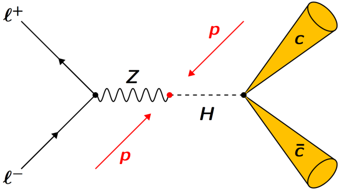
$Z(\ell^+\ell^-)H(c\bar{c})$ Search

IoP Annual APP and HEPP Conference, 26/3/18

Elliot Reynolds



Overview

Motivation

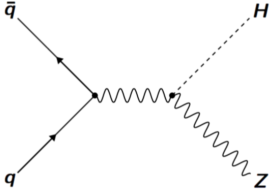
- With a SM BR of **2.9%**, $\boldsymbol{H \rightarrow c\bar{c}}$ is the SM process with the largest Yukawa coupling to lack experimental evidence
- The smallness of y_c^{SM} makes it highly susceptible to modifications from new physics

Aims

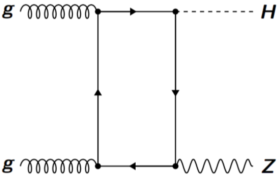
- Use ATLAS **2015+16** data ($\mathbf{36.1\ fb^{-1}}$) to set direct limit on $\boldsymbol{Z(\ell\ell)H(c\bar{c})}$
- Focus on associated production with $Z(\ell\ell)$ due to high $S/\sqrt{B}$, simple background composition & low exposure to experimental uncertainties
- Pioneer use of new Run 2 c-tagging algorithm

Present Constraints

- Indirect bounds on unobserved Higgs decays from global fits impose $\boldsymbol{BR(H \rightarrow c\bar{c}) < 20\%}$[†]
- Run 1 ATLAS $H \rightarrow J/\psi\gamma$ search provides a limit of about $\boldsymbol{220 \times y_c^{SM}}$, with mild theoretical assumptions[‡]



$pp \rightarrow ZH$ dominated $q\bar{q} \rightarrow ZH$ processes, $\sigma \approx 0.76$ pb at $\sqrt{s} = 13$ TeV



Smaller contributions from $gg \rightarrow ZH$, but harder p_T^H, $\sigma \approx 0.12$ pb at $\sqrt{s} = 13$ TeV

[†] arXiv:1310.7029

[‡] arXiv:1501.03276

Data, Selection and Categorisation

Data and Trigger

36.1 fb^{-1} of **13 TeV** data, collected during 2015 and 2016 using a **single electron or muon trigger**

$Z \rightarrow \ell^+\ell^-$ Selection

- Exactly **2 same flavour leptons** (e or μ), passing loose identification, impact parameter and isolation requirements
- Require **opposite charges** (μ only)
- Both leptons $\boldsymbol{p_T > 7\ \mathrm{GeV}}$, with at least one $\boldsymbol{p_T > 27\ \mathrm{GeV}}$ and $\boldsymbol{|\eta| < 2.5}$
- $\boldsymbol{81\ \mathrm{GeV} < m_{\ell\ell} < 101\ \mathrm{GeV}}$
- $\boldsymbol{p_T^Z > 75\ \mathrm{GeV}}$

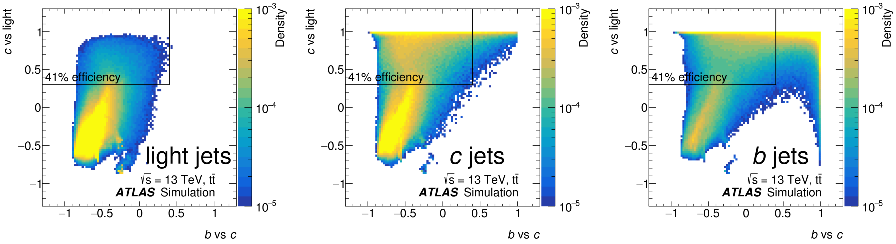
$H \rightarrow c\bar{c}$ Selection

- At least **2 jets** with $\boldsymbol{|\eta| < 2.5}$ and $\boldsymbol{p_T > 20\ \mathrm{GeV}}$
- Leading jet $\boldsymbol{p_T > 45\ \mathrm{GeV}}$
- $H \rightarrow c\bar{c}$ candidate formed from two highest p_T jets
- Dijet $\Delta R_{c\bar{c}}$ requirement on $H \rightarrow c\bar{c}$ jets which varies with p_T^H
- At least **one $\boldsymbol{H \rightarrow c\bar{c}}$ jet $\boldsymbol{c}$-tagged**

Event Categorisation

Events divided into **4 categories**, each with $H \rightarrow c\bar{c}$ candidates from 1 or 2 c-tagged jets, and p_T^Z above or below 150 GeV

c-Tagging



light-jets **c-jets** ***b*-jets**

Figure: arXiv:1802.04329

- BDT-based discriminant built using low-level *b*-tagging variables
- BDTs trained to separate *c*-jets from *b*-jets (x-axes), and from light-jets (y-axes)
- Rectangular cuts in 2D discriminant space optimised for analysis
- *c*-jet efficiency of **41%** for a light-jet rejection of **10** and a *b*-jet rejection of **4**
- Efficiency calibrated in data
- Uncertainties of 5% for *b*-jets, 20% for light-jets, and 25% for *c*-jets
- 'Truth-tagging', parameterised in p_T and $|\eta|$, applied to simulated events to preserve statistics

Backgrounds and Simulation

Backgrounds

- Background dominated by $\boldsymbol{Z + \text{jets}}$
- Smaller contributions from ZZ, ZW and $t\bar{t}$
- $W + \text{jets}$, WW, single-top and multi-jet shown to be negligible ($< 0.5\%$)
- $ZH(b\bar{b})$ treated as a background, and constrained to SM expectation

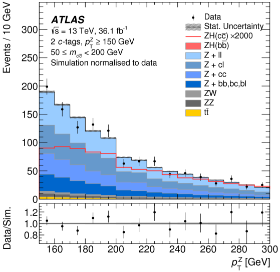


Figure: arXiv:1802.04329

Simulation

Process	Generator	Parton Shower	Cross-section (QCD)
$q\bar{q} \to ZH$	Powheg-BOX v2	Pythia 8	NNLO
$gg \to ZH$	Powheg-BOX v2	Pythia 8	NLO+NLL
$Z + \text{jets}$	Sherpa 2.2.1	Sherpa	NNLO
ZW, ZZ	Sherpa 2.2.1	Sherpa	NLO
$t\bar{t}$	Powheg-BOX v2	Pythia 8	NNLO+NNLL

$\Delta R_{c\bar{c}}$ Control Distributions

- Post-fit $\Delta R_{c\bar{c}}$ control distributions
- More control distributions in backup slides
- Good data-MC agreement observed in all post-fit control distributions

1 *c*-tag **2 *c*-tags**

$p_T^Z / \mathrm{GeV} > 150$

$75 < p_T^Z / \mathrm{GeV} < 150$

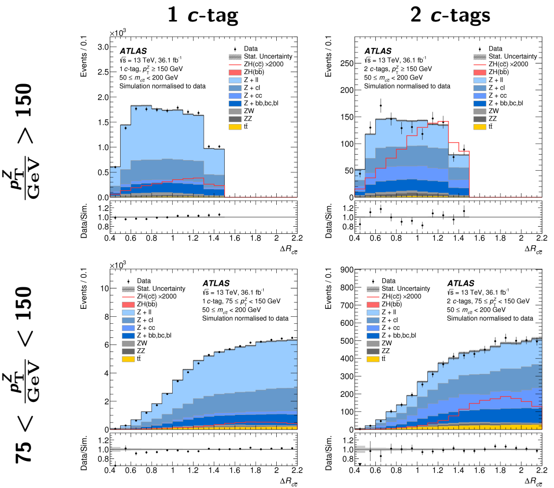


Figure: arXiv:1802.04329

Quantifying the Search Results

Statistical Model Overview

- Simultaneous likelihood fit performed in all 4 event categories
- $\boldsymbol{m_{c\bar{c}}}$ used as observable
- **Signal yield** used as parameter of interest
- $Z + \text{jets}$ background normalisation free in fit
- All other background yields constrained to theory expectations
- The ZV production rate was measured to cross-check the analysis methods

Implementation of Systematic Uncertainties

- Uncertainties modelled as nuisance parameters in fit, constrained using auxiliary measurements
- Grouped uncertainty breakdown performed

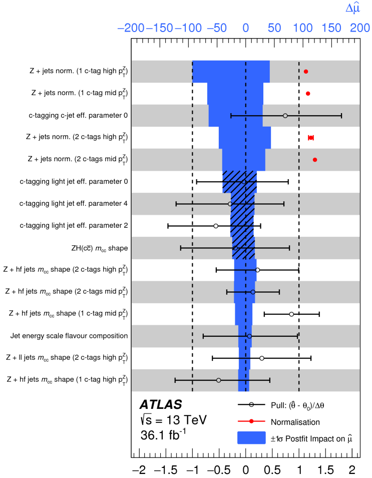


Figure: arXiv:1802.04329

Systematic Uncertainties

Signal and Background Modelling

- $ZH(c\bar{c}/b\bar{b})$ and ZZ/ZW normalisation uncertainties from theory
- $t\bar{t}$ normalisation uncertainty from data/MC ratio in $e+\mu$ CR
- Acceptance and shape uncertainties from MC generator comparisons

Table: arXiv:1802.04329

Source	$\sigma/\sigma_{\mathrm{tot}}$
Statistical	**49%**
$Z+\mathrm{jets}$ Normalisation	31%
Systematic	**87%**
Flavour Tagging	73%
Background Modelling	47%
Lepton, Jet and Lumi.	28%
Signal Modelling	28%
MC statistical	6%

Experimental Uncertainties

- **Leptons:** Trigger, reconstruction, identification, track to vertex association (μ-only) and isolation scale factor uncertainties; with energy/momentum scale and resolution uncertainties
- **Jets:** Energy scale, resolution, and jet vertex tagging scale factor uncertainties
- **Flavour-Tagging:** Eigen-vector reduction, resulting in 11 NPs for fit
- **Miscellaneous:** Luminosity and pileup reweighting uncertainties

Post-Fit $m_{c\bar{c}}$ Distributions

- No significant upward fluctuation observed
- Best fit signal strength value: $\hat{\mu} = -69^{+73}_{-129}$
- Data used to set 95% CL CL_s upper limit on signal strength
- Post-fit $Z + \text{jets}$ normalisation parameters between 1.1 and 1.3

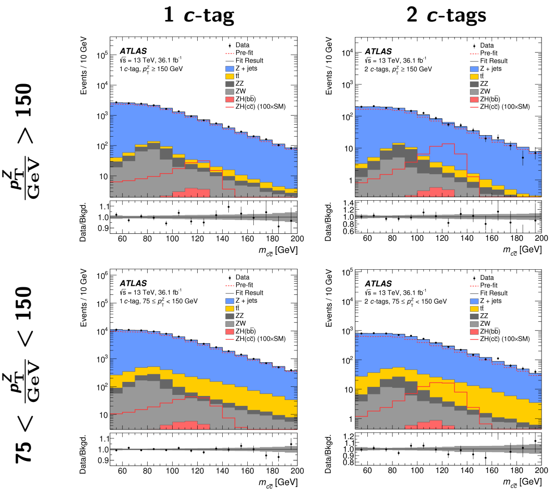


Figure: arXiv:1802.04329

Results

ZV Validation Measurement

- Observed significance of ZV peak $\mathbf{1.4\sigma}$, compatible with the SM expectation of $\mathbf{(2.2 \pm 0.9)\sigma}$

Limits on ZH(c$\bar{\text{c}}$) Production

- 95% CL CL_s upper limit on $pp \rightarrow ZH(c\bar{c})$ production set at **107**x the SM expectation (**2.7 *pb***)
- Worlds tightest direct constraint on $H \rightarrow c\bar{c}$!

Consistency and Robustness Checks

- Compatibility p-value between fits with $\mu_{ZH(c\bar{c})}$ (un)correlated between categories is 66%
- Limit robust against modified $ZH(b\bar{b})$ rates, with variations between 0 and 2x the SM expectation causing a ±5% shift in the limit

Table: arXiv:1802.04329

Process	Post-fit $\mathcal{A} \times \epsilon$ [%]			
	1 c-tag		2 c-tags	
	$75 < p_{\mathrm{T}}^{Z} < 150\,GeV$	$p_{\mathrm{T}}^{Z} > 150\,GeV$	$75 < p_{\mathrm{T}}^{Z} < 150\,GeV$	$p_{\mathrm{T}}^{Z} > 150\,GeV$
$ZH(c\bar{c})$	2.2	1.3	0.5	0.3
$ZH(b\bar{b})$	1.7	1.0	0.2	0.1

Summary and Prospects

Summary

- First use of new *c*-tagging algorithms to perform search for $ZH(c\bar{c})$
- Methods validated through ZV-based cross-check
- $pp \rightarrow ZH(c\bar{c})$ production above **107**x the SM expectation excluded at 95% CL!

Prospects

- Factor of **~ 2**[†] drop in total uncertainty possible through use of other VH channels ($W(\ell\nu)$ and $Z(\nu\bar{\nu})$)
- Further gains (**$\sim$ 7%**[†]) possible through use of BDT-based analysis strategy, or splitting event categories by jet multiplicity
- *c*-tagging performance improving rapidly, with next generation of algorithms utilising advanced 'deep-learning' techniques
- Improved statistical power at HL-LHC should reduce statistical and systematic uncertainties

---

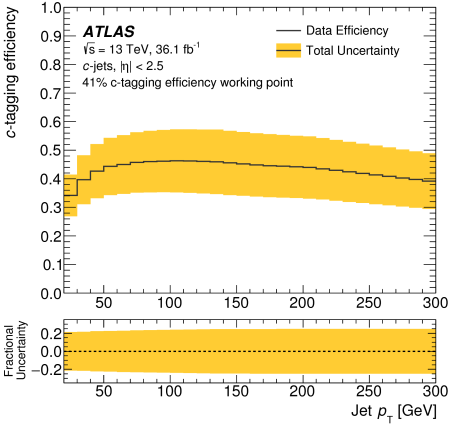
[†] arXiv:1708.03299v2

Backup Slides

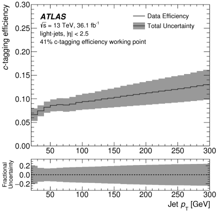
- *c*-tagging efficiencies
- Linear post-fit $m_{c\bar{c}}$ distributions
- Background flavour compositions:
  1. $Z + \mathrm{jets}$
  2. Diboson
- Post-fit control distributions:
  1. p_T^Z
  2. $p_T^{lead\ jet}$
  3. $p_T^{sublead\ jet}$

c-Tagging Efficiencies

c-jets



light-jets



b-jets

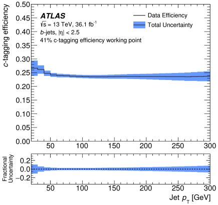


Figure: arXiv:1802.04329

Post-Fit $m_{c\bar{c}}$ Distributions (Linear)

- Data consistent with background-only hypothesis
- Best fit signal strength value: $\hat{\mu} = -69^{+73}_{-129}$
- Data used to set 95% CL CL_s upper limit on signal strength
- Post-fit $Z + \mathrm{jets}$ normalisation parameters between 1.1 and 1.3

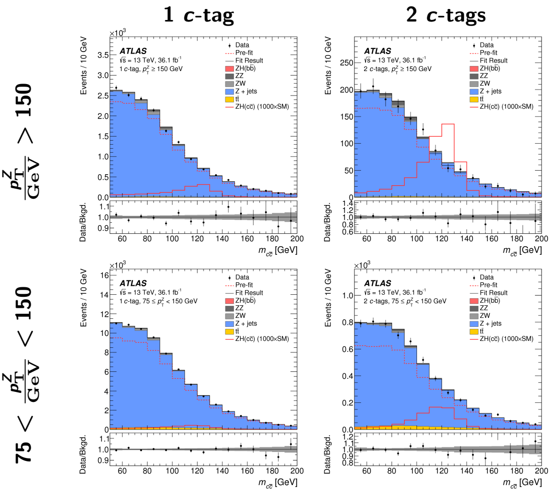


Figure: arXiv:1802.04329

$Z + \text{jets}$ Flavour Composition

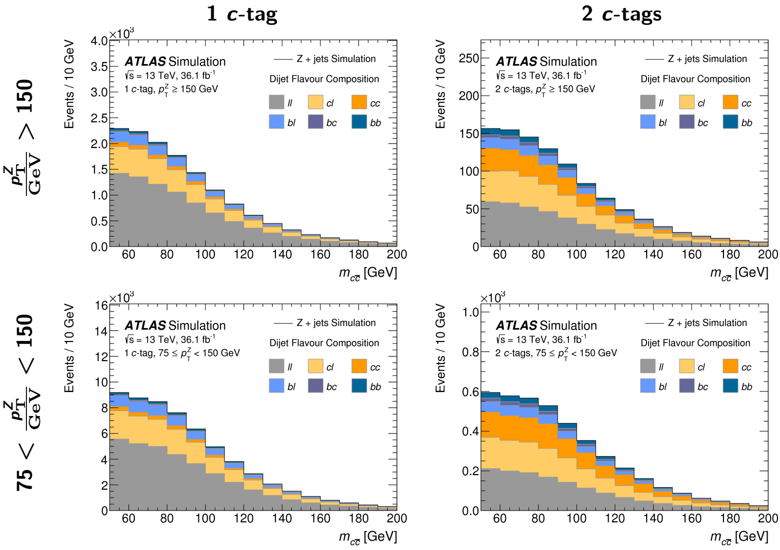


Figure: arXiv:1802.04329

Diboson Flavour Composition

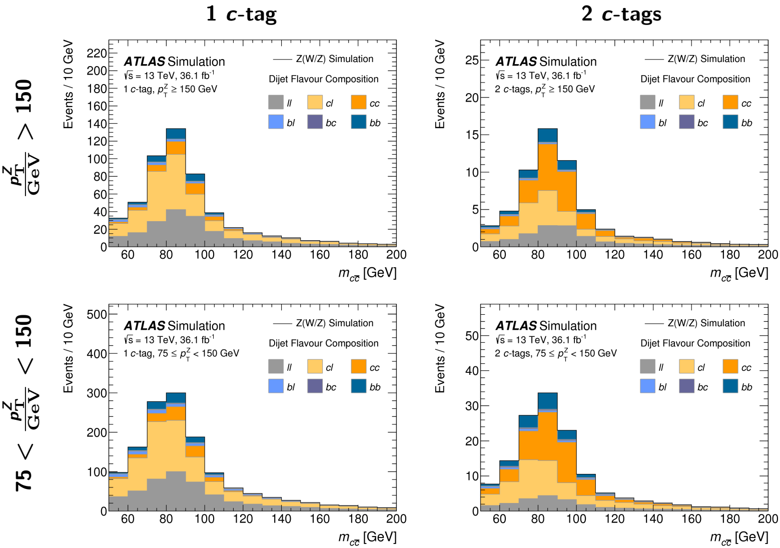


Figure: arXiv:1802.04329

p_T^Z Distributions

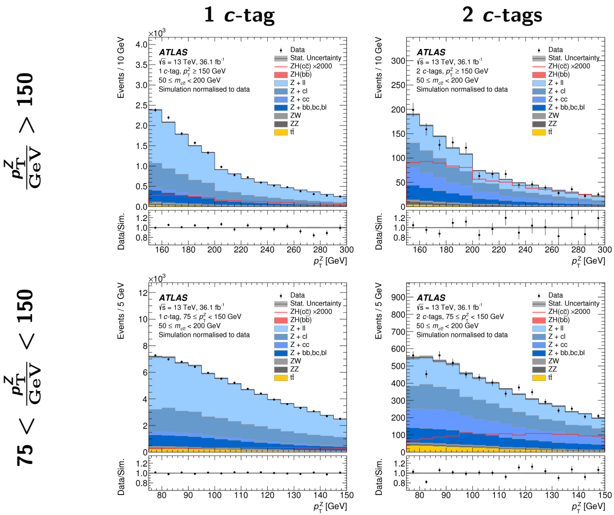


Figure: arXiv:1802.04329

$p_T^{lead\ jet}$ Distributions

1 c-tag | 2 c-tags

$\frac{p_T^Z}{\text{GeV}} > 150$

$75 < \frac{p_T^Z}{\text{GeV}} < 150$

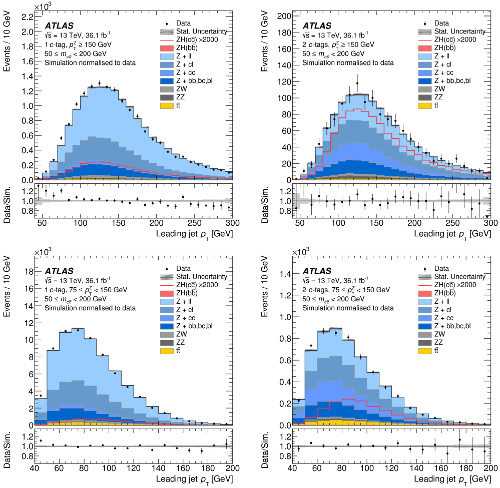


Figure: arXiv:1802.04329

$p_T^{sublead\ \ jet}$ Distributions

1 c-tag **2 c-tags**

$\frac{p_T^Z}{\mathrm{GeV}} > 150$

$75 < \frac{p_T^Z}{\mathrm{GeV}} < 150$

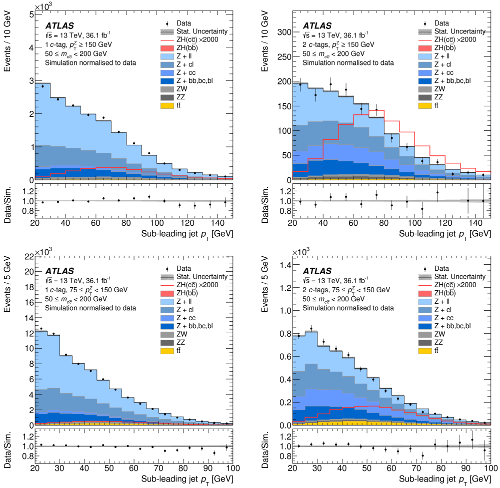


Figure: arXiv:1802.04329