COFUND. A project supported by the European Union

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

UCL HEP Seminar

16th November 2018

Andy Chisholm
(CERN and University of Birmingham)

“Yukawa” couplings between the Higgs (ϕ) and fermion (ψ) fields are possible:

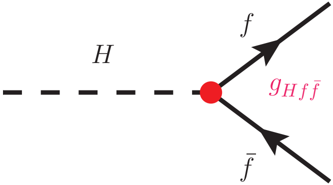
$$\mathcal{L}_{\text{fermion}} = -y_f \cdot \left[\bar{\psi}_L \phi \psi_R + \bar{\psi}_R \bar{\phi} \psi_L\right]$$

If ϕ has a non-zero VEV, expansion leads to (where h is the physical Higgs field):

$$\mathcal{L}_{\text{fermion}} = \underbrace{-\frac{y_f v}{\sqrt{2}} \cdot \bar{\psi}\psi}_{\text{mass term}} - \underbrace{\frac{y_f}{\sqrt{2}} \cdot h\bar{\psi}\psi}_{\text{Yukawa coupling term}}$$

Results in Higgs–fermion coupling proportional to the fermion mass ($g_{Hf\bar{f}} = m_f/v$)

- Gauge invariant fermion mass terms in SM ✓
- y_f “predicted” in SM given knowledge of v and m_f ($v \approx 246$ GeV from EW observables) ✓
- Offers no fundamental insight into the observed fermion mass hierarchy ✗

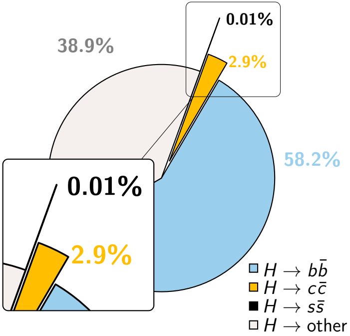


While Yukawa couplings provide concrete predictions for $Hf\bar{f}$ interactions, they fail to describe the origin of the fermion mass hierarchy i.e. why is $m_t/m_e \approx \mathcal{O}(10^5)$!?

Physics beyond the SM is clearly required to explain the fermion mass hierarchy!

Why is the charm quark Yukawa coupling important?

- The smallness of the SM charm (c) quark coupling ($y_c = \frac{\sqrt{2}m_c(m_H)}{v} \approx 4 \times 10^{-3}$) make possible **modifications from potential new physics easier to spot**
- $H \to c\bar{c}$ decays constitute the **largest part of the SM prediction for Γ_H for which we have no experimental evidence**
- **We only have experimental evidence for 3rd generation Yukawa couplings!**
- Many BSM models **predict modifications to 1st and 2nd generation fermion Higgs couplings alone**, with SM-like couplings to 3rd



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 $\mathcal{B}(H \to c\bar{c}) < 20\%$, while global fits indirectly bound Γ_H leading to $y_c/y_c^{SM} < 6$, **assuming SM 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)

Several methods to study the $Hc\bar{c}$ coupling at the LHC have been proposed in the literature, the most promising (in my opinion) are:

Idea 1 - Exclusive $H \to J/\psi\,\gamma$ decays

- Rare exclusive radiative Higgs boson decays to vector mesons are sensitive to the $Hq\bar{q}$ couplings (arXiv:1503.00290)
- The $H \to J/\psi\,\gamma$ decay has been proposed as a clean probe of the $Hc\bar{c}$ coupling, though decay width "only" evolves as $(\text{const.} + y_c)^2$ ($\text{const.} \gg y_c$)
- ATLAS pioneered searches in this channel during Run 1 (arXiv:1501.03276)

Idea 2 - Associated production of a Higgs boson and charm quark

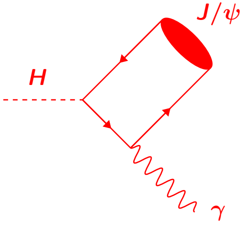
- Tree level sensitivity to $Hc\bar{c}$ coupling (arXiv:1507.02916, arXiv:1606.09253)
- Use jet c-tagging to identify charm quark signature and a suitably "clean" Higgs decay (e.g. $H \to \gamma\gamma$)
- Alternatively, study p_T^H distribution to look for potential shape modifications...

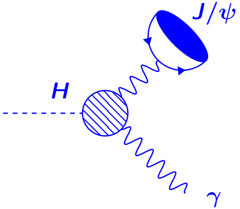
Idea 3 - Inclusive $H \to c\bar{c}$ decays (The focus of this seminar...)

- Inclusive $H \to c\bar{c}$ decays are directly sensitive to the $Hc\bar{c}$ coupling, with the decay width evolving as $\Gamma_{H\to c\bar{c}} \propto y_c^2$
- Use double jet c-tagging and focus on VH ($V = W, Z$) production with leptonic V decays to mitigate the large multi-jet backgrounds

The radiative decay $H \to J/\psi\,\gamma$ could provide a clean probe of the $Hc\bar{c}$ coupling at the LHC

- **Interference** between **direct** ($H \to c\bar{c}$) and **indirect** ($H \to \gamma\gamma^*$) contributions
- **Direct** (upper diagram) amplitude provides sensitivity to the **magnitude and sign** of the $Hc\bar{c}$ coupling
- **Indirect** (lower diagram) amplitude provides dominant contribution to the width, not sensitive to $Hc\bar{c}$ coupling
- Very rare decays in the SM, but **rate dominated by "indirect" component**, sensitivity to $Hc\bar{c}$ coupling somewhat diluted

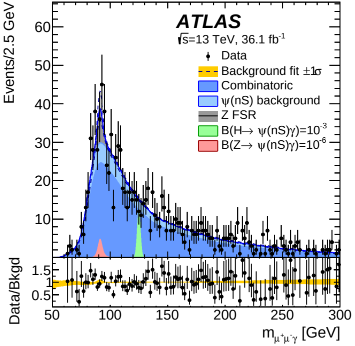


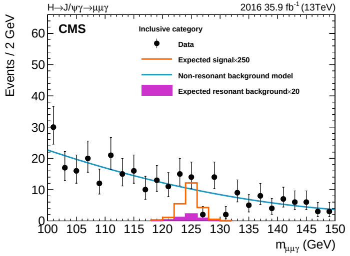


$$\Gamma = \left|C_{\mathrm{I}} - C_{\mathrm{D}} \cdot \frac{y_c}{y_c^{SM}}\right|^2 \times 10^{-7}\ \mathrm{MeV}\ (C_{\mathrm{I}} \approx 10, C_{\mathrm{D}} \approx 1)$$

$$\mathcal{B}(H \to J/\psi\,\gamma) = (2.99 \pm 0.16) \times 10^{-6}$$

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





Recently both ATLAS and CMS updated their searches for $H \to J/\psi\,\gamma$ decays with 36 fb^{-1} of $\sqrt{s} = 13$ TeV Run 2 data

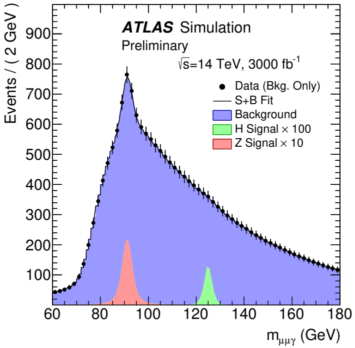
- Both search for $H \to J/\psi\,\gamma$ with $J/\psi \to \mu^+\mu^-$ using a "cut-based" analysis
- Sensitive to branching fractions around two orders of magnitude away from SM prediction
- Limits corresponds to $|y_c/y_c^{SM}| \approx 100$ (when considered relative to $H \to \gamma\gamma$ to remove Γ_H dependence)

Expt.	95% CL upper limit on $\mathcal{B}(H \to J/\psi\,\gamma)$		
	Expected	**Observed**	**Obs./$\mathcal{B}_{SM}$**
ATLAS†	$(3.0^{+1.4}_{-0.8}) \times 10^{-4}$	3.5×10^{-4}	**117×**
CMS‡	$(5.2^{+2.4}_{-1.6}) \times 10^{-4}$	7.6×10^{-4}	**253×**

† **Phys. Lett. B 786 (2018) 134 (arXiv:1807.00802)**

‡ **Submitted to EPJC (arXiv:1810.10056)**

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



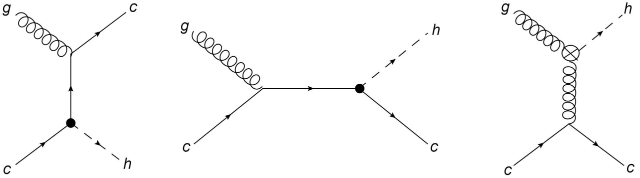
	Expected branching ratio limit at 95% CL		
	$\mathcal{B}(H \to J/\psi\gamma)$ [10^{-6}]		$\mathcal{B}(Z \to J/\psi\gamma)$ [10^{-7}]
	Cut Based	Multivariate Analysis	Cut Based
300 fb^{-1}	185^{+81}_{-52}	153^{+69}_{-43}	$7.0^{+2.7}_{-2.0}$
3000 fb^{-1}	55^{+24}_{-15}	44^{+19}_{-12}	$4.4^{+1.9}_{-1.1}$

	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

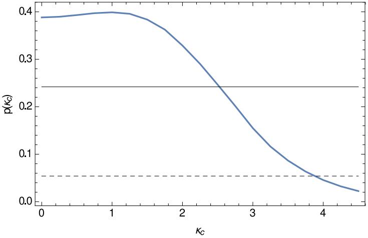
- Optimistic scenario with MVA analysis still only sensitive to $\mathcal{B}(H \to J/\psi\,\gamma)$ at **15× SM value with 3000 fb^{-1}**

New ideas likely required to reach SM sensitivity in a HL-LHC scenario with this channel!

The production of Higgs boson in association with a charm quark is directly sensitive to the charm quark Yukawa coupling



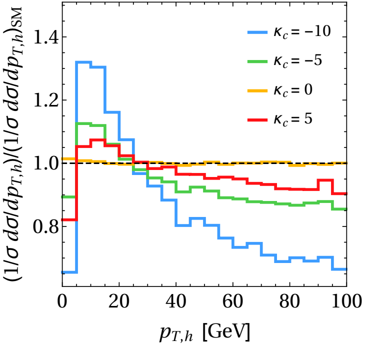
↑ **Examples of "direct" (left and centre) and "indirect" (right) $cg \to Hc$ diagrams (from arXiv:1507.02916)**



↑ **Expected p-value as a function of $\kappa_c = y_c/y_c^{SM}$ (from arXiv:1507.02916)**

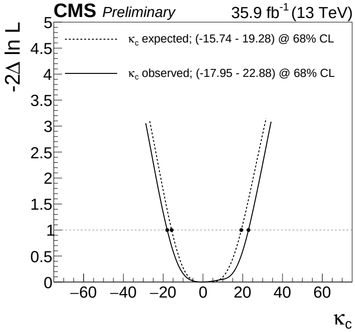
- t-channel diagram (left) is expected to dominate the cross-section and is sensitive to the $Hc\bar{c}$ coupling, highly sensitive channel!
- No experimental measurements yet, though the sensitivity at the HL-LHC has been surveyed in the literature (arXiv:1507.02916)
- Assuming a data sample of $3\,\text{ab}^{-1}$ at $\sqrt{s} = 14$ TeV, $\mathcal{O}(1)$ constraints on y_c/y_c^{SM} are expected to be obtained...

Idea 2 - Associated Higgs boson + charm quark production



Effect of modified y_c on p_T^H from $cg \to Hc$ diagrams

(Phys. Rev. Lett. 118, 121801 (2017), arXiv:1606.09253)



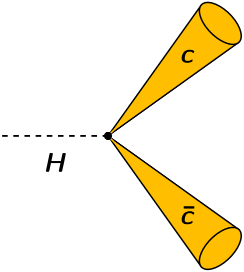
Bound on y_c/y_c^{SM} from Run 2 CMS data

(CMS-PAS-HIG-17-028)

- In the case of a modified Higgs coupling to heavy quarks $Q = c, b$, the shape of the inclusive p_T^H spectrum would change due to the modified $gQ \to HQ$ contribution
- Recently, CMS used their measured p_T^H distribution from $H \to \gamma\gamma$ and $H \to 4\ell$ accounting for dependence on y_c (and y_b)
- Considering only shape variation (no assumption on Γ_H, less model dependent) and profiling y_b/y_b^{SM}, obtain constrain of $-18 < y_c/y_c^{SM} < 23$ at 68% CL

Motivation

- The branching fraction for $H \to c\bar{c}$ decays is around 2.9% for a SM Higgs boson with $m_H = 125$ GeV
- In comparison to the $H \to J/\psi\,\gamma$ decay, this is a huge rate! Furthermore, it scales directly with y_c^2...
- In $\sqrt{s} = 13$ TeV pp collisions, one expects around 1600 $H \to c\bar{c}$ decays in every $1\,\text{fb}^{-1}$ of data!
- **But**, how can we hope to separate $H \to c\bar{c}$ from the **HUGE** jet background at the LHC?



Strategy

- Charm quark initiated jets (c-jet) will typically contain a c-hadron, though most of the jets produced in LHC pp collisions will not...
- If we can exploit the presence of a c-hadron within the jet, we can hope to separate c-jets from light flavour (u, d, s, g) and b-jets (which also have a unique signature)
- Focus on production channels involving leptons or large $E_{\mathrm{T}}^{\mathrm{miss}}$ (e.g. $Z(\ell\ell, \nu\nu)H$ and/or $W(\ell\nu)H$), to reduce the jet backgrond

Part I - Charm jet tagging with ATLAS

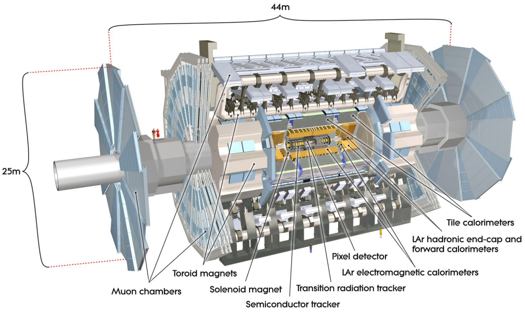
Introduction

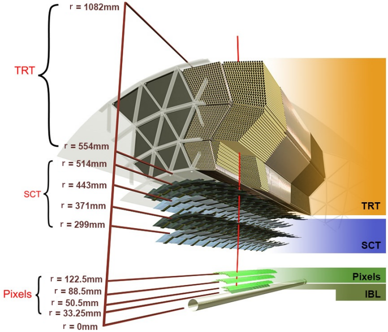
- Jets containing either c- or b-hadrons can be "tagged" by virtue of the unique properties of the heavy flavour hadrons
- These techniques are collectively known as jet "flavour tagging" and only differ in the fine details if one is interested to "tag" c-jets or b-jets
- **I will describe how these techniques are implemented within the ATLAS experiement** ("flavour tagging" can mean different things to different collider experiments)

Jet Labelling Conventions

- ***b*-jet:** Jets containing a b-hadron
- ***c*-jet:** Jets containing a c-hadron but no b-hadron
- **Light flavour jet:** Jets containing no b or c-hadrons (originating from u, d, s quark and gluon fragmentation)

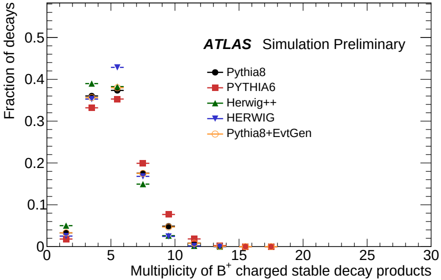
General purpose detector, well suited to studying heavy flavour jets

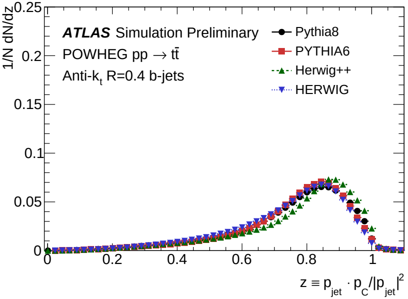




- **Inner Detector (ID):** Silicon Pixels and Strips (SCT) with Transition Radiation Tracker (TRT) $|\eta| < 2.5$ and (new for Run 2) Insertable B-Layer (IBL)
- **LAr EM Calorimeter:** Highly granular + longitudinally segmented (3-4 layers)
- **Had. Calorimeter:** Plastic scintillator tiles with iron 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)

- **Lifetime:** Long enough to lead to a measureable decay length (around 5mm for a 50 GeV boost)
- **Mass:** Weakly decaying *b*-hadrons have masses around 5 GeV, leading to high decay product multiplicities (average of 5 charged particles per decay)
- **Fragmentation:** Much harder than jets initiated by other species (*b*-hadrons carry around 75% of jet energy, on average)

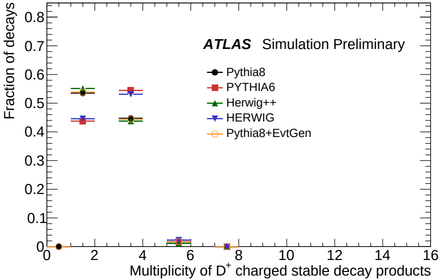


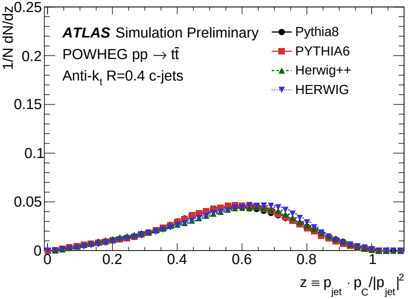


Left: Mean charged multiplicity in B^+ mesons decays

Right: *b*-quark fragmentation function

- **Lifetime:** Shorter than the *b*-hadrons by around a factor of 2-3, still enough for measureable decay length (around 1-3mm for a 50 GeV boost)
- **Mass:** Weakly decaying *c*-hadrons have masses around 2 GeV, around 2–3$\times$ lower than *b*-hadrons (mean of $\approx$ 2 charged particles per decay)
- **Fragmentation:** Softer than *b*-jets, but still harder than jets initiated by light species (*c*-hadrons carry around 55% of jet energy, on average)

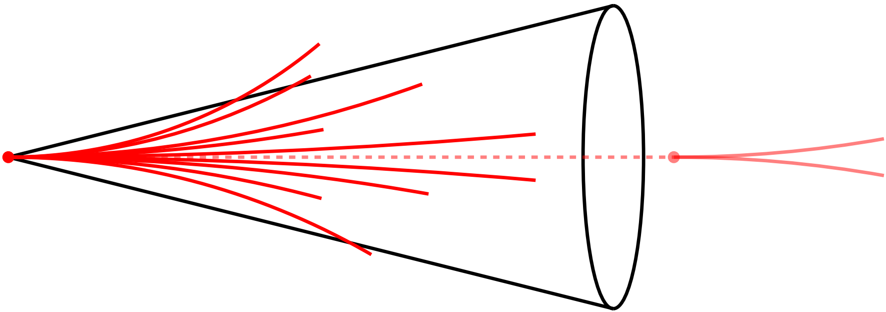




Left: Mean charged multiplicity in D^+ mesons decays

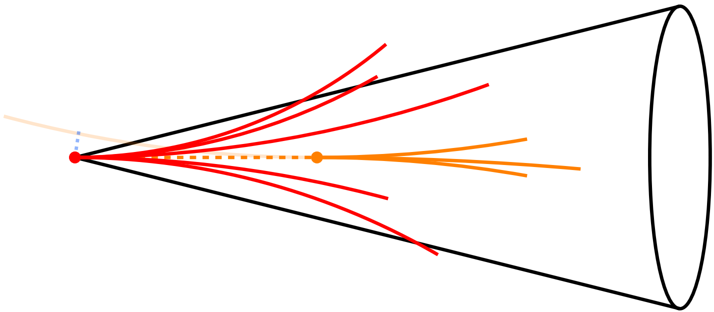
Right: *c*-quark fragmentation function

Anatomy of a light flavour (u, d, s) jet



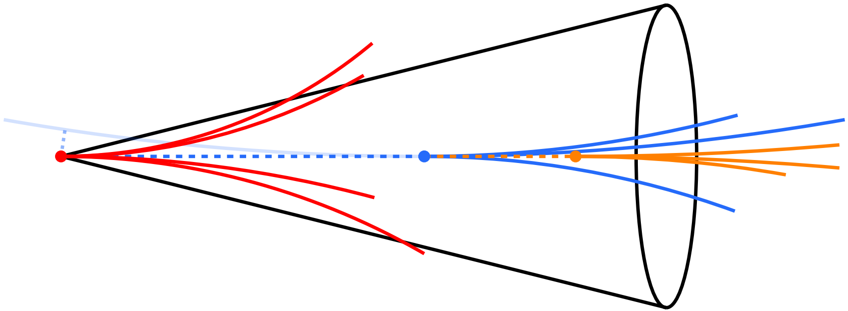
Typical Experimental Signature

- **Light-quarks hadronise into many light hadrons which share the jet energy**
- **Tracks from this vertex often have impact parameters consistent with zero**
- **Long-lived light hadrons (e.g. K_S^0, Λ^0) can be produced, though they are more likely to decay very far (many cm) from the primary *pp* vertex**



Typical Experimental Signature

- ***c*-quark fragments into a *c*-hadron which carries around half of the jet energy**
- ***c*-hadron decay vertex often displaced from the primary *pp* vertex by a few mm**
- **Tracks from this vertex can often have large impact parameters**



Typical Experimental Signature

- ***b*-quark fragments into a *b*-hadron which carries most of the jet energy**
- **Most *b*-hadrons ($\approx$ 90%) decay into *c*-hadrons**
- ***b*-hadron decay vertex often displaced from the primary *pp* vertex by a few mm**
- **Subsequent *c*-hadron decay vertex often displaced by a further few mm**
- **Tracks from both of these vertices often have large impact parameters**

Charm tagging is not new, many experiments at high energy ($\sqrt{s} \gg m_{B\bar{B}}$) colliders (e.g. Sp$\bar{\text{p}}$S, Tevatron, SLD, LEP, HERA) have built "charm taggers" which tend to fall within the following classes:

"Exclusive" charm jet tagging

- Focus on the full reconstruction of exclusive c-hadron decay chains (e.g. $D^{\star\pm} \to D^0(K^-\pi^+)\pi^\pm$) or leptons from semi-leptonic c-hadron decays
- ✓ Can often provide a very pure sample of jets containing c-hadrons
- ✗ The efficiency is typically low $\mathcal{O}(1\%)$, limited by the c-hadron branching fractions of interest

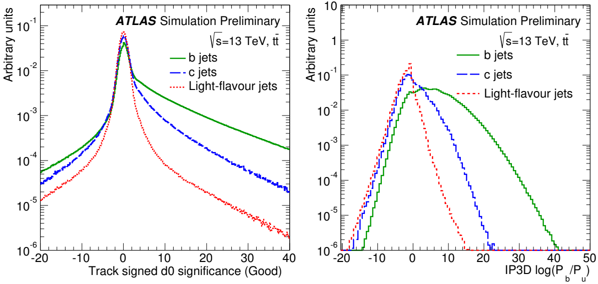
"Inclusive" charm jet tagging

- An alternative approach is to to exploit more "inclusive" observables, such as track impact parameters or secondary vertices
- ✓ The efficiency of this approach is typically very high $\mathcal{O}(10\%)$)
- ✗ The c-jet purity is often lower than these "traditional" approaches
- More suited for use with machine learning (ML) techniques

ATLAS have developed an "inclusive" c-tagging algorithm based on several "low level" taggers combined into a "high level" tagger using ML techniques

The signed IPs of tracks associated to jets are powerful jet flavour distriminants:

- Exploit "sign" of impact parameter: positive if track point of closest approach to PV is downstream of plane defined by the PV and jet axis
- Tracks from *b*-hadrons tend to have highly significant (IP/σ_{IP}) positive IPs, while most tracks from the PV have a narrow, symmetric distribution
- ✓ Very inclusive and highly efficient
- ✗ Relies upon accurate measurement of jet axis, sensitive to "mis-tag" high IP tracks from V^0 decays or material interactions, IP/σ_{IP} difficult to model in detector simulation

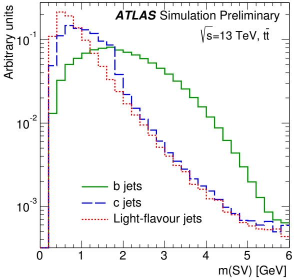


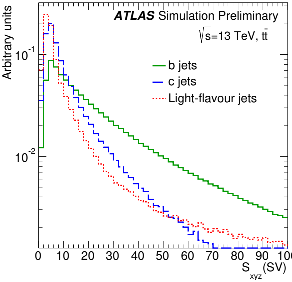
Left: Transverse IP significance distribution **Right: likelihood ratio discriminant based on 3D IPs of tracks**

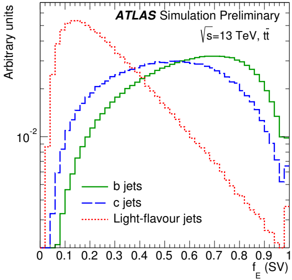
ATLAS Low Level Taggers: 2 - Secondary Vertices (SV)

Exploit expectation of a secondary vertex from either *b* or *c*-hadron decays:

- Attempt to reconstruct a secondary vertex from high IP tracks associated with jet
- Use invariant mass of tracks at SV to discriminate *b* or *c*-hadron decay vertices from V^0 decays or material interations
- Exploit hard c/b-jet fragmentation, SV should carry a large fraction of jet energy
- ✓ SV found in up to ≈ 80% of *b*-jets but only a few % of light flavour jets
- ✗ Degraded light jet rejection as jet p_T increases, careful considerations to mitigate "tagging" of material interactions required



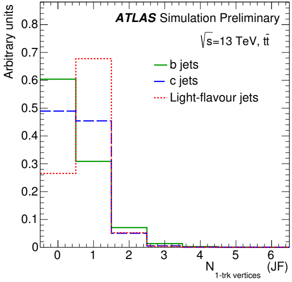


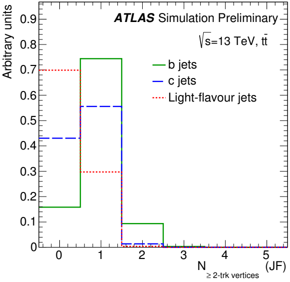


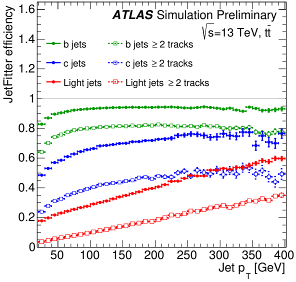
Left: Inv. mass of tracks at SV **Centre: 3D SV decay length significance** **Right: Energy fraction of SV tracks**

Exploit common occurance of cascade decay chain; *b*-hadron $\rightarrow$ *c*-hadron:

- Use Kalman filter to search for common axis on which three vertices lie: primary (pp) $\rightarrow$ secondary (b-hadron) $\rightarrow$ tertiary (c-hadron)
- Can then look for "1 track vertices" with decay chain axis
- ✓ Addition of 1 track vertices improves efficiency, constraint to decay chain axis improves separation power of SV based discriminants
- ✗ Degraded performance for c/b-hadron vertices as jet p_T increases, high fake rate for 1 track vertices (increases light jet "mis-tag" rate)





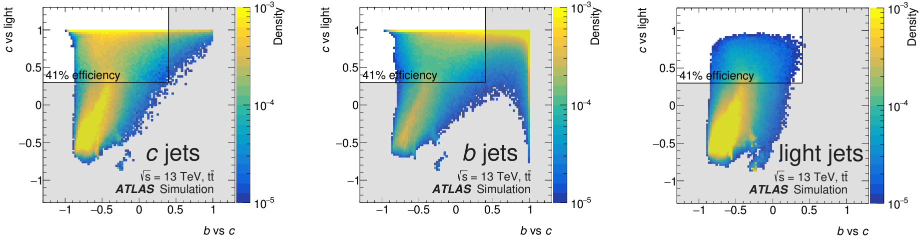


Left: Multiplicity of 1 track vertices **Centre: Multiplicity of 2+ track vertices** **Right: Reco. efficiency vs. jet p_T**

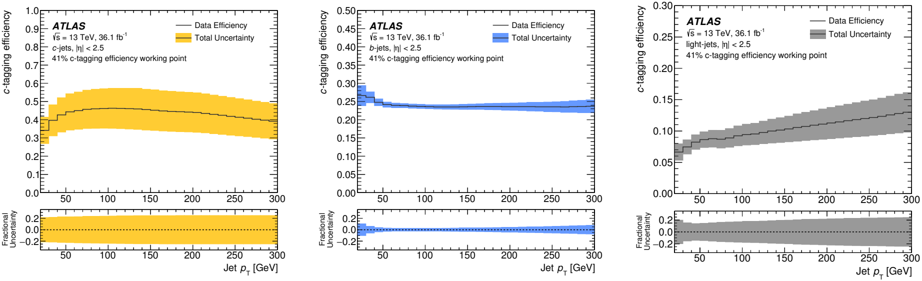
Combine approaches to exploit all features of *c*/*b*-jets and mitigate the shortcomings of the individual methods:

- ✓ Benefit from the advantages of all basic techniques/algorithms
- ✗ Complex sensitivity to convolution of all detector and physics modelling issues relies strongly on "calibration" in data (see next slide)
- Use the output of the three basic approaches as input to a boosted decision tree (BDT) to build two discriminants, one trained to separate *c*-jets from *b*-jets (x-axis), another to separate *c*-jets from light-jets (y-axis)

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



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



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

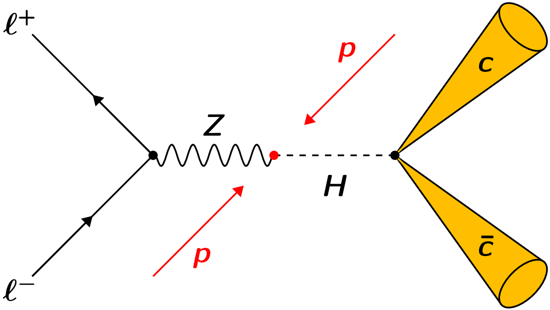
_c_-tagging efficiency for _b_-, _c_- and light flavour jets measured in data ↑

- Working point for $H \to c\bar{c}$ exhibits a c-jet tagging efficiency of around 40%
- Rejects b-jets by around a factor $4\times$ and light jets by around a factor $10\times$
- Efficiency calibrated in data with samples of b-jets from $t \to Wb$ decays and c-jets from $W \to cs, cd$ decays (in $t\bar{t}$ events)
- Typical total relative uncertainties of around **25%**, **5%** and **20%** for ***c*-**, ***b*-** and **light** jets, respectively

Part II - Search for $H \rightarrow c\bar{c}$ decays with ATLAS

How can we use the "charm tagger" to search for $H \rightarrow c\bar{c}$ decays?

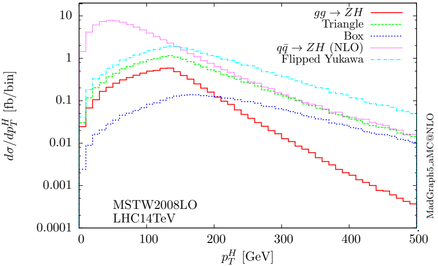
Given the success of the W/Z associated production channel in observing $H \to b\bar{b}$ decays[†], this channel is an obvious first candidate for a $H \to c\bar{c}$ search



- **Focus on ZH production** with $Z \to e^+e^-$ and $Z \to \mu^+\mu^-$ decays for first ATLAS analysis: **Phys. Rev. Lett. 120 (2018) 211802, arXiv:1802.04329**
- Low exposure to experimental uncertainties, main backgrounds from $Z+$jets, $Z(W/Z)$ and $t\bar{t}$
- Pioneer use of **new c-tagging algorithm** developed by ATLAS for Run 2 to identify the experimental signature of an inclusive $H \to c\bar{c}$ decay

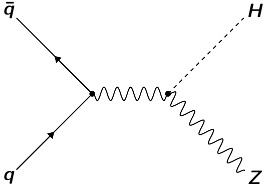
† ATLAS: Phys. Lett. B 786 (2018) 59 CMS: Phys. Rev. Lett. 121 (2018) 121801

- In $\sqrt{s} = 13$ TeV pp collisions, Higgs boson production in association with a Z boson represents around 1.6% of the inclusive production rate
- The cross-section is dominated by the $q\bar{q} \to ZH$ process, with total cross-section $\sigma_{q\bar{q}} \approx 0.76\,\text{pb}$
- Smaller contributions from $gg \to ZH$, with total cross-section $\sigma_{gg} \approx 0.12\,\text{pb}$, though it exhibits a harder p_T^H spectrum below ≈ 150 GeV

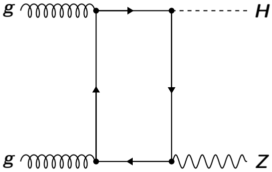


↑ **p_T^H distribution for $q\bar{q}$ and gg initiated ZH production (from arXiv:1503.01656)**

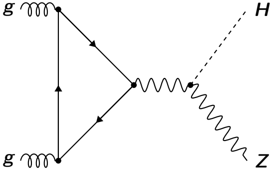
Representative Feynman diagrams for $q\bar{q}/gg \to ZH$ processes →



$q\bar{q} \to ZH$ diagram



$gg \to ZH$ "box" diagram



$gg \to ZH$ "triangle" diagram

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

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

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

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

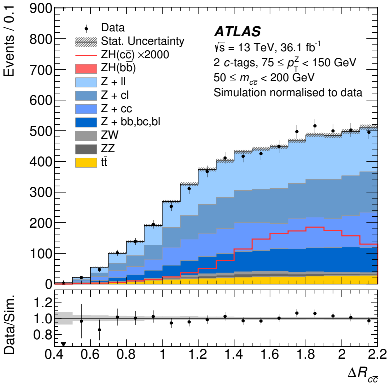
- Consider anti-k_T $R = 0.4$ calorimeter jets with $|\eta| < 2.5$ and $p_T > 20$ GeV
- At least two jets with leading jet $p_T > 45$ GeV
- Form $H \to c\bar{c}$ candidate from the two highest p_T jets in an event
- At least one c-tagged jet from $H \to c\bar{c}$ candidate
- Dijet angular separation 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

Signal and Background Modelling

Background Modelling

- Background dominated by $Z+\text{jets} \rightarrow$ (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 $W+\text{jets}$, WW, single-top and multi-jet



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

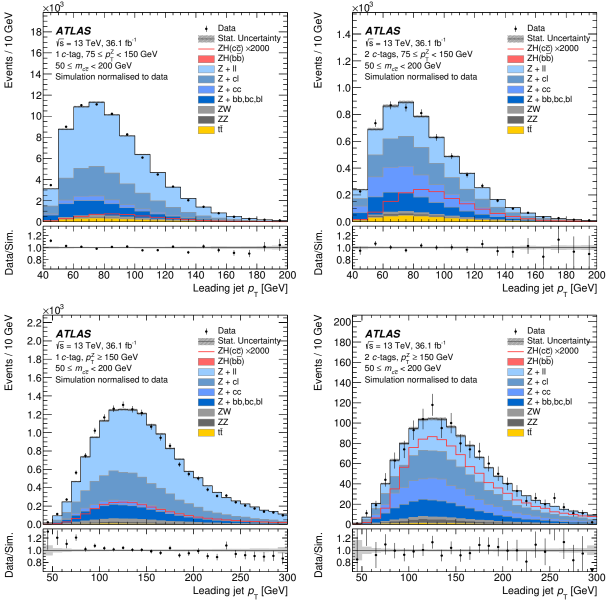
- Normalised with LHC Higgs XS WG YR4 recommendations (arXiv:1610.07922)
- $ZH(b\bar{b})$ treated as background normalised to SM expectation (with th. uncertainty)

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)
$Z+\text{jets}$	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

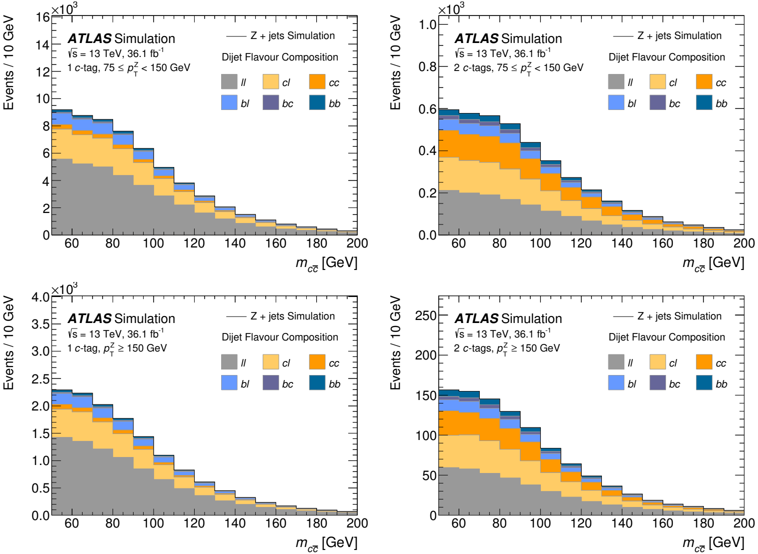
↓ Left: 1 c-tag events



↑ Right: 2 c-tag events

$Z+$ jets flavour composition after c-tagging

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



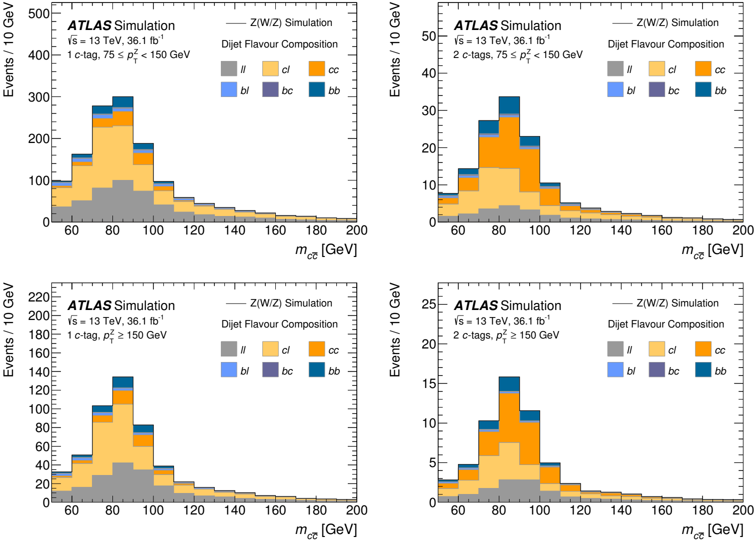
↓ Left: 1 c-tag events

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

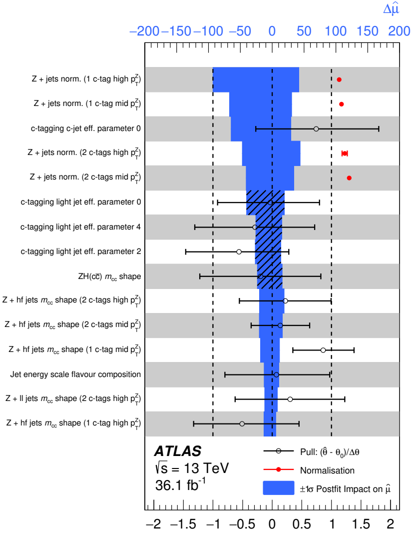
Statistical Model

- 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

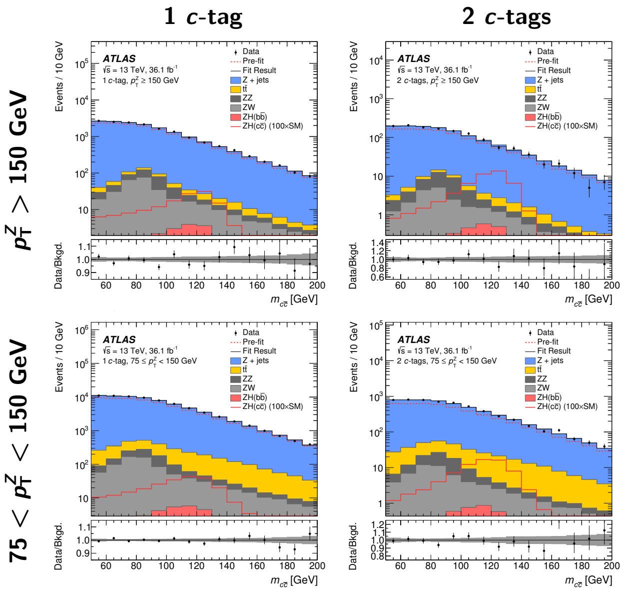
Sensitivity dominated by systematic uncertainties, clear that these uncertainties should be reduced in order to fully exploit a larger dataset in the future



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 norm.) to reduce with a larger dataset



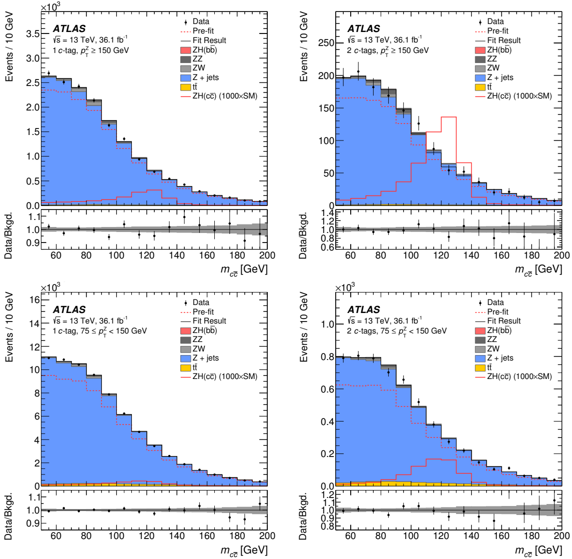
- No significant evidence for $ZH(c\bar{c})$ production
- Data consistent with background only hypothesis

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**

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**$\times$ $(150^{+80}_{-40}$ expected) the SM expectation

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

⚠ **None of the following interpretation is sanctioned by ATLAS, responsibility lies solely with me!** However, everything is calculated using *published information alone*...

Ultimate goal is derive a model independent constraint on $Hc\bar{c}$ coupling, best way to do this is to exploit synergy with $ZH, H \to b\bar{b}$ channel

- Consider the ratio of $\mu_{ZH(c\bar{c})}/\mu_{ZH(b\bar{b})}$ for the $Z \to \ell^+\ell^-$ channel
- Sensitive to ratio κ_c/κ_b and <u>independent</u> of model dependent assumption on Γ_H
- Assume production is <u>identical</u> between $ZH(c\bar{c})$ and $ZH(b\bar{b})$ (i.e. selection phase space, categories etc.), leading to <u>perfect cancellation</u> of production cross-sections

$$\mu_{ZH(c\bar{c})} = \frac{\Gamma_{H\to c\bar{c}}}{\Gamma^{\text{SM}}_{H\to c\bar{c}}} \cdot \frac{\Gamma^{\text{SM}}_H}{\Gamma_H} \cdot \frac{\sigma(pp \to ZH)}{\sigma^{\text{SM}}(pp \to ZH)} = \kappa_c^2 \cdot \frac{\Gamma^{\text{SM}}_H}{\Gamma_H} \cdot \frac{\sigma(pp \to ZH)}{\sigma^{\text{SM}}(pp \to ZH)}$$

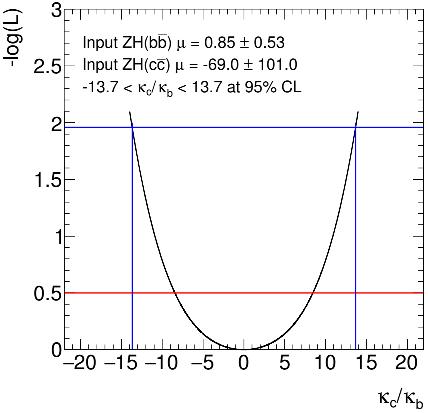
$$\mu_{ZH(b\bar{b})} = \frac{\Gamma_{H\to b\bar{b}}}{\Gamma^{\text{SM}}_{H\to b\bar{b}}} \cdot \frac{\Gamma^{\text{SM}}_H}{\Gamma_H} \cdot \frac{\sigma(pp \to ZH)}{\sigma^{\text{SM}}(pp \to ZH)} = \kappa_b^2 \cdot \frac{\Gamma^{\text{SM}}_H}{\Gamma_H} \cdot \frac{\sigma(pp \to ZH)}{\sigma^{\text{SM}}(pp \to ZH)}$$

$$\frac{\mu_{ZH(c\bar{c})}}{\mu_{ZH(b\bar{b})}} = \left(\frac{\kappa_c}{\kappa_b}\right)^2$$

- For now, consider systematic uncertainties for $ZH(c\bar{c})$ and $ZH(b\bar{b})$ as <u>uncorrelated</u>

What is the current sensitivity to κ_c/κ_b?

- Consider existing $ZH(c\bar{c})$ result and "combine" with recent ATLAS 80 fb^{-1} $Z(\ell\ell)H(b\bar{b})$ measurement†
- Small differences in selection and categories, but production cancellation hypothesis likely not too bad
- Treatment of systematics as un-correlated should give a more conservative constraint on κ_c/κ_b

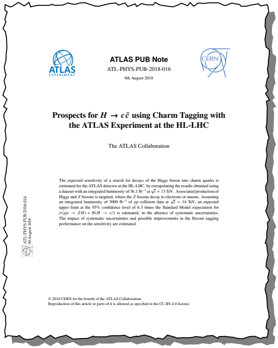


Existing results offer constraint at the level of $|\kappa_c/\kappa_b| < 14$ at 95% CL

- This is only possible when considering combination with $ZH(b\bar{b})$, not enough constraint (even with assumption for Γ_H) with $ZH(c\bar{c})$ analysis alone

† **Phys. Lett. B 786 (2018) 134 (arXiv:1807.00802)**

ATL-PHYS-PUB-2018-016


ATLAS PUB Note

ATL-PHYS-PUB-2018-016

8th August 2018

Prospects for $H \to c\bar{c}$ using Charm Tagging with the ATLAS Experiment at the HL-LHC

The ATLAS Collaboration

The expected sensitivity of a search for decays of the Higgs boson into charm quarks is estimated for the ATLAS detector at the HL-LHC, by extrapolating the results obtained using a dataset with an integrated luminosity of 36.1 fb^{-1} at $\sqrt{s}$ = 13 TeV. Associated production of Higgs and Z bosons is targeted, where the Z bosons decay to electrons or muons. Assuming an integrated luminosity of 3000 fb^{-1} of pp collision data at $\sqrt{s}$ = 14 TeV, an expected upper limit at the 95% confidence level of 6.3 times the Standard Model expectation for $\sigma(pp \to ZH) \times \mathcal{B}(H \to c\bar{c})$ is estimated, in the absence of systematic uncertainties. The impact of systematic uncertainties and possible improvements in the flavour tagging performance on the sensitivity are estimated.

ATL-PHYS-PUB-2018-016
08 August 2018

© 2018 CERN for the benefit of the ATLAS Collaboration.
Reproduction of this article or parts of it is allowed as specified in the CC-BY-4.0 license.

What sensitivity can we expect for a HL-LHC scenario with a $\sqrt{s} = 14$ TeV 3000 fb^{-1} dataset?

- A projection of the existing $Z(\ell\ell)H, H \to c\bar{c}$ analysis was prepared for the upcoming HL-LHC physics yellow report
- Generally very simiar to the Run 2 analysis, with several minor changes (described below)

Similarities

- Consider $Z(\ell\ell)H$ channel only (no addition of $W(\ell\nu)H$ or $Z(\nu\nu)H$)
- Identical event selection, categorisation and fit procedure

Differences

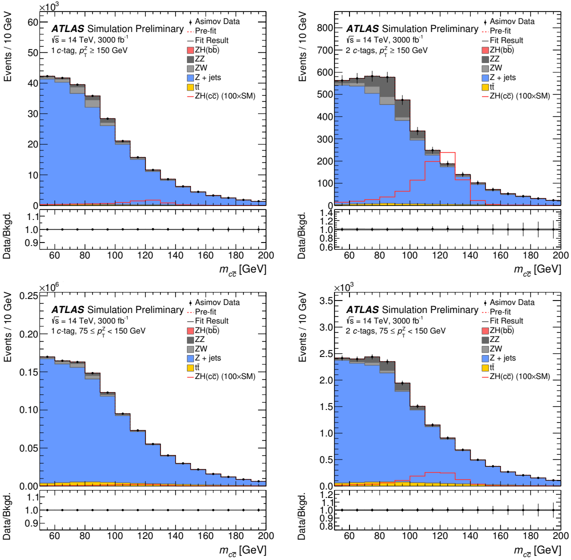
- Move to a tighter c-tagging working point (18% c-jet, 5% b-jets, 0.5% light jets)
- Don't consider systematic uncertainties (though their effect is estimated)

Prospects for $Z(\ell\ell)H, H \rightarrow c\bar{c}$ at the HL-LHC

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

$p_T^Z > 150$ GeV

$75 < p_T^Z < 150$ GeV



- Result of fit to expected ("Asimov") dataset for 3000 fb^{-1}
- Background composition (in terms of "process") very simiar
- Di-jet flavour composition now more c-jet enriched (you can't see that from these plots)

Projected Results

- Expected limit on $Z(\ell\ell)H, H \rightarrow c\bar{c}$ production at **$6.3\times$ SM prediction** at 95% CL (c.f. 150× expected for 36.1 fb^{-1} at 13 TeV)
- ⚠Corresponds to **around $|\kappa_c/\kappa_b| < 3$** (with naive scaling of ATLAS Run 2 $ZH(b\bar{b})$ result based on luminosity only)

Things to remember

- Limit deteriorates by up to +36% with the inclusion of systematic uncertainties (estimated from Run 2 analysis)
- Projection considers the **$Z(\ell\ell)H$ channel alone!** (sensitivity of $W(\ell\nu)H$ and $Z(\nu\nu)H$ channels at least as good)

⚠**As before, this is NOT an ATLAS result, but my estimate based on public information alone**

Summary

- Search for $pp \to ZH, H \to c\bar{c}$ production with c-tagging techniques provides limit of **$\sigma(pp \to ZH) \times \mathcal{B}(H \to c\bar{c}) < 2.7$ pb (110× SM expectation)** at 95% CL
- **Corresponds (roughly) to constraint of $|\kappa_c/\kappa_b| < 14$**, when considered within the context the latest ATLAS $ZH, H \to b\bar{b}$ measurement
- Limit expected to improve to **6× SM expectation for nominal HL-LHC scenario**
- This inclusive channel is more sensitive to the $Hc\bar{c}$ coupling than the $H \to J/\psi\,\gamma$ decay, but comparable to approaches based on modified $gc \to Hc$ production
- Clear that **no single approach can yet claim it will manage to probe the $Hc\bar{c}$ coupling down to the SM prediction** by the end of the LHC era

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

- Large gains in sensitivity possible with multivariate techniques and other VH channels ($W(\ell\nu)$ and $Z(\nu\nu)$)
- Performance of c-tagging is developing rapidly, next generation algorithms already exploit advanced ML techniques (ATL-PHYS-PUB-2017-013), huge scope for innovation!
- Much to gain (e.g. sensitivity to κ_c/κ_b) from synchronisation with $VH(b\bar{b})$ channel

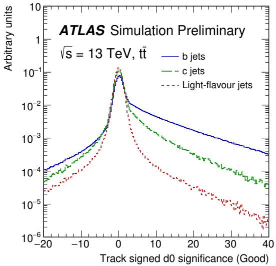
Thank you for your attention!

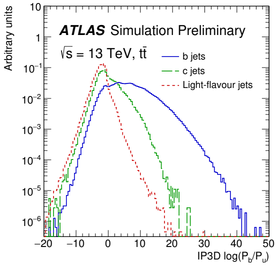
Additional Slides

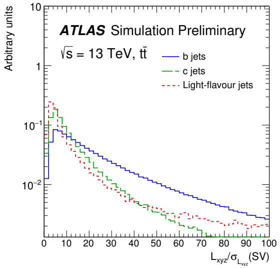
Background Composition

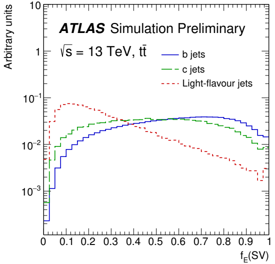
Sample	Yield, 50 $GeV < m_{c\bar{c}} <$ 200 GeV			
	1 c-tag		2 c-tags	
	$75 \leq p_{\mathrm{T}}^{Z} < 150\,GeV$	$p_{\mathrm{T}}^{Z} \geq 150\,GeV$	$75 \leq p_{\mathrm{T}}^{Z} < 150\,GeV$	$p_{\mathrm{T}}^{Z} \geq 150\,GeV$
$Z+$jets	69400 ± 500	15650 ± 180	5320 ± 100	1280 ± 40
ZW	750 ± 130	290 ± 50	53 ± 13	20 ± 5
ZZ	490 ± 70	180 ± 28	55 ± 18	26 ± 8
$t\bar{t}$	2020 ± 280	130 ± 50	240 ± 40	13 ± 6
$ZH(b\bar{b})$	32 ± 2	19.5 ± 1.5	4.1 ± 0.4	2.7 ± 0.2
$ZH(c\bar{c})$ (SM)	-143 ± 170 (2.4)	-84 ± 100 (1.4)	-30 ± 40 (0.7)	-20 ± 29 (0.5)
Total	72500 ± 320	16180 ± 140	5650 ± 80	1320 ± 40
Data	72504	16181	5648	1320

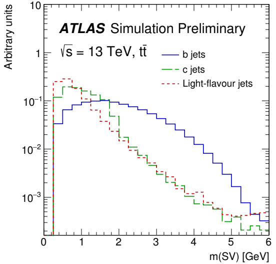
Examples of *c*-tagging input variables

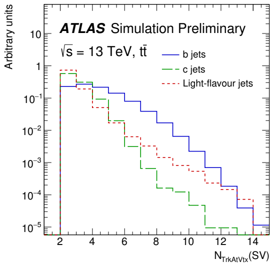








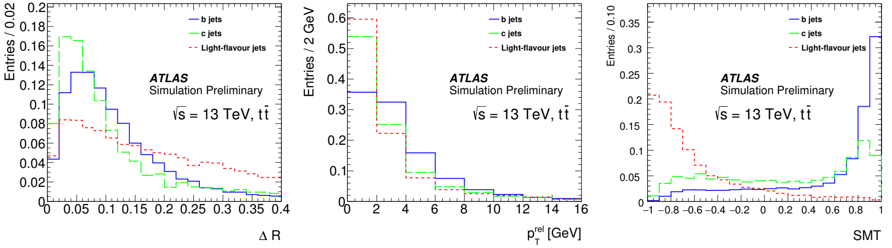




More details in ATL-PHYS-PUB-2016-012

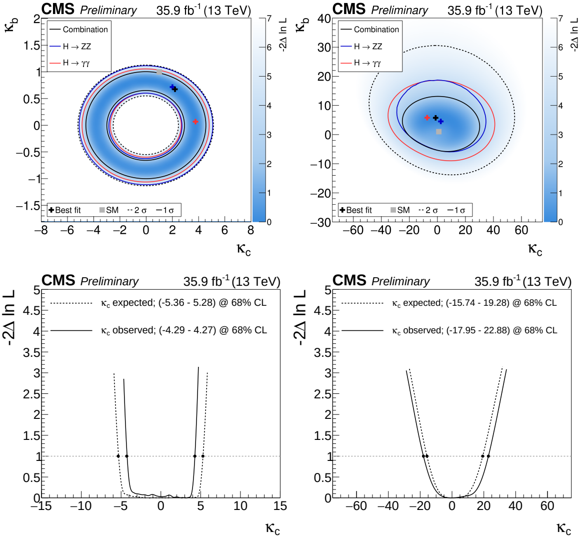
Exploit the large branching fractions for the semi-leptonic c/b hadron decays and the clean "muon-in-jet" experimental signature:

- Expect much higher rate of muons within b/c-jets, relative to light flavour jets, due to the decays $B \to \mu\nu X$ and $B \to DX \to \mu\nu X'$ ($\mathcal{B}$ of around 10% each)
- ✓ Complementary to SV and IP based taggers, different c/b hadron properties exploited and ATLAS detector components employed
- ✗ Light flavour jet backgrounds from muons produced in π/K decays in flight difficult to model in simulation

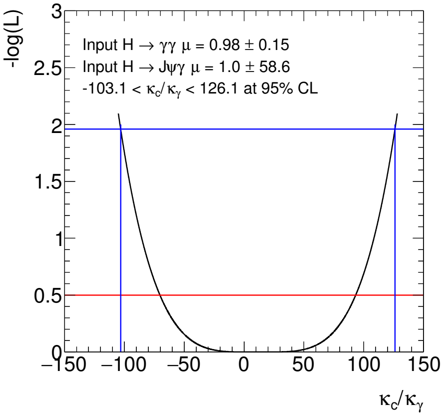


Left: ΔR of muon w.r.t. jet axis **Centre: p_T of muon relative to the jet axis** **Right: BDT built from muon observables**

Top: κ_c vs. κ_b **Bottom: κ_c, profiling κ_b**



Left: Normalisation + shape information **Right: Only shape information**



- Consider the ratio of signal strength measurements for $H \to J/\psi\,\gamma$ w.r.t. $H \to \gamma\gamma$
- Dependence on Γ_H and $\sigma(pp \to H)$ (approximately) cancels in this ratio, sensitive to κ_c/κ_γ
- Figure above based on ATLAS Run 2 $H \to J/\psi\,\gamma$ search and latest $H \to \gamma\gamma$ measurement (arXiv:1802.04146)

⚠ **This is NOT an ATLAS result, but my estimate based on public information alone**

Focus on the experimentally clean $J/\psi \to \mu^+\mu^-$ decays and target high rate inclusive H production

Trigger and Data Sample

- Dedicated photon + single muon triggers implemented to identify distinctive event topology
- Collected **$36.1\,\mathrm{fb}^{-1}$ $\sqrt{s} = 13\,\mathrm{TeV}$ pp dataset** during the 2015 and 2016 LHC runs

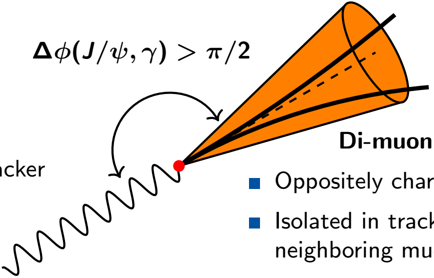
J/ψ Selection

- Require $m_{\mu^+\mu^-}$ loosely consistent with J/ψ mass
- **Minimum $p_{\mathrm{T}}^{J/\psi}$ requirement** varying with $m_{J/\psi\,\gamma}$ from 34 – 54.4 GeV, depending on channel (to optimise both H and Z searches)

$\Delta\phi(J/\psi, \gamma) > \pi/2$

$p_{\mathrm{T}}^{\mu\,\mathrm{lead}} > 18$ GeV

$p_{\mathrm{T}}^{\mu\,\mathrm{sub\text{-}lead}} > 3$ GeV



Photon Selection

- "Tight" photon ID requirements
- Isolated in both tracker and calorimeter

$p_{\mathrm{T}}^{\gamma} > 35$ GeV

Di-muon Selection

- Oppositely charged pair of muons
- Isolated in tracker (accounting for neighboring muon track)
- $L_{xy}/\sigma_{L_{xy}} < 3$ to reject $b \to J/\psi X$