Optimizing Cuts in HWW decay channel

Noah Reidy

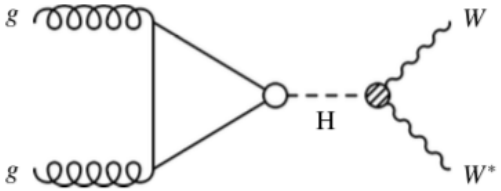
University of Oklahoma

July 29, 2019

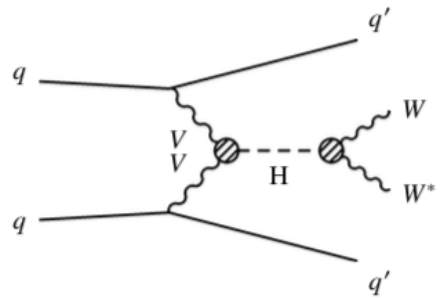
Advisor: Dr.Strauss

HWW Research Overview

- Higgs decay channel to two W bosons
- Second most likely was for the higgs to decay
- Search for the higgs through two production modes
  - gluon gluon fusion (ggF)
  - vector boson fusion (VBF)



(a) ggF production

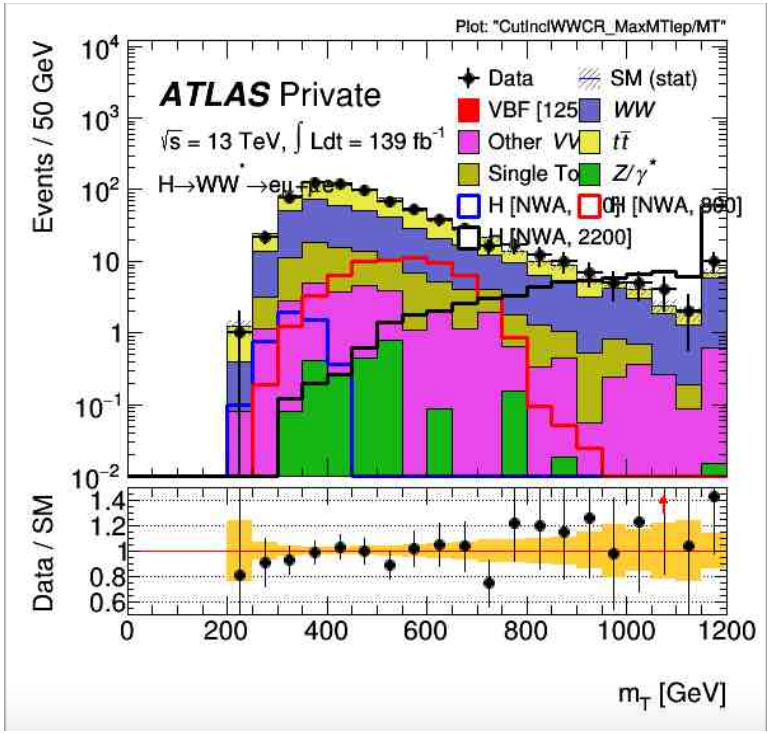


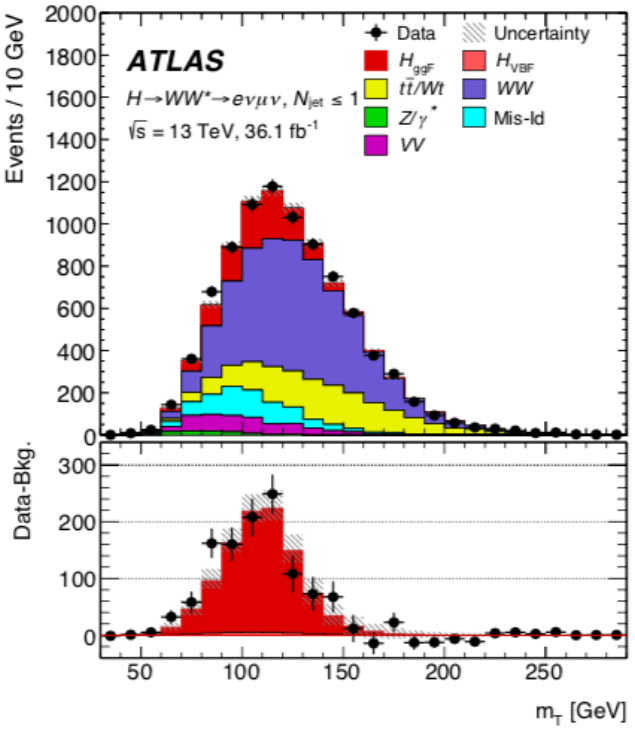
(b) VBF production

Overview

- Optimize Cuts for potential high mass Higgs particles
- Look at two different scans in the 1 and 2 jet regions
  - VBF and GGF as signal with the whole background folder
  - VBF as signal and ggf as background
- The first looks at where the vbf is compared with ggf in the regions
- The second looks at optimal cuts in the regions for different masses

Histogram Examples





Previous Work and Goal

- A Graduate student at CERN has done work probing different mass regions on high mass signal using Grid Scanner
- Found that the cuts in those regions had room for improvement
- Expand his work to 1 and 2 Jet specific regions
- Calculate significance for cut in different mass regions

Observables

- Split the two regions into two different scans:
  - 1 Jet: DPhill, min(DEtalj), and JetEta_1J
  - 2 Jet: Use DPhill, Mjj, and DYjj
- DPhill was highly correlated with 1 and 2 jet in Boosted Decision Trees
- The other variables were ones we cut on previously

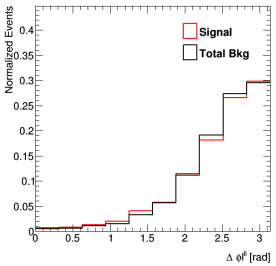
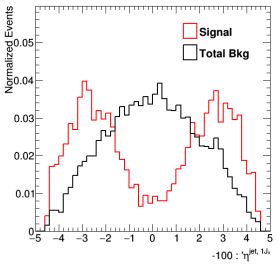
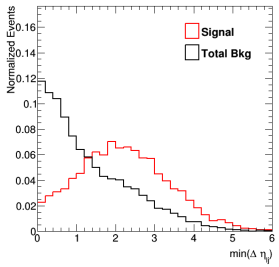
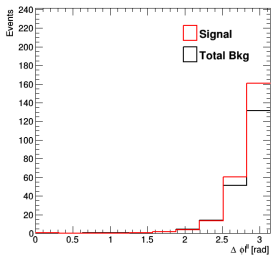
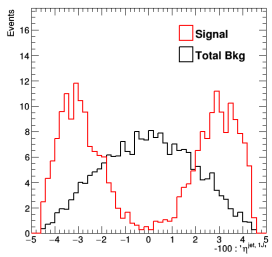
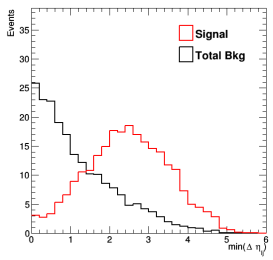
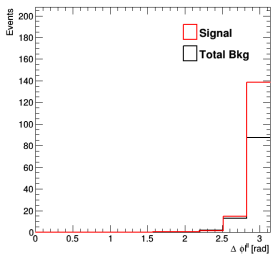
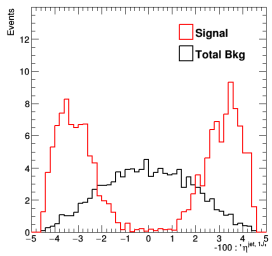
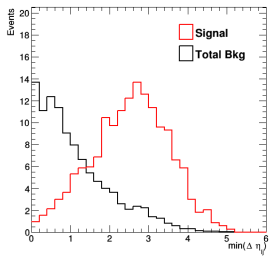
Previous Cuts

SR_{ggF}	SR_{VBF1J}	SR_{VBF2J}
	Common selections	
	$N_{b\text{-tag}} = 0$	
	$\lvert\Delta\eta_{\ell\ell}\rvert < 1.8$	
	$m_{\ell\ell} > 55\,\text{GeV}$	
	$p_{\text{T}}^{\ell,\text{lead}} > 45\,\text{GeV}$	
	$p_{\text{T}}^{\ell,\text{sublead}} > 30\,\text{GeV}$	
	veto if $p_{\text{T}}^{\ell,\text{other}} > 15\,\text{GeV}$	
	$\max(m_{\text{T}}^{W}) > 50\,\text{GeV}$	
ggF phase space	VBF1J phase space	VBF2J phase space
Inclusive in N_{jet} but excluding VBF1J and VBF2J phase	$N_{\text{jet}} = 1$ and $\lvert\eta_j\rvert > 2.4$, $\min(\lvert\Delta\eta_{j\ell}\rvert) > 1.75$	$N_{\text{jet}} \geq 2$ and $m_{jj} > 500\,\text{GeV}$, $\lvert\Delta y_{jj}\rvert > 4$

Seperation Scan: VBF as Signal; GGF as signal

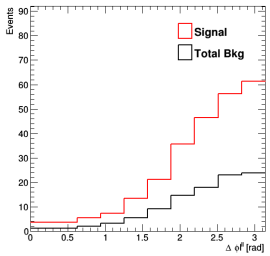
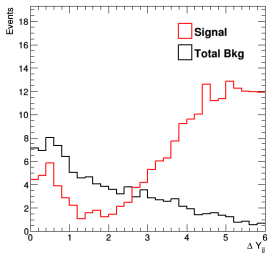
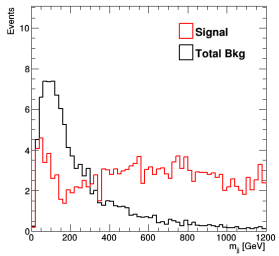
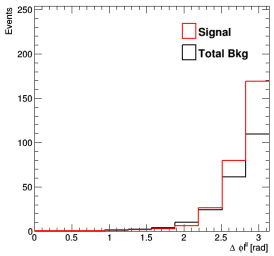
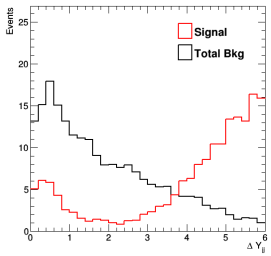
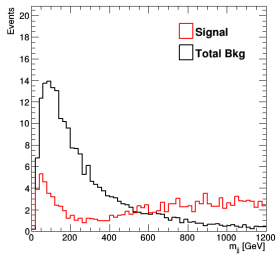
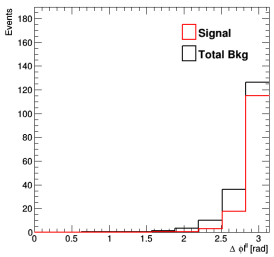
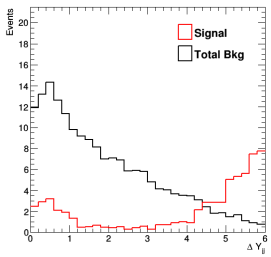
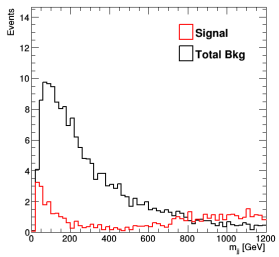
1 Jet

VBF signal, GGF Background

	DPhill	JetEta_1J	Min(DEtalj)
400mh			
1000mh			
2200mh			
Cut	NA	2.4	1.75

2 Jet

VBF signal, GGF Background

	DPhill	DYjj	Mjj
400mh			
1000 mh			
2200mh			
Cuts	NA	4	500

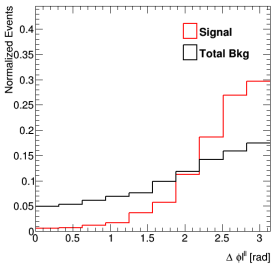
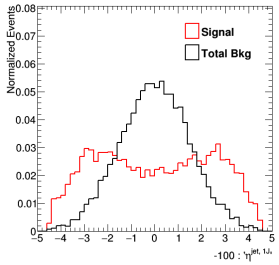
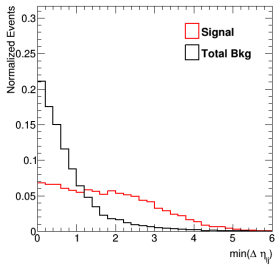
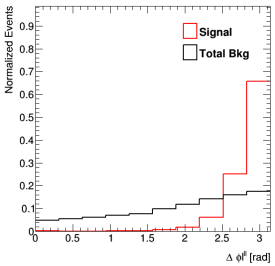
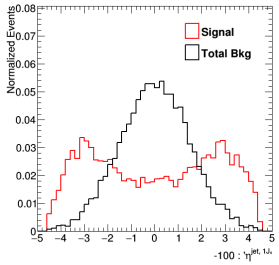
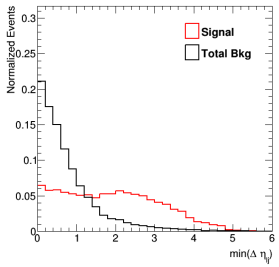
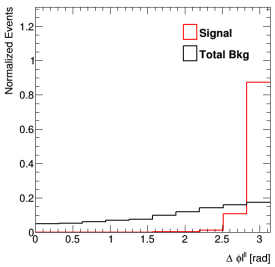
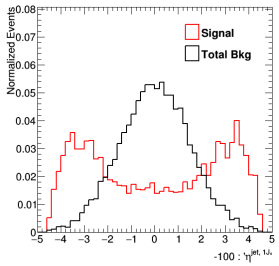
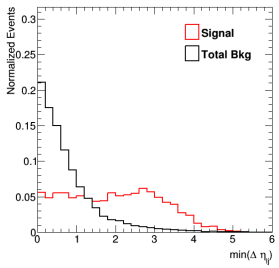
Analysis

- For both the 1 and 2 jet cases the separation is found in all the variables excluding DPhill
- The cuts gridscanner output confirms that the seperation cuts we have are doing a good job

VBF+GGF as signal; Total Background

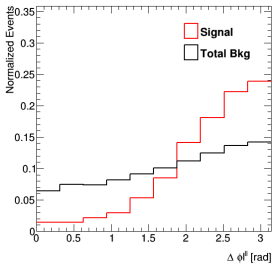
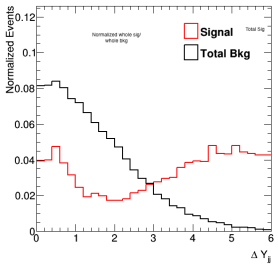
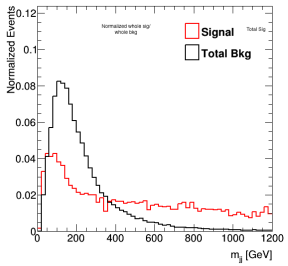
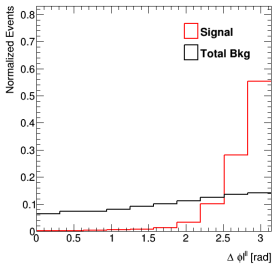
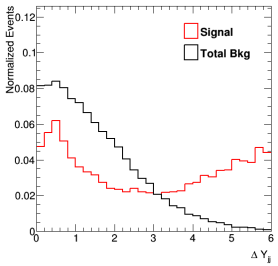
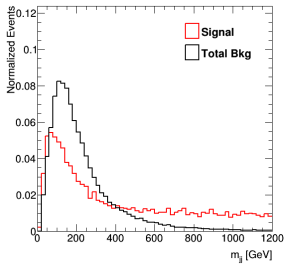
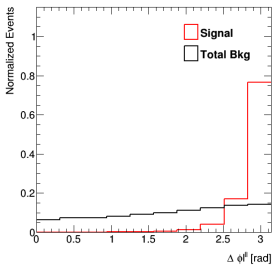
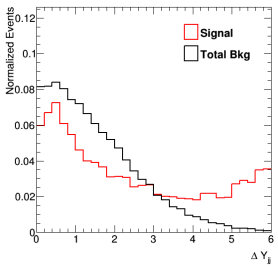
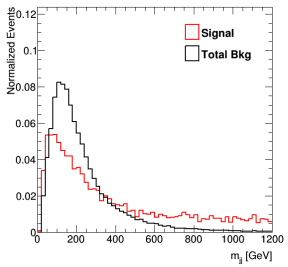
1 Jet

VBF+GGF signal, total Background

	DPhill	JetEta_1J	Min(DEtalj)
400mh			
1000 mh			
2200mh			
Cut	NA	2.4	1.75

2 Jet

VBF+GGF signal, total Background

	DPhill	DYjj	Mjj
400mh			
1000 mh			
2200mh			
Cuts	NA	4	500

1 Jet

- Top 3 Significance Points

Mass	DPhill	Jet Eta 1J	Min (DEatlj)	Signal	Bkg	$\frac{S}{\sqrt{s+b}}$
400mh	1.256	1.4	1.6	78.726	2181.99	1.655
	1.57	1.4	1.6	76.5782	2062.5	1.655
	1.884	1.4	1.6	72.7546	1882.72	1.645
1000mh	2.512	1.8	1.8	99.2386	869.378	3.188
	2.512	1.8	1.6	107.025	1021.49	3.185
	2.512	1.4	1.6	112.636	1143.47	3.178
2200mh	2.826	1.8	1.8	64.0784	492.309	2.716
	2.826	1.6	1.8	65.2822	518.594	2.701
	2.826	2.0	1.8	63.1343	486.553	2.692

2 Jet

- Top 3 Significance Points

Mass	DPhill	DYjj	Mjj	Signal	Bkg	$\frac{S}{\sqrt{s+b}}$
400mh	1.57	5.6	1080	68.6752	391.27	3.202
	1.57	5.6	1060	69.3697	405.015	3.184
	1.57	5.6	1120	67.0018	375.828	3.183
1000mh	2.512	5.4	1200	122.774	143.914	7.518
	2.512	5.6	1060	124.796	152.345	7.496
	2.512	5.6	1200	116.499	125.16	7.49
2200mh	2.826	6	1060	62.7709	42.8455	6.107
	2.826	6	1080	62.5349	42.7821	6.093
	2.826	6	1100	62.1799	42.8509	6.067

Implications

- There seemed to be a flaw in the way that we utilize Gridscanner
  - DYjj and Mjj cuts cut out too much or all of the region
  - Mjj does not make much sense because of high it cuts out
- The standard cuts seem to match up better with the plots that GridScanner puts out
- We determined that while gridscanner gives value that are numerically correct they do not make a lot of sense when applied
- The given cuts cut out nearly all statistics
  - For example 20 signal samples with 100 background is better than 1 signal sample with 0 background

Further Work

- The program scans over the whole mass region and gives out the cuts for the max significance
- Isolate different mass regions prior to running gridscanner in hopes of getting better optimized cuts