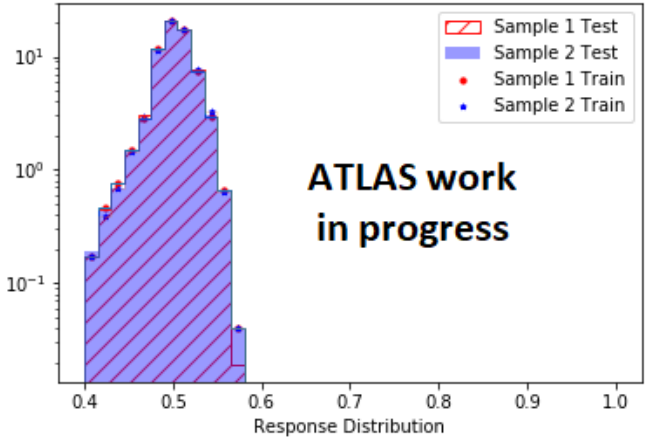



Figure: 250 WH125 events

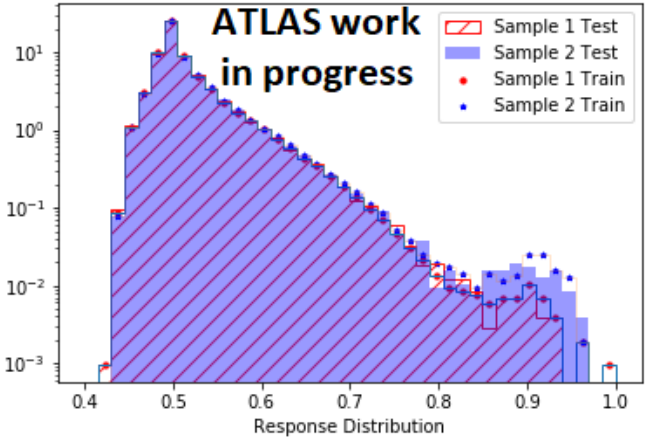



Figure: 1000 WH125 events

# Response Distribution: m_yy cuts

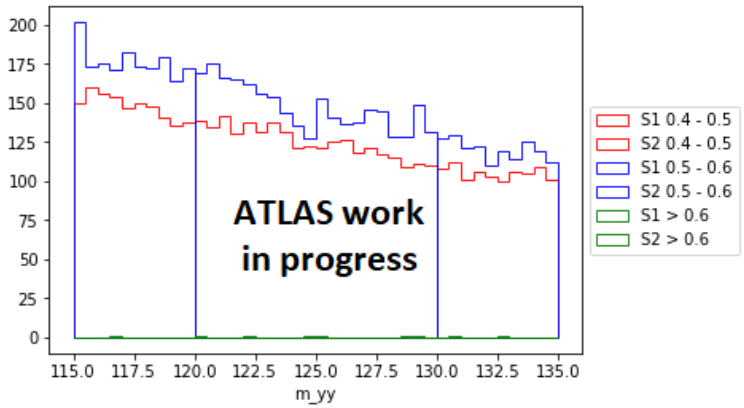



Figure: 10 WH125 events

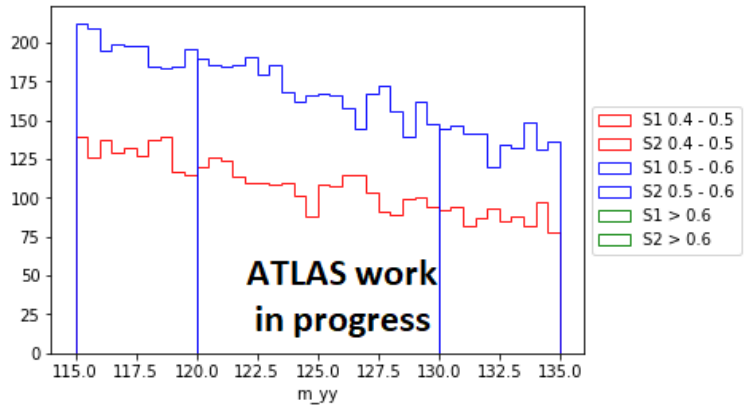



Figure: 100 WH125 events

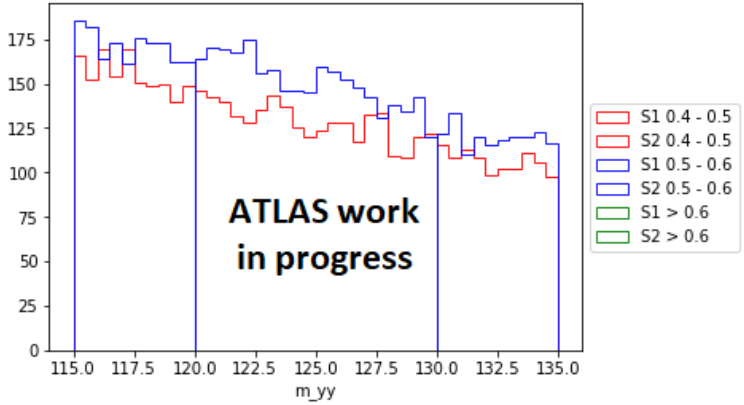



Figure: 250 WH125 events

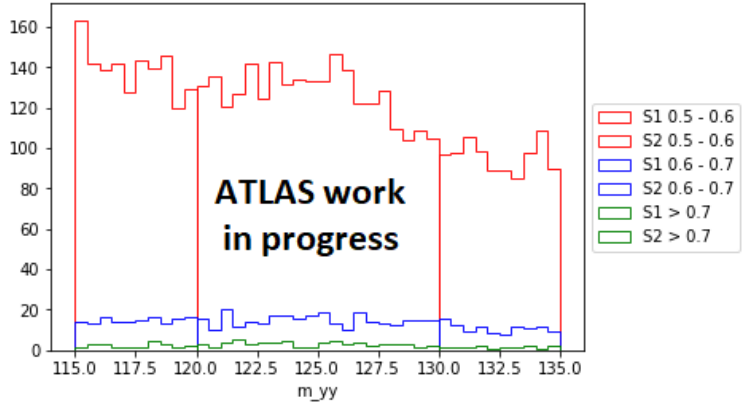



Figure: 1000 WH125 events

# ROC Curves

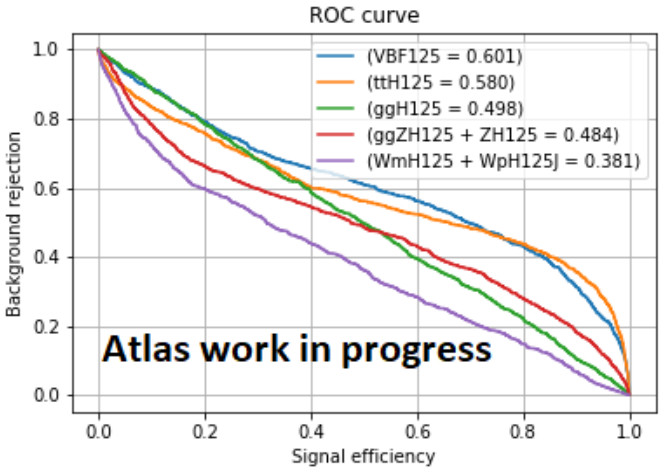



Figure: 10 events

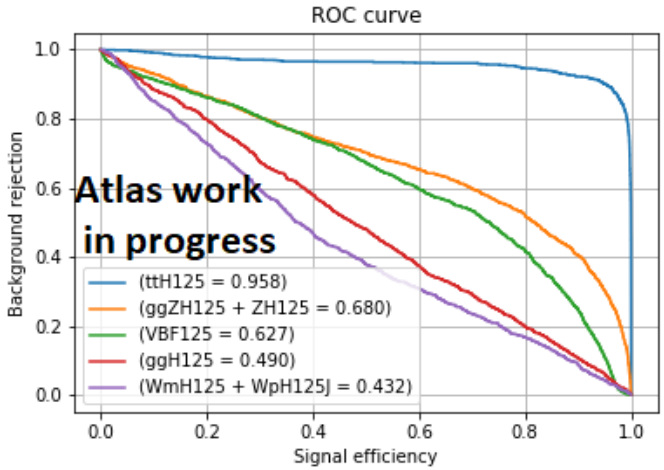



Figure: 100 events

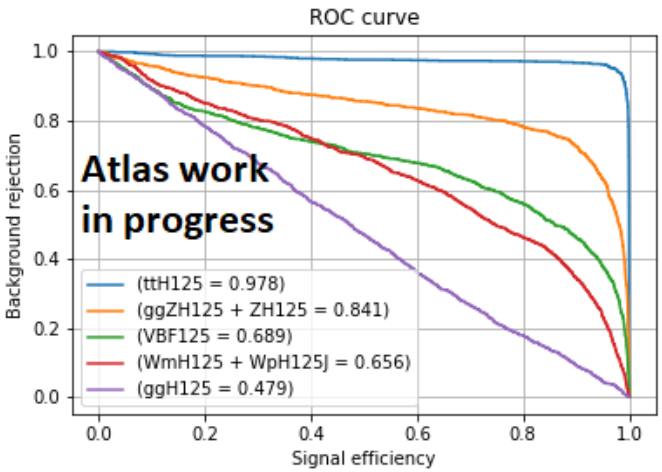



Figure: 250 events

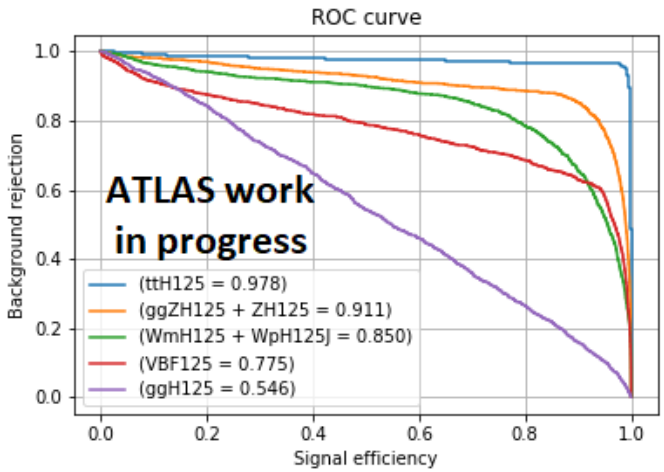
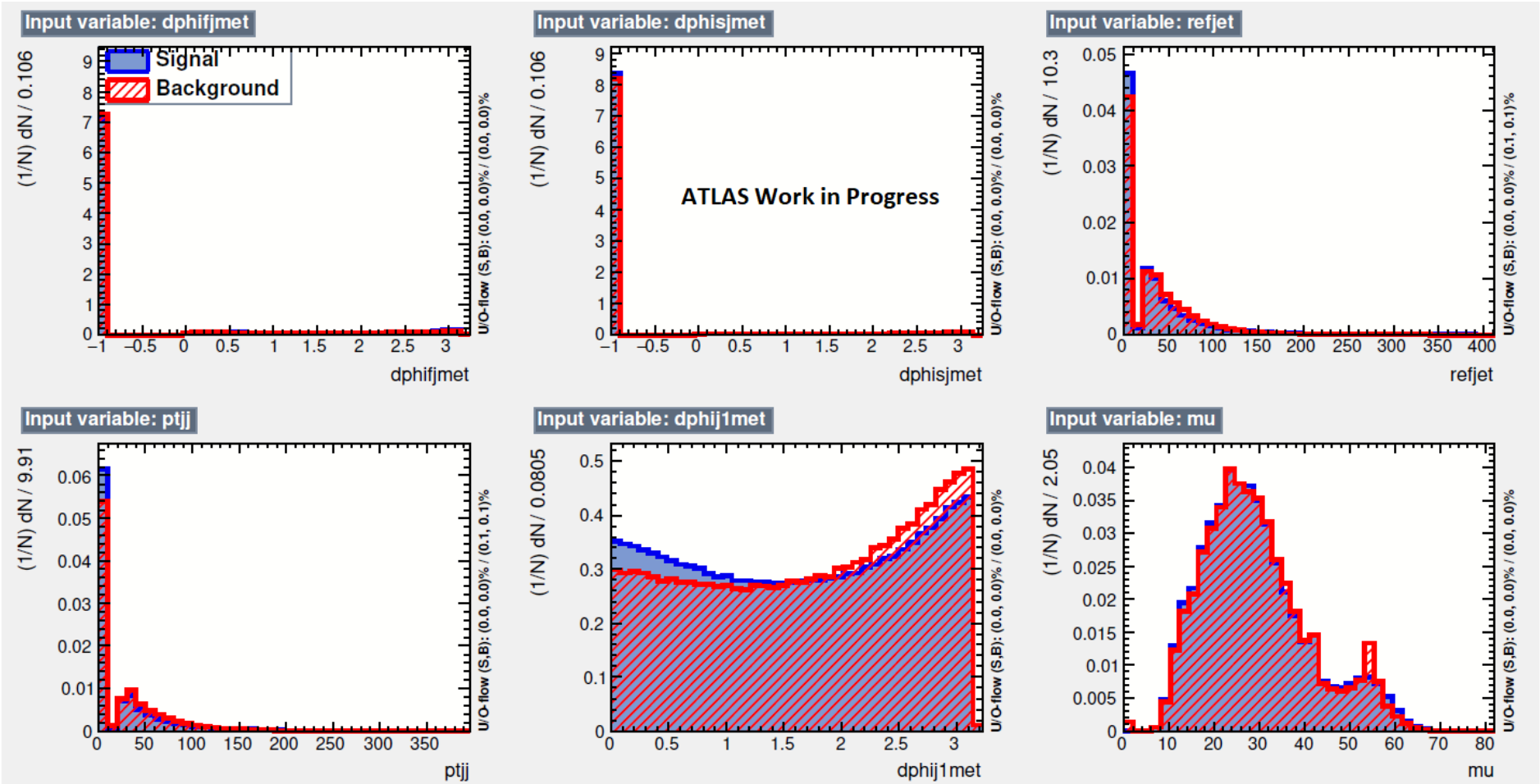
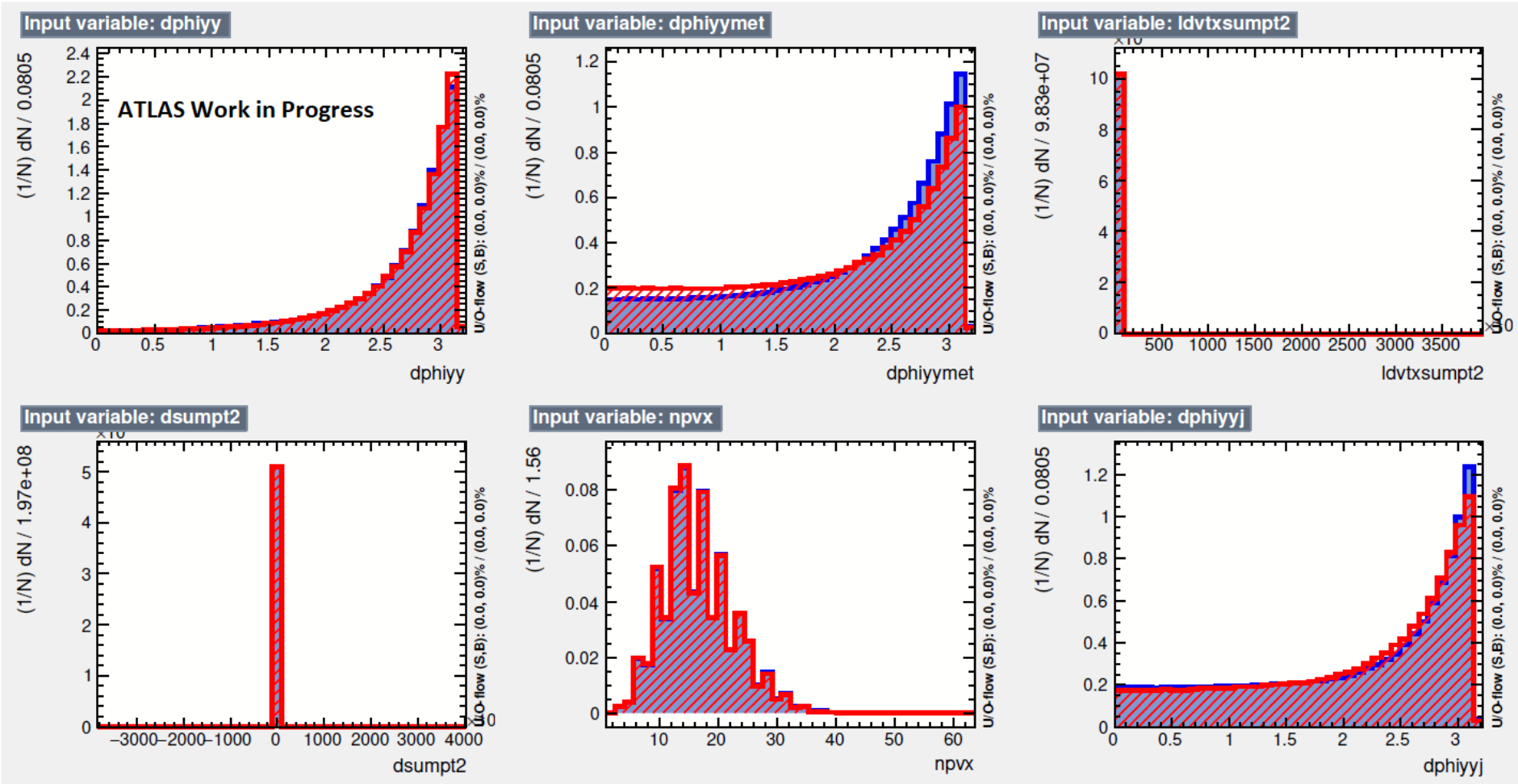



Figure: 1000 events

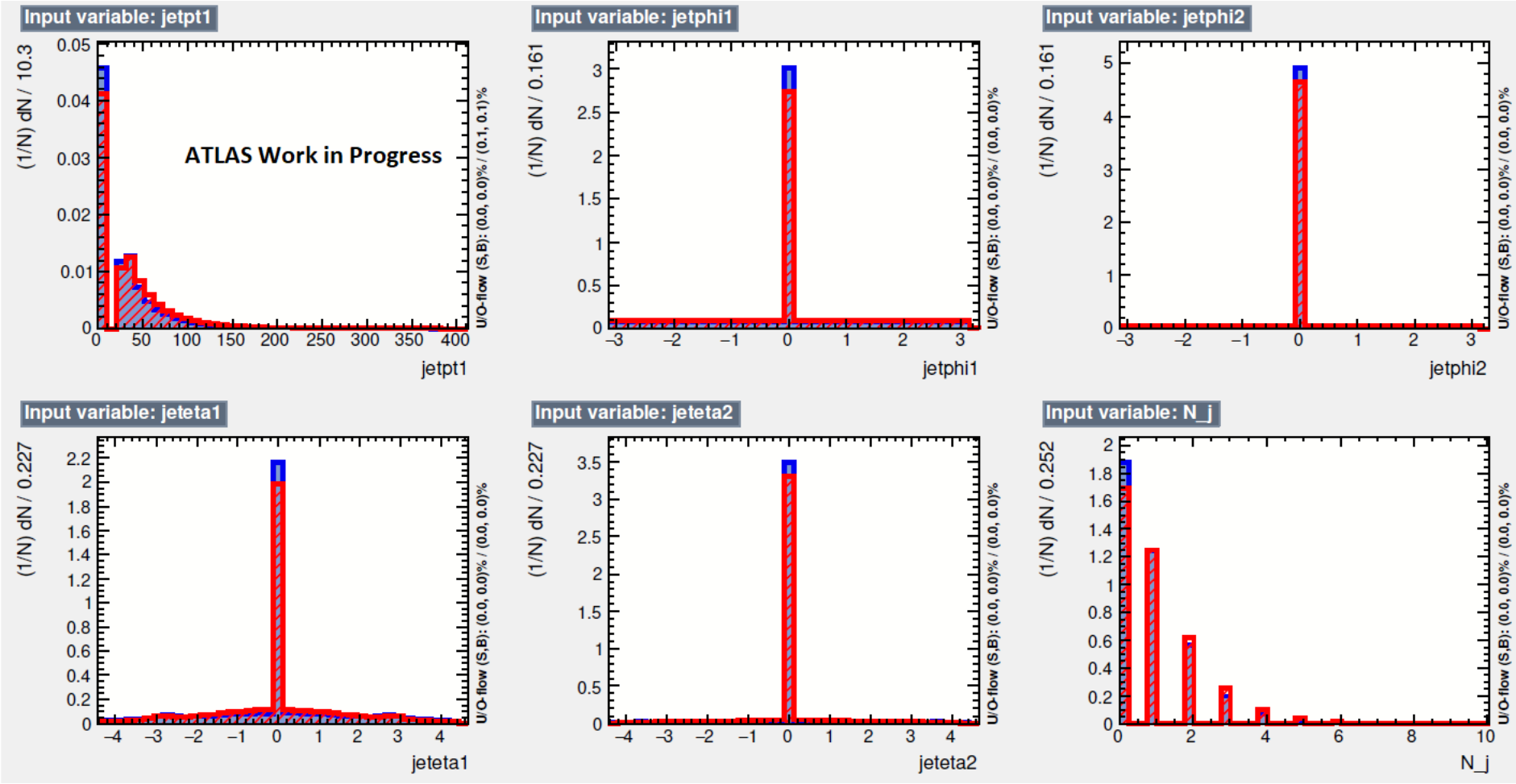
# 1D Distributions: Background vs ggH Signal



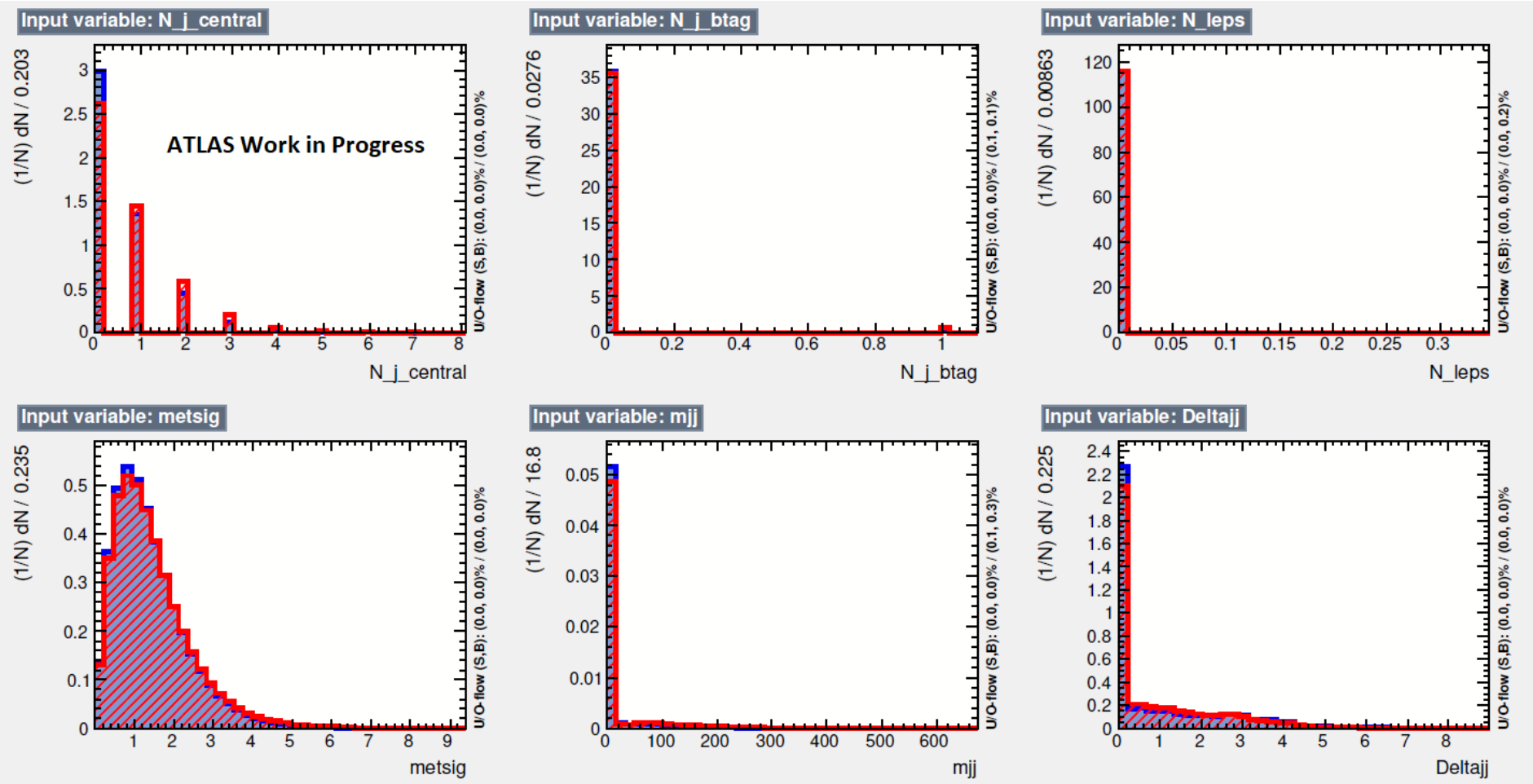
# 1D Distributions: Background vs ggH Signal



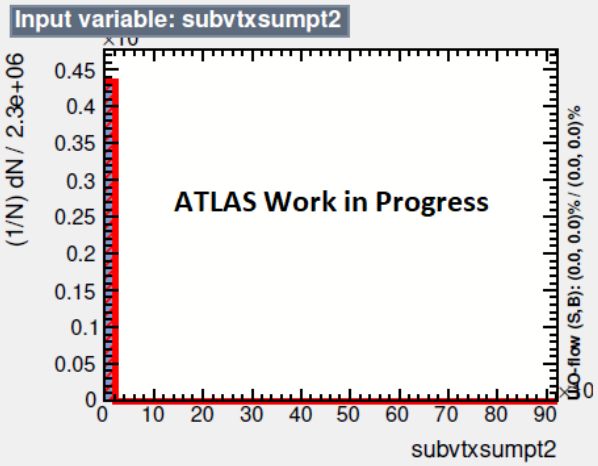
# 1D Distributions: Background vs ggH Signal



# 1D Distributions: Background vs ggH Signal



# 1D Distributions: Background vs ggH Signal



# Variables Ranking: ggH and VBFH

| Rank | Variable | Separation |
|---|---|---|
| 1 | N_j_central | 8.923e-03 |
| 2 | ptjj | 8.042e-03 |
| 3 | mjj | 7.363e-03 |
| 4 | jeteta1 | 7.134e-03 |
| 5 | dphiyymet | 6.834e-03 |
| 6 | refjet | 5.620e-03 |
| 7 | jetpt1 | 4.878e-03 |
| 8 | Deltajj | 4.876e-03 |
| 9 | jeteta2 | 4.777e-03 |
| 10 | N_j | 4.101e-03 |
| 11 | mu | 3.387e-03 |
| 12 | jetphi2 | 2.650e-03 |
| 13 | dphij1met | 2.311e-03 |
| 14 | jetphi1 | 2.055e-03 |
| 15 | dphiyyj | 1.122e-03 |
| 16 | metsig | 1.038e-03 |
| 17 | dphisjmet | 9.643e-04 |
| 18 | dphifjmet | 6.986e-04 |
| 19 | N_j_btag | 5.689e-04 |
| 20 | dphiyy | 3.364e-04 |
| 21 | npvx | 1.115e-04 |
| 22 | dsumpt2 | 1.608e-06 |
| 23 | subvtxsumpt2 | 1.129e-06 |
| 24 | ldvtxsumpt2 | 7.397e-07 |
| 25 | N_leps | 0.000e+00 |

***ggH***

| Rank | Variable | Separation |
|---|---|---|
| 1 | Deltajj | 3.069e-01 |
| 2 | jetpt1 | 2.836e-01 |
| 3 | refjet | 2.716e-01 |
| 4 | N_j | 2.232e-01 |
| 5 | jeteta1 | 2.208e-01 |
| 6 | dphiyy | 2.108e-01 |
| 7 | dphifjmet | 2.006e-01 |
| 8 | jeteta2 | 1.808e-01 |
| 9 | jetphi1 | 1.639e-01 |
| 10 | jetphi2 | 1.454e-01 |
| 11 | dphiyyj | 1.104e-01 |
| 12 | ptjj | 1.045e-01 |
| 13 | N_j_central | 5.364e-02 |
| 14 | mjj | 2.743e-02 |
| 15 | dphiyymet | 1.360e-02 |
| 16 | mu | 3.461e-03 |
| 17 | dphij1met | 3.349e-03 |
| 18 | dphisjmet | 3.014e-03 |
| 19 | N_j_btag | 6.605e-04 |
| 20 | metsig | 3.616e-04 |
| 21 | npvx | 2.056e-04 |
| 22 | dsumpt2 | 6.968e-06 |
| 23 | ldvtxsumpt2 | 4.513e-06 |
| 24 | subvtxsumpt2 | 2.700e-06 |
| 25 | N_leps | 0.000e+00 |

***VBFH***

# Variables Ranking: ttH, ZH and WH

| Rank | Variable | Separation |
|---|---|---|
| 1 | N_j_central | 7.069e-01 |
| 2 | N_j | 6.946e-01 |
| 3 | mjj | 6.267e-01 |
| 4 | jetpt1 | 6.122e-01 |
| 5 | N_j_btag | 6.117e-01 |
| 6 | jetphi2 | 4.931e-01 |
| 7 | jeteta2 | 4.710e-01 |
| 8 | ptjj | 4.679e-01 |
| 9 | refjet | 4.662e-01 |
| 10 | dphiyy | 3.338e-01 |
| 11 | jeteta1 | 2.482e-01 |
| 12 | jetphi1 | 2.270e-01 |
| 13 | N_leps | 1.843e-01 |
| 14 | Deltajj | 1.749e-01 |
| 15 | metsig | 1.264e-01 |
| 16 | dphisjmet | 5.148e-02 |
| 17 | dphifjmet | 4.481e-02 |
| 18 | dphiyymet | 2.161e-02 |
| 19 | dphiyyj | 1.653e-02 |
| 20 | mu | 3.326e-03 |
| 21 | dphij1met | 2.439e-03 |
| 22 | npvx | 2.906e-04 |
| 23 | ldvtxsumpt2 | 1.472e-05 |
| 24 | dsumpt2 | 2.096e-06 |
| 25 | subvtxsumpt2 | 5.323e-33 |

***ttH***

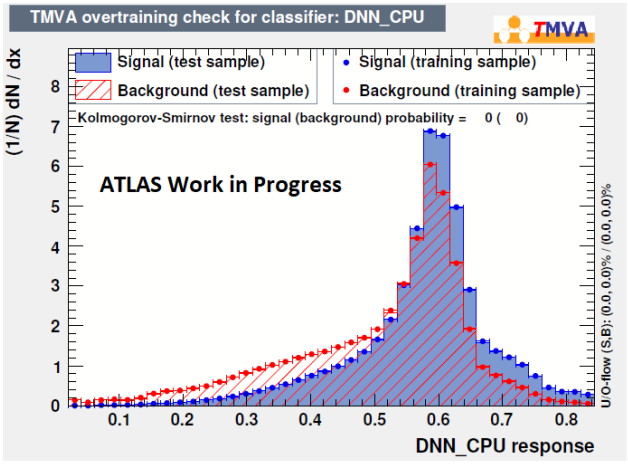
| Rank | Variable | Separation |
|---|---|---|
| 1 | dphiyy | 1.971e-01 |
| 2 | refjet | 1.255e-01 |
| 3 | ptjj | 1.249e-01 |
| 4 | N_j_central | 1.207e-01 |
| 5 | jetpt1 | 1.187e-01 |
| 6 | mjj | 1.157e-01 |
| 7 | N_j | 1.086e-01 |
| 8 | jeteta2 | 9.170e-02 |
| 9 | jetphi2 | 9.018e-02 |
| 10 | jeteta1 | 7.384e-02 |
| 11 | jetphi1 | 6.416e-02 |
| 12 | Deltajj | 6.351e-02 |
| 13 | metsig | 5.760e-02 |
| 14 | N_j_btag | 4.070e-02 |
| 15 | N_leps | 3.315e-02 |
| 16 | dphiyyj | 1.965e-02 |
| 17 | dphisjmet | 5.590e-03 |
| 18 | dphiyymet | 4.219e-03 |
| 19 | mu | 3.375e-03 |
| 20 | dphifjmet | 2.071e-03 |
| 21 | dphij1met | 1.270e-03 |
| 22 | npvx | 2.560e-04 |
| 23 | dsumpt2 | 9.966e-07 |
| 24 | subvtxsumpt2 | 7.010e-07 |
| 25 | ldvtxsumpt2 | 5.677e-07 |

***ZH***

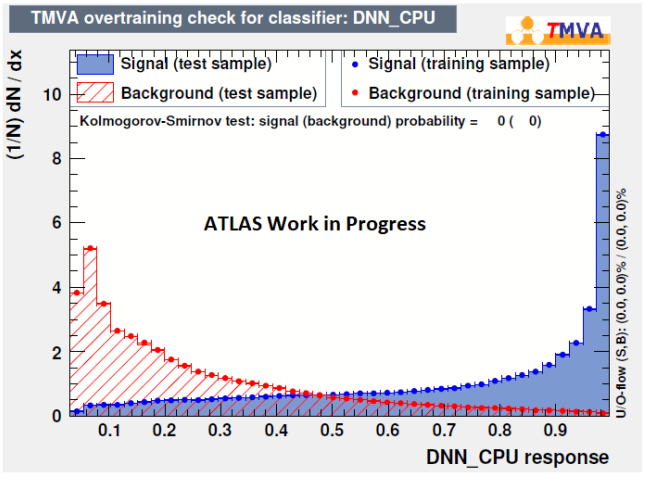
| Rank | Variable | Separation |
|---|---|---|
| 1 | dphiyy | 1.930e-01 |
| 2 | refjet | 1.277e-01 |
| 3 | ptjj | 1.238e-01 |
| 4 | jetpt1 | 1.140e-01 |
| 5 | N_j_central | 1.133e-01 |
| 6 | mjj | 1.041e-01 |
| 7 | N_j | 1.021e-01 |
| 8 | N_leps | 9.160e-02 |
| 9 | jeteta2 | 8.048e-02 |
| 10 | jetphi2 | 7.900e-02 |
| 11 | jeteta1 | 7.673e-02 |
| 12 | jetphi1 | 6.681e-02 |
| 13 | Deltajj | 6.375e-02 |
| 14 | metsig | 5.274e-02 |
| 15 | dphiyyj | 2.128e-02 |
| 16 | dphisjmet | 5.099e-03 |
| 17 | mu | 3.396e-03 |
| 18 | N_j_btag | 2.860e-03 |
| 19 | dphifjmet | 2.026e-03 |
| 20 | dphiyymet | 1.245e-03 |
| 21 | dphij1met | 1.086e-03 |
| 22 | npvx | 2.731e-04 |
| 23 | dsumpt2 | 1.693e-05 |
| 24 | subvtxsumpt2 | 9.723e-06 |
| 25 | ldvtxsumpt2 | 8.373e-06 |

***WH***

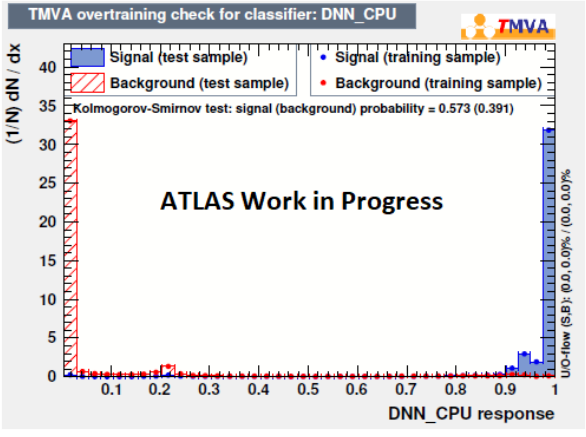
# Supervised Learning Response Distribution



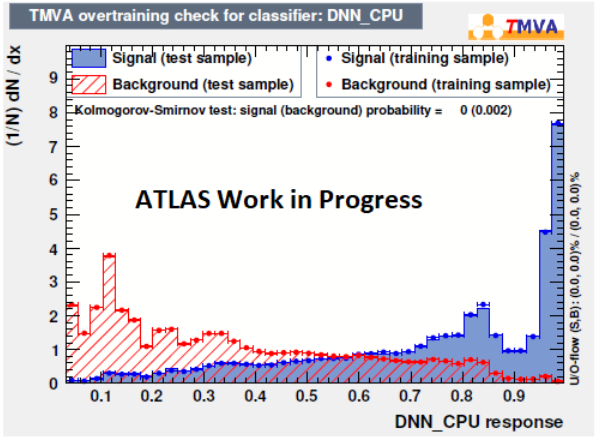

*ggH*



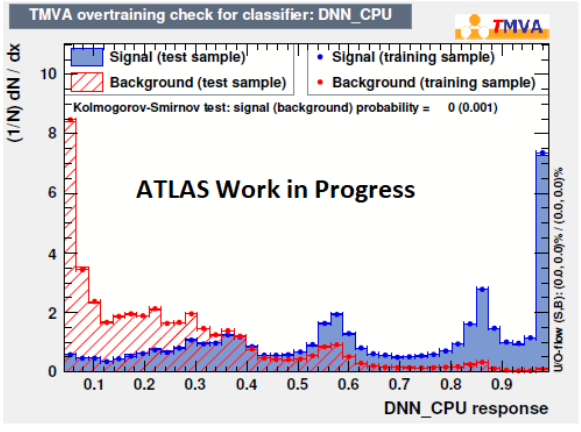

*VBFH*




*ttH*

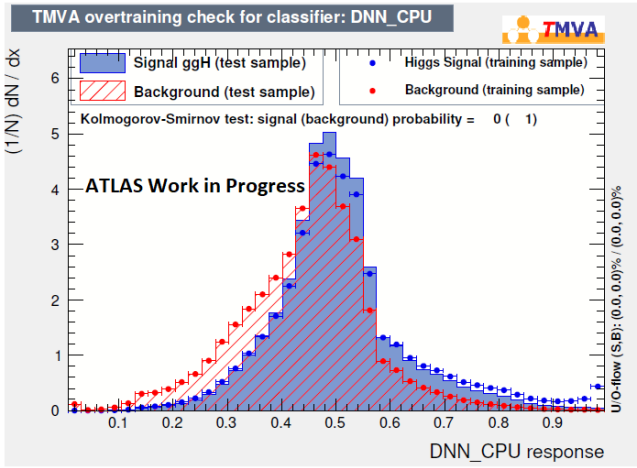



*WH*

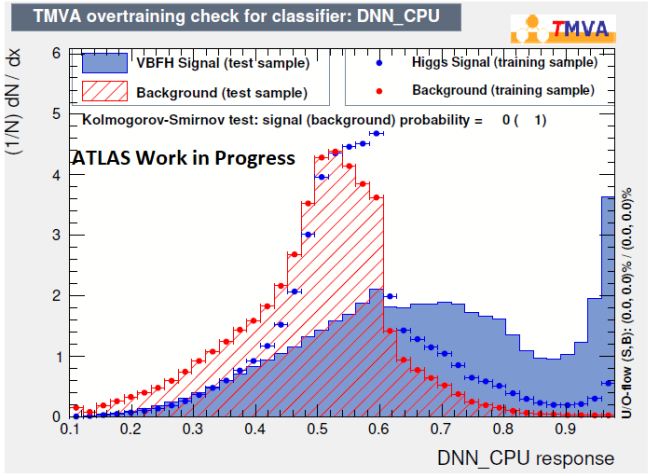



*ZH*

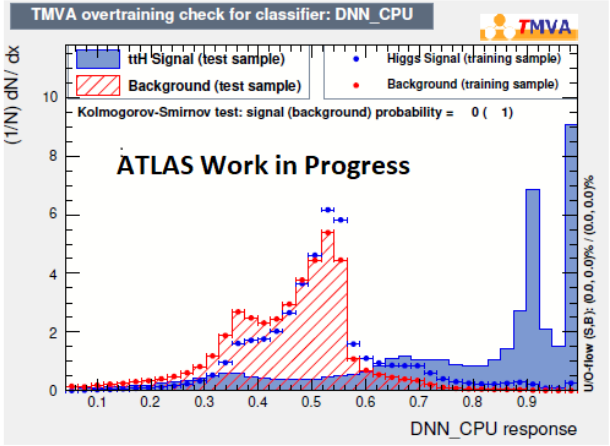
# Unlabeled-Supervised Learning Response Distribution



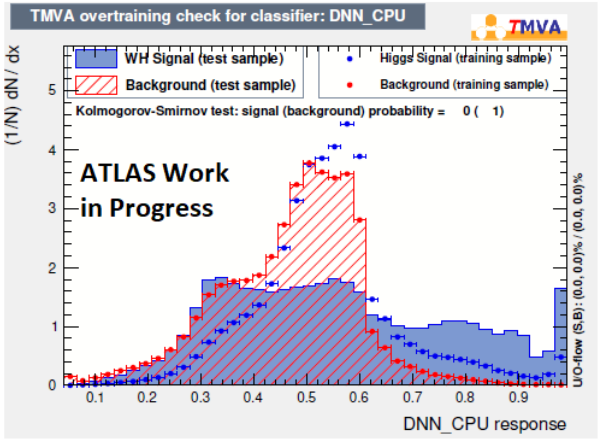

*ggH*



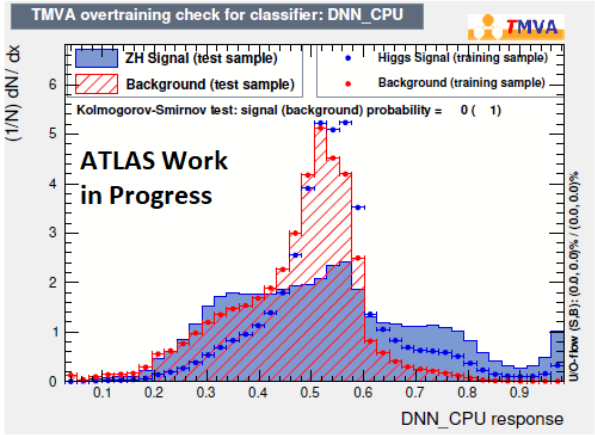

*VBFH*




*ttH*

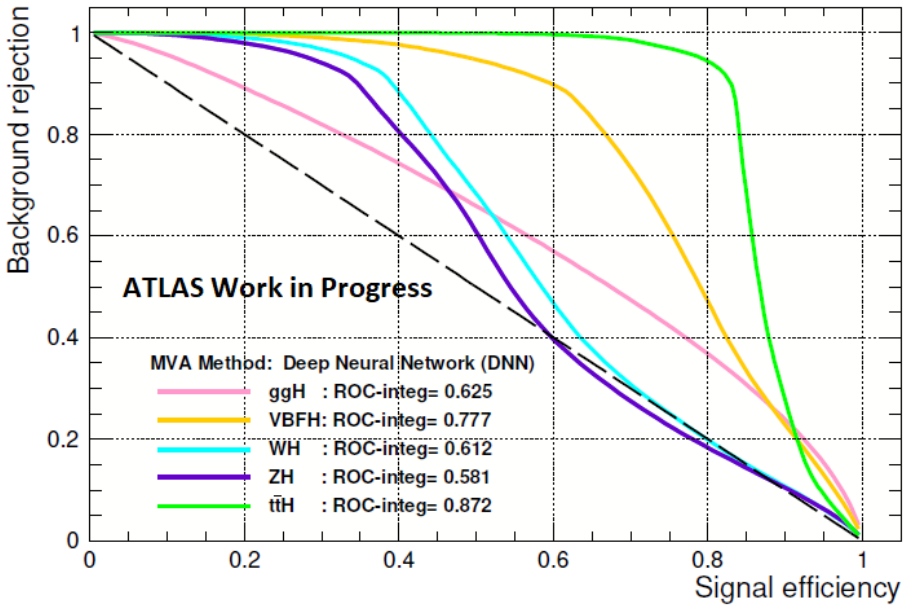



*WH*

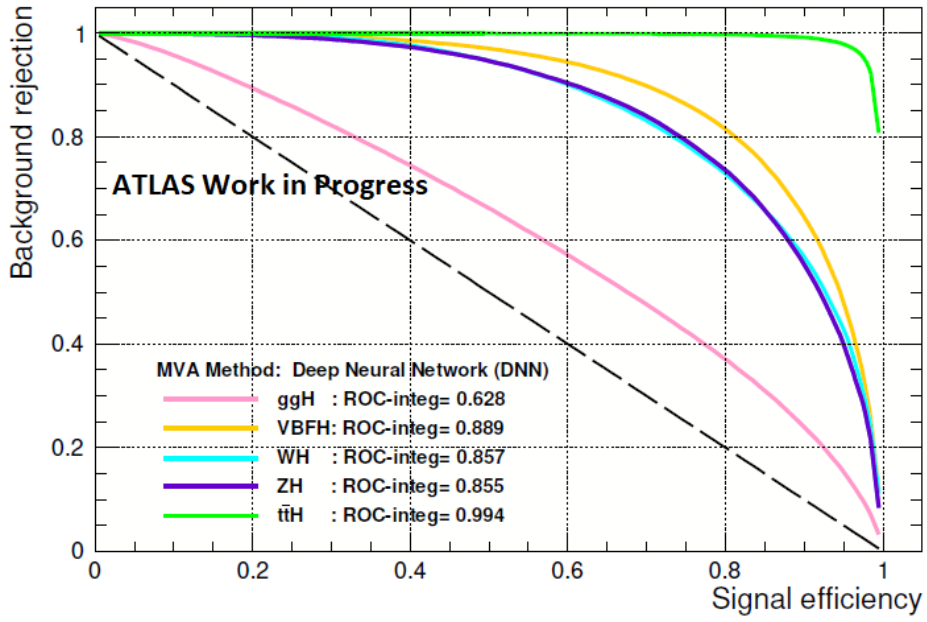



*ZH*

# ROC curves comparison

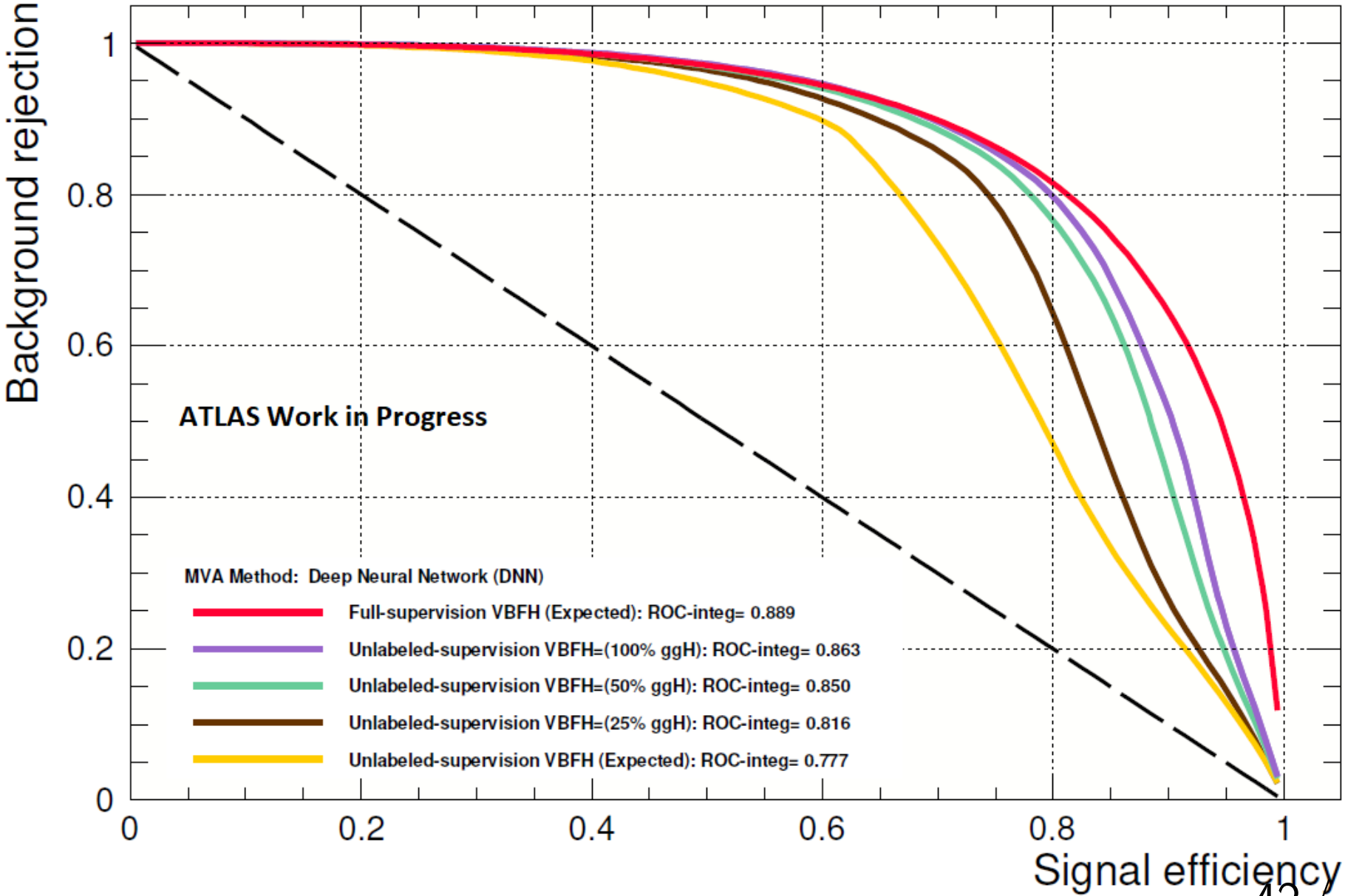



***Unlabeled-supervision***

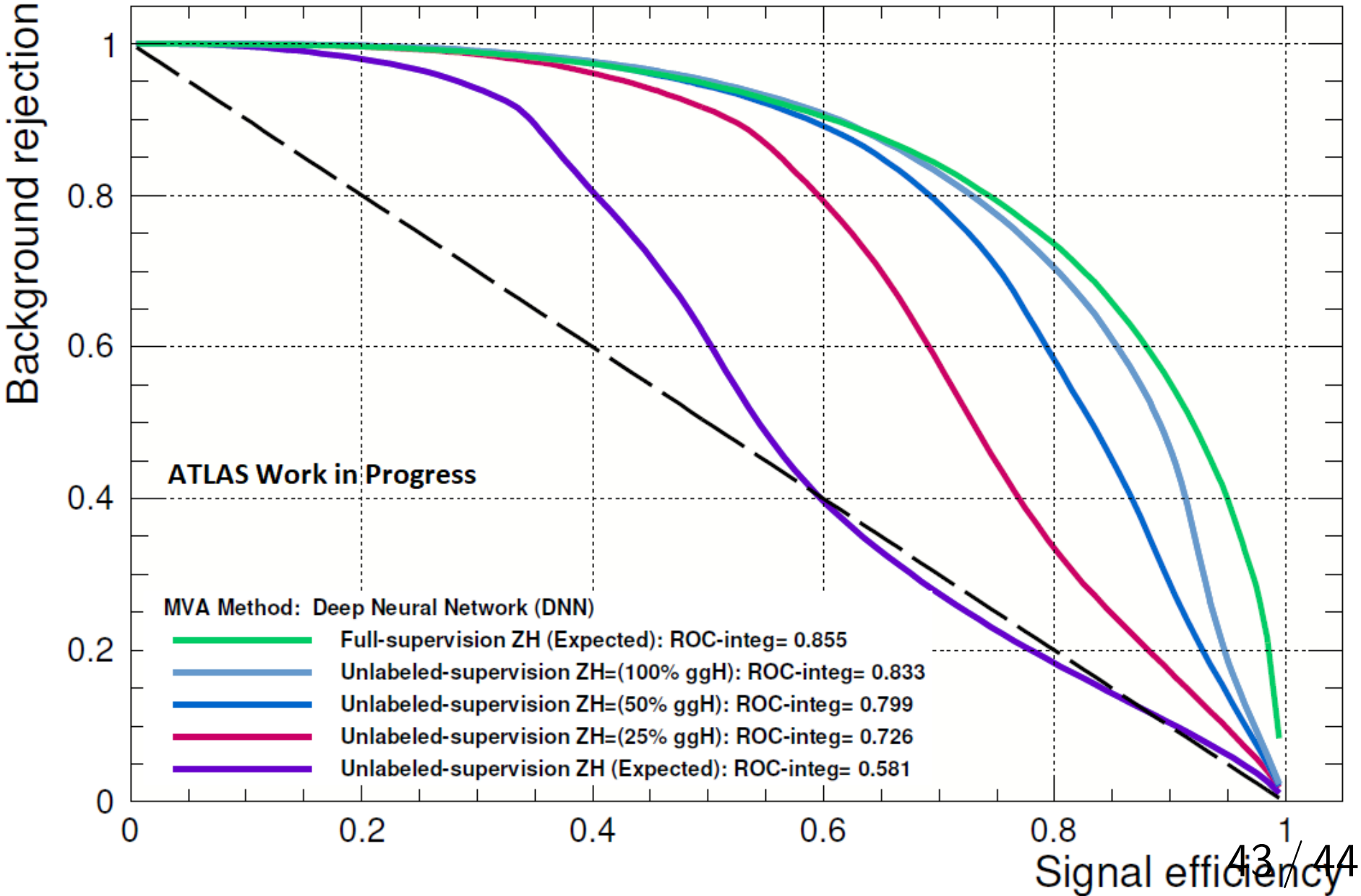



***Full-supervised***

# **Unlabeled-Supervision:** Statistical effect on VBF



# **Unlabeled-Supervision:**Statistical effect on ZH



# Conclusion

- There is good separation power for both weak supervision and supervised learning
- Weak supervision shows a good improvement for both VBHF and ZH when you increase the statistics
- The weak supervision can be applied to the real data since we do not know what is signal and what is background