Extrapolation of ATLAS sensitivity to $H \rightarrow b\bar{b}$ and $H \rightarrow c\bar{c}$ decays in VH production at the HL-LHC

The ATLAS Collaboration

The expected sensitivity of the measurements of Higgs boson decays into $b\bar{b}$ and $c\bar{c}$ with the ATLAS detector at the HL-LHC is estimated by extrapolating the results obtained using a dataset of $\sqrt{s} = 13$ TeV pp collisions corresponding to an integrated luminosity of 139 fb^{-1}. The associated production of a Higgs boson with either a W or a Z boson, targeting Higgs boson decays to b- or c-quarks and W or Z boson decays to leptons, is presented. A HL-LHC scenario in which a dataset of $\sqrt{s} = 14$ TeV pp collisions corresponding to an integrated luminosity of 3000 fb^{-1} is assumed. The projection of the existing measurements is performed accounting for both statistical and systematic uncertainties. For the measurements of the $ZH, H \rightarrow b\bar{b}$ and $WH, H \rightarrow b\bar{b}$ processes, projected uncertainties are estimated to be 7% and 8% respectively. Fiducial cross-sections for the $VH, H \rightarrow b\bar{b}$ processes are measured as a function of the vector boson transverse momentum in kinematic fiducial volumes. The estimated total uncertainties vary from 7% in the high transverse momentum regions to 18% in the low transverse momentum regions. For Higgs boson decays to $c\bar{c}$, produced in association with a W or a Z boson, an expected upper limit of 6.4 times the predicted cross-section times branching fraction for the process is estimated, corresponding to an expected constraint on the charm quark Yukawa coupling modifier $|\kappa_c| < 3.0$, at the 95% confidence level. The combination of the measurements of the $VH, H \rightarrow b\bar{b}$ and $VH, H \rightarrow c\bar{c}$ processes leads to an expected constraint on the ratio of coupling modifiers $|\kappa_c/\kappa_b| < 2.7$ at 95% confidence level.

Revised based on the version released on 22 October 2021 (corrected Figure 8 to include the correct plot for the 2-lepton channel).

1 Introduction

Since the discovery of the Higgs boson in 2012 by the ATLAS and CMS collaborations [1–8], many of its production and decay modes have been observed. For Standard Model Higgs boson with a mass of $m_H = 125$ GeV, the most copious decay is that into a $b\bar{b}$ pair, with an expected branching fraction of 58.2%. The experimental study of this decay channel at the LHC is challenging due to large multi-jet background. Higgs boson production in association with a vector boson has been found to represent the most sensitive production mode to measure $H \rightarrow b\bar{b}$ decays at the LHC. Using this channel, both the ATLAS [9] and CMS [10] collaborations have observed the $H \rightarrow b\bar{b}$ decay using data collected at centre-of-mass energies of 7 TeV, 8 TeV and 13 TeV during Runs 1 and 2 of the LHC. Recently, the ATLAS collaboration improved this measurement using the full Run 2 dataset of 139 fb^{-1} at 13 TeV [11], where $H \rightarrow b\bar{b}$ decays in association with a W or Z boson are measured with observed (expected) significances of 4.0 (4.1) and 5.3 (5.1) standard deviations above the background only hypothesis, respectively. ATLAS also used the same dataset to perform differential measurements of VH production in kinematic fiducial volumes defined in the simplified template cross-section (STXS) framework [12, 13] as a function of the vector boson transverse momentum. The cross-section measurements were found to be in agreement with the Standard Model expectation, with uncertainties ranging from 30% in the high momentum regions to 85% in the low momentum regions. Most recently, the STXS measurement was also performed in a combination of the resolved and boosted $VH, H \rightarrow b\bar{b}$ analyses [14].

The branching fraction for the $H \rightarrow c\bar{c}$ decay is predicted in the SM, for $m_H = 125$ GeV, to be 2.89% [15], around 20 times smaller than that of the $H \rightarrow b\bar{b}$ decay. No experimental evidence for this decay has been obtained, though direct searches for the $VH, H \rightarrow c\bar{c}$ process have been performed by both the ATLAS [16] and CMS [17] collaborations using 13 TeV pp collision datasets corresponding to 36.1 fb^{-1} and 35.9 fb^{-1}, respectively. More recently, ATLAS utilised the full Run 2 pp dataset of 139 fb^{-1} at 13 TeV to set an improved observed (expected) upper limit of 26 (31) times the predicted cross-section times branching fraction on the $VH, H \rightarrow c\bar{c}$ process [18]. This limit was also interpreted within the kappa framework [19, 20], leading to an observed (expected) constraint on the charm Yukawa coupling modifier $|\kappa_c| < 8.5$ (12.4) at the 95% confidence level (CL).

This note describes the extrapolation of the current ATLAS $VH, H \rightarrow b\bar{b}$ and $VH, H \rightarrow c\bar{c}$ analyses to a HL-LHC scenario defined by a dataset of 14 TeV pp collisions corresponding to an integrated luminosity of 3000 fb^{-1}. The analysis strategies used follow exactly the full Run 2 analyses.

2 Input analyses and extrapolation procedure

The following section provides a brief overview of the $VH, H \rightarrow b\bar{b}$ and $VH, H \rightarrow c\bar{c}$ analyses performed using the 139 fb^{-1} $\sqrt{s} = 13$ TeV pp collision dataset, in addition to details on the setup used for the extrapolation of the two analyses.

2.1 Overview of the $VH, H \rightarrow b\bar{b}$ analysis

The measurement of Higgs boson decays to b-quarks targets a Higgs boson produced in association with a W or Z boson, in order to suppress multi-jet background, and the details of the measurement are outlined in Ref. [11]. The ATLAS Run 2 pp collision dataset at 13 TeV is used to perform measurements of

Figure 1: The simplified template cross-section regions used for measurements and the corresponding reconstructed $VH, H \rightarrow b\bar{b}$ analysis regions. The analysis is not sensitive to the regions WH, $p_{\mathrm{T}}^{W,\,\mathrm{t}} < 150$ GeV and ZH, $p_{\mathrm{T}}^{Z,\,\mathrm{t}} < 75$ GeV, and their cross-sections are fixed to the SM prediction within their theoretical uncertainties. All leptonic decays of the vector bosons (including $Z \rightarrow \tau\tau$ and $W \rightarrow \tau\nu$, which are extrapolated from the electron and muon channel measurements) are considered for the STXS definition.

Higgs boson decays into b-quark pairs, separately in the WH and ZH production modes. Additionally, cross-section measurements are conducted in the reduced VH, $V \rightarrow$ leptons stage-1.2 STXS region scheme [12, 13] described in Ref. [21] and summarised in Figure 1.

Events are categorised based on the decay of the vector boson in the 0-, 1- and 2-lepton channels, exploring the $ZH \rightarrow \nu\nu b\bar{b}$, $WH \rightarrow \ell\nu b\bar{b}$ and $ZH \rightarrow \ell\ell b\bar{b}$ decay modes respectively (where $\ell = e, \mu$). To reduce the level of background, an important component of the $VH, H \rightarrow b\bar{b}$ analysis involves the identification of jets which originate from the hadronisation of a b-quark (b-jets). Details on the reconstruction of jets and other physics objects can be found in [11].

Jets are identified as b-jets (b-tagged) using a multivariate discriminant (MV2) [22], which was tuned to produce an average efficiency of 70% for b-jets in simulated $t\bar{t}$ events, which corresponds to light-flavour (u-, d-,s-quark and gluon) jet and c-jet misidentification efficiencies of 0.3% and 12.5%, respectively. The events in the $VH, H \rightarrow b\bar{b}$ analysis regions are required to have exactly two b-tagged jets. In order to increase the available Monte Carlo (MC) statistics the b-tagging requirement is not applied directly to c- and light-jets, but rather events are instead weighted by the probability of being b-tagged (*truth-flavour tagging*). After requiring the presence of two b-tagged jets the main backgrounds to the $VH, H \rightarrow b\bar{b}$ signal are V + jets processes, particularly those involving b-jets, and $t\bar{t}$ production.

The analysis regions are defined using the value of the transverse momentum of the reconstructed vector boson, p_{T}^{V} and jet multiplicity. The p_{T}^{V} categories include: $75 < p_{\mathrm{T}}^{V} < 150$ GeV (in the 2-lepton channel only), $150 < p_{\mathrm{T}}^{V} < 250$ GeV or $p_{\mathrm{T}}^{V} > 250$ GeV. Each of these categories are split into jet multiplicity categories, two or three (three or more in the 2-lepton channel) jets. This results in a total of 14 analysis regions, which are further split into signal and control regions. The control regions are defined as either having low or high ΔR[1] (with the boundaries varying with p_{T}^{V}) between the two leading b-jets, resulting in an enrichment in either V + jets or $t\bar{t}$. These categories are included in the analysis as yield-only control regions. The remaining events with intermediate ΔR enter the signal regions.

---

[1] $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$

A multivariate discriminant in the form of a boosted decision tree (BDT) is trained to improve the sensitivity of the analysis by discriminating the VH signal from background processes. A profile likelihood fit is performed to the BDT discriminant, and systematic uncertainties are included as nuisance parameters (NPs) in the fit. Fits are performed in three different configurations which vary the number of parameters of interest (POIs). First, a fit is performed for the signal strength, $\mu_{VH}^{b\bar{b}}$, that multiplies the cross-sections times branching fractions of the $VH, H \rightarrow b\bar{b}$ process. The resulting measurement is:

$$\mu_{VH}^{b\bar{b}} = 1.02_{-0.17}^{+0.18} = 1.02_{-0.11}^{+0.12}(\text{stat.})_{-0.13}^{+0.14}(\text{syst.})$$

Secondly, the signal strengths of the WH and ZH production modes are measured separately, measuring:

$$\mu_{WH}^{b\bar{b}} = 0.95_{-0.25}^{+0.27} = 0.95 \pm 0.18(\text{stat.})_{-0.18}^{+0.19}(\text{syst.})$$

$$\mu_{ZH}^{b\bar{b}} = 1.08_{-0.24}^{+0.25} = 1.08 \pm 0.17(\text{stat.})_{-0.15}^{+0.18}(\text{syst.})$$

The correlation coefficient between the two POIs is found to be 2.7%. Finally, a fit with 5 parameters of interest is performed, measuring the signal cross-section multiplied by the $H \rightarrow b\bar{b}$ and $V \rightarrow$ leptons branching fractions in the five STXS regions, with uncertainties between 30% (low p_T^V regions) to 85% (high p_T^V regions) on the POIs.

2.2 Overview of the $VH, H \rightarrow c\bar{c}$ analysis

The ATLAS search for Higgs boson decays into a charm quark pair shares many characteristics with the $VH, H \rightarrow b\bar{b}$ analysis performed on the same $\sqrt{s} = 13$ TeV pp collision dataset corresponding to an integrated luminosity of 139 fb^{-1}. Events where the Higgs boson is produced in the VH production mode are targeted and categorised by the decay of the vector boson into 0-, 1- or 2-lepton channels. The main difference with respect to the $VH, H \rightarrow b\bar{b}$ analysis is the jet flavour tagging and Higgs boson reconstruction strategy. $H \rightarrow c\bar{c}$ candidates are formed from the two highest p_T jets in the event, irrespective of flavour tagging considerations. Jets are tagged as containing either b- or c-hadrons using two discriminants resulting from the multivariate tagging algorithms, MV2 and DL1 [22]. A dedicated c-tagging operating point based on the DL1 discriminant, combined with a b-tag veto requirement using MV2 at a 70% working point, was optimised to have an average efficiency of 27% to tag c-jets, with b- and light-jet misidentification efficiencies of 8% and 1.6%, respectively. Events are required to have either one or two jets c-tagged with this dedicated operating point. Any additional jets in the event are also required not to be b-tagged. The b-tag veto requirement ensures that the $VH, H \rightarrow b\bar{b}$ and $VH, H \rightarrow c\bar{c}$ signal regions are orthogonal. In order to maximise the statistical power of the MC samples available, the c-tagging requirement is not directly applied to the diboson, V + jets or top-quark samples. Instead, events are weighted by the probabilities, parameterised in jet p_T and η, for each jet to be c-tagged, based on its flavour label, to obtain predictions for one and two c-tagged jet events (*truth-flavour tagging*). To achieve good closure between this technique and directly applying the c-tagging discriminant, small corrections and corresponding uncertainties are applied, based on the ΔR (for V+jets only) and total event yields (for all samples).

Events are categorised in terms of the number of c-tagged jets, the overall jet multiplicity and p_T^V, where a $p_T^V > 150$ GeV category is defined for all channels and a low p_T^V region ($75 < p_T^V < 150$ GeV) is included

in the 2-lepton channel only. The major backgrounds in the $VH, H \to c\bar{c}$ analysis come from V + jets and $t\bar{t}$ events. Dedicated control regions are defined for all major backgrounds. For V + jets, two sets of control regions are defined, by either requiring high ΔR between the two $H \to c\bar{c}$ candidate jets (high ΔR control region) or requiring both $H \to c\bar{c}$ candidate jets to fail the combined c-tag and b-tag veto requirement (0 c-tag control region). For the 0- and 1-lepton channels, dedicated control regions for the $t\bar{t}$ backgrounds are formed from three jet events in which the third jet is b-tagged. For the 2-lepton channel, the $t\bar{t}$ control region is formed from $e^{\pm}\mu^{\mp}$ events and implemented as yield-only control regions.

A profile likelihood fit is performed to all signal and control regions, using the invariant mass of the two leading jets (m_{cc}) as the signal to background discriminant. Measurements of three POIs are extracted simultaneously: $\mu_{VH}^{c\bar{c}}$, μ_{VW}^{cq}, and $\mu_{VZ}^{c\bar{c}}$, which correspond to signal strengths that multiply the cross-sections times branching fractions of the $VH, H \to c\bar{c}$, $VW, W \to cq$ and $VZ, Z \to c\bar{c}$ processes, respectively. Systematic uncertainties are included as nuisance parameters, including experimental uncertainties, uncertainties due to the theoretical modelling of signal and background, or uncertainties due to limited MC sample statistics. The observed (expected) significances of the $VW, W \to cq$ and $VZ, Z \to c\bar{c}$ processes are 3.8 (4.6) and 2.6 (2.2) standard deviations, respectively. An upper limit of 26 (31^{+12}_{-8}) is observed (expected) on the $VH, H \to c\bar{c}$ signal strength at 95% CL. Within the context of the kappa framework [19, 20], the $VH, H \to c\bar{c}$ signal strength can be reparametrised in terms of the charm quark coupling modifiers κ_c, while setting all other couplings to their SM predictions:

$$\mu_{VH}^{c\bar{c}}(\kappa_i) = \frac{\kappa_c^2}{B(H \to c\bar{c})\kappa_c^2 + (1 - B(H \to c\bar{c}))} \tag{1}$$

This parametrisation is used to derive the observed (expected) constraint of $|\kappa_c| < 8.5$ (12.4) at 95% CL.

2.3 Combination of the $VH, H \to b\bar{b}$ and $VH, H \to c\bar{c}$ analyses

The two separate analyses are statistically combined by means of a simple combination procedure. The statistical analysis setup used for the separate analyses is unchanged with the exception of allowing the $VH, H \to b\bar{b}$ and $VH, H \to c\bar{c}$ POIs to affect the corresponding contributions in both analysis regions. A simultaneous fit to the $VH, H \to b\bar{b}$ and $VH, H \to c\bar{c}$ analysis regions is performed to extract the signal strengths for both the $VH, H \to b\bar{b}$ and $VH, H \to c\bar{c}$ processes. Experimental nuisance parameters related to jets, leptons, E_T^{miss} or luminosity are correlated between the two analyses, while all remaining experimental and modelling uncertainties are kept separate in the combined fit, owing to differences in their implementation or parameterisation in the two analyses. In the $VH, H \to c\bar{c}$ analysis regions the $VW, W \to cq$ and $VZ, Z \to c\bar{c}$ processes are left to float freely, in agreement with the standalone analysis. In addition to a fit in which the two signal strength parameters are measured simultaneously, it is also possible to place a constraint on the ratio of Higgs coupling modifiers κ_c/κ_b, while profiling κ_b and setting all other coupling modifiers to their SM expectations, by reparametrising the signal strengths as:

$$\mu_{VH}^{b\bar{b}} = \frac{\kappa_b^2}{1 - B(H \to b\bar{b}) - B(H \to c\bar{c}) + B(H \to c\bar{c})\kappa_c^2 + B(H \to b\bar{b})\kappa_b^2} \tag{2}$$

$$\mu_{VH}^{c\bar{c}} = \frac{\kappa_c^2}{1 - B(H \to b\bar{b}) - B(H \to c\bar{c}) + B(H \to c\bar{c})\kappa_c^2 + B(H \to b\bar{b})\kappa_b^2} \tag{3}$$

2.4 Extrapolation setup

The extrapolation from the Run 2 analyses to the expectation for the HL-LHC is performed by scaling the signal and background expectations to an increased integrated luminosity and center-of-mass energy. First, a scale factor is applied to the normalisation of the signal and background predictions for all processes in all categories, which accounts for the increase in integrated luminosity from 139 fb^{-1} to 3000 fb^{-1}. Next, scale factors are applied to the different processes to account for the increase in $\sqrt{s}$ from 13 to 14 TeV, which are derived using the expected cross-sections [20]. These scale factors adopt values between 1.10 and 1.18, depending on the process and a summary of these scale factors can be found in table 1.

Table 1: Summary of the p_{T}^V inclusive scale factors applied to the different signal and background processes.

$qq \to WH$ ($H \to c\bar{c}/b\bar{b}$)	1.10
$qq \to ZH$ ($H \to c\bar{c}/b\bar{b}$)	1.11
$gg \to ZH$ ($H \to c\bar{c}/b\bar{b}$)	1.18
$t\bar{t}$ $gg \to ZZ$	1.16
$qq \to VV$ V+jets single top	1.10

Experimental and theory uncertainties are included in the extrapolation, following the prescriptions used in the Run 2 analyses, but the size of the uncertainties are scaled to account for reductions in their statistical components and potential improvements in analysis techniques associated with the larger available dataset [23, 24]. A summary of the scale factors applied to the systematic uncertainties is shown in Table 2. Experimental uncertainties related to $E_{\mathrm{T}}^{\mathrm{miss}}$ are scaled by 1/2 and the luminosity uncertainty is reduced to 1% (0.58 of the value at Run 2). Flavour tagging uncertainties are scaled by a factor of 1/2, except for those related to light-jets in the $VH, H \to c\bar{c}$ analysis, which are left unscaled. All other experimental uncertainties are kept at their Run 2 levels as their impact on the analysis is small. Theory and background uncertainties on the signal and background processes are also scaled by a factor of 1/2. Uncertainties associated with the finite size of simulated samples are not considered. In the case of the $VH, H \to c\bar{c}$ analysis, this also includes uncertainties relating to *truth-tagging*. This decision is motivated by the expected improvements in the ATLAS simulation infrastructure and reconstruction software, which would allow samples to be generated of the sufficient size to largely negate these uncertainties.

$VH, H \to b\bar{b}$ fit setup: In the case of the extrapolation of the $VH, H \to b\bar{b}$ measurement, two separate fit configurations, in which the POI choices are varied, are performed to an Asimov dataset. Firstly, a fit with two POIs, namely the signal strengths of the $ZH, H \to b\bar{b}$ and $WH, H \to b\bar{b}$ processes, is performed. Secondly, a fit including five STXS POIs is performed, as described in section 2.1.

$VH, H \to c\bar{c}$ fit setup: For the extrapolation of the $VH, H \to c\bar{c}$ search, a fit is performed on an Asimov dataset, determining the expected signal strength of the $VH, H \to c\bar{c}$ process, while the $VZ, Z \to c\bar{c}$ and $VW, W \to cq$ signal strengths are free to float. In addition to the signal strength, an upper limit at 95% CL is determined, and the expected constraint on κ_c is computed.

Table 2: Summary of scale factors applied to systematic uncertainties in the extrapolation. The luminosity uncertainty is reduced from 1.7% to 1.0%, resulting in a scale factor of 0.58.

Uncertainties	Scale Factor
$E_{\mathrm{T}}^{\mathrm{miss}}$	0.5
Lepton	1
Jet	1
Flavour tagging c-, b- and τ-jets	0.5
Flavour tagging light-jets (MV2c10 in $VH(bb)$)	0.5
Flavour tagging light-jets (DL1 in $VH(cc)$)	1.0
Luminosity	0.58
Signal modelling	0.5
Background modelling	0.5
MC statistics	0
Truth-tagging uncertainties ($VH, H \rightarrow c\bar{c}$ only)	0

Combined fit setup: Using the individual analysis definitions described above, a simultaneous fit to all $VH, H \rightarrow b\bar{b}$ and $VH, H \rightarrow c\bar{c}$ regions is performed. First, the expected signal strengths for $VH, H \rightarrow b\bar{b}$ and $VH, H \rightarrow c\bar{c}$ are determined. Then a fit to $|\kappa_c/\kappa_b|$ is performed, with κ_b being allowed to float in the fit.

3 Results

3.1 Extrapolation of the $VH, H \rightarrow b\bar{b}$ analysis

3.1.1 Signal strength measurements

The signal strength measurements for the $ZH, H \rightarrow b\bar{b}$ and $WH, H \rightarrow b\bar{b}$ processes, derived from a fit to an Asimov dataset, extrapolated to $3000\,\mathrm{fb}^{-1}$ and $\sqrt{s} = 14\,\mathrm{TeV}$, are found to be:

$$\mu_{WH}^{b\bar{b}} = 1.00 \pm 0.08 = 1.00 \pm 0.04(\mathrm{stat.}) \pm 0.07(\mathrm{syst.})$$
$$\mu_{ZH}^{b\bar{b}} = 1.00 \pm 0.07 = 1.00 \pm 0.03(\mathrm{stat.}) \pm 0.06(\mathrm{syst.})$$

The correlation coefficient between the two POIs is found to be +19%, caused by the experimental and modelling uncertainties. The results are also illustrated in Figure 2, in which the statistical and systematic uncertainties are shown separately.

Table 3 shows the effect of the different groups of systematic uncertainties on the WH and ZH signal strengths. The expected sensitivity is limited by systematic uncertainties alone, and the largest single contribution comes from the signal modelling uncertainties. One notable feature is that the signal theory uncertainties have a significantly larger contribution for the ZH process, due to the larger uncertainties on the $gg \rightarrow ZH$ contribution [20]. Background modelling uncertainties have a smaller effect, and increasing or decreasing the background modelling uncertainties by a factor of two has at most a 10% relative effect on
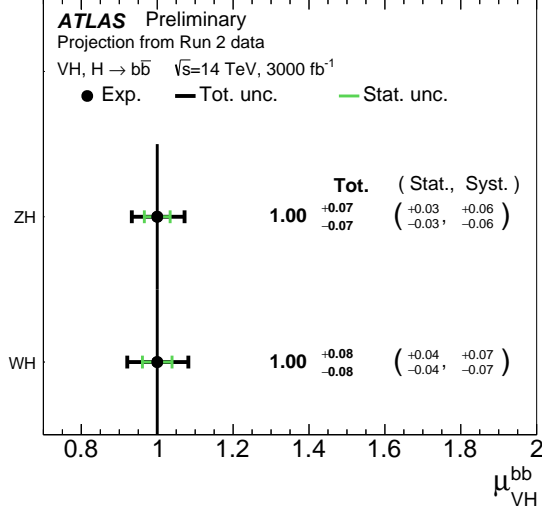
Figure 2: Fitted values of the Higgs boson signal strength $\mu_{VH}^{b\bar{b}}$, for $m_H = 125\,\text{GeV}$ and for WH and ZH, from a fit to an Asimov dataset scaled to $3000\,\text{fb}^{-1}$ and $\sqrt{s} = 14\,\text{TeV}$. The individual $\mu_{VH}^{b\bar{b}}$ values for the $(W/Z)H$ processes are obtained from a simultaneous fit with the signal strength for each of the WH and ZH processes floating independently. The correlation coefficient between the signal strength of the WH and ZH processes is $+19\%$.

the total uncertainty of both signal strength parameters. Experimental uncertainties represent contributions of a similar size, with jet and flavour tagging uncertainties having the largest impact. MC statistical uncertainties are not included in these results and adding them would have a relative impact on the total uncertainty of around 2%. This was estimated by performing a fit with scaled MC statistical uncertainties, assuming that the data-to-MC luminosity ratio stays the same as for Run 2.

3.1.2 Cross-section measurements

The expected measurements of cross-sections times $H \to b\bar{b}$ and $V \to$ leptons branching fractions at the HL-LHC, $\sigma \times B$, are summarised in Figure 3, along with the correlations between the different STXS parameters of interest. The relative uncertainties on the parameters vary between 8% in the highest p_{T}^V region and 18% in the lowest p_{T}^V region. In the low p_{T}^V categories, the experimental and background modelling uncertainties are the most important, whereas the signal modelling uncertainties have a larger impact at high p_{T}^V. MC statistical uncertainties are not included in these results and adding them would have no impact in the low p_{T}^V regions and up to 5% relative impact in the high p_{T}^V regions. This was estimated by performing a fit with scaled MC statistical uncertainties, assuming that the data-to-MC luminosity ratio stays the same as for Run 2.

Table 3: Breakdown of contributions to the uncertainty in the fitted values of $\mu_{ZH}^{b\bar{b}}$ and $\mu_{WH}^{b\bar{b}}$ from a fit to an Asimov dataset scaled to 3000 fb^{-1} and $\sqrt{s} = 14$ TeV. The sum in quadrature of different sources of uncertainty may differ from the total due to correlations. In the case that the up and down systematic variations have different values, the mean of the absolute values is shown.

| Source of uncertainty		$\Delta\mu_{ZH}^{b\bar{b}}$	$\Delta\mu_{WH}^{b\bar{b}}$
Total		0.070	0.081
Statistical		0.034	0.039
Systematics		0.063	0.070
Statistical uncertainties			
Data statistics only		0.031	0.037
$t\bar{t}\,e\mu$ control region		0.006	0.003
Floating normalisations		0.017	0.028
Theoretical and modelling uncertainties			
Signal		0.047	0.031
Z+jets		0.017	0.010
W+jets		0.004	0.022
single top		0.005	0.012
$t\bar{t}$		0.007	0.017
Diboson		0.020	0.027
Multi-Jet		< 0.001	0.001
Experimental uncertainties			
Jets		0.022	0.032
Leptons		0.006	0.011
$E_{\mathrm{T}}^{\mathrm{miss}}$		0.006	0.005
Pile-up and luminosity		0.009	0.009
	b-jets	0.018	0.009
Flavour tagging	c-jets	0.004	0.035
	light-jets	0.006	0.009
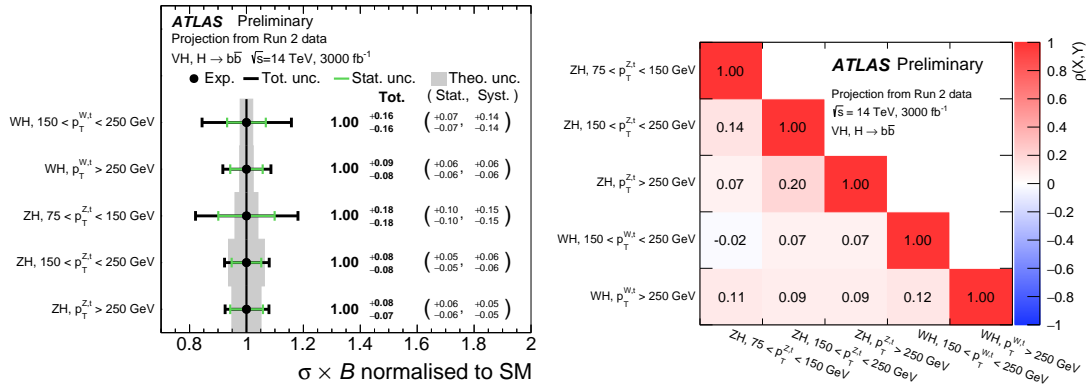
Figure 3: Best-fit values and uncertainties for the $VH, H \to b\bar{b}$ signal strength in the reduced STXS scheme (left) and their correlations (right). The results are determined in a fit to an Asimov dataset scaled to 3000 fb^{-1} and $\sqrt{s} = 14$ TeV. The theory uncertainties are shown as a grey band and reduced by a factor 1/2 with respect to the Run 2 analysis.

3.2 Extrapolation of the $VH, H \rightarrow c\bar{c}$ analysis

3.2.1 Signal strengths constraints

The results of a fit for the $VH, H \rightarrow c\bar{c}$ signal strength, extrapolated to the HL-LHC scenario, are shown in Figure 4, along with the corresponding 95% CL limits. Results are shown for separate fits in which POIs for the individual lepton channels are either separated or combined. In the combined fit, the best fit value is:

$$\mu_{VH}^{c\bar{c}} = 1.0 \pm 3.2 = 1.0 \pm 2.0(\text{stat.})^{+2.6}_{-2.5}(\text{syst.})$$

The expected upper limit on $\mu_{VH}^{c\bar{c}}$ of 6.4 times the SM prediction at 95% CL is derived. Table 4 shows the contribution of different groups of uncertainties. The statistical and systematic uncertainties are of similar size. The largest single contribution to systematic uncertainties comes from the modelling of the Z + jets background.

The scaling of the signal and background uncertainties has a relative impact of around 10% when decreasing or increasing the systematic uncertainty scale factors by a factor 2. Due to the relatively low c-tagging efficiency, statistical uncertainties related to the available MC sample size have a non-negligible impact on the $VH, H \rightarrow c\bar{c}$ analysis. While these uncertainties were neglected for the basic extrapolation presented here, including the *truth-tagging* uncertainties would degrade the expected limit by 4%. Including the MC statistical uncertainties would degrade the expected limit by 5%. This was estimated by performing a fit with scaled MC statistical uncertainties, assuming that the data-to-MC luminosity ratio stays the same as for Run 2.

Flavour tagging has a significant impact on the sensitivity of the $VH, H \rightarrow c\bar{c}$ analysis, and due to the improved resolution of the replacement inner detector at the HL-LHC (ITk), a better flavour tagging performance can be expected [23–28]. Using the existing c-tagging algorithm based on a deep-learning neural network (DL1) and requiring a 27% c-jet efficiency operating point, initial studies suggest that
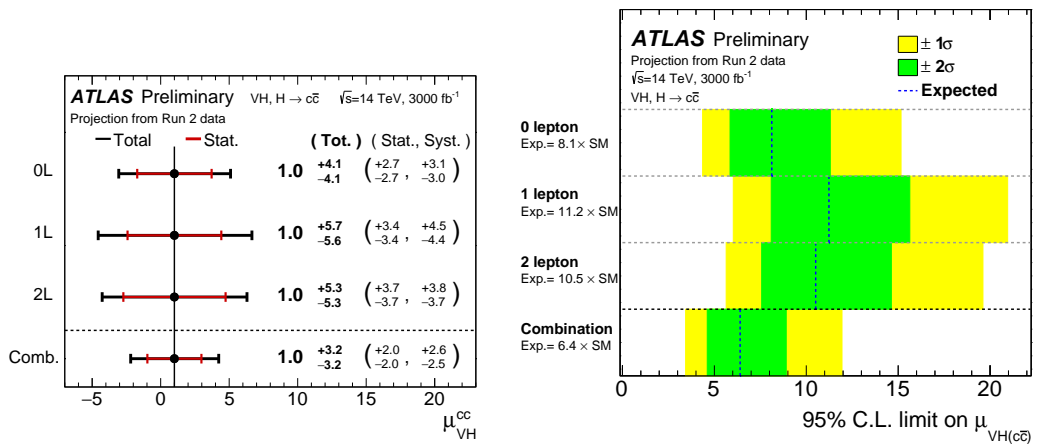

Figure 4: Best fit signal strengths (left) and limits (right) on the $VH, H \rightarrow c\bar{c}$ signal strength in the different lepton channels from a fit to an Asimov dataset scaled to 3000 fb^{-1} and $\sqrt{s} = 14$ TeV. The per-channel results are obtained using a five-POI fit, in which each channel has a separate $VH, H \rightarrow c\bar{c}$ parameter of interest.

Table 4: Breakdown of contributions to the uncertainty in the fitted values of $\mu_{VH}^{c\bar{c}}$ from a fit to an Asimov dataset scaled to $3000\,\text{fb}^{-1}$ and $\sqrt{s} = 14\,\text{TeV}$. The sum in quadrature of different sources of uncertainty may differ from the total due to correlations. In the case that the up and down systematic variations have different values, the mean of the absolute values is shown.

Source of uncertainty		$\Delta\mu_{VH}^{c\bar{c}}$
Total		3.21
Statistical		1.97
Systematics		2.53
Statistical uncertainties		
Data statistics only		1.59
Floating normalisations		0.95
Theoretical and modelling uncertainties		
$VH, H \rightarrow c\bar{c}$		0.27
Z+jets		1.77
Top-quark		0.96
W+jets		0.84
Diboson		0.34
$VH, H \rightarrow b\bar{b}$		0.29
Multi-Jet		0.09
Experimental uncertainties		
Jets		0.59
Leptons		0.20
$E_{\text{T}}^{\text{miss}}$		0.18
Pile-up and luminosity		0.19
Flavour tagging	c-jets	0.61
	b-jets	0.16
	light-jets	0.51
	τ-jets	0.19

the b- and light flavour jet rejection factors are expected to improve by factors of 1.5 and 3 respectively. This performance improvement is expected to improve the $VH, H \rightarrow c\bar{c}$ limit by 10-15%, where the improvement in light flavour jet rejection has a larger impact. Improved c-tagging performance is foreseen in the future thanks to the installation of ITk and new machine learning approaches.

3.2.2 κ_c interpretation

The re-interpretation of the $VH, H \rightarrow c\bar{c}$ signal strength within the κ framework allows an estimate of the sensitivity to the coupling between the Higgs boson and the charm quark, in terms of the coupling modifier κ_c. The resulting expected constraints on κ_c, at 68% and 95% CL, for the HL-LHC scenario are summarised in Table 5, and the corresponding profile likelihood scans are shown in Figure 5. The extrapolated re-interpretation, considering all three lepton channels, results in an expected constraint of $|\kappa_c| < 3.0$ at 95% CL.

Table 5: Expected pre-fit Asimov 68% CL and 95% CL limits on κ_c in the individual channel fits and in the combination for the extrapolation to 3000 fb^{-1} and $\sqrt{s} = 14$ TeV.

Channel	Expected Limit	
	68% CL	95% CL
0-lepton	[-2.4, 2.4]	[-3.5, 3.5]
1-lepton	[-2.8, 2.8]	[-4.3, 4.3]
2-lepton	[-2.7, 2.7]	[-4.1, 4.1]
Combination	[-2.2, 2.2]	[-3.0, 3.0]
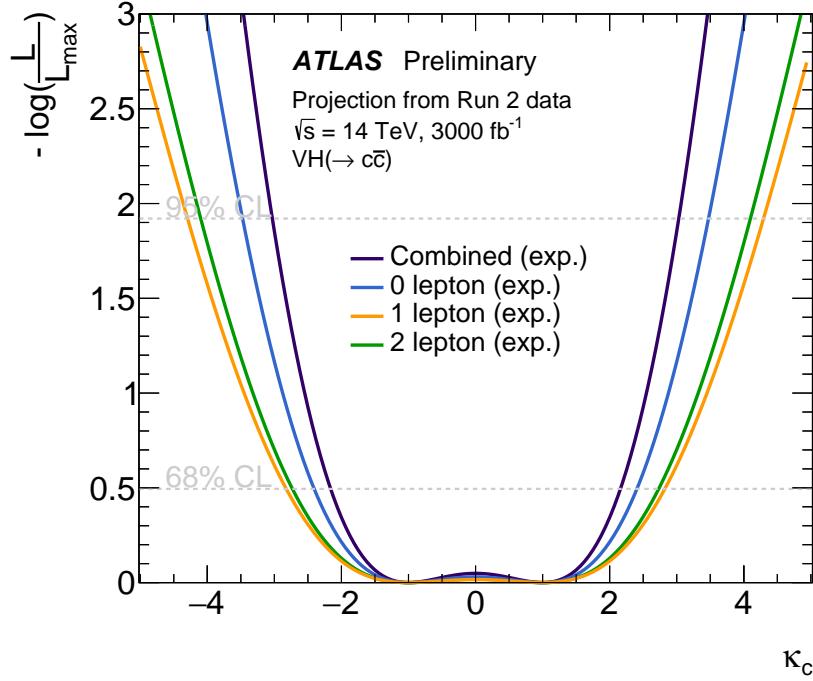
Figure 5: Expected profile likelihood scans for κ_c at the HL-LHC. The per-channel likelihoods are obtained using a three POI fit, in which each channel has a separate $VH, H \rightarrow c\bar{c}$ parameter of interest.

3.3 Extrapolation of combination

Using the previously defined standalone analysis strategies, a combined fit to the $VH, H \rightarrow b\bar{b}$ and $VH, H \rightarrow c\bar{c}$ signal strength parameters is performed, resulting in expected best fit values of:

$$\mu_{VH}^{b\bar{b}} = 1.00 \pm 0.06,$$
$$\mu_{VH}^{c\bar{c}} = 1.00 \pm 3.20,$$

The correlation coefficient between the $VH, H \rightarrow b\bar{b}$ and $VH, H \rightarrow c\bar{c}$ signal strength parameters is -11%. The resulting likelihood scan is shown in 6. Reinterpreting the signal strengths in terms of κ_b and κ_c results
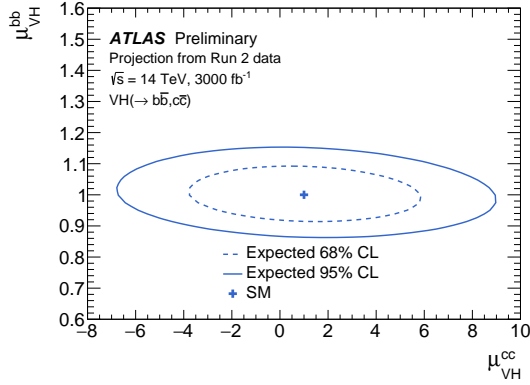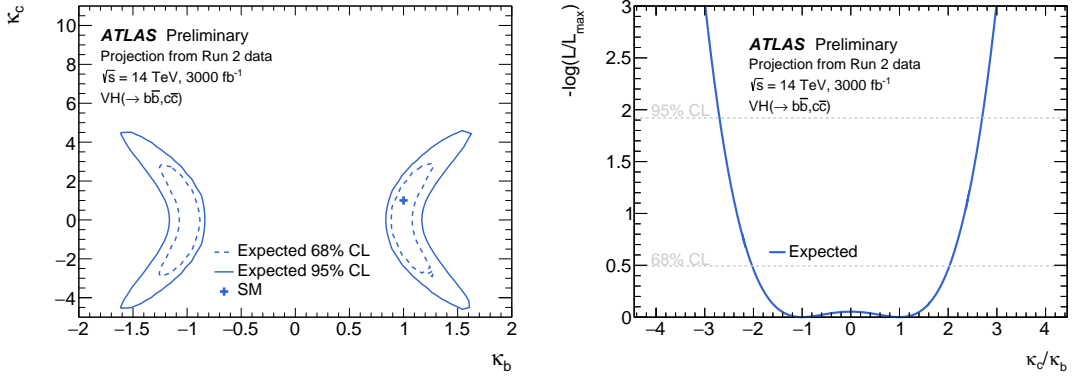
12

Figure 6: Expected profile likelihood scans for the $VH, H \rightarrow b\bar{b}/c\bar{c}$ combination extrapolated to a dataset of $3000\,\text{fb}^{-1}$ at $\sqrt{s} = 14\,\text{TeV}$. A two-dimensional scan is shown for a simultaneous fit to $\mu_{VH}^{c\bar{c}}$ and $\mu_{VH}^{b\bar{b}}$.

Figure 7: Expected profile likelihood scans for the $VH, H \rightarrow b\bar{b}/c\bar{c}$ combination extrapolated to a dataset of $3000\,\text{fb}^{-1}$ at $\sqrt{s} = 14\,\text{TeV}$. A two-dimensional scan is shown for a simultaneous fit to κ_c and κ_b (left). A one-dimensional scan is shown for a fit to κ_c/κ_b (right), where κ_b is also left to float freely.

in the likelihood scans shown in Figure 7 and provides an expected constraint of $|\kappa_c/\kappa_b| < 2.7$ at 95% CL.

4 Conclusions

The expected sensitivities of the measurements of Higgs boson decays to $b\bar{b}$ and $c\bar{c}$ in the VH production mode have been estimated for the ATLAS detector at the HL-LHC. Based on the results obtained using a dataset of $139\,\text{fb}^{-1}$, the analyses have been extrapolated to an expected pp collision dataset of $3000\,\text{fb}^{-1}$ at a center-of-mass energy of $14\,\text{TeV}$, taking into account both statistical and systematic uncertainties. For the measurements of the $ZH, H \rightarrow b\bar{b}$ and $WH, H \rightarrow b\bar{b}$ processes uncertainties of 7% and 8% can be expected. The uncertainties on cross-section measurements of $VH, H \rightarrow b\bar{b}$ decays in kinematic fiducial volumes have been estimated to vary between 8%, in regions of high transverse momentum of the vector boson, and 18% in the low momentum regions. In the search for $VH, H \rightarrow c\bar{c}$ decays a 95% CL upper limit on the signal strength of 6.4 times the predicted cross-section times branching fraction for the Higgs can be expected, which can be interpreted in the kappa framework as a constraint on the charm Yukawa

coupling modifier of $|\kappa_c| < 3.0$, at the 95% CL. The combination of the two measurements provides an expected constraint of $|\kappa_c/\kappa_b| < 2.7$ at the 95% CL at HL-LHC.

Appendix

A Additional material for the $VH, H \to b\bar{b}$ extrapolation

A.1 p_T^V distributions

Figure 8 shows the p_T^V distributions for the 2 jets and 2 b-tag categories in the 0-, 1- and 2-lepton channels of the $VH, H \to b\bar{b}$ analysis after performing the scaling of the templates to $3000\,\text{fb}^{-1}$ and 14 TeV, and scaling the systematic uncertainties.
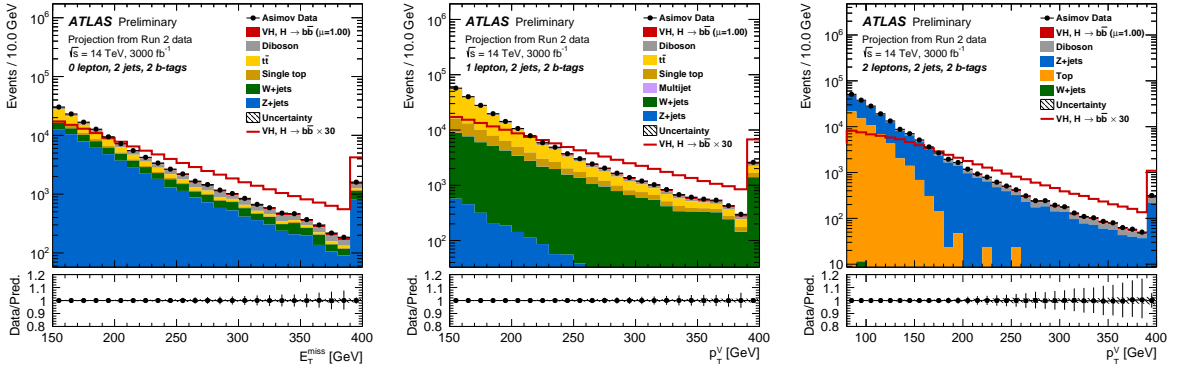

Figure 8: Post-fit distributions from an unconditional fit to a pre-fit Asimov dataset p_T^V distributions for events with two jets and two b-tags in the 0-lepton (left), 1-lepton (centre), and 2-lepton (right) channels. The MC templates are scaled to $3000\,\text{fb}^{-1}$ and 14 TeV, and the corresponding Asimov dataset is shown in black.

A.2 Cross-section measurements

A.2.1 Breakdown of uncertainties

Table 6 shows the effect of the different groups of systematic uncertainties on the fitted values of the $VH, H \to b\bar{b}$ signal strength in the reduced STXS scheme. In the lower p_T bins the experimental and background modelling uncertainties have the largest impact, whereas in the the higher p_T bins the signal modelling uncertainties are more important.

A.3 Results without scaling of systematics

In addition to the nominal results, an extrapolation of the $VH, H \to b\bar{b}$ analysis was performed without the scaling of any systematic uncertainties. For the fit to WH and ZH this results in uncertainties of 12% and 11% on the WH and ZH signal strengths, respectively. In the fit to the different STXS cateogories, the uncertainties vary between 24% in the low p_T^V and 10% in the high p_T^V regions.

Table 6: Breakdown of contributions to the uncertainty on the fitted values of the $VH, H \rightarrow b\bar{b}$ signal strength in the reduced STXS scheme from a fit to an Asimov dataset scaled to 3000 fb^{-1} and $\sqrt{s} = 14$ TeV. The sum in quadrature of different sources of uncertainty may differ from the total due to correlations. In the case that the up and down systematic variations have different values, the mean of the absolute values is shown.

| Source of uncertainty $\quad p_T^{V,t}$	WH 150–250 GeV	WH >250 GeV	ZH 75–150 GeV	ZH 150–250 GeV	ZH >250 GeV
Total	0.157	0.085	0.180	0.079	0.077
Statistical	0.068	0.057	0.099	0.052	0.057
Systematics	0.141	0.063	0.150	0.059	0.051
Statistical uncertainties					
Data statistics only	0.063	0.054	0.082	0.059	0.056
$t\bar{t}\,e\mu$ control region	0.018	0.004	0.050	0.012	0.004
Floating normalisations	0.063	0.023	0.080	0.030	0.019
Theoretical and modelling uncertainties					
Signal	0.020	0.027	0.023	0.028	0.036
Z+jets	0.023	0.009	0.103	0.019	0.019
W+jets	0.049	0.018	0.021	0.008	0.003
single top	0.043	0.009	0.015	0.012	0.005
$t\bar{t}$	0.062	0.018	0.024	0.008	0.008
Diboson	0.024	0.028	0.038	0.020	0.015
Multi-Jet	0.001	< 0.001	0.001	< 0.001	< 0.001
Experimental uncertainties					
Jets	0.076	0.031	0.083	0.031	0.019
Leptons	0.015	0.013	0.022	0.005	0.009
E_T^{miss}	0.058	0.012	0.058	0.008	0.007
Pile-up and luminosity	0.015	0.011	0.023	0.009	0.09
Flavour tagging $\quad$ b-jets	0.020	0.008	0.071	0.026	0.010
c-jets	0.061	0.029	0.003	0.005	0.003
light-jets	0.018	0.007	0.007	0.005	0.007
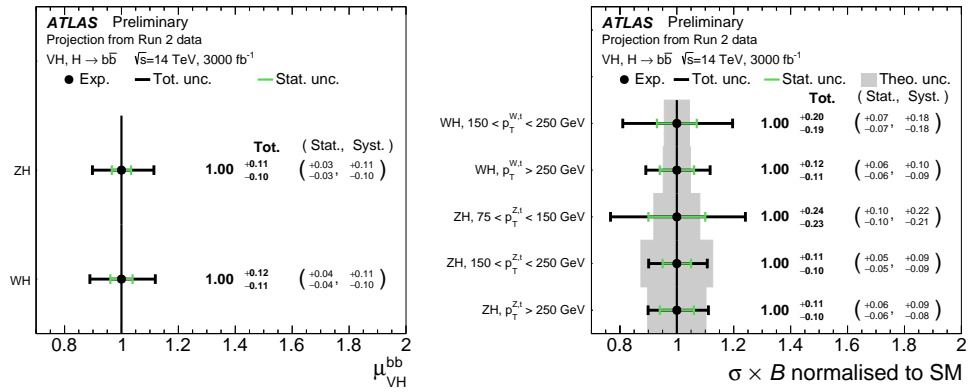
Figure 9: Best-fit values and uncertainties for the $VH, H \rightarrow b\bar{b}$ signal strength for WH and ZH (left) in the reduced STXS scheme (right). The results are determined in a fit to an Asimov dataset scaled to 3000 fb^{-1} and $\sqrt{s} = 14$ TeV and without any scaling applied to the systematic uncertainties. For the STXS fit the theory uncertainties are shown as a grey band at the same level as in the Run 2 analysis.

B Additional material for the $VH, H \to c\bar{c}$ extrapolation

B.1 Mass distributions

Figure 10 shows the m_{cc} distributions in a selection of analysis regions of the $VH, H \to c\bar{c}$ analysis after performing the scaling of the templates to 3000 fb^{-1} and 14 TeV, and scaling the systematic uncertainties.
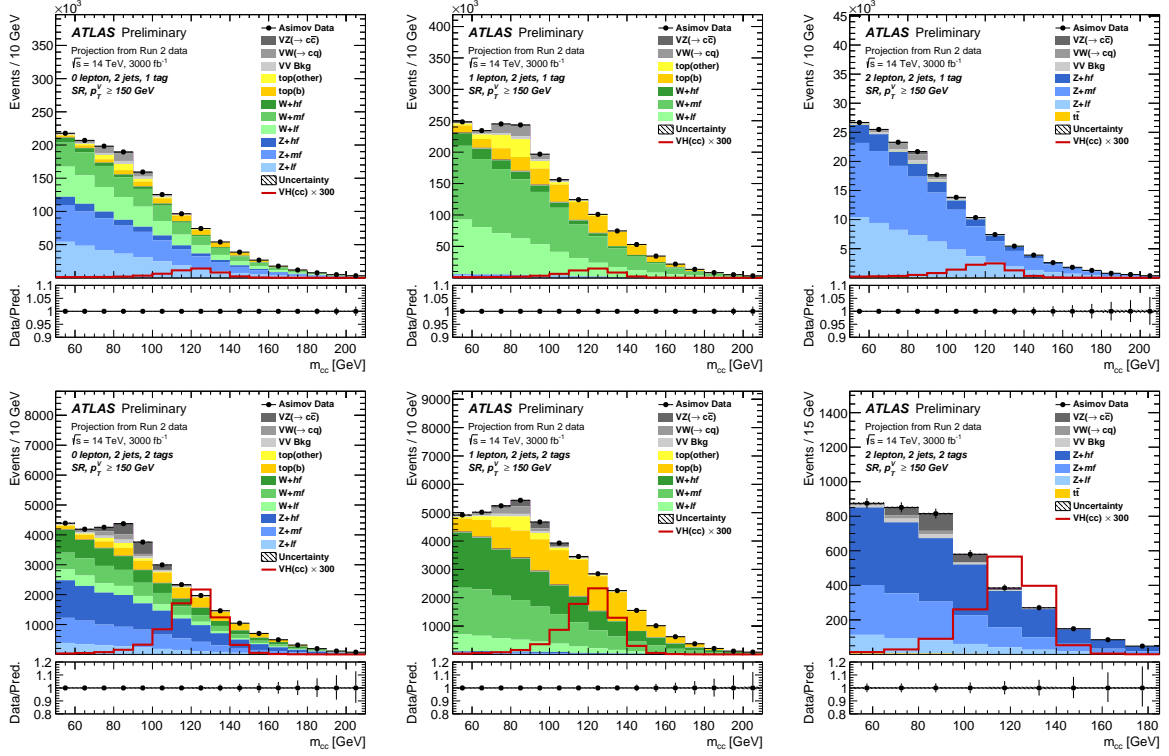

Figure 10: Post-fit distributions from an unconditional fit to a pre-fit Asimov dataset for six selected signal regions out of 44 analysis regions, with two jets and $p_T^V > 150$ GeV for the 0-lepton (left), 1-lepton (centre), and 2-lepton (right) channels, with 1 c-tag (top) and 2 c-tags (bottom). The MC templates are scaled to 3000 fb^{-1} and 14 TeV, and the corresponding Asmiov dataset is shown in black.

B.2 Results without scaling of systematics

In addition to the nominal results, an extrapolation of the $VH, H \to c\bar{c}$ analysis was performed without the scaling of any systematic uncertainties. The resulting signal strengths and limits are shown in Figure 11. In this case, the systematic uncertainties are more sizeable than the statiscal uncertainties, and in the combined fit the expected upper limit is 7.4 times the SM prediction at 95% CL.
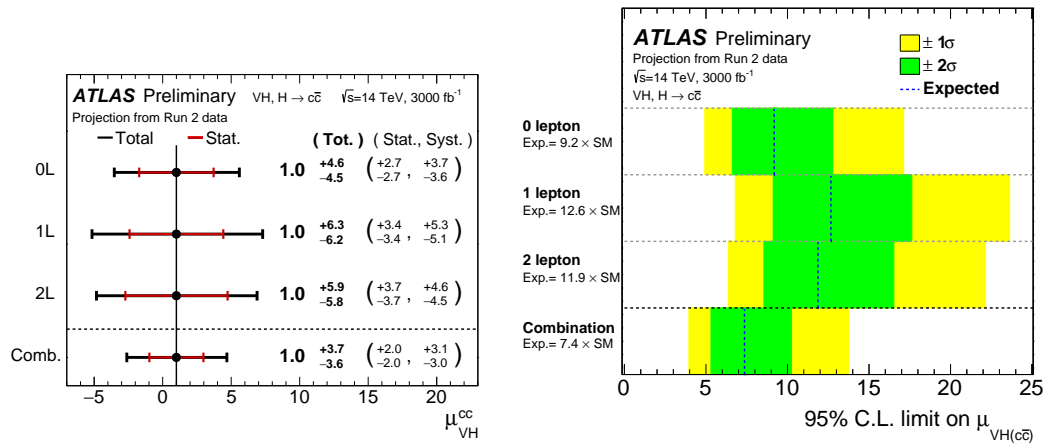
17

Figure 11: Best fit signal strengths (left) and limits (right) on the $VH, H \rightarrow c\bar{c}$ signal strength in the different lepton channels from a fit to an Asimov dataset scaled to $3000\,\text{fb}^{-1}$ and $\sqrt{s} = 14\,\text{TeV}$, without any scaling applied on the systematic uncertainties.

References

[1] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321 (cit. on p. 2).

[2] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964) 132 (cit. on p. 2).

[3] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508 (cit. on p. 2).

[4] G. Guralnik, C. Hagen and T. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585 (cit. on p. 2).

[5] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, Phys. Rev. **145** (1966) 1156 (cit. on p. 2).

[6] T. Kibble, *Symmetry Breaking in Non-Abelian Gauge Theories*, Phys. Rev. **155** (1967) 1554 (cit. on p. 2).

[7] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, arXiv: 1207.7214 [hep-ex] (cit. on p. 2).

[8] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30, arXiv: 1207.7235 [hep-ex] (cit. on p. 2).

[9] ATLAS Collaboration, *Observation of H → b̄b decays and VH production with the ATLAS detector*, Phys. Lett. B **786** (2018) 59, arXiv: 1808.08238 [hep-ex] (cit. on p. 2).

[10] CMS Collaboration, *Observation of Higgs Boson Decay to Bottom Quarks*, Phys. Rev. Lett. **121** (2018) 121801, arXiv: 1808.08242 [hep-ex] (cit. on p. 2).

[11] ATLAS Collaboration, *Measurements of WH and ZH production in the H → b̄b decay channel in pp collisions at* 13 *TeV with the ATLAS detector*, Eur. Phys. J. C **81** (2021) 178, arXiv: 2007.02873 [hep-ex] (cit. on pp. 2, 3).

[12] S. Badger et al., *Les Houches 2015: Physics at TeV Colliders Standard Model Working Group Report*, (2016), arXiv: 1605.04692 [hep-ph], URL: https://lss.fnal.gov/archive/2016/conf/fermilab-conf-16-175-ppd-t.pdf (cit. on pp. 2, 3).

[13] N. Berger et al., *Simplified Template Cross Sections - Stage 1.1*, (2019), arXiv: 1906.02754 [hep-ph] (cit. on pp. 2, 3).

[14] *Combination of measurements of Higgs boson production in association with a W or Z boson in the b̄b decay channel with the ATLAS experiment at* $\sqrt{s}$ = 13 *TeV*, tech. rep., All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2021-051: CERN, 2021, URL: http://cds.cern.ch/record/2782535 (cit. on p. 2).

[15] A. Djouadi, J. Kalinowski and M. Spira, *HDECAY: A program for Higgs boson decays in the Standard Model and its supersymmetric extension*, Comput. Phys. Commun. **108** (1998) 56, arXiv: hep-ph/9704448 (cit. on p. 2).

[16] ATLAS Collaboration,
      Search for the Decay of the Higgs Boson to Charm Quarks with the ATLAS Experiment,
      Phys. Rev. Lett. **120** (2018) 211802, arXiv: 1802.04329 [hep-ex] (cit. on p. 2).

[17] CMS Collaboration, *A search for the standard model Higgs boson decaying to charm quarks*,
      JHEP **03** (2020) 131, arXiv: 1912.01662 [hep-ex] (cit. on p. 2).

[18] *Direct constraint on the Higgs-charm coupling from a search for Higgs boson decays to charm
      quarks with the ATLAS detector*, tech. rep., CERN, 2021,
      URL: https://cds.cern.ch/record/2771724 (cit. on p. 2).

[19] LHC Higgs Cross Section Working Group,
      Handbook of LHC Higgs Cross Sections: 3. Higgs Properties, CERN-2013-004 (2013),
      arXiv: 1307.1347 [hep-ph] (cit. on pp. 2, 5).

[20] D. de Florian et al.,
      Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector, (2016),
      arXiv: 1610.07922 [hep-ph] (cit. on pp. 2, 5–7).

[21] ATLAS Collaboration, *Measurement of VH, $H \to b\bar{b}$ production as a function of the vector-boson
      transverse momentum in 13 TeV pp collisions with the ATLAS detector*, JHEP **05** (2019) 141,
      arXiv: 1903.04618 [hep-ex] (cit. on p. 3).

[22] ATLAS Collaboration, *ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$
      events in pp collisions at $\sqrt{s}$ = 13 TeV*, Eur. Phys. J. C **79** (2019) 970,
      arXiv: 1907.05120 [hep-ex] (cit. on pp. 3, 4).

[23] ATLAS Collaboration,
      Expected performance of the ATLAS detector under different High-Luminosity LHC conditions,
      ATL-PHYS-PUB-2021-023, 2021, URL: http://cds.cern.ch/record/2776650
      (cit. on pp. 6, 10).

[24] ATLAS Collaboration, *Expected performance of the ATLAS detector at the High-Luminosity LHC*,
      ATL-PHYS-PUB-2019-005, 2019, URL: https://cds.cern.ch/record/2655304
      (cit. on pp. 6, 10).

[25] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Pixel Detector*, 2017,
      URL: https://cds.cern.ch/record/2285585 (cit. on p. 10).

[26] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Strip Detector*, 2017,
      URL: http://cds.cern.ch/record/2257755 (cit. on p. 10).

[27] ATLAS Collaboration, *Expected b-tagging Performance with the upgraded ATLAS Inner Tracker
      Detector at the High-Luminosity LHC*, ATL-PHYS-PUB-2020-005, 2020,
      URL: https://cds.cern.ch/record/2713377 (cit. on p. 10).

[28] ATLAS Collaboration, *Expected tracking and related performance with the updated ATLAS Inner
      Tracker layout at the High-Luminosity LHC*, ATL-PHYS-PUB-2021-024, 2021,
      URL: https://cds.cern.ch/record/2776651 (cit. on p. 10).