Optimisation of VBF tagging for Higgs $\rightarrow \gamma\gamma$

Rob JB Cantrill
Robert.Cantrill@cern.ch
Pedro Teixeira-Dias
Ricardo Goncalo

IoP Conference 2012:
Queen Mary, University of London

Search for a Standard Model Higgs Boson

- Analysis – Standard Model Higgs Boson.
- Currently working with an ATLAS Higgs sub-group analysing 2011 data.
- But today – presenting a study I've completed on the 2010 data.
- The Higgs Mechanism is believed to give particles their masses through Electroweak symmetry breaking.
- If a Standard Model Higgs boson exists and is light it can decay into two photons, which we can search for using the ATLAS detector.

Higgs Decay Modes

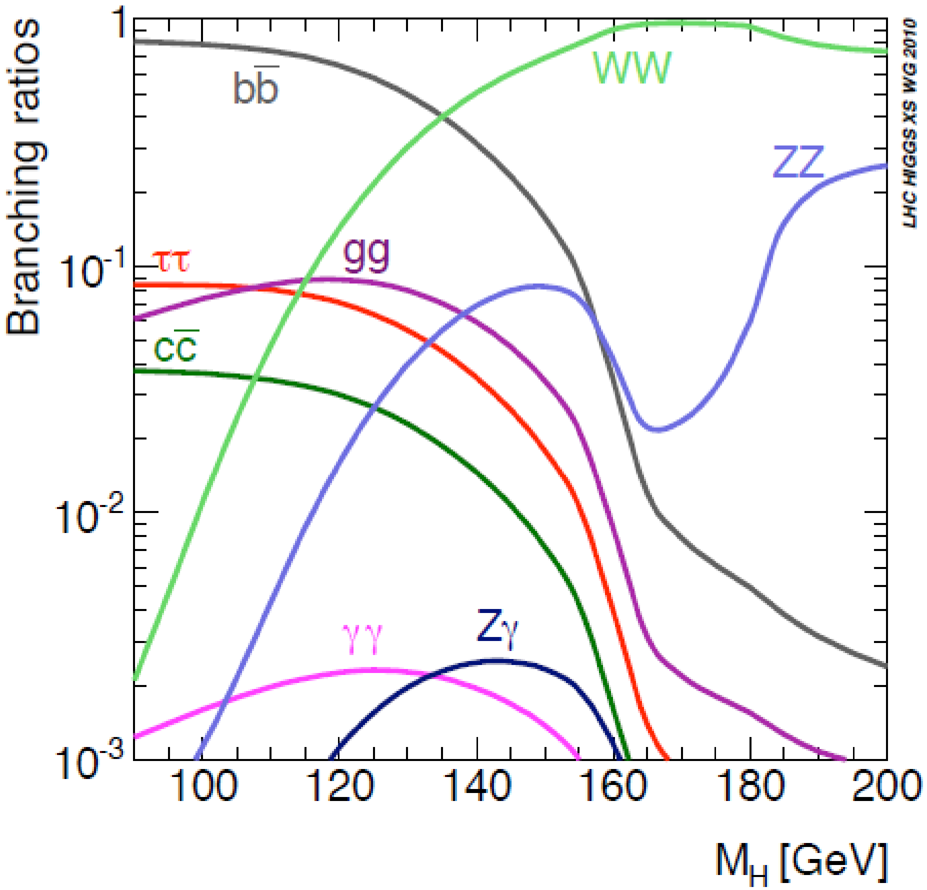


Fig. 35: SM Higgs branching ratios as a function of the Higgs-boson mass.

arXiv:1101.0593

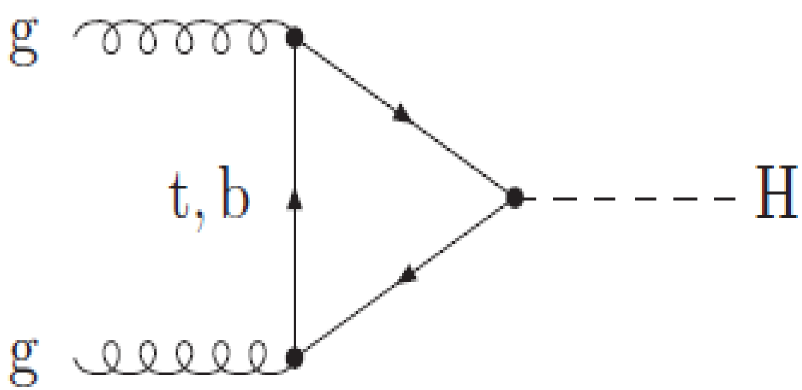
- Feasible chance of picking up $H \rightarrow \gamma\gamma$ events.
- Branching fraction to $\gamma\gamma$ is small compared to b-quark anti b-quark pair.
- Many LHC collision can produce b quark jets.
- Whereas a diphoton event is less likely.
- Background: for comparison

$\sigma_{\gamma\gamma} \sim 100$pb arXiv:1109.5141

$\sigma_{bb} \sim 10000$pb ATLAS-CONF-2011-056

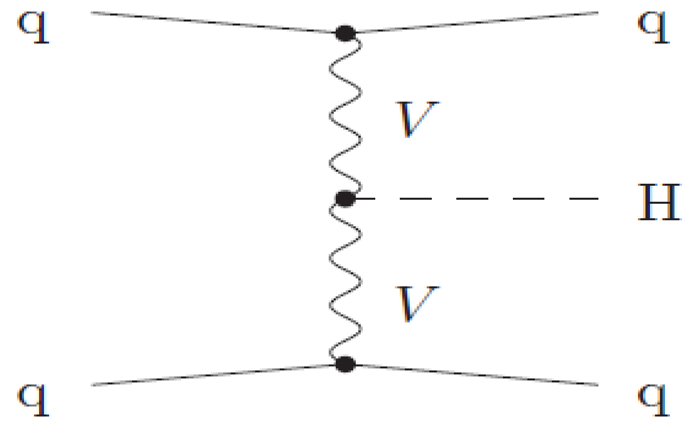
Higgs Production Mechanisms

σ=16.6pb (mH = 120GeV/c2)



arXiv:1101.0593

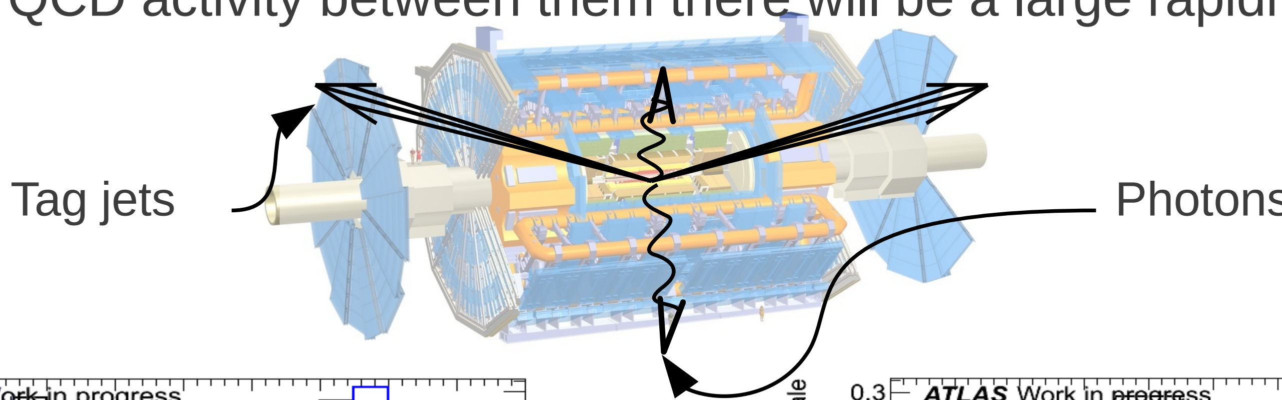
σ=1.3pb (mH = 120GeV/c2)

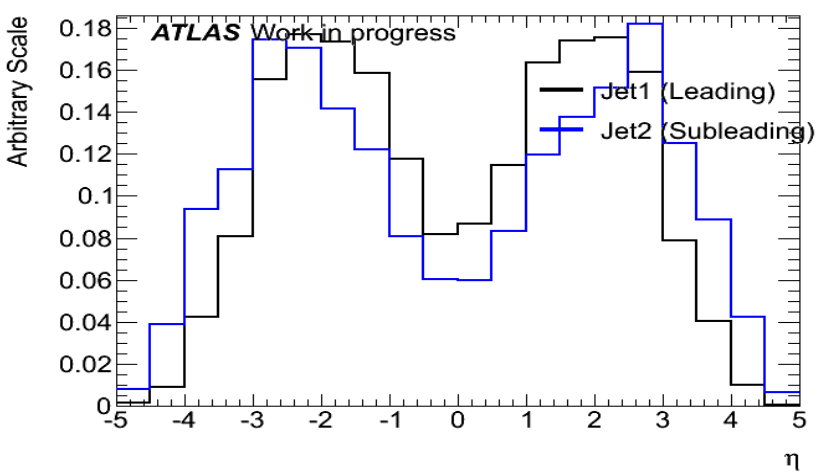


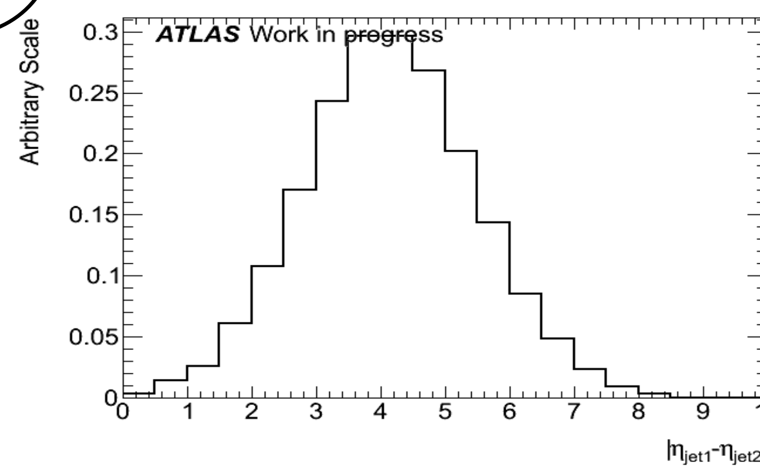
- Most common production is gluon-gluon fusion (left).
- Vector Boson Fusion (VBF) is the second most common mechanism (right). Where two Ws or Zs are radiated from the proton which fuse.
- VBF is the focus for this study.

VBF Jets

- Although cross section is ~ 10 times smaller than gg fusion, we can get good signal to background separation because of the two distinctive tag jets in the VBF production.
- The tag jets are the remnants of the protons so they show up two forward jets in the detector with large transverse momenta (p_T).
- Due to no QCD activity between them there will be a large rapidity gap.







Original proposed event selection

- The event selection was originally proposed in the ATLAS CSC notes when the LHC planned to run at a centre of mass energy, $\sqrt{s}$, of 14TeV.
- Optimisation was done with Monte Carlo simulation, for a VBF Higgs Boson signal with mass 120 GeV/c^2.

Original proposed event selection

ATLAS Work in progress

Cuts	Event Selection Efficiency at 14 TeV	Event Selection Efficiency at 7 TeV
1. Two leading photons	not stated	52.4%
2. $p_{T,\gamma 1} > 50$ GeV/c $p_{T,\gamma 2} > 25$ GeV/c	37.0%	41.9%
3. $p_{T,jet1} > 40$ GeV/c $p_{T,jet2} > 20$ GeV/c	not stated	67.8%
4. $\eta_{jet1} . \eta_{jet2} < 0$	not stated	56.6%
5. $\lvert \eta_{jet1} - \eta_{jet2} \rvert > 3.6$	40.8%	36.5%
6. Two leading photons between tag jets	39.5%	34.6%
7. $M_{jet1,jet2} > 500$ GeV/c^2	32.7%	24.1%
8. Veto third jet $p_T > 20$ GeV/c and $\lvert \eta \rvert < 3.2$	29.2%	21.2%

- Loss of signal efficiency with $\sqrt{s}$ = 7 TeV.
- Efficiencies for cuts 3-8 are relative to the number of events after cut 2 to demonstrate loss in signal efficiency for the jet selection.
- Motivation for re-optimisation of the tag jet selection cuts.

Data background sample

- Insufficient MC to simulate VBF background
- Use data with jet events as background sample.
- 0.06 VBF H$\rightarrow \gamma\gamma$ events in 21 pb^{-1} of the 2010 data that was used. No danger of signal contamination.
- Advantage – Nature 'knows' the correct cross sections.
- Tight $\gamma\gamma$ selection would leave to few events to study, so we decided to loosen the cuts on $\gamma\gamma$.
- Only interested in optimising the tag jet selection.

Event Selection

ATLAS Work in progress

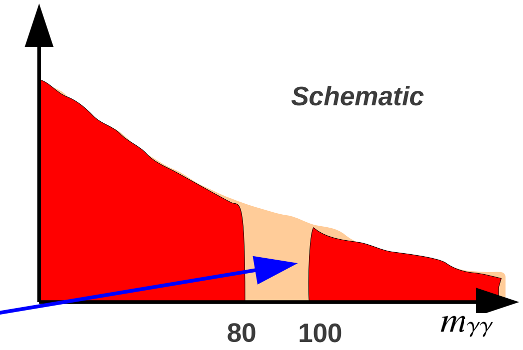
Event Selection		Events After Cuts	
		Signal	Background
i.	Total	29978	257929
ii.	Reject events which contain bad quality jets [14].	29875	254400
iii.	Egamma good runs list.	29875	238476
iv.	At least one primary vertex with at least three tracks.	28871	238455
v.	Events with 2 loose photons outside the crack, which aren't in dead regions of the calorimeter. Failing this, events must contain at least 2 loose electrons, outside the crack, which aren't in dead regions of the calorimeter with a $p_T > 25$ GeV/c.	20049	12748
vi.	The two leading p_T jets have $p_T > 20$ GeV/c.	13509	2085
vii.	The leading jets are in opposite hemispheres of the detector is not necessary to re-optimise this cut.	11002	882

- We insist that there is $\gamma\gamma$ or *ee* in each event.
- Jets that mimic VBF jets recoil off the hard scatter.
- We allow *ee* events to gain more statistics
- Z + jet events may not be suitable to use for background that will mimic the VBF jets.
- We must check by analysing the jet distributions.

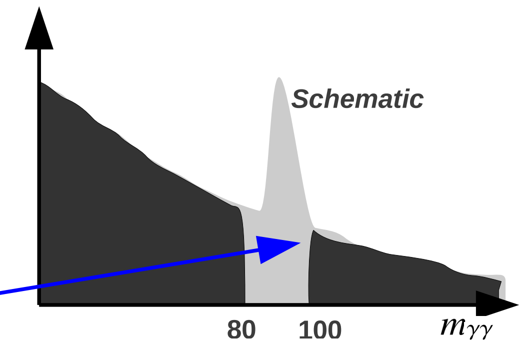
Check on *ee* + jets

- Test four scenarios

  (1) $\gamma\gamma$ events.



  (2) $\gamma\gamma$ events where invariant mass, $m_{\gamma\gamma} < 100\ \mathrm{GeV}/c^2$ and $m_{\gamma\gamma} > 80\ \mathrm{GeV}/c^2$ are removed.

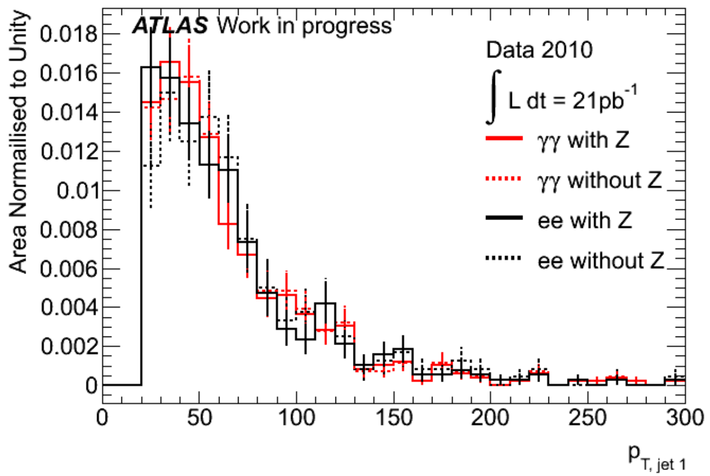


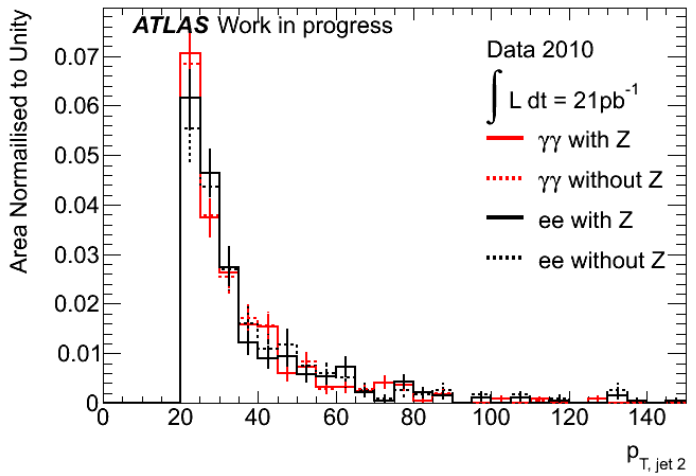
  (3) *ee* events.

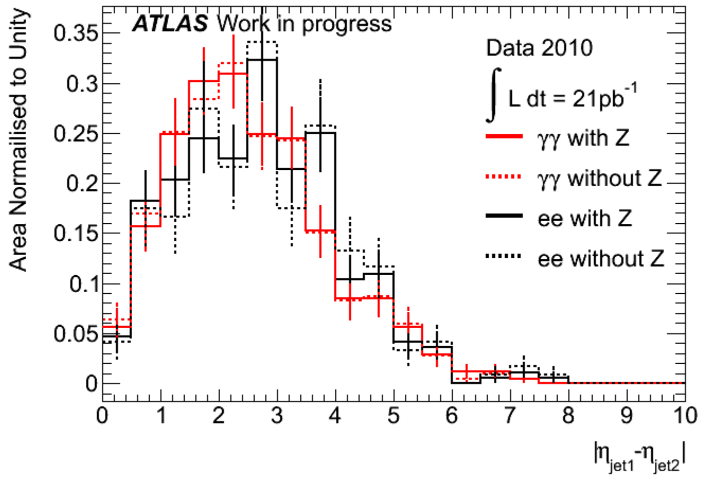
  (4) *ee* events where invariant mass, $m_{ee} < 100\ \mathrm{GeV}/c^2$ and $m_{ee} > 80\ \mathrm{GeV}/c^2$ are removed.

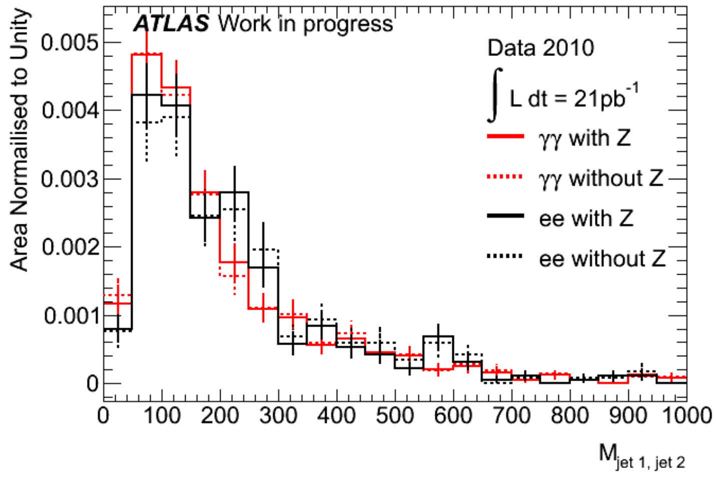
- This will identify any effect of Z + jets on the jet distributions.

Jet distributions

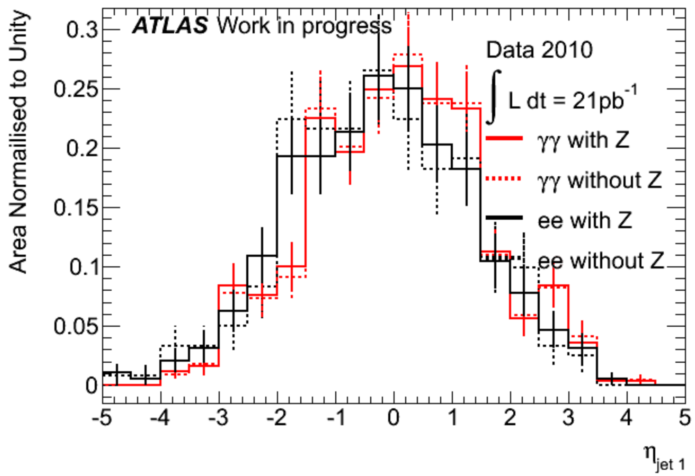


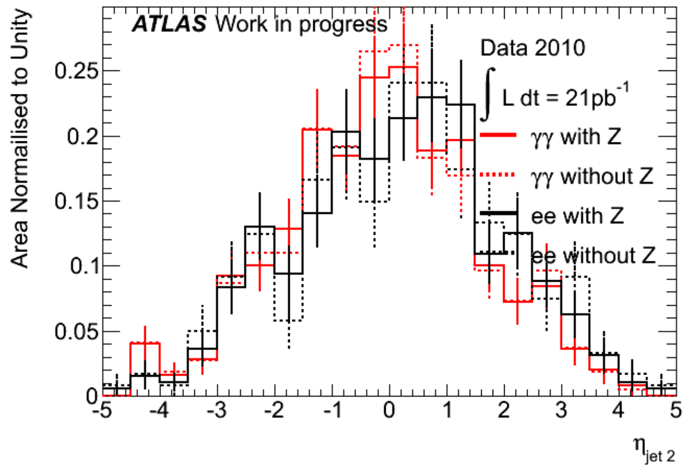






Jet distributions



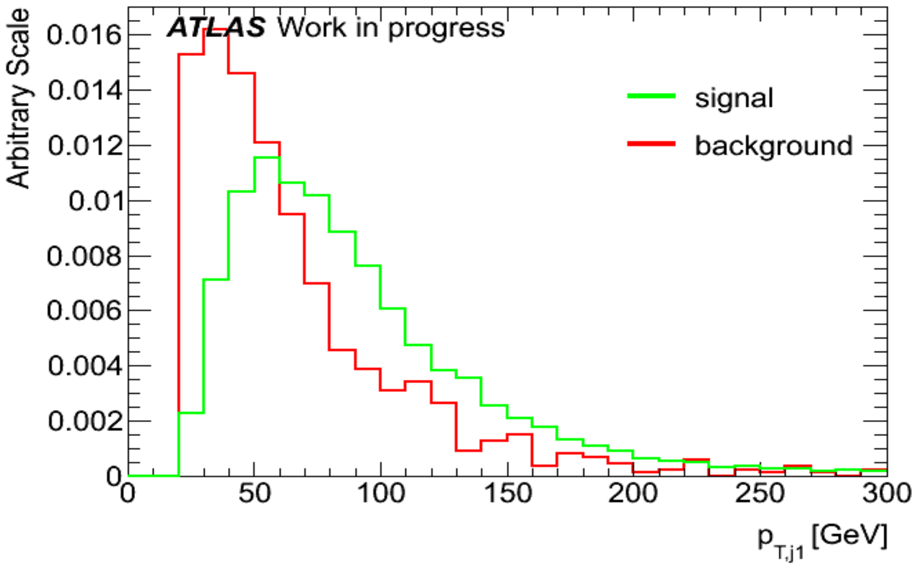


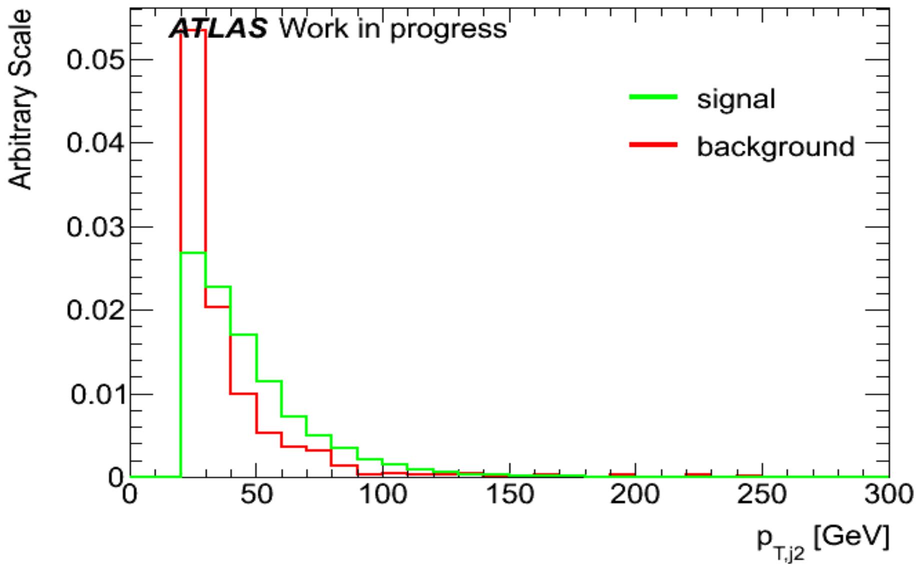
- There was no obvious difference between the four separate scenarios.
- Decided to use *ee* events in addition to $\gamma\gamma$ events.

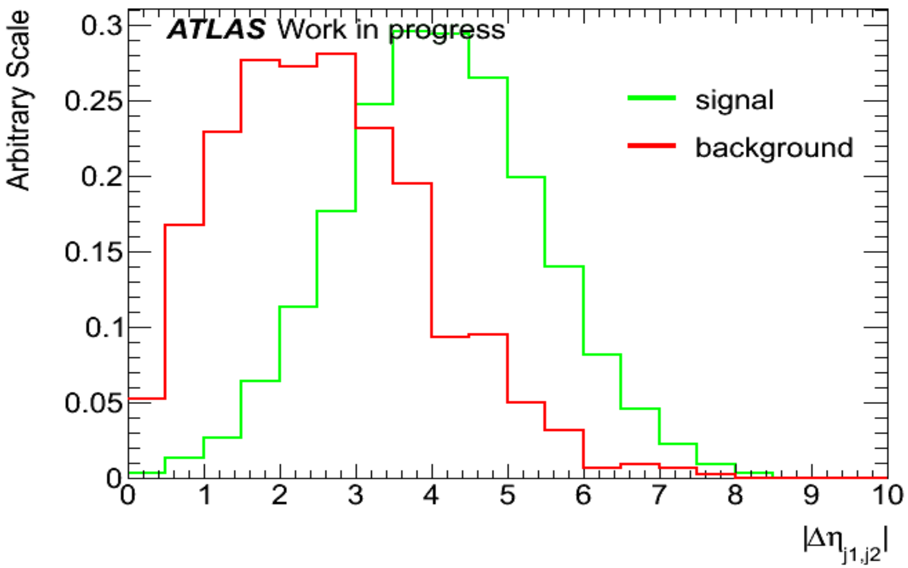
Optimisation

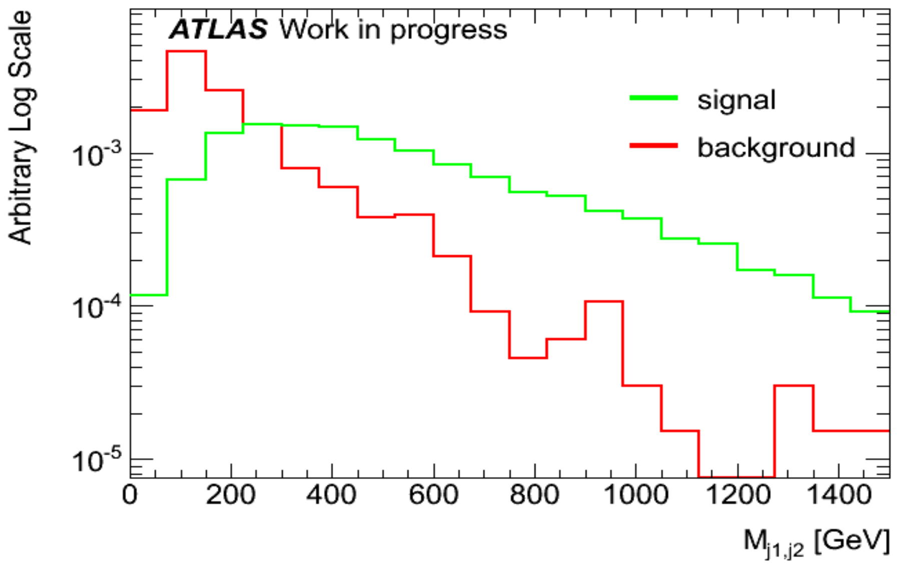
- Decided to optimise the following cuts.
  - p_T leading jet ($p_T$jet 1).
  - p_T subleading jet (p_T jet 2).
  - Pseudorapidity difference between the jets ($|\eta_{jet1} - \eta_{jet2}|$)
  - Invariant mass of the two jets ($M_{jet1,jet2}$)
- Used a multivariate analysis tool (TMVA) to find the optimum cuts on the above 4 variables.
- Both signal and background were each divided into two independent samples for the purpose of testing and training.

Variables to be optimised

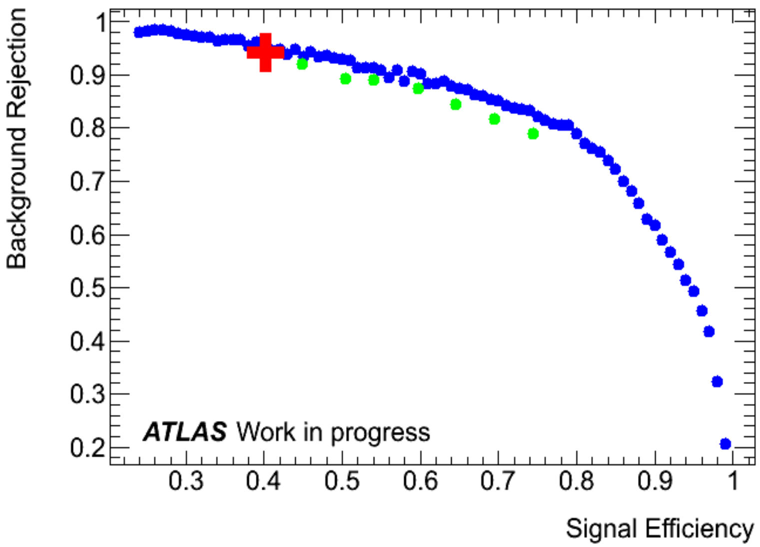


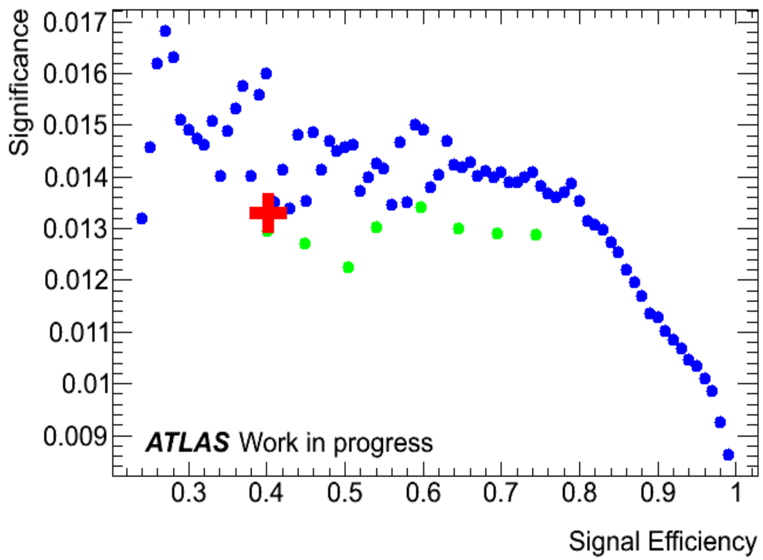






Results





- Left – signal efficiency vs background rejection.
  - Optimal position is in the far top right hand corner of the curve.
  - Each point represents a set of cuts.
  - Blue points are from the training sample. Green points are same cuts applied to the test sample.
  - Nominal cuts (red cross) can be improved.
- Right – expected significance vs signal efficiency
  - We see that the significance does not alter by a noticeable amount if we increased the signal efficiency.
  - expected significance, Z for a Poisson counting experiment.

$$Z = \sqrt{2((s+b)ln(1+s/b)-s)}$$

arXiv:1007.1727

ATLAS Work in progress

Results

Efficiencies		Cut Variables			
Signal [%]	Background [%]	$p_{T,jet1}$ [GeV/c]	$p_{T,jet2}$ [GeV/c]	$\lvert\eta_{jet1} - \eta_{jet2}\rvert$	$M_{jet1,jet2}$ [GeV/c^2]
Improved Cuts					
54.8	11.1	30	20	2.5	425
60.6	13.5	45	20	3.1	300
67.3	17.4	25	20	3.1	300
69.3	18.1	30	20	3.1	275
76.8	23.2	30	20	1.8	275
Nominal					
40.3	5.8	40	20	3.6	500

- A few suggestion of cuts which yield a high signal efficiency of jet selection.
- Efficiencies are relative to the number of *ee* or $\gamma\gamma$ events that have two jets in opposite hemispheres.
- Suggestions that the cuts should be looser relative to those at $\sqrt{s}$ = 14 TeV.
- p_T cut on the second jet is not tightened further due to low discrimination power of that variable.

Outlook

- Applying experience with jet studies with the current ATLAS Higgs search for $\gamma\gamma$ decay.
- Currently looking into systematic uncertainty of the jet p_T to see how big an effect this has on the jet selection.
- VBF looks set to be a useful contribution to the Higgs search.
- Re-optimisation should definitely be considered to maximise our signal sensitivity.