Search for associated Higgs production in 8 TeV pp collisions with the 20.3 fb^{-1} of ATLAS data in the $W(\mathrm{H} \to WW^* \to lvqq) \to lvlvqq$

Joe Taenzer

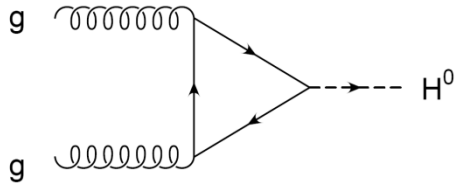
CAP Congress 2014

June 17th, 2014

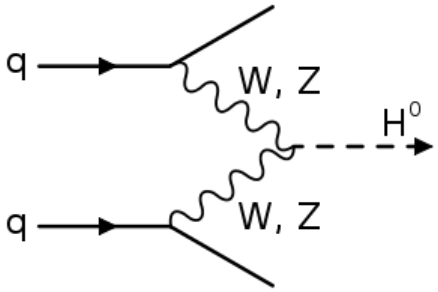
ATLAS EXPERIMENT

UNIVERSITY OF TORONTO

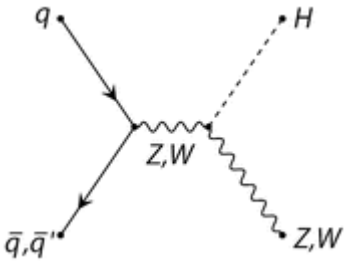
ATLAS Higgs Searches



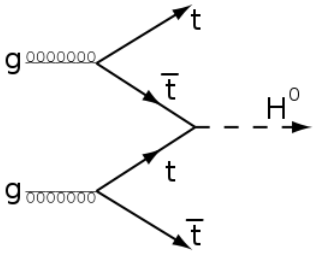
gluon-gluon fusion
σ ~ 19.5 pb



Vector boson fusion
σ ~ 1.6 pb

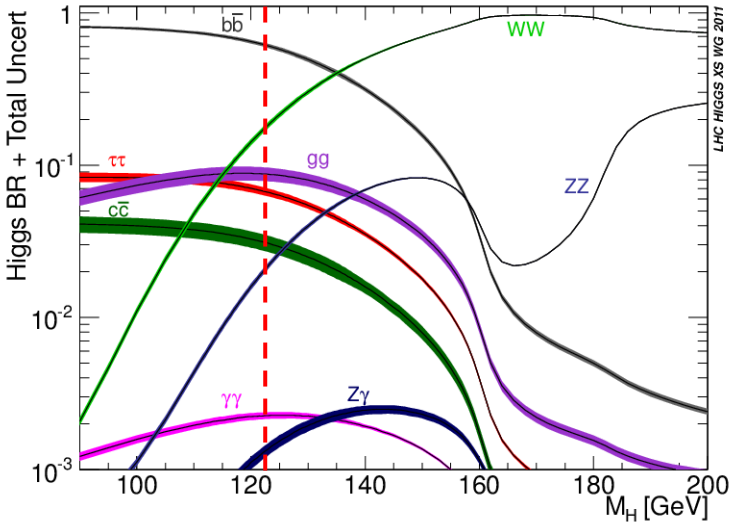


W/Z bremsstrahlung
σ ~ 0.7 pb / 0.4 pb



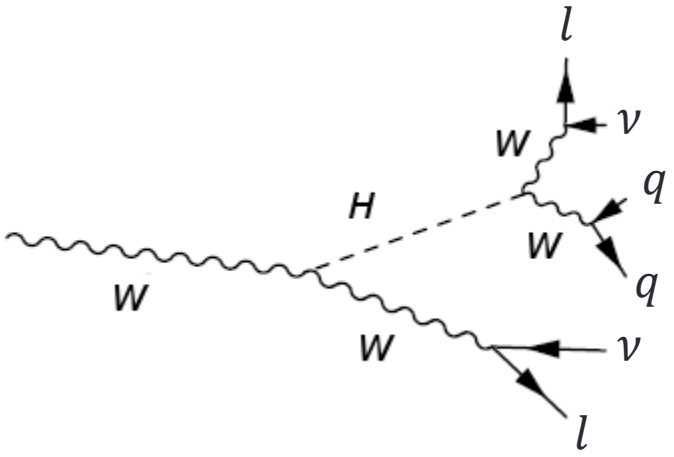
$t\bar{t}$ fusion
σ ~ 0.7 pb / 0.4 pb

- Higgs boson candidate observed by both ATLAS and CMS in the ggF and VBF production modes and WW*, $\gamma\gamma$, ZZ*, decay modes
- Measurements of spin, CP, cross-section indicate the Higgs candidate is compatible with the standard model Higgs boson with m_H ~ 125 GeV
- **Associated production searches can add increased precision in measurements of WH/ZH couplings, Higgs spin and CP.**
- **WH → WW is sensitive to pure Higgs to W coupling**



Analysis Strategy

- Process: $W(H \rightarrow WW^*) \rightarrow lvlvqq$
  Associated W boson and one W boson from the Higgs decay leptonically.



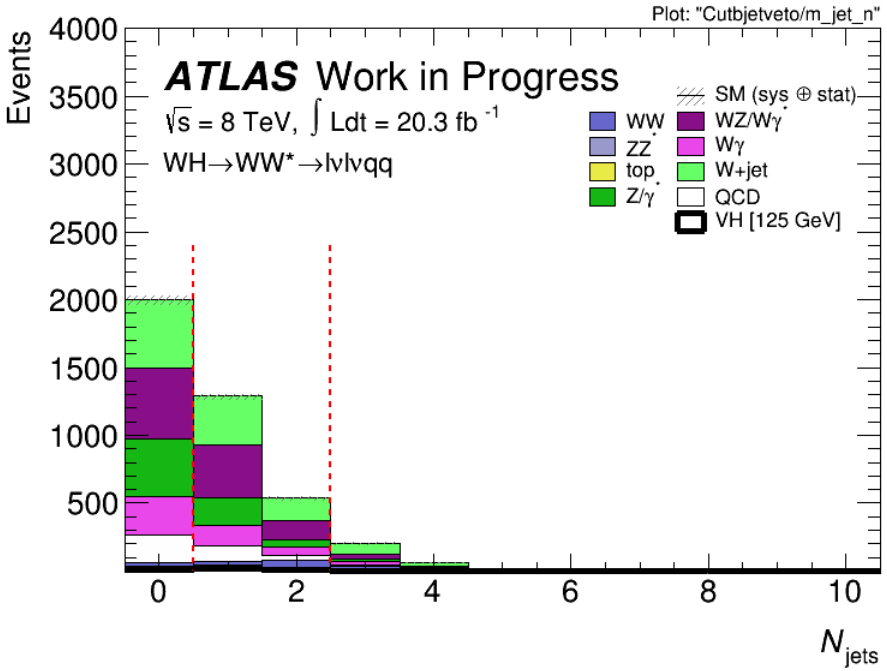
Analysis is still blinded, expected results are from MC simulated events.

- Final state signature in the detector is two leptons, missing energy (neutrinos), jets.
- **Require both leptons have the same sign** to suppress SM background processes such as Drell-Yan (DY), W^+W^-, $t\bar{t}$, etc.
- Require high missing transverse energy (${E_T}^{miss}$) to subtract backgrounds with no natural source of missing energy.
- Allow events with only one jet, in case the virtual W boson decays hadronically, as the decay products of the virtual W boson are likely to be soft and may be lost in reconstruction.

Pre-selection and backgrounds

Preselection removes events that don't match the signal final state, e.g. Matching of final state objects to primary vertex, exactly two leptons, lepton-lepton and lepton-jet overlap removal, remove events containing b-tagged jets, lepton transverse moment (p_T) and η acceptance.

Full list of pre-selection cuts in the backup slides!

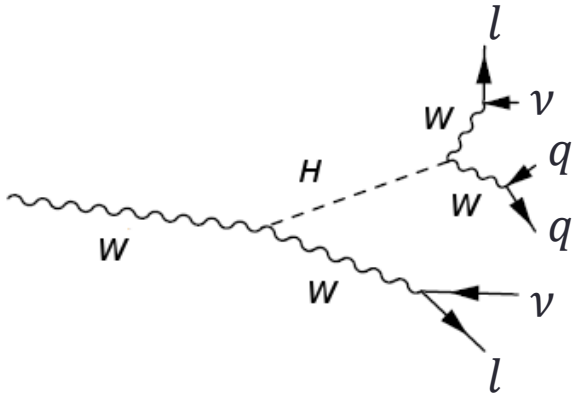


Major backgrounds after preselection based on Monte Carlo (MC) simulated events:

WZ (27% of total background)
W+jets (28% of total background)
Wγ (13% of total of background)
Z+jets (17% of total background)

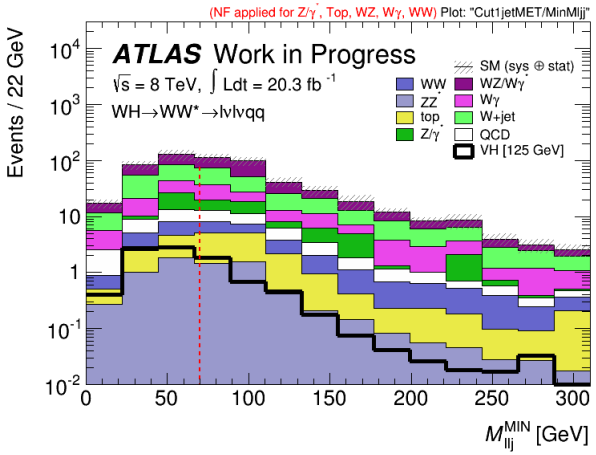
Other backgrounds (QCD dijet, WW, ZZ*, $t\bar{t}$, single top, etc.) make small contributions of order 1-5%.

Kinematic Variables



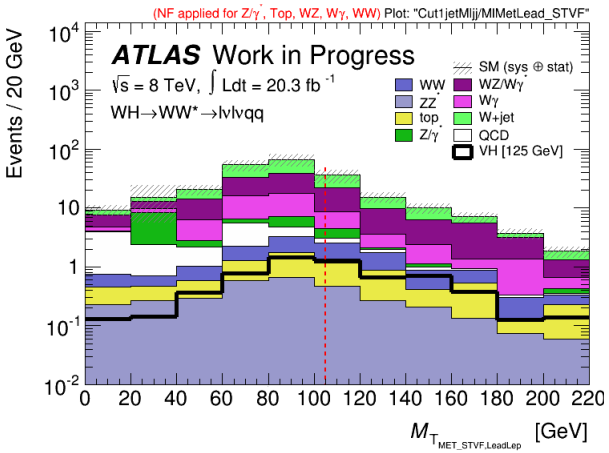
$M_{ljj}{}^{Min}/M_{lj}{}^{Min}$

Invariant mass of jets and a lepton. The lepton that minimizes the quantity is used (Higgs candidate lepton)



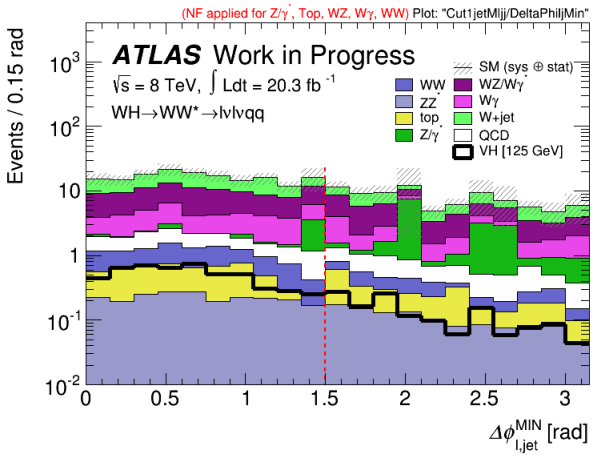
$M_T{}^{Lead}$

Transverse mass of missing energy and leading lepton.



$\Delta\varphi_{lj}{}^{Min}$

Opening angle between the Higgs candidate lepton and a jet. In events with more than one jet, the jet that minimizes the quantity is used.



Signal Region definition

- Signal region selection was optimized to obtain the best 95% CL_s limit for signal with m_H = 125 GeV.

1 jet SR	2 jet SR	Motivation
E_T^{miss} > 45 GeV	E_T^{miss} > 50 GeV	Further reduce Z, ZZ*, DY, QCD backgrounds
M_{lj}^{Min} < 70 GeV	M_{ljj}^{Min} < 115 GeV	Higgs mass discriminant
M_{ll} > 55 GeV	-	Required for orthogonality to other H → WW analyses
M_T^{Lead} > 105 GeV	-	Reduce W+jets background
$\Delta\varphi_{lj}^{Min}$ < 1.5		Neutral current discriminant
Sub-leading lepton pT > 20 GeV (μμ)		Reduce jet fakes backgrounds

Control Regions

- Use regions pure in one background process (or several similar processes) to assess data/MC agreement, and to derive a normalization factor (NF) to correct the cross-section.

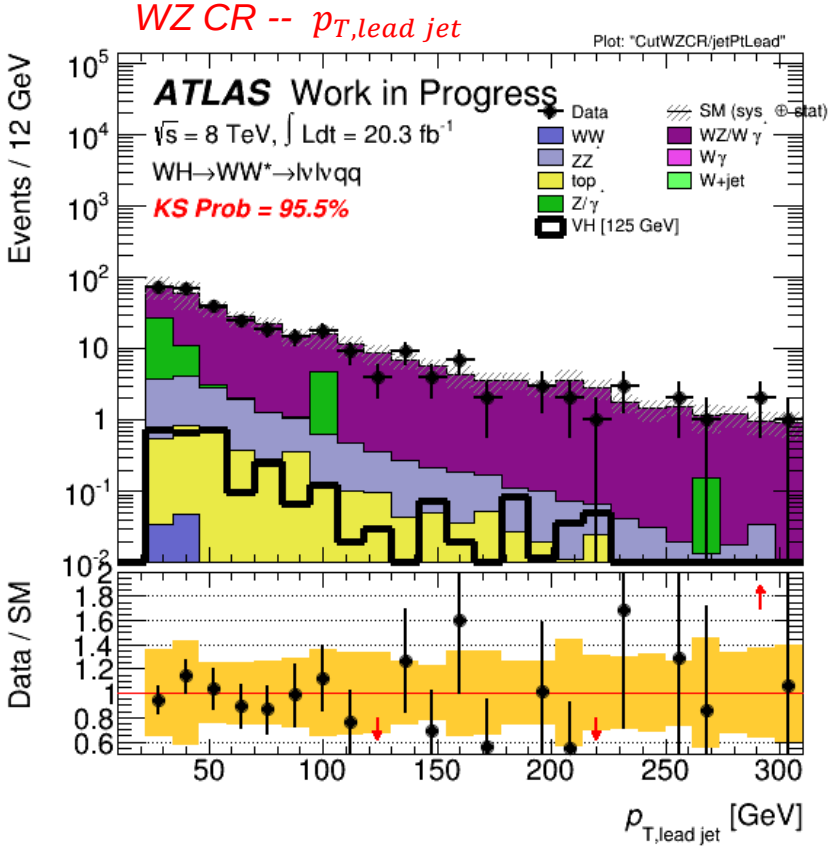
ATLAS Work in Progress	
CR	**NF**
WZ	0.97 ± 0.3 (stat+syst)
Wγ	1.07 ± 0.3 (stat+syst)
Z+jets	0.86 ± 0.1 (stat+syst)
W^+W^-	0.7 ± 0.4 (stat+syst)
Top	1.06 ± 0.2 (stat+syst)

NFs are obtained by simultaneously fitting of all 5 CRs, finding the normalization which best satisfies data/MC = 1 for all CRs.

WZ Control Region

Definition:

- Require exactly three leptons.
- Require one same flavour, opposite sign lepton pair with invariant mass within 15 GeV of the Z boson mass.
- Normal signal region cuts otherwise, stopping after the $E_T{}^{miss}$ cut stage.



ATLAS Work in Progress				
	WZ events	**Total Bkg**	**Data**	**Data/MC**
Before NFs	272 ± 4 ↓ x 0.97	320 ± 20 ↓ x All NFs	314	0.96 ± 0.07
After NFs	263 ± 4	310 ± 15	314	1.00 ± 0.07

Data/MC validation (1)

- Want to verify Data/MC agreement without unblinding the analysis.
- Establish sideband region by reversing an SR cut ($\Delta\varphi_{lj}{}^{Min}$ >= 1.5) and apply NFs derived from CRs.

ATLAS Work in Progress

Cut	Signal	Total Bkg.	Data	Data/MC
Pre-selection	21.9 ± 0.5	3380 ± 40	3463	1.03 ± 0.02
1 jet sideband				
Njets = 1	5.1 ± 0.3	750 ± 20	794	1.05 ± 0.05
$E_T{}^{miss}$ > 45 GeV	3.0 ± 0.2	310 ± 10	292	0.93 ± 0.06
$M_{lj}{}^{Min}$ < 70 GeV	1.4 ± 0.1	87 ± 8	70	0.8 ± 0.1
$M_{\ell\ell}$ > 55 GeV	0.8 ± 0.1	51 ± 8	38	0.7 ± 0.2
$M_T{}^{Lead}$ > 105 GeV	0.8 ± 0.1	28 ± 6	22	0.78 ± 0.25
2 jet sideband				
Njets = 2	2.9 ± 0.2	336 ± 8	377	1.12 ± 0.06
$E_T{}^{miss}$ > 50 GeV	1.7 ± 0.1	154 ± 4	157	1.02 ± 0.09
$M_{ljj}{}^{Min}$ < 115 GeV	0.38 ± 0.05	15 ± 1	10	0.68 ± 0.22

Uncertainties are statistics only

NFs applied here

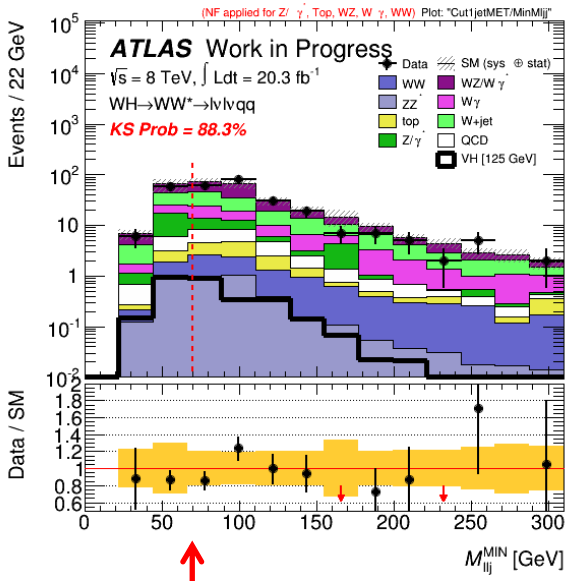
NFs applied here

Data/MC agreement is quite good if you consider approx. 15%-20% systematic uncertainty on the background estimate.

Data/MC validation (2)

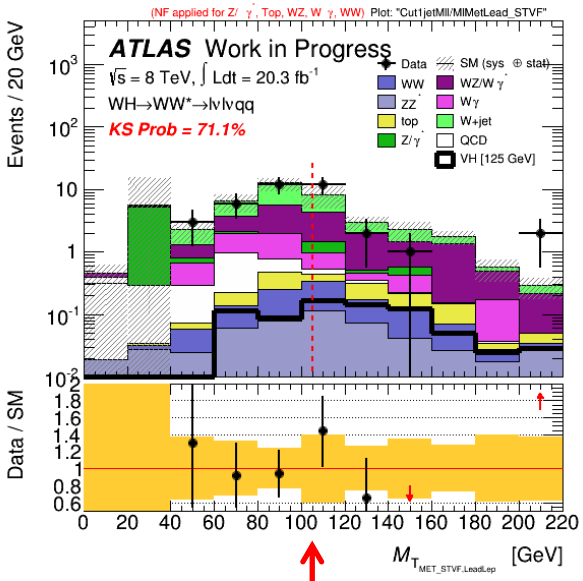
- Plots of main kinematic variables (1 jet sideband):

$M_{lj}{}^{Min}$ at $E_T{}^{miss}$ cut stage



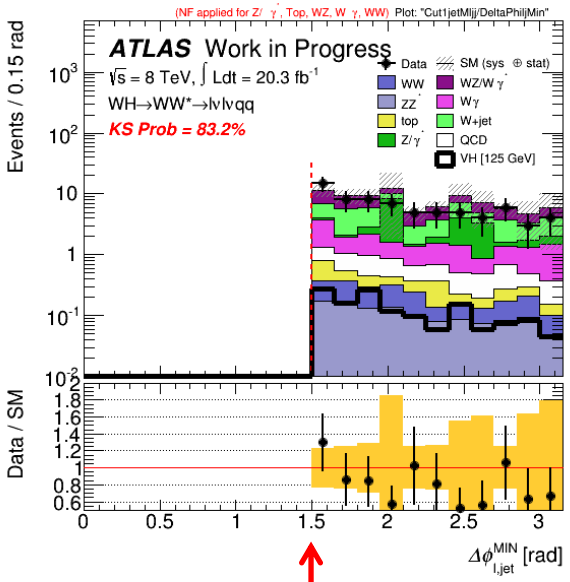
Cut: $M_{lj}{}^{Min} <$ 70 GeV

$M_T{}^{Lead}$ at $M_{lj}{}^{Min}$ cut stage



Cut: $M_T{}^{Lead} >$ 105 GeV

$\Delta\varphi_{lj}{}^{Min}$ at $M_{lj}{}^{Min}$ cut stage



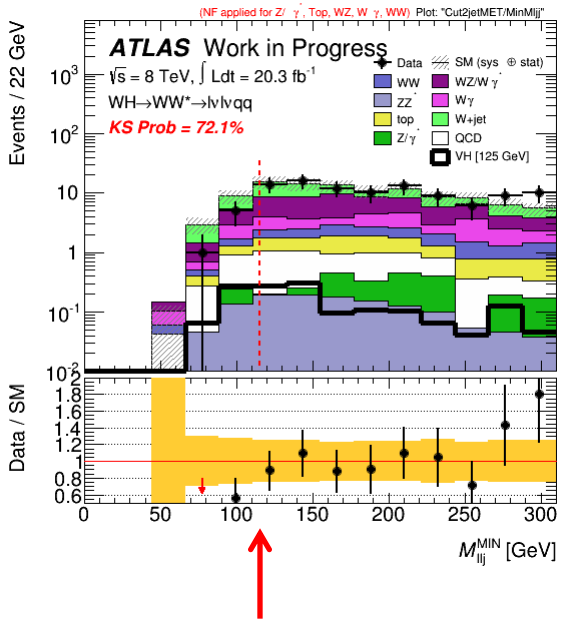
Cut: $\Delta\varphi_{lj}{}^{Min} \geq$ 1.5

Outside of a few bins, any data/MC disagreement is well covered by systematic + statistics uncertainties (yellow band in the ratio plots).

Data/MC validation (3)

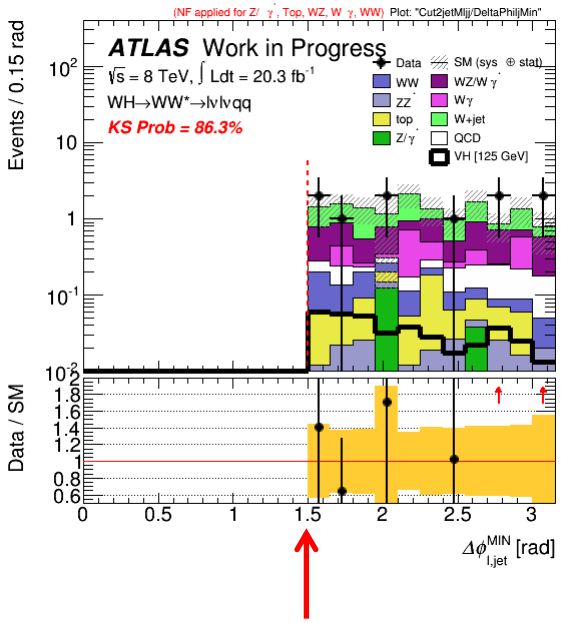
- Plots of main kinematic variables (2 jet sideband):

$M_{lj}{}^{Min}$ at $E_T{}^{miss}$ cut stage



Cut: $M_{lj}{}^{Min} < 115$ GeV

$\Delta\varphi_{lj}{}^{Min}$ at $M_{lj}{}^{Min}$ cut stage



Cut: $\Delta\varphi_{lj}{}^{Min} \geq 1.5$

Outside of a few bins, any data/MC disagreement is well covered by systematic + statistics uncertainties (yellow band in the ratio plots).

Systematic Uncertainties

Detector uncertainties

WZ theory uncertainties

ATLAS Work in Progress	
Systematic	**% of Total bkg in 1 jet SR (2 jet SR)**
Jet fake rate (inclusive)	9.7 (14)
Jet Energy Scale (inclusive)	3.1 (6.1)
Electron scale and efficiency	1.6 (1.4)
and many more small uncertainties (< 1%)	
WZ PDF	2.3 (1.8)
WZ QCD scale up	6.1 (5.1)
WZ QCD scale down	2.0 (4.1)

WH theory uncertainties

ATLAS Work in Progress	
Systematic	**% of WH signal in 1 jet SR (2 jet SR)**
WH PDF	2.3 (3.3)
WH QCD scale	1.3 (1.1

Expected Results

- Expected MC simulated events surviving our signal region selection:

ATLAS Work in Progress									
Signal (VH)	**WW**	**ZZ***	**WZ/Wγ***	**Wγ**	**top**	**Z+jets**	**W+jets**	**QCD dijet**	**Total bkg.**
1 jet signal region [m_H = 125 GeV]									
1.9±0.2	1.9±0.1	0.70±0.04	15.2±0.9	5.0±0.9	1.7±0.3	0	14±1	0.6±0.1	39±2
2 jet signal region [m_H = 125 GeV]									
1.1±0.1	1.9±0.1	0.18±0.02	6.4±0.7	2.9±0.6	1.2±0.2	0	6.8±0.9	0.75±0.09	20±2

- Expected 95% CL_s fit: **6.7 x SM (0.17σ)**
- Analysis is not sensitive enough for direct observation at 8 TeV, but there's still hope for LHC Run 2!
- For Run1, these 8 TeV results will be combined with a 7 TeV analysis and other W/Z(H → WW*) channels at both 7 and 8 TeV to improve sensitivity.

$W(H \to WW^* \to lvlv) \to lvlvlv$
$W(H \to WW^* \to lvlv) \to qqlvlv$
$Z(H \to WW^* \to lvlv) \to lvlvll$

} *Other W/Z(H → WW*) channels that are close to being finalized*

BACKUP

Pre-selection

Pre-selection cuts are used to eliminate events that don't resemble signal events, to keep the size of the data set manageable. Objects (e.g. leptons, jets) are also defined at the pre-selection stage.

Event selection:

- Reconstructed vertex with at least two associated tracks (primary vertex).
- Exactly two good leptons (electrons or muons).
- One reconstructed lepton matched to the lepton trigger.

Lepton selection:

- Lepton p_T > 22 GeV (leading), 15 GeV (sub-leading)
- Both leptons passing stringent identification criteria.
- Electron $|\eta| < 2.47$, muon $|\eta| < 2.5$.
- Leptons must be isolated from other energy deposits in the calorimeter.
- Lepton-lepton overlap removal

Jet selection:

- Anti-k_T algorithm with radius R=0.4.
- Jet matched to primary vertex.
- Jet p_T > 25 GeV, Jet $|\eta| < 2.5$.
- Jet p_T > 30 GeV, Jet $|\eta| >= 2.5$.
- No b-tagged jets
- Lepton-jet overlap removal

Other selections:

- Same sign leptons
- M_{ll} > 12 GeV (ee/μμ), 10 GeV (eμ)
- $|M_{ll} - M_Z|$ > 15 GeV (ee/μμ)
- Etmiss > 25 GeV

Background/Signal Modeling

Process	Generator
WH/ZH	Pythia 8
ggF/VBF	Powheg Box + Pythia 8
WZ	Powheg Box + Pythia 8
WZ (4EW coupling)	Sherpa
$t\bar{t}$/Wt/single top	Powheg Box + Pythia 6
$t\bar{t}$+W/Z	Madgraph + Pythia 6
qq→W^+W^-	Powheg Box + Pythia 6
gg→W^+W^-	gg2WW + Herwig
W±W±jj	Sherpa
Zγ	Sherpa
Z/γ^*	Alpgen + Herwig
ZZ*	Powheg Box + Pythia 8
WWW/WWZ/WZZ	Magraph + Pythia 6

Jet fakes background strategy

- W+jets and QCD backgrounds contaminate the signal region via jets faking leptons.
- Fakes are not well modeled in Monte Carlo (MC) simulation, so use a data driven method to derive a fake factor.
- Fake factor is derived in data from a Z+jets sample, decomposed into opposite sign (OS) and same sign (SS) components.

Fake factor, $f_l = \frac{N_{id}}{N_{anti-id}}$, where

N_{id} = Number of jets passing normal lepton ID

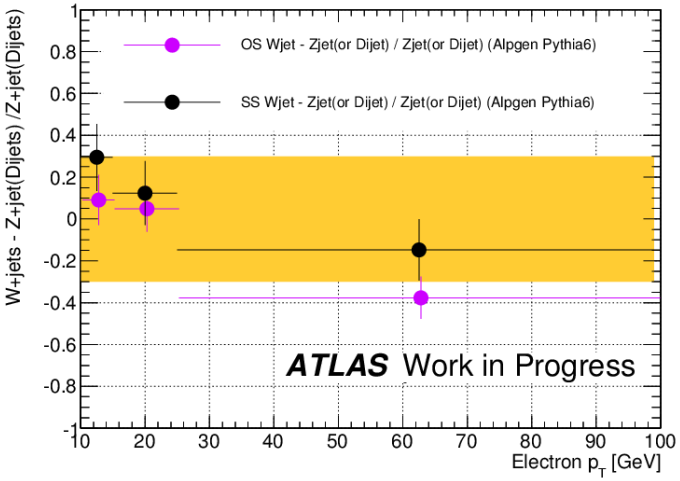
$N_{anti\text{-}id}$ = Number of jets passing fake enriched lepton ID

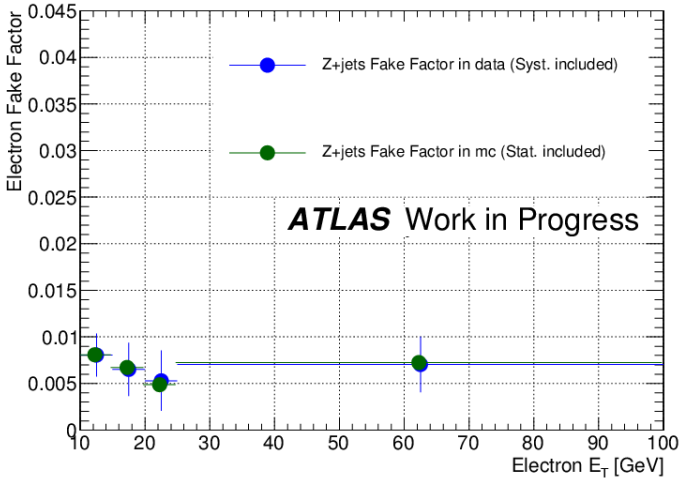
From data

$$N^{\text{W+jet(OS)}}_{\text{id+id}} = f^{OS}_{W+jets} \cdot N^{\text{W+jet(OS)}}_{\text{id+anti-id}} = f^{incl.}_{Z+jets} \cdot \frac{f^{OS}_{W+jets}}{f^{incl.}_{Z+jets}} N^{\text{W+jet(OS)}}_{\text{id+anti-id}}$$

$$N^{\text{W+jet(SS)}}_{\text{id+id}} = f^{SS}_{W+jets} \cdot N^{\text{W+jet(SS)}}_{\text{id+anti-id}} = f^{incl.}_{Z+jets} \cdot \frac{f^{SS}_{W+jets}}{f^{incl.}_{Z+jets}} N^{\text{W+jet(SS)}}_{\text{id+anti-id}}$$

From MC



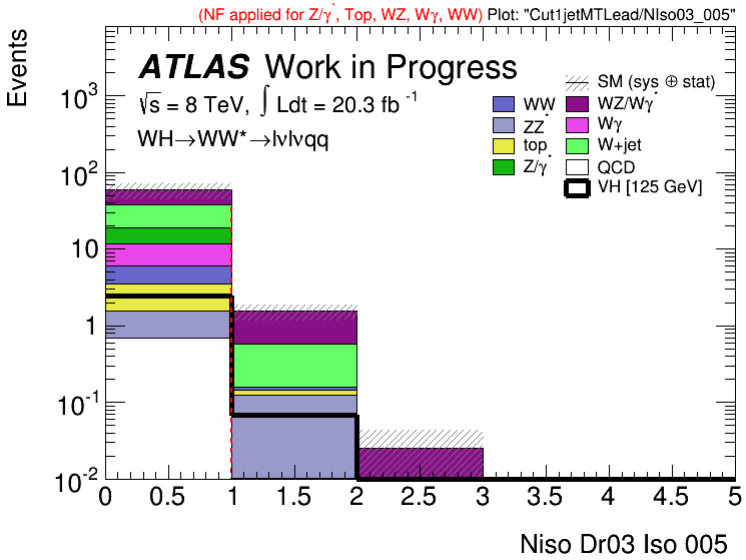


W+jets and Z+jets fake factors in MC (left image)

Z+jets fake factors in data and MC (right image)

WZ background R&D

- WZ contaminates our SR primarily via a lost lepton (after fully leptonic decay), so a candidate variable for removing WZ events is the *number of isolated tracks not associated with a reconstructed lepton*.
- Isolation: define a cone around the track and set a threshold on the fraction of energy in the cone that isn't carried by the track itself.
- The size of the cone and the fraction of the total track energy it contains can be varied to discriminate between leptons and jets. Therefore, an isolated track not already associated with a lepton is a good candidate for the lost third lepton in WZ events!



Cone size ΔR = 0.3 witht 5% threshold

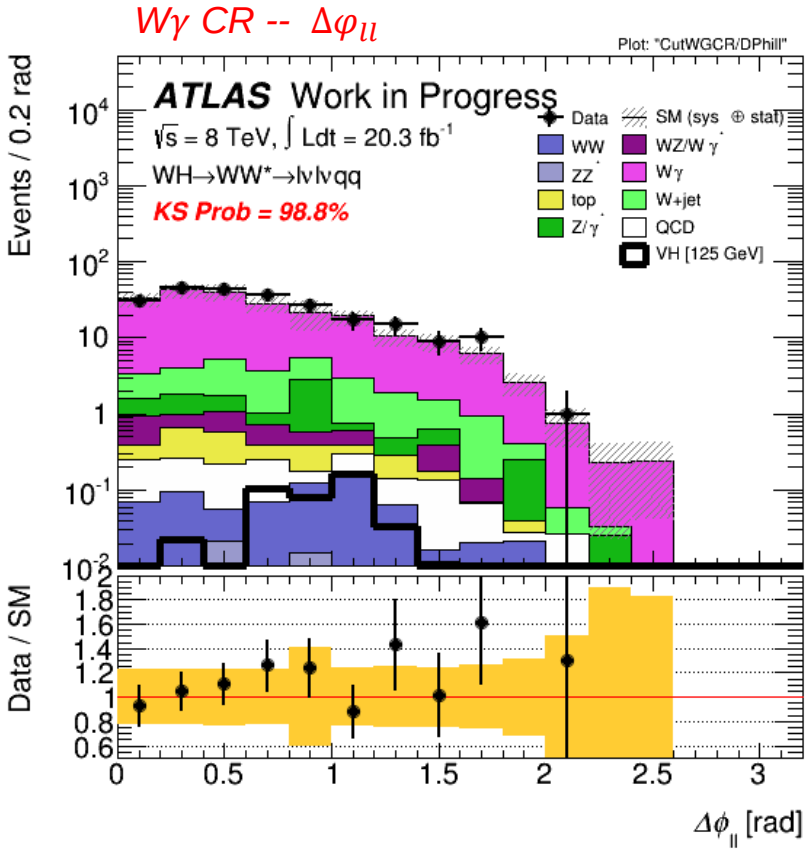
Unfortunately, despite the efficacity of this variable in discriminating WZ events, cutting on it removes too much signal and it ends up degrading our sensitivity.

Not used in current analysis, but could be useful in the LHC Run2 data when signal statistics will be higher!

$W\gamma$ Control Region

Definition:

- Require at least one lepton from conversion, with no hits in the b-layer.
- Normal SR cuts up to the MET cut stage, followed by
- $\Delta\varphi_{ll} < 2.5$
- $p_{T.ll} > 30\ GeV$
- $M_{ll} < 50\ GeV$

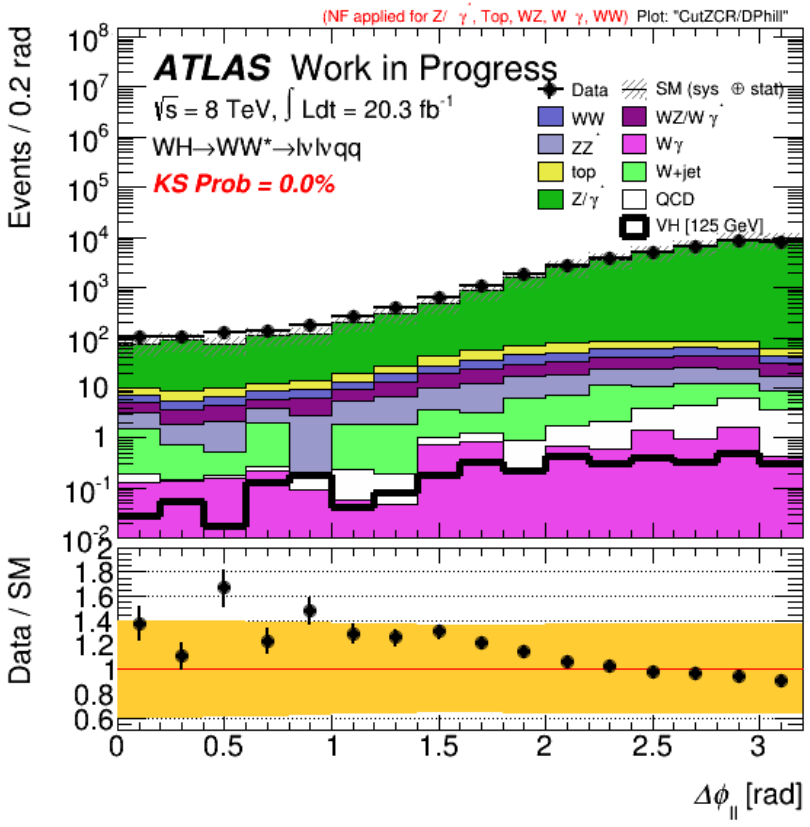


ATLAS Work in Progress				
	Wγ events	**Total Bkg**	**Data**	**Data/MC**
Before NFs	184 ± 5 ↓ x 1.07	215 ± 6 ↓ x All NFs	235	1.09 ± 0.08
After NFs	198 ± 6	227 ± 6	235	1.04 ± 0.07

Z+jets Control Region

Definition:

- Opposite sign leptons
- Invert Z veto $|M_{ll} - M_Z| <= 15$ GeV (ee/μμ)
- $55 \text{ GeV} < M_{ll} < 85$ GeV (eμ)
- Normal SR cuts otherwise, stopping at the $M_{ljj}^{\ Min}/M_{lj}^{\ Min}$ cut stage.



ATLAS Work in Progress				
	Z+jets events	**Total Bkg**	**Data**	**Data/MC**
Before NFs	46500 ± 400	47400 ± 400	40038	0.84 ± 0.01
After NFs	39900 ± 400	40700 ± 400	40038	0.98 ± 0.01

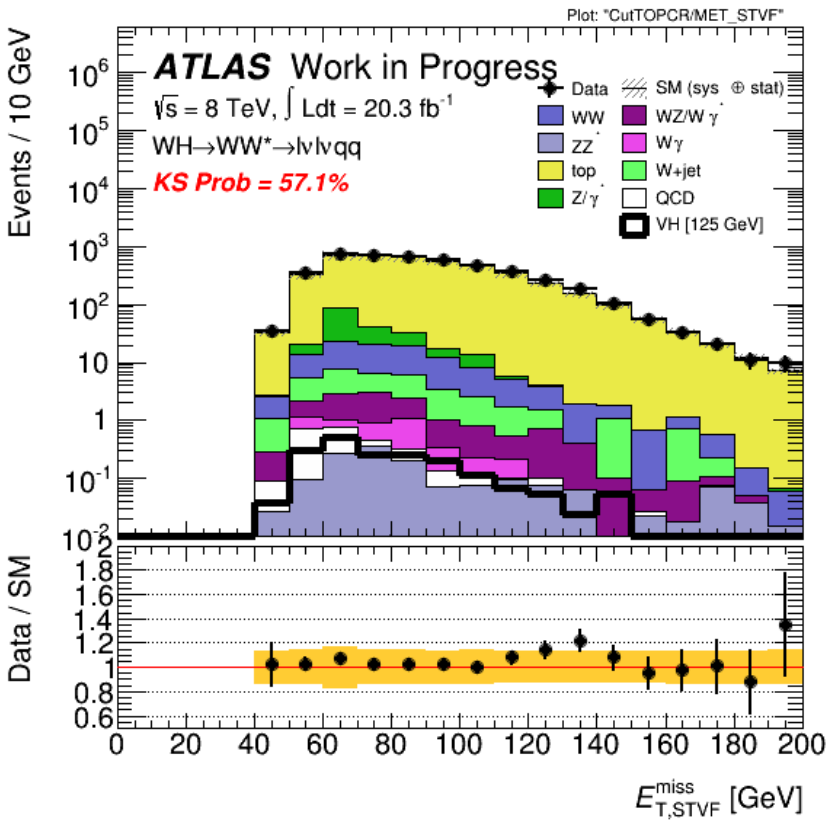
x 0.86 (Z+jets events: Before NFs → After NFs)

x All NFs (Total Bkg: Before NFs → After NFs)

Top Control Region

Definition:

- Opposite sign leptons
- Require at least 1 b-tagged jet
- E_T^{miss} > 60 GeV (ee/μμ)
- Normal SR cuts otherwise, stopping at the M_T^{Lead} cut stage.

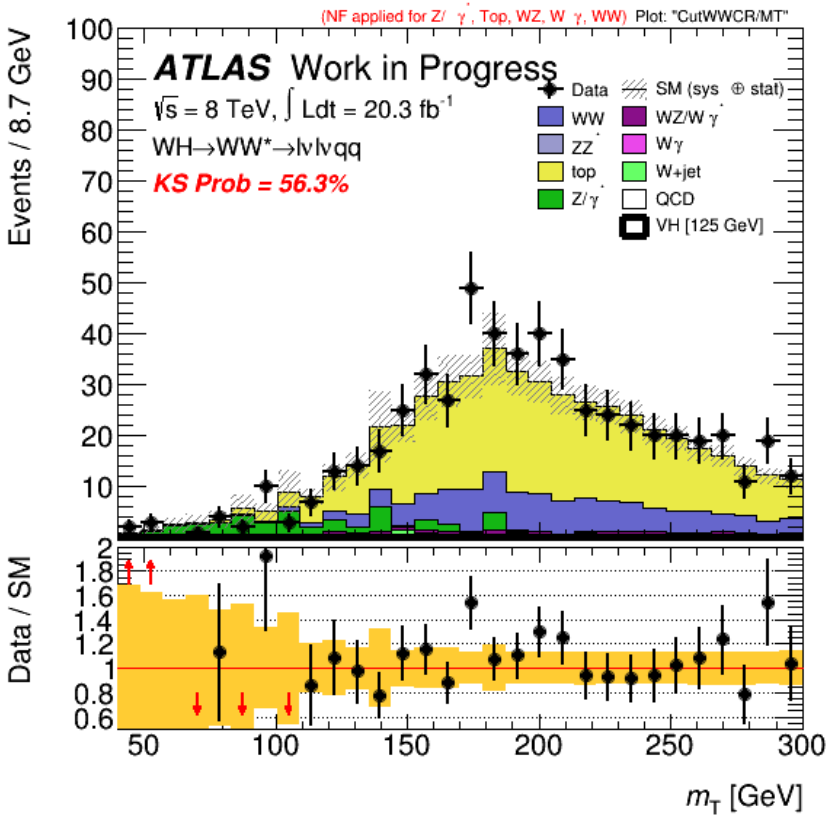


ATLAS Work in Progress				
	Top events	**Total Bkg**	**Data**	**Data/MC**
Before NFs	4156 ± 7 ↓ x 1.06	4400 ± 20 ↓ x All NFs	4604	1.05 ± 0.02
After NFs	4405 ± 8	4600 ± 10	4604	1.00 ± 0.02

W^+W^- Control Region

Definition:

- Opposite sign leptons
- E_T^{miss} > 80 GeV
- Normal SR cuts otherwise, stopping at the M_T^{Lead} cut stage.



ATLAS Work in Progress				
	WW events	**Total Bkg**	**Data**	**Data/MC**
Before NFs	174 ± 2 ↓ x 0.7	650 ± 10 ↓ x All NFs	621	0.96 ± 0.04
After NFs	128 ± 1	610 ± 10	621	1.01 ± 0.04

Charge flip uncertainty

- Several background processes contaminate our SR via a charge flip, i.e. an electron is reconstructed with the incorrect sign (this is a negligible effect for muons).
- The charge flip rate is assessed in both MC simulated events and data events using a pure Z → ee + jets region, and the relative difference between the rates in MC and data is taken as a systematic uncertainty.

ATLAS Work in Progress

$$\left(\frac{SS}{OS}\right)_{data,Zregion} = 0.000046 \qquad \left(\frac{SS}{OS}\right)_{MC,Zregion} = 0.000053$$

Table 112: Results for SS and OS dielectron pair events in the Z region for MC and data.

Region	process	OS electron pair	SS electron pair	ratio SS/OS
Z region	Data	5.39407e+06 ± 2322	24804 ±157	0.0046
	MC	5.24237e+06 ± 4647	28650 ±330	0.0055
	data/MC	1.029 ±0.001	0.866 ± 0.011	-
Z region +1 jet	Data	870763 ± 933	4356 ± 66	0.0050
	MC	852388 ± 1918	5043±141	0.0059
	data/mc	1.022 ± 0.003	0.863 ± 0.027	-
Z region +2 jets	Data	212759 ± 461	1194 ± 35	0.0056
	MC	215423 ± 974	1390 ± 74	0.0065
	ratio	0.988 ±0.005	0.860 ± 0.052	-