COFUND. A project supported by the European Union

Searching for Higgs boson decays to charm quark pairs with charm jet tagging at ATLAS

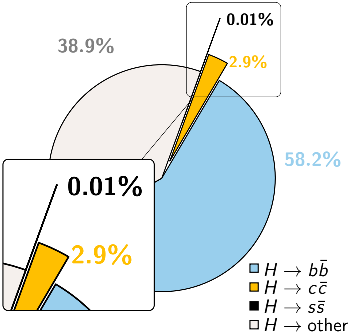
CMS Flavour Tagging Workshop, IIHE Brussels

11th April 2018

Andy Chisholm (CERN)
on behalf of the ATLAS collaboration

Why is the charm quark Yukawa coupling important?

- The smallness of the charm (c) quark coupling ($y_c = \frac{\sqrt{2}m_c(m_H)}{v} \approx 4 \times 10^{-3}$) make it **highly susceptible to modifications from potential new physics**
- $H \to c\bar{c}$ decays constitute the **largest part of the SM prediction for Γ_H for which we have no experimental evidence**
- To date, we **only have experimental evidence for 3rd generation Yukawa couplings!**



Cartoon of SM 125 GeV $H \to q\bar{q}$ branching fractions, $H \to u\bar{u}/d\bar{d}$ too small to show!

What are the existing indirect constraints?

- Constraints on unobserved Higgs decays impose around $\mathcal{B}(H \to c\bar{c}) < 20\%$, global fits to LHC data indirectly bound Γ_H leading to $y_c/y_c^{SM} < 6$, **assuming SM Higgs production and no BSM decays** (arXiv:1310.7029, arXiv:1503.00290)
- Direct bound of around $\Gamma_H < 1$ GeV from $H \to \gamma\gamma$ and $H \to 4\ell$ lineshapes impose around $y_c/y_c^{SM} < 120$, **but this is model independent** (arXiv:1503.00290)

How can we constrain these couplings in a more direct way?

How can we constrain the charm Yukawa couplings in a more direct way?

Exclusive $H \to J/\psi\,\gamma$ decays

- Exclusive radiative Higgs decays to J/ψ are sensitive to $Hc\bar{c}$ couplings, very rare in SM with $\mathcal{B}\,(H \to J/\psi\,\gamma) = (2.8 \pm 0.2) \times 10^{-6}$ (arXiv:1407.6695)
- Both ATLAS (arXiv:1501.03276) and CMS (arXiv:1507.03031) have searched for such decays, both leading to limits of $\mathcal{B}\,(H \to J/\psi\,\gamma) < 1.5 \times 10^{-3}$
- Implies bound on charm Yukawa coupling of $y_c/y_c^{SM} < 220$ at 95% CL (arXiv:1503.00290)
- **Side Note:** The analogous decays $H \to \phi\,\gamma$ and $H \to \rho\,\gamma$ are sensitive to the light quark Yukawa couplings (see arXiv:1712.02758)

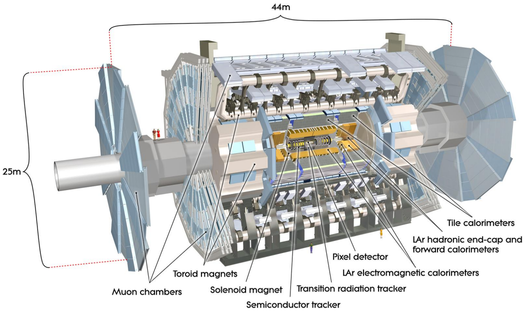
Kinematic distributions in inclusive production

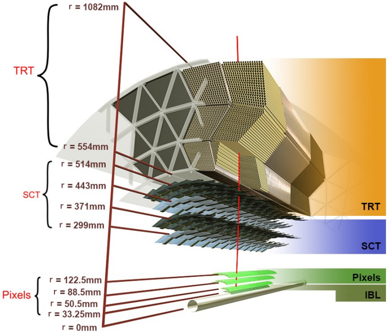
- Modifications to the heavy quark $\mathcal{Q} = c, b$ Yukawa couplings could change the shape of the inclusive p_T^H spectrum due to enhanced $g\mathcal{Q} \to H\mathcal{Q}$ contribution (arXiv:1606.09621)
- p_T^H well measured in the $H \to \gamma\gamma$ and $H \to 4\ell$ channels, which imposes a 95% CL bound of $-16 < y_c/y_c^{SM} < 18$, based on Run 1 ATLAS + CMS results (arXiv:1606.09253)

Inclusive $H \to c\bar{c}$ decays

- Study inclusive $H \to c\bar{c}$ decays with c-tagged jets, direct sensitivity to $Hc\bar{c}$ coupling
- First search from LHCb, though only sensitive to $\sim 5000\times$ SM rate (LHCb-CONF-2016-006)
- Recent ATLAS search for $Z(\ell\ell)H(c\bar{c})$ production (arXiv:1802.04329), **focus of this talk!**

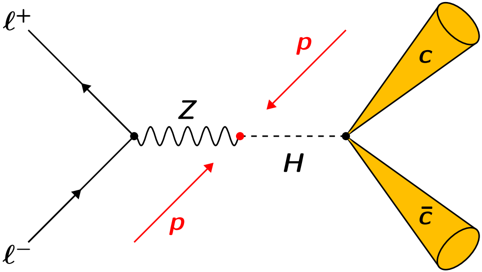
General purpose detector, well suited to studying heavy flavour jets

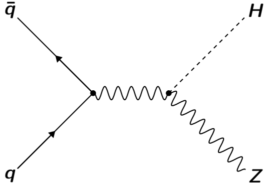




- **Inner Detector (ID):** Silicon Pixels and Strips (SCT) with Transition Radiation Tracker (TRT) $|\eta| < 2.5$ and (new for Run 2) Insertable B-Layer (IBL)
- **LAr EM Calorimeter:** Highly granular + longitudinally segmented (3-4 layers)
- **Had. Calorimeter:** Plastic scintillator tiles with iron absorber (LAr in fwd. region)
- **Muon Spectrometer (MS):** Triggering $|\eta| < 2.4$ and Precision Tracking $|\eta| < 2.7$
- **Jet Energy Resolution:** Typically $\sigma_E/E \approx 50\%/\sqrt{E(\text{ GeV})} \oplus 3\%$
- **Track IP Resolution:** $\sigma_{d_0} \approx 60\,\mu\text{m}$ and $\sigma_{z_0} \approx 140\,\mu\text{m}$ for $p_T = 1$ GeV (with IBL)

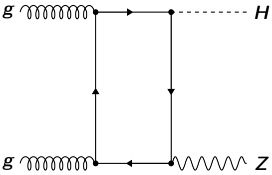
Given the success of the W/Z associated production channel in providing evidence for $H \to b\bar{b}$ decays†, this channel is an obvious first candidate for a $H \to c\bar{c}$ search





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

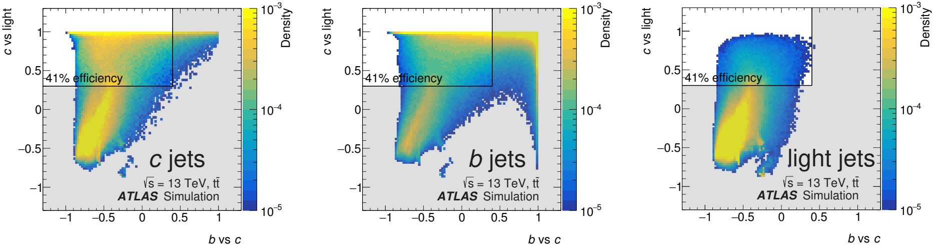
- **Focus on ZH production** with $Z \to e^+e^-$ and $Z \to \mu^+\mu^-$ decays for first ATLAS analysis
- Low exposure to experimental uncertainties, main backgrounds from $Z +$ jets, $Z(W/Z)$ and $t\bar{t}$
- Pioneer use of **new c-tagging algorithms** developed by ATLAS for Run 2 to identify the experimental signature of an inclusive $H \to c\bar{c}$ decay



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

† **ATLAS: arXiv:1708.03299 CMS: arXiv:1708.04188**

New c-tagging algorithms developed by ATLAS for Run 2!

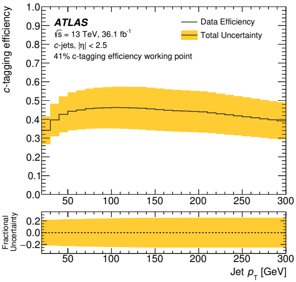


***c*-jets** **b-jets** **light flavour (u, d, s, g) jets**

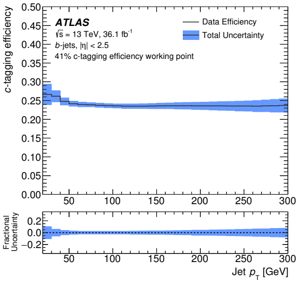
- Multivariate discriminant(s) built from input variables from low-level b-tagging algorithms (e.g. track impact parameter likelihood, secondary vertex finder)
- Trained with the same input variables used by the standard ATLAS Run 2 b-tagging algorithm (see ATL-PHYS-PUB-2015-022 for details)
- Implemented as two BDT discriminants, one trained to separate c-jets from b-jets (x-axis), another to separate c-jets from light-jets (y-axis)

"c-tag" jets by making a cut in the 2D discriminant space, working point optimised for $ZH, H \rightarrow c\bar{c}$ is shown in the rectangular selection (shaded region rejected)

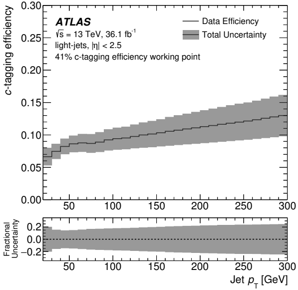
Introduction to *c*-jet tagging: Performance



_c_-jets



_b_-jets



light flavour (u, d, s, g) jets

Efficiency of *c*-tagging algorithm for *b*-, *c*- and light flavour jets measured in data ↑

- Working point for $ZH, H \to c\bar{c}$ exhibits a *c*-jet tagging efficiency of around 40%
- Rejects *b*-jets by around a factor 4× and light jets by around a factor 10×
- Efficiency calibrated in data with samples of *b*-jets from $t \to Wb$ (ATLAS-CONF-2014-004) and *c*-jets from $W \to cs, cd$ in $t\bar{t}$ events (ATLAS-CONF-2018-001)
- Typical total relative uncertainties of around **20%**, **5%** and **20%** for ***c*-**, ***b*-** and **light** jets, respectively

Use a $\sqrt{s} = 13$ TeV pp collision sample collected during 2015 and 2016 corresponding to an integrated luminosity of 36.1 fb^{-1}

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

- Trigger with lowest available p_{T} single electron or muon triggers
- Exactly two same flavour reconstructed leptons (e or μ)
- Both leptons $p_{\mathrm{T}} > 7$ GeV and at least one with $p_{\mathrm{T}} > 27$ GeV
- Require opposite charges (dimuons only)
- $81 < m_{\ell\ell} < 101$ GeV
- $p_{\mathrm{T}}^{Z} > 75$ GeV

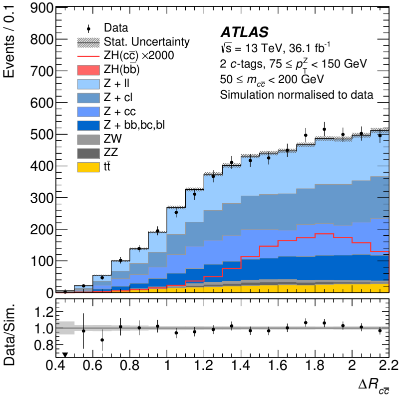
$H \to c\bar{c}$ Selection

- Consider anti-k_{T} $R = 0.4$ calorimeter jets with $|\eta| < 2.5$ and $p_{\mathrm{T}} > 20$ GeV
- At least two jets with leading jet $p_{\mathrm{T}} > 45$ GeV
- Form $H \to c\bar{c}$ candidate from the two highest p_{T} jets in an event
- At least one c-tagged jet from $H \to c\bar{c}$ candidate
- Dijet angular separation ΔR_{jj} requirement which varies with p_{T}^{Z}

Split events into 4 categories (with varying S/B) based on $H \to c\bar{c}$ candidates with 1 or 2 c-tags and p_{T}^{Z} above/below 150 GeV

Background Modelling

- Background dominated by $Z +$ jets $\rightarrow$ (enriched in heavy flavour jets)
- Smaller contributions from $ZZ(q\bar{q})$, $ZW(q\bar{q}')$ and $t\bar{t}$
- Negligible ($< 0.5\%$) contributions from $W +$ jets, WW, single-top and multi-jet



Simulation of $ZH(c\bar{c}/b\bar{b})$

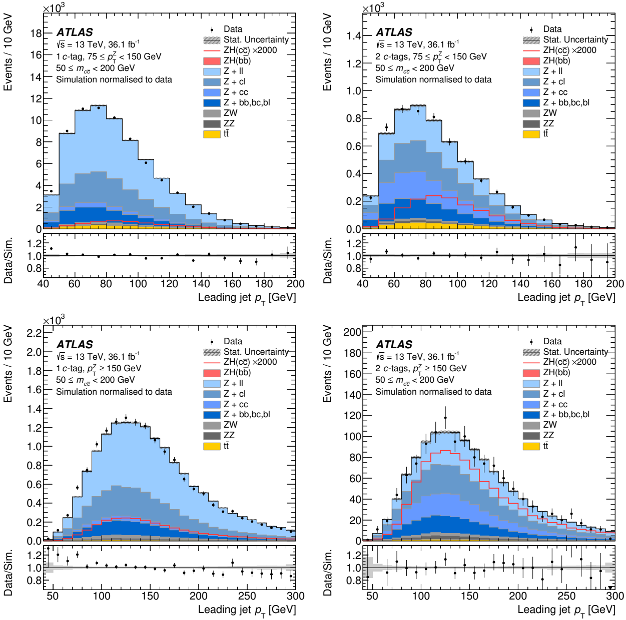
- Normalised with LHC Higgs XS WG YR4 recommendations (arXiv:1610.07922)
- $ZH(b\bar{b})$ treated as background normalised to SM expectation (with $\sigma \times \mathcal{B}$ uncertainty)

Process	MC Generator	Normalisation Cross section
$q\bar{q} \rightarrow ZH(c\bar{c}/b\bar{b})$	Powheg+GoSaM+MiNLO+Pythia8	NNLO (QCD) NLO (EW)
$gg \rightarrow ZH(c\bar{c}/b\bar{b})$	Powheg+Pythia8	NLO+NLL (QCD)
$Z +$ jets	Sherpa 2.2.1	NNLO
ZZ and ZW	Sherpa 2.2.1	NLO
$t\bar{t}$	Powheg+Pythia8	NNLO+NNLL

The nominal MC generators used to model the signal and backgrounds

Background composition after c-tagging

↓ Left: 1 c-tag events

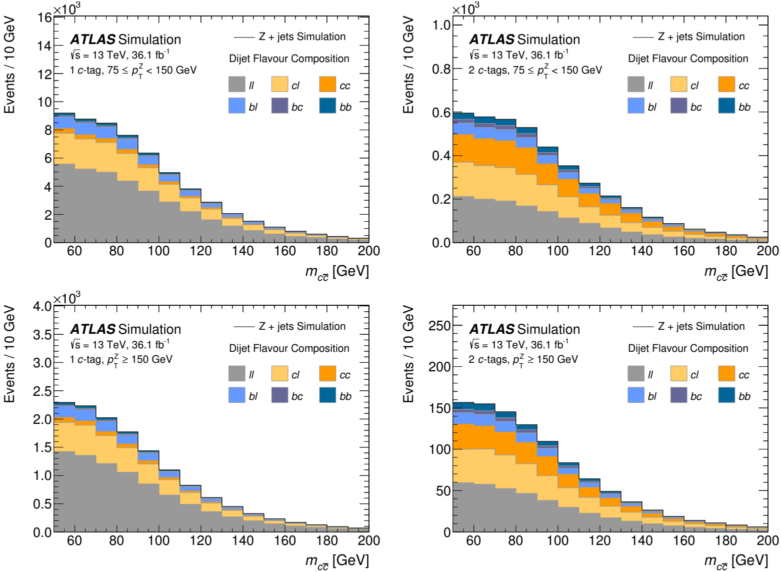


↑ Right: 2 c-tag events

$Z+$ jets flavour composition after c-tagging

Flavour composition of the $Z+$ jets sample enriched with c-jets

↓ Left: 1 c-tag events

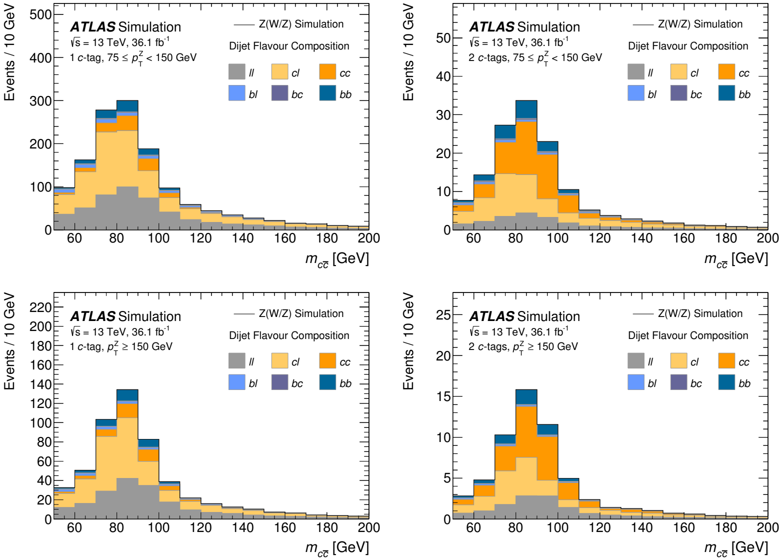


↑ Right: 2 c-tag events

ZZ and ZW flavour composition after c-tagging

c-tagged ZZ and ZW production enriched in $Z \to c\bar{c}$ and $W \to cs, cd$ decays

↓ Left: 1 c-tag events



↑ Right: 2 c-tag events

Statistical Model

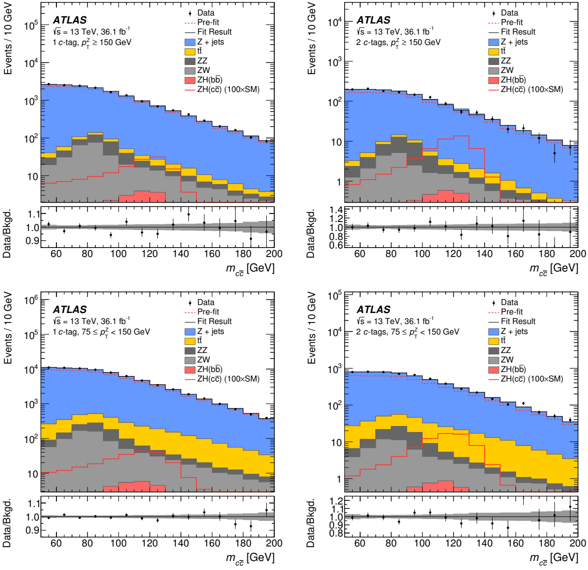
- Use the $H \to c\bar{c}$ candidate invariant mass $m_{c\bar{c}}$ as S/B discriminant
- Perform simultaneous binned likelihood fit to 4 categories within region $50 < m_{c\bar{c}} < 200$ GeV
- $ZH(c\bar{c})$ signal parameterised with free signal strength parameter, μ, common to all categories
- $Z +$ jets background determined directly from data with separate free normalisation parameter for each of the four categories

Systematic Uncertainties

- Included in the fit model as constrained nuisance parameters which parametrize the constraints from auxiliary measurements (e.g. lepton/jet calibrations)
- Experimental uncertainties associated with luminosity, c-tagging, lepton and jet performance are all included in the model
- Normalisation, acceptance and $m_{c\bar{c}}$ shape uncertainties associated with signal and background simulation are also included

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

$p_T^Z > 150$ GeV / **$75 < p_T^Z < 150$ GeV**



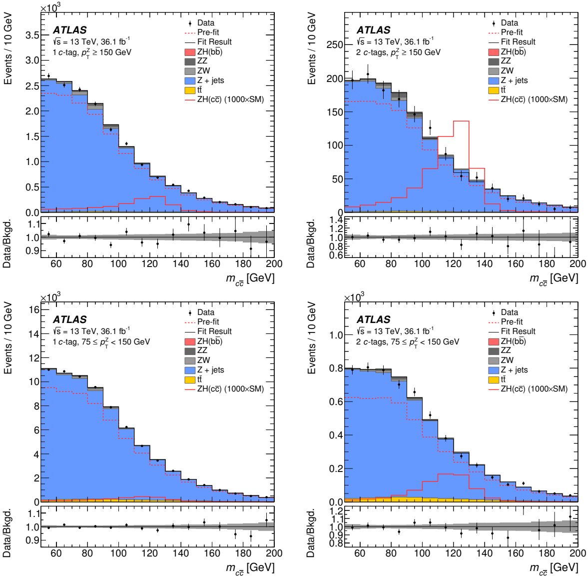
- No significant evidence for $ZH(c\bar{c})$ production
- Data consistent with background only hypothesis

SM expected number of $ZH(c\bar{c})$ events
1 c-tag $75 < p_T^Z < 150$ GeV **2.1**
1 c-tag $p_T^Z > 150$ GeV **1.2**
2 c-tags $75 < p_T^Z < 150$ GeV **0.5**
2 c-tags $p_T^Z > 150$ GeV **0.3**

1 c-tag | 2 c-tags

$p_T^Z > 150$ GeV

$75 < p_T^Z < 150$ GeV

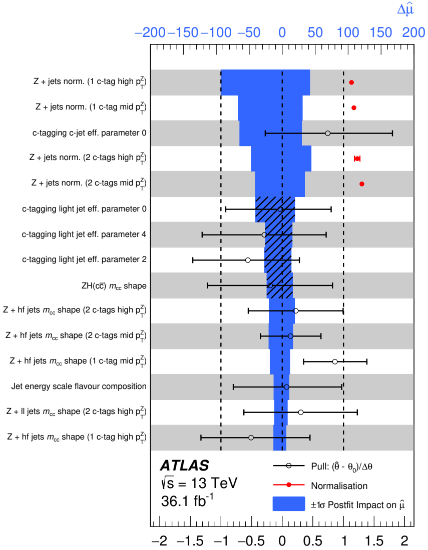


- No significant evidence for $ZH(c\bar{c})$ production
- Data consistent with background only hypothesis

SM expected number of $ZH(c\bar{c})$ events
1 c-tag $75 < p_T^Z < 150$ GeV **2.1**
1 c-tag $p_T^Z > 150$ GeV **1.2**
2 c-tags $75 < p_T^Z < 150$ GeV **0.5**
2 c-tags $p_T^Z > 150$ GeV **0.3**

Understanding the Sensitivity

Sensitivity dominated by systematic uncertainties, clear that these uncertainties should be reduced in order to fully exploit a larger dataset in the future



Source	$\sigma/\sigma_{\text{tot}}$
Statistical	49%
Floating $Z+$jets Normalisation	31%
Systematic	87%
Flavour Tagging	73%
Background Modeling	47%
Lepton, Jet and Luminosity	28%
Signal Modeling	28%
MC statistical	6%

Note: correlations between nuisance parameters within groups leads to $\sum_i \sigma_i^2 \neq \sigma_{\text{syst.}}^2$

- *c*-tagging uncertainties and background modelling (particularly $Z+$jets $m_{c\bar{c}}$ shape) have the dominant impact
- However, we can expect many of these uncertainties (e.g. $Z+$jets normalisation) to reduce with a larger dataset

Cross check with ZV production

- To validate background modelling and uncertainty prescriptions, measure production rate of the sum of ZZ and ZW relative to the SM expectation
- Observe (expect) ZV production with significance of 1.4σ (2.2σ)
- Measure ZV signal strength of $0.6^{+0.5}_{-0.4}$, consistent with SM expectation

Limits on $ZH(c\bar{c})$ production

95% CL CL_s upper limit on $\sigma(pp \to ZH) \times \mathcal{B}(H \to c\bar{c})$ [pb]			
Observed	Median Expected	Expected $+1\sigma$	Expected -1σ
2.7	3.9	6.0	2.8

- No evidence for $ZH(c\bar{c})$ production with current dataset (as expected)
- Upper limit of $\boldsymbol{\sigma(pp \to ZH) \times \mathcal{B}(H \to c\bar{c}) < 2.7}$ **pb** set at 95% CL, to be compared to an SM value of 2.55×10^{-2} pb
- Corresponds to **110×** the SM expectation

World's most stringent direct constraint on inclusive $H \to c\bar{c}$ decays!

Summary

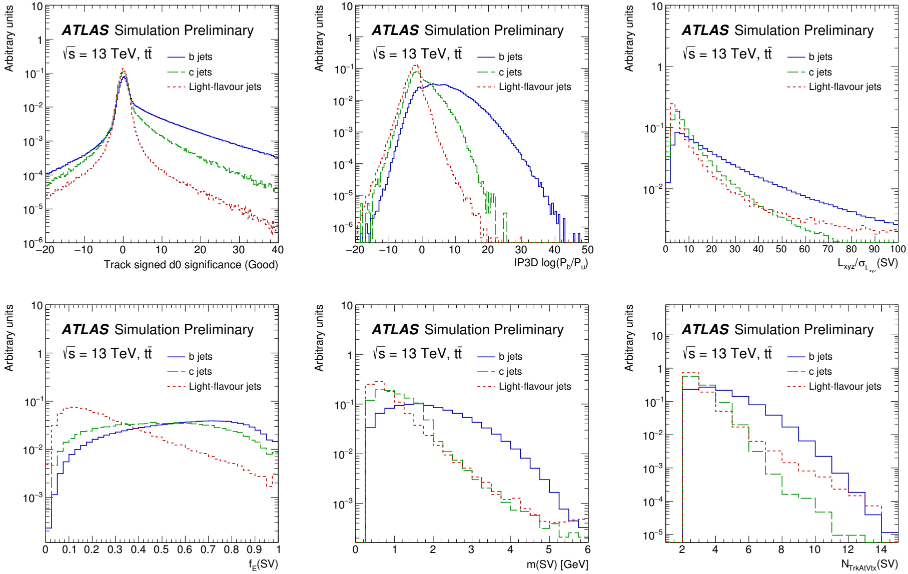
- Search for $\boldsymbol{ZH(c\bar{c})}$ production exploiting new c-tagging techniques provides limit of $\boldsymbol{\sigma(pp \to ZH) \times \mathcal{B}(H \to c\bar{c}) < 2.7\,\text{pb}}$ excluding **110× SM expectation**
- Demonstrates that this inclusive channel is likely more sensitive to the charm quark Yukawa coupling than the exclusive $H \to J/\psi\,\gamma$ channel
- Not yet able to compete with constraints obtained from interpreting measurements of Higgs boson kinematic distributions in terms of modified $gc \to Hc$ production
- Clear that **no single approach can yet claim it will manage to probe the charm quark Yukawa coupling down to the SM prediction** by the end of the LHC era
- Likely that multiple approaches will be required, this channel will become ever more important as larger datasets are collected!

What next for inclusive $H \to c\bar{c}$ decays?

- Large gains in sensitivity possible with multivariate techniques and other VH channels (e.g. $W(\ell\nu)/Z(\nu\nu)$) or a dedicated search/category in the high p_T^H boosted regime
- If future c-tagging algorithms can reach the performance of today's b-tagging, one could probably expect to observe $H \to c\bar{c}$ decays by the end of the LHC programme!
- Performance of c-tagging is developing rapidly, next generation algorithms already exploit advanced ML techniques (ATL-PHYS-PUB-2017-013), huge scope for innovation!

Additional Slides

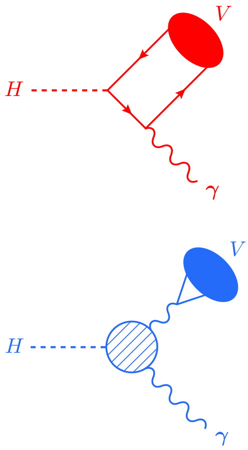
Examples of *c*-tagging input variables



More details in ATL-PHYS-PUB-2016-012

$H \to \mathcal{Q}\,\gamma$ decays could provide a clean probe of the charm and light quark couplings

- $\mathcal{Q}$ is a vector ($J^{PC} = 1^{--}$) light meson or quarkonium state such as $V = J/\psi,\ \phi,\ \rho(770)$
- **Interference** between **direct** ($H \to q\bar{q}$) and **indirect** ($H \to \gamma\gamma^{*}$) contributions
- **Direct** (upper diagram) amplitude provides sensitivity to the **magnitude and sign** of the $Hq\bar{q}$ couplings (i.e. $\mathcal{Q} = J/\psi$ sensitive to $Hc\bar{c}$ coupling)
- **Indirect** (lower diagram) amplitude provides dominant contribution to the width, not sensitive to Yukawa couplings
- Very rare decays in the SM!



$$\mathcal{B}\,(H \to J/\psi\,\gamma) = (2.8 \pm 0.2) \times 10^{-6} \quad \ddagger$$
$$\mathcal{B}\,(H \to \phi\,\gamma) = (2.3 \pm 0.1) \times 10^{-6} \quad \dagger$$
$$\mathcal{B}\,(H \to \rho\,\gamma) = (1.7 \pm 0.1) \times 10^{-5} \quad \dagger$$

More details: † JHEP 1508 (2015) 012 (arXiv:1505.03870) and ‡ Phys. Rev. D 90, 113010 (2014) (arXiv:1407.6695)

First search for such rare Higgs decays was performed by ATLAS with Run 1 dataset

PRL **114**, 121801 (2015) PHYSICAL REVIEW LETTERS week ending 27 MARCH 2015

Search for Higgs and Z Boson Decays to $J/\psi\gamma$ and $\Upsilon(nS)\gamma$ with the ATLAS Detector

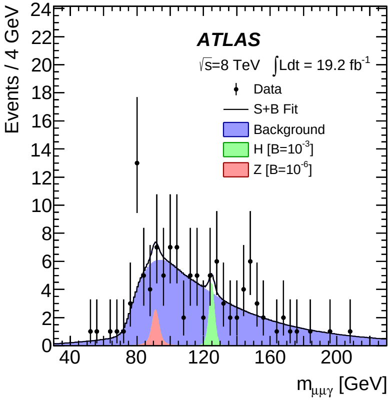
G. Aad *et al.**
(ATLAS Collaboration)

(Received 15 January 2015; published 26 March 2015)

A search for the decays of the Higgs and Z bosons to $J/\psi\gamma$ and $\Upsilon(nS)\gamma$ ($n = 1, 2, 3$) is performed with pp collision data samples corresponding to integrated luminosities of up to 20.3 fb^{-1} collected at $\sqrt{s} = 8$ TeV with the ATLAS detector at the CERN Large Hadron Collider. No significant excess of events is observed above expected backgrounds and 95% C.L. upper limits are placed on the branching fractions. In the $J/\psi\gamma$ final state the limits are 1.5×10^{-3} and 2.6×10^{-6} for the Higgs and Z boson decays, respectively, while in the $\Upsilon(1S, 2S, 3S)\gamma$ final states the limits are $(1.3, 1.9, 1.3) \times 10^{-3}$ and $(3.4, 6.5, 5.4) \times 10^{-6}$, respectively.

DOI: 10.1103/PhysRevLett.114.121801 PACS numbers: 14.80.Bn, 13.38.Dg, 14.70.Hp, 14.80.Ec

- Studied quarkonium decays, in particular $H \to J/\psi\,\gamma$ (with $J/\psi \to \mu^+\mu^-$)
- Similar limit subsequently found by CMS†
- **First direct information** on decay modes sensitive to the $Hc\bar{c}$ coupling
- Interpreted as $Hc\bar{c}$ coupling limit of $y_c/y_c^{SM} < 220$ at 95% CL‡ (assuming dependence on $\sigma(pp \to H)/\Gamma_H$ is removed by considering ratio with $H \to 4\ell$ rate)



Branching fraction limit (95% CL):

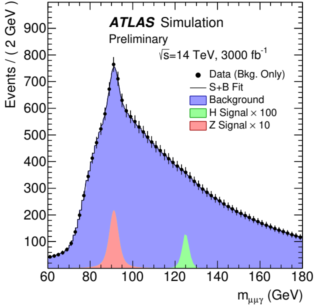
$$\mathcal{B}\,(H \to J/\psi\,\gamma) < 1.5 \times 10^{-3}$$

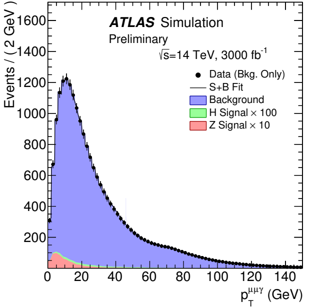
Around 500× the SM expectation

† Phys. Lett. B753 (2016) 341 (arXiv:1507.03031)

‡ Phys. Rev. D92, 033016 (2015) (arXiv:1503.00290)

Prospects for $H \to J/\psi\,\gamma$ in a HL-LHC scenario





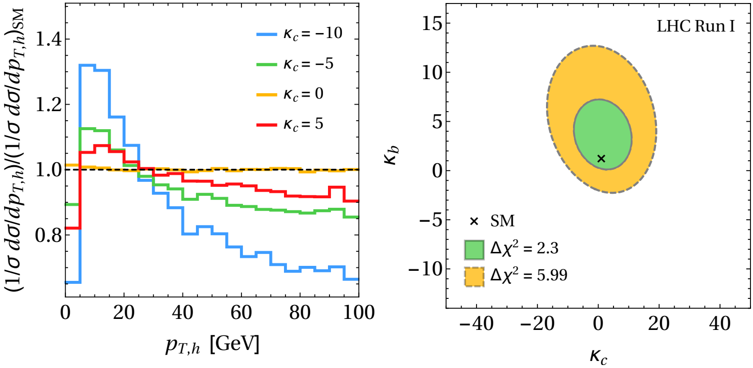
Run 1 $H \to J/\psi\,\gamma$ analysis projected to $\sqrt{s} = 14$ TeV scenario with 300(0) fb^{-1}

	Expected branching ratio limit at 95% CL		
	$\mathcal{B}(H \to J/\psi\gamma)$ [10^{-6}]		$\mathcal{B}(Z \to J/\psi\gamma)$ [10^{-7}]
	Cut Based	Multivariate Analysis	Cut Based
300 fb^{-1}	185^{+81}_{-52}	153^{+69}_{-43}	$7.0^{+2.7}_{-2.0}$
3000 fb^{-1}	55^{+24}_{-15}	44^{+19}_{-12}	$4.4^{+1.9}_{-1.1}$

	Standard Model expectation	
	$\mathcal{B}(H \to J/\psi\gamma)$ [10^{-6}]	$\mathcal{B}(Z \to J/\psi\gamma)$ [10^{-7}]
	2.9 ± 0.2	0.80 ± 0.05

- Optimistic scenario with MVA analysis still only sensitive to $\mathcal{B}(H \to J/\psi\,\gamma)$ **15× SM value with 3000 fb^{-1}**
- New ideas likely required to reach SM sensitivity in a HL-LHC scenario!

More details in <u>ATL-PHYS-PUB-2015-043</u>



↑ **Left: Effect of modified κ_c on p_T^H from $cg \to Hc$ diagrams Right: bounds from Run 1 data (both from arXiv:1606.09253)**

- In the case of a modified heavy quark $\mathcal{Q} = c, b$ Yukawa coupling, the shape of the inclusive p_T^H spectrum would change due to the modified $g\mathcal{Q} \to H\mathcal{Q}$ contribution
- p_T^H can be measured in the $H \to \gamma\gamma$ and $H \to 4\ell$ channels, which imposes a 95% CL bound of $-16 < y_c/y_c^{SM} < 18$ (arXiv:1606.09253, based on ATLAS+CMS Run 1)
- Projecting to HL-LHC scenario with $3\,\text{ab}^{-1}$, bound evolves to $-0.6 < y_c/y_c^{SM} < 3.0$