PHYSICAL REVIEW D **93,** 013001 (2016)

Prospects for measuring the Higgs boson coupling to light quarks

Gilad Perez,[1,*] Yotam Soreq,[1,†] Emmanuel Stamou,[1,‡] and Kohsaku Tobioka[1,2,3,§]

[1]*Department of Particle Physics and Astrophysics, Weizmann Institute of Science, Rehovot 7610001, Israel*
[2]*Theory Center, High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801, Japan*
[3]*Raymond and Beverly Sackler School of Physics and Astronomy, Tel-Aviv University, Tel-Aviv 6997801, Israel*

(Received 16 June 2015; published 5 January 2016)

We discuss the prospects to probe the light-quark Yukawa couplings to the Higgs boson. The Higgs coupling to the charm quark can be probed both via inclusive and exclusive approaches. On the inclusive frontier, we use our recently proposed method together with published experimental studies for the sensitivity of the Higgs coupling to bottom quarks to find that the high-luminosity LHC can be sensitive to modifications of the charm Yukawa of the order of a few times its standard model (SM) value. We also present a preliminary study of this mode for a 100 TeV hadronic machine (with similar luminosity) and find that the bound can be further improved, possibly within the reach of the expected signal in the SM. On the exclusive frontier, we use the recent ATLAS search for charmonia and photon final state. This study yields the first measurement of the background relevant to these modes. Using this background measurement we project that at the high-luminosity LHC, unless the analysis strategy is changed, the sensitivity of the exclusive final state to the charm Yukawa to the charm Yukawa will be rather poor, of the order of 50 times the SM coupling. We then use a Monte-Carlo study to rescale the above backgrounds to the $h \to \phi\gamma$ case and obtain a much weaker sensitivity to the strange Yukawa, of order of 3000 times the SM value. We briefly speculate what would be required to improve the prospects of the exclusive modes.

DOI: 10.1103/PhysRevD.93.013001

I. INTRODUCTION

Now that the Higgs particle has been discovered [1,2] the standard model (SM) is complete. It has a minimal scalar sector of electroweak (EW) symmetry breaking and is a theory consistent up to very high scales. Furthermore, the SM enjoys a set of accidental (exact and approximate) symmetries leading to: baryon–lepton number conservation, suppression of processes involving flavor-changing neutral currents and CP violation. Nevertheless, the flavor sector of the SM has a very particular structure. The Higgs couplings depend linearly on the masses, which implies that most of the Yukawas are small and hierarchical, leading to the SM flavor puzzle. However, at present there is no strong direct evidence for the validity of this particular structure. For instance, it is not impossible that the masses of the first two generation fermions originate from a different source of EW symmetry breaking thus leading to deviations from the simple SM relation between fermion masses and their coupling to the Higgs.

With new physics it is actually easy to obtain enhancements or suppressions in the strengths of Higgs to light-quark interactions. Furthermore, as the Higgs is rather light, within the SM it can only decay to particles that interact very weakly with it, with the dominant decay to $b\bar{b}$. A deformation of the Higgs couplings to the lighter SM particles, say for the charm or other quarks (see Refs. [3–12]), could compete with the Higgs–bottom coupling and would lead to a dramatic change of the Higgs phenomenology [13].

Our knowledge of the Higgs Yukawa couplings is mainly on the third-generation charged fermions. Though not yet fully conclusive, it is consistent with the SM Higgs mechanism of fermion-mass generation [14–19]. Regarding the first two generations, at present, we only have a rather weak upper bound on the corresponding signal strengths of muons and electrons [20,21]

$$\mu_\mu \leq 7, \qquad \mu_e \leq 4 \times 10^5, \tag{1}$$

at 95% confidence level (CL) where $\mu_f \equiv \frac{\sigma_h}{\sigma_h^{\rm SM}} \frac{{\rm BR}_{f\bar{f}}}{{\rm BR}_{f\bar{f}}^{\rm SM}}$ with σ_h standing for the Higgs-production cross section, ${\rm BR}_X \equiv {\rm BR}(h \to X)$ and the SM script indicating the SM case. In addition, in Ref. [22] we recasted the ATLAS [15] and CMS [18] studies of $Vh(b\bar{b})$ and obtained a first direct bound on the charm signal strength,

$$\mu_c \leq 270, \tag{2}$$

at 95% CL. These bounds are very weak, yet they are sufficient to exclude Higgs-coupling universality to quarks and charged leptons.

Let us summarize the current status of the theoretical and experimental activity relevant to probing Higgs to

*gilad.perez@weizmann.ac.il
†yotam.soreq@weizmann.ac.il
‡emmanuel.stamou@weizmann.ac.il
§kohsakut@post.tau.ac.il

light-quark couplings. On the theoretical frontier, it was demonstrated in Refs. [13,22] that inclusive charm-tagging enables the LHC experiments to constrain the charm Yukawa coupling. Furthermore, it was shown that the Higgs–charm coupling may be probed by looking at exclusive decay modes involving a c-$\bar{c}$ vector meson and a photon [23]. This makes the charm Yukawa coupling rather special among the light quarks as it can be probed both with inclusive and exclusive approaches. A similar mechanism, based on exclusive decays to light-quark states and gauge bosons $\gamma/W/Z$, was shown to yield a potential access to the Higgs–light-quark couplings [24]. (See also Refs. [25–28] for studies of exclusive EW gauge-boson decays and new-physics searches.) On the experimental side, ATLAS recently published two searches for supersymmetry [29,30], which employ charm-tagging (c-tagging) [31]. ATLAS further published an analysis that focuses on Higgs decays to quarkonia (e.g. J/ψ, Υ) plus a photon final states [32].[1]

In this paper, we discuss the prospects to probe light-quark Yukawa couplings to the Higgs boson at LHC run II, the high-luminosity LHC (HL-LHC) and possible future colliders. We will demonstrate that the HL-LHC can reach a sensitivity for the charm Yukawa up to few times the SM value by using the inclusive method recently proposed by us [22]. We also present a preliminary study of this mode for a 100 TeV hadronic machine assuming similar luminosity as at the HL-LHC. We shall find that the sensitivity can be further improved, possibly probing the signal expected in the SM. On the exclusive frontier we shall use the recent ATLAS result on the search for charmonia plus a photon final state [32]. This study yields the first measurement of the background relevant to these modes. It is dominated by a jet converted to photon or a real photon plus charmonia production. Given this background measurement, we project that, unless the analysis strategy is changed, the sensitivity of the exclusive final state to the charm Yukawa at the HL-LHC will be rather poor, of the order of 50 times the SM coupling. We then use a PYTHIA simulation to rescale the above backgrounds to the $h \to \phi\gamma$ case and obtain a much weaker sensitivity to the strange Yukawa of order of 3000 times the SM value.

In the following section we discuss the prospects for probing the charm Yukawa coupling using charm tagging. In Sec. III we focus on the exclusive case. For completeness, we also briefly describe in Sec. II C the corresponding status for e^+e^- machines. We conclude in Sec. IV.

II. INCLUSIVE ANALYSIS

We begin by estimating the future sensitivity of the LHC and a future 100 TeV pp collider to probe the $h \to c\bar{c}$ signal strength, μ_c, and the charm Yukawa, $\kappa_c \equiv y_c/y_c^{\rm SM}$, via the inclusive method proposed in Ref. [22]. The method takes advantage of the fact that the signal strength in searches for $h \to b\bar{b}$ requires two b-tagged jets. In this way, the same analyses are also sensitive to $h \to c\bar{c}$ events because charm jets (c-jets) pass the tag criteria with a non-negligible rate. To account for such events, the $h \to b\bar{b}$ signal strength, μ_b, is extended

$$\mu_b \equiv \frac{\sigma_h {\rm BR}_{b\bar{b}}}{\sigma_h^{\rm SM}{\rm BR}_{b\bar{b}}^{\rm SM}} \to \frac{\sigma_h {\rm BR}_{b\bar{b}}\epsilon_{b_1}\epsilon_{b_2} + \sigma_h {\rm BR}_{c\bar{c}}\epsilon_{c_1}\epsilon_{c_2}}{\sigma_h^{\rm SM}{\rm BR}_{b\bar{b}}^{\rm SM}\epsilon_{b_1}\epsilon_{b_2} + \sigma_h^{\rm SM}{\rm BR}_{c\bar{c}}^{\rm SM}\epsilon_{c_1}\epsilon_{c_2}} = \left(\mu_b + \frac{{\rm BR}_{c\bar{c}}^{\rm SM}}{{\rm BR}_{b\bar{b}}^{\rm SM}}\frac{\epsilon_{c_1}\epsilon_{c_2}}{\epsilon_{b_1}\epsilon_{b_2}}\mu_c\right)\Big/\left(1 + \frac{{\rm BR}_{c\bar{c}}^{\rm SM}}{{\rm BR}_{b\bar{b}}^{\rm SM}}\frac{\epsilon_{c_1}\epsilon_{c_2}}{\epsilon_{b_1}\epsilon_{b_2}}\right). \quad (3)$$

Here, ϵ_b and ϵ_c are the efficiencies to tag jets originating from bottom and charm quarks, respectively. The subscripts 1 and 2 refer to the efficiency of tagging the first and second jet, respectively and ${\rm BR}_{c\bar{c}}^{\rm SM}/{\rm BR}_{b\bar{b}}^{\rm SM} \simeq 5\%$ [34]. For brevity, we define the ratio of tagging efficiencies, $\epsilon_{c/b}^2 \equiv (\epsilon_{c_1}\epsilon_{c_2})/(\epsilon_{b_1}\epsilon_{b_2})$.

An analysis employing a single jet-tagger, such as medium b-tagging, constrains only a linear combination of μ_b and μ_c. To be able to separately obtain μ_c we need at least two analyses with different ratios, $\epsilon_{c/b}^2$, in order to break the degeneracy in Eq. (3). The sensitivity is best when in the employed tagger $\epsilon_{c/b}^2$ is large meaning that many c-jets are being tagged. This situation is realized if we combine b-tagging with *charm tagging* (c-tagging), which is a jet-tagger optimized for c-jets. For the b-tagger we will always consider the medium b-tagging working point as in Ref. [35]. For c-tagging we shall use what ATLAS already employed at run I [29–31] and refer to it as c-tagging I. Moreover, given the recent installation of the new Insertable B-Layer (IBL) subdetector [36] in the ATLAS detector, the capability for c-tagging is expected to be much improved. Related to this, Ref. [37] shows possible improvement of the life-time resolution of $B_s^0 \to J/\psi\phi$ decay by 30%, thanks to the IBL. Thus, we consider two additional c-taggings, referred to as c-tagging II and c-tagging III. All tagging efficiencies are summarized in Table I in which ϵ_l denotes the efficiency to tag a light jet.

c-tagging uses almost the same experimental information as b-tagging. As a result jets that are c-tagged may but also may not pass the b-tagging criteria [31,38]. The actual experiments can employ b- and c-tagging simultaneously, but for our analysis it is not possible to fully take into

TABLE I. The tagging efficiencies for the four jet-taggers used in our analysis.

	ϵ_b	ϵ_c	ϵ_l
b-tagging	70%	20%	1.25%
c-tagging I	13%	19%	0.5%
c-tagging II	20%	30%	0.5%
c-tagging III	20%	50%	0.5%

[1]Note added: during the reviewing process of this article the CMS collaboration presented results on a similar analysis [33].

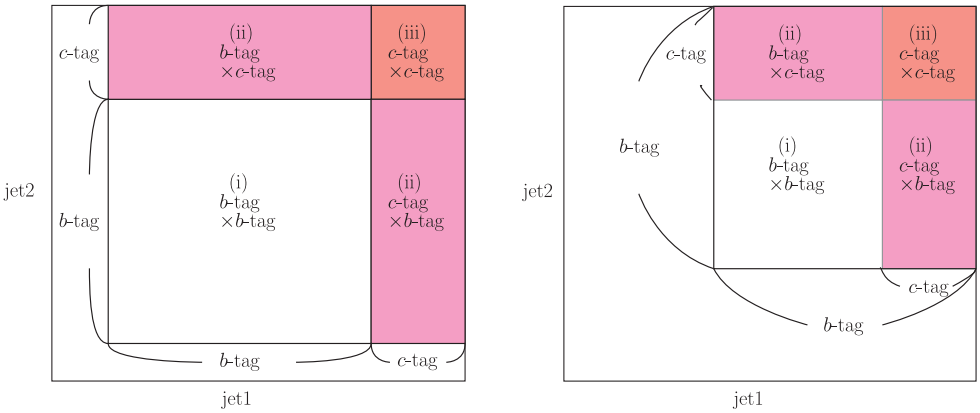

FIG. 1. Uncorrelated and correlated scenario in the left and right panel, respectively. c-tagged jets are a subset of b-tagged jets in the correlated scenario, while in the uncorrelated scenario b- and c-tagging cover different jets.

account this correlation. Therefore, whenever possible, we study the following two extreme scenarios.

Uncorrelated scenario b- and c-tagging are uncorrelated and possible to employ simultaneously. In this case, if a jet is b-tagged, the jet is never c-tagged, and vice versa, i.e., there is no overlap between b- and c-tagged jets.

Correlated scenario c-tagging is fully correlated with b-tagging and is a tighter version of b-tagging. In this case, c-tagged jets are always also b-tagged, but the opposite is not necessary, i.e., c-tagged jets are a subset of b-tagged jets.

The actual situation is expected to be something between the following two scenarios. However, we will show in Sec. II A that the final results in the two scenarios are similar. We define three categories by combining taggers: (i) two jets are b-tagged, (ii) one is b-tagged and one is c-tagged, (iii) two jets are c-tagged. To avoid double counting in categories (i) and (ii) of the correlated scenario, the c-tagged jets are removed from b-tagged jets. A schematic picture of the three categories for the scenarios is shown in Fig. 1. In our analysis we use RunDec [39] to compute the quark masses at the Higgs mass using the inputs from PDG [40] finding $m_s = 53.2$ MeV, $m_c = 0.612$ GeV and $m_b = 2.78$ GeV.

A. LHC 14 TeV

For LHC run II and HL-LHC, we base our study on the dedicated ATLAS analysis for the future measurement of μ_b based on the Higgs production associated with W/Z bosons at LHC 14 TeV with 300 fb^{-1} and 3000 fb^{-1} [35]. The analysis requires two b-tagged jets, i.e., it employs a single tagger, which is insufficient to disentangle μ_b and μ_c. To discuss the future sensitivities, we thus need to estimate the number of signal and background events in the three categories for the correlated and uncorrelated scenario once also c-tagging is employed. We utilize the Monte-Carlo (MC) studies presented in Figs. 3–6 of Ref. [35] in the following way. These figures provide the number of events in each bin for signal and for each background after applying all cuts and requiring two b-tagged jets. Let us consider a bin of signal or a specific background that originally has an x- and a y-jet, where $x, y = b, c, l$ (real b-jet, c-jet, light-jet), and the number provided is N.[2] We then obtain the number of events for categories (i)–(iii) in uncorrelated and correlated scenarios as below.

Uncorrelated scenario:

$$\begin{aligned} N^{(\mathrm{i})} &= N, \\ N^{(\mathrm{ii})} &= \frac{\epsilon_x^{(b\text{-tag})}\epsilon_y^{(c\text{-tag})} + \epsilon_x^{(c\text{-tag})}\epsilon_y^{(b\text{-tag})}}{\epsilon_x^{(b\text{-tag})}\epsilon_y^{(b\text{-tag})}} N, \\ N^{(\mathrm{iii})} &= \frac{\epsilon_x^{(c\text{-tag})}\epsilon_y^{(c\text{-tag})}}{\epsilon_x^{(b\text{-tag})}\epsilon_y^{(b\text{-tag})}} N, \end{aligned} \tag{4}$$

Correlated scenario:

$$\begin{aligned} N^{(\mathrm{i})} &= N - N^{(\mathrm{ii})} - N^{(\mathrm{iii})}, \\ N^{(\mathrm{ii})} &= \frac{\epsilon_x^{(b\text{-tag})}\epsilon_y^{(c\text{-tag})} + \epsilon_x^{(c\text{-tag})}\epsilon_y^{(b\text{-tag})}}{\epsilon_x^{(b\text{-tag})}\epsilon_y^{(b\text{-tag})}} N - 2N^{(\mathrm{iii})}, \\ N^{(\mathrm{iii})} &= \frac{\epsilon_x^{(c\text{-tag})}\epsilon_y^{(c\text{-tag})}}{\epsilon_x^{(b\text{-tag})}\epsilon_y^{(b\text{-tag})}} N. \end{aligned} \tag{5}$$

The rescaling is done on a bin-by-bin basis. After rescaling different background differently ($N \to B_X$ with X denoting the type of background), we obtain the total background for each category, $B^{(\mathrm{i,ii,iii})}$, by summing all backgrounds, $B^{(\mathrm{i})} = \sum_X^{\mathrm{all}} B_X^{(\mathrm{i})}$ and analogously for (ii) and (iii). The expected signal for each category is straightforward, $N \to S$.

To obtain the future sensitivity we then follow the statistical procedure described in Ref. [22]. Given the

[2]For instance, $x = y = b$ for $h \to b\bar{b}$ signal and $t\bar{t}$ background; $x = y = c$ for $W + c\bar{c}$ background; $(x = b, y = l)$ for single top background.

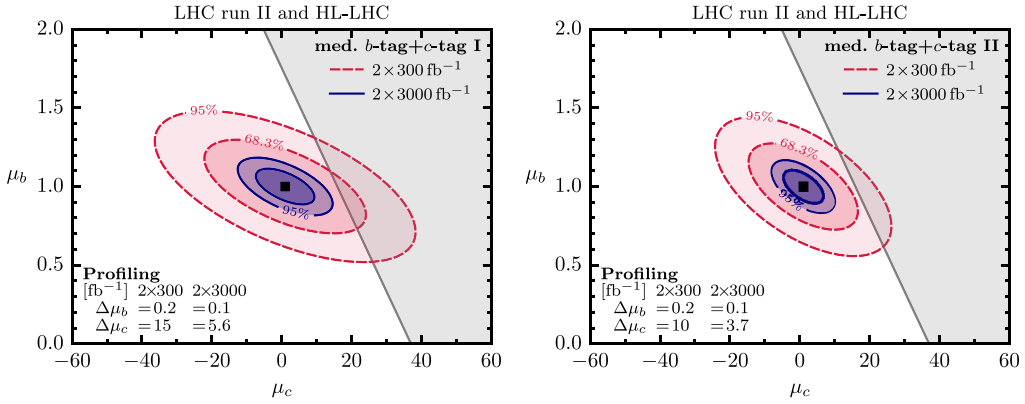

FIG. 2. 300 fb^{-1} and 3000 fb^{-1} prospects for the signal strengths at the LHC, for $h \to b\bar{b}$ and $h \to c\bar{c}$ based on b- and c-tagging for the uncorrelated scenario employing c-tagging I (left panel) and c-tagging II (right panel). The grey shaded region is unphysical unless Higgs production is modified with respect to the SM case. The profiled likelihood ratio [41] is used for the respective 1-σ uncertainty of $\Delta\mu_b$ and $\Delta\mu_c$.

expectation of signal and background, we construct a likelihood function of μ_c and μ_b based on the Poisson probability-distribution function, and use the likelihood ratio for parameter estimates. In Fig. 2 we present the future reach for the signal strengths of $h \to b\bar{b}$ and $h \to c\bar{c}$ in the uncorrelated scenario by combining b-tagging with c-tagging I (left panel) and II (right panel). Note that the correlated scenario cannot be defined in the case of c-tagging II and III, see Appendix for details and Fig. 6 therein for the c-tagging I result of the correlated scenario. We obtain the expected uncertainty on μ_c (μ_b) by profiling μ_b (μ_c). We list the 1-σ ranges for μ_c and μ_b for different scenarios and employed c-tagging assuming the total luminosity of 2×300 fb^{-1} and 2×3000 fb^{-1} expected at LHC run II and HL-LHC in Table II. The sensitivities for c-tagging I in the correlated and uncorrelated scenario are similar, so we conclude that these results represent well the actual future reach.

The translation of the constraints of the charm and bottom signal strengths to the Yukawa couplings themselves requires some caution. If we assume that Higgs production is not modified with respect to the SM, the signal strengths are given by $\mu_c = \mathrm{BR}_{c\bar{c}}/\mathrm{BR}_{c\bar{c}}^{\mathrm{SM}}$ and $\mu_b = \mathrm{BR}_{b\bar{b}}/\mathrm{BR}_{b\bar{b}}^{\mathrm{SM}}$. In the extreme case in which the Higgs decays solely to charms and bottoms, $\mathrm{BR}_{c\bar{c}} + \mathrm{BR}_{b\bar{b}} = 1$ holds and the two rates are linearly dependent. As long as the measured values of μ_c and μ_b are consistent with this hypothesis, an arbitrary large value of $\kappa_c \equiv y_c/y_c^{\mathrm{SM}}$ is allowed with some $\kappa_b \equiv y_b/y_b^{\mathrm{SM}}$. This corresponds to a flat direction in the κ_c–κ_b plane. In other words, as long as the experimental result is consistent with

$$\mu_c \mathrm{BR}_{c\bar{c}}^{\mathrm{SM}} + \mu_b \mathrm{BR}_{b\bar{b}}^{\mathrm{SM}} > 1, \tag{6}$$

one cannot constrain κ_c and κ_b assuming only SM Higgs production. We illustrate this case with the grey shaded region in Fig. 2. If this region overlaps with the allowed regions of μ_c–μ_b (colored ellipses), it means that there is a flat direction in the κ_c–κ_b if SM Higgs production is assumed.

A charm Yukawa much larger than in the SM enhances the Higgs production in the Vh production channel; for $\kappa_c \sim \mathcal{O}(100)$ it is twice as large as the SM expectation [22]. This mechanism enabled us to obtain a direct constraint

TABLE II. 1-σ uncertainties after profiling [41] for the signal strengths, μ_b and μ_c, and the Yukawa-coupling modifications, κ_b and κ_c at the LHC with $\sqrt{s} = 14$ TeV. The results for different scenarios (uncorrelated and correlated) and different c-tagging are shown for a total luminosity of 2×300 fb^{-1} and 2×3000 fb^{-1}. The error in κ_b and κ_c is asymmetric; we only show the upper bound. The 95% CL regions for κ_b and κ_c are given in the last two columns.

$\mathcal{L}$	$\sqrt{s} = 14$ TeV	$\Delta\mu_b$	$\Delta\mu_c$	$\Delta\kappa_b$	$\Delta\kappa_c$	κ_b @ 95% CL	$\lvert\kappa_c\rvert$ @ 95% CL
2×300 fb^{-1}	Correlated c-tagging I	± 0.22	± 15	$+1.3$	$+8.6$	[0.67, 7.07]	<37
	Uncorrelated c-tagging I	± 0.20	± 15	$+1.5$	$+9.4$	[0.69, 7.16]	<38
	Uncorrelated c-tagging II	± 0.18	± 10	$+0.5$	$+4.1$	[0.70, 4.70]	<21
	Uncorrelated c-tagging III	± 0.17	± 5.8	$+0.29$	$+2.2$	[0.70, 1.90]	<6.0
2×3000 fb^{-1}	Correlated c-tagging I	± 0.084	± 5.6	$+0.20$	$+2.1$	[0.84, 1.57]	<5.5
	Uncorrelated c-tagging I	± 0.075	± 5.6	$+0.20$	$+2.1$	[0.85, 1.60]	<5.6
	Uncorrelated c-tagging II	± 0.069	± 3.7	$+0.12$	$+1.4$	[0.86, 1.30]	<3.7
	Uncorrelated c-tagging III	± 0.065	± 2.0	$+0.087$	$+0.82$	[0.87, 1.18]	<2.5

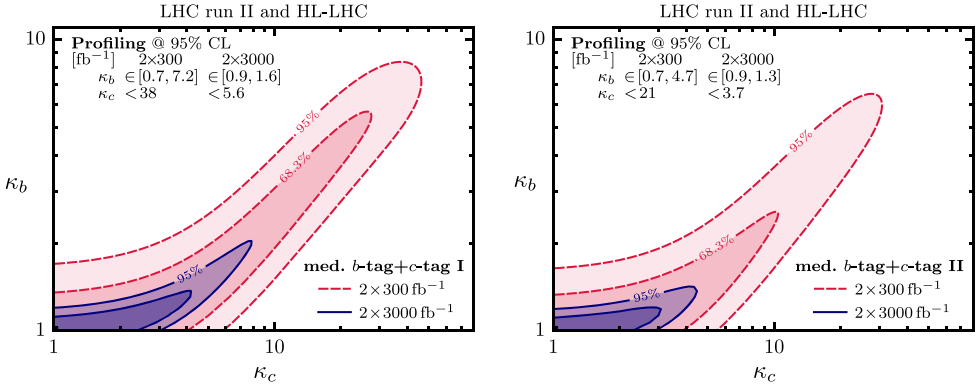

FIG. 3. 300 fb^{-1} and 3000 fb^{-1} prospects for probing κ_b and κ_c at the LHC, with $h \to b\bar{b}$ and $h \to c\bar{c}$ based on b- and c-tagging for the uncorrelated scenario employing c-tagging I (left panel) and c-tagging II (right panel). All other Higgs couplings are assumed to be like in the SM. The profiled likelihood ratio [41] is used for the respective reach on κ_b and κ_c.

on κ_c already with the available 8 TeV data set [22]. For the 14 TeV projection, as seen in Fig. 2, considering non-SM production is essential to constrain κ_c with 300 fb^{-1}, while for the high-luminosity stage its effect is minor. Details of non-SM Vh production at 14 TeV are discussed in the Appendix.

In the analysis for the prospects for couplings, we float only κ_b and κ_c freely and assume the other couplings stay as in the SM, in particular, $\kappa_V = 1$. In Fig. 3 we show the expected future reach in the κ_c–κ_b plane taking into account non-SM Vh production for the uncorrelated scenario employing c-tagging I and II. We obtain the expected upper bound on κ_c (κ_b) by profiling over κ_b (κ_c) [41]. The 1-σ uncertainties, as well as the 95% CL ranges for κ_b and κ_c for the different cases are listed in Table II.

Finally, we compare the projected reach of the direct charm-Yukawa measurement to the indirect bound from the global analysis of the Higgs couplings. For $\kappa_c < 5$ the effects of non-SM Higgs production due to large charm Yukawa are negligible and the constraint on κ_c can be deduced from the bound on the untagged Higgs decays. The ATLAS projection for the HL-LHC is $\mathrm{BR}_{\mathrm{untagged}} < 10\%$ at 95% CL (without theoretical uncertainties) [42]. It can be interpreted as $\kappa_c \lesssim 2$. This upper bound is comparable to our projection with c-tagging III, see Table III.

B. pp collider with $\sqrt{s} = 100$ TeV

In this section we perform a first study of the sensitivity reach of a 100 TeV pp collider in measuring the charm-quark Yukawa via the inclusive rate. As a byproduct we obtain also the sensitivity of a bottom-Yukawa measurement at 100 TeV. For such a machine, there exists no detailed, fully realistic study for the $h \to b\bar{b}$ prospects like the one of ATLAS for 14 TeV [35], which we employed for our 14 TeV study. For this reason, we investigate the 100 TeV reach by simulating the signal and main backgrounds at the leading-order (LO) parton level using MadGraph 5.2 [43] and multiplying with the inclusive k-factors.

The main difficulty remains to find a way to reduce background, while keeping as many signal $h \to c\bar{c}$ events as possible. To this end, we follow two orthogonal directions. First, we look into the boosted-Higgs regime in which the Higgs has $p_T(h) > 350$ GeV. In this case we rely on available jet-substructure techniques to extract the signal and reduce the $t\bar{t}$ background that dominates in this kinematic configuration. Second, we "unboost" the Higgs by binning in H_T. This way the S/B ratio for $h \to c\bar{c}$ is large in lower H_T bins as the main background, $t\bar{t}$, typically has higher H_T than the signal.

We shall find that the sensitivity reach for the bottom Yukawa is not significantly different in the two cases, as in both there are enough $h \to b\bar{b}$ events. For the charm Yukawa, however, the "unboosted" analysis appears more promising, due to the fact that it accepts a larger fraction of the rather rare signal events. Given that the capabilities of a future 100 TeV collider and the advancements with respect to current experiments are currently not well known, the fact that our projections will be based on LO simulations suffices. However, it is important to note that we expect significantly better results from realistic studies that employ multiple bins with increased S/B ratio. For instance, the projected uncertainty on $\Delta\mu_b$ from Ref. [35] would be approximately a factor of 2 larger without binning. Furthermore, in Ref. [35] the sensitivity of a purely cut-based analysis was compared to the one obtained employing multivariate techniques. In the latter the uncertainty is decreased by roughly 25%. This gives us confidence that the results presented here are conservative and there is room for improvements in the future.

1. Boosted-Higgs analysis

The field of searching for boosted massive particles and jet-substructure is very rich and we shall not attempt to describe it here in any detail (see e.g., [44] for a recent review). Instead we focus on one specific method to study the sensitivity to the Higgs couplings to bottom and charm

TABLE III. 1-σ uncertainties after profiling [41] for the signal strengths, μ_b and μ_c, and the Yukawa-coupling modifications, κ_b and κ_c, for a boosted Higgs using jet substructure at a pp collider with $\sqrt{s} = 100$ TeV. The results for different scenarios (uncorrelated and correlated) and different c-tagging are shown for a total luminosity of 3000 fb^{-1} and 2×3000 fb^{-1}. The 95% CL regions for κ_b and κ_c are given in the last two columns.

$\mathcal{L}$	Boosted $\sqrt{s} = 100$ TeV	$\Delta\mu_b$	$\Delta\mu_c$	$\Delta\kappa_b$	$\Delta\kappa_c$	κ_b @ 95% CL	$\lvert\kappa_c\rvert$ @ 95% CL
3000 fb^{-1}	Correlated c-tagging I	±0.048	±5.7	$+0.22$	$+2.2$	[0.89, 1.72]	<6.1
	Uncorrelated c-tagging I	±0.067	±5.8	$+0.22$	$+2.2$	[0.86, 1.71]	<6.1
	Uncorrelated c-tagging II	±0.063	±4.0	$+0.14$	$+1.5$	[0.87, 1.34]	<4.0
	Uncorrelated c-tagging III	±0.059	±2.2	$+0.089$	$+0.90$	[0.87, 1.19]	<2.6
2×3000 fb^{-1}	Correlated c-tagging I	±0.034	±4.0	$+0.14$	$+1.6$	[0.92, 1.36]	<4.1
	Uncorrelated c-tagging I	±0.048	±4.1	$+0.14$	$+1.6$	[0.89, 1.36]	<4.1
	Uncorrelated c-tagging II	±0.044	±2.8	$+0.090$	$+1.1$	[0.90, 1.20]	<3.1
	Uncorrelated c-tagging III	±0.042	±1.6	$+0.061$	$+0.67$	[0.90, 1.13]	<2.2

quarks at a 100 TeV collider. The sensitivity to the $h \to b\bar{b}$ decay mode with the Higgs being boosted and produced in association with a leptonically decaying W at the LHC with $\sqrt{s} = 8$ and 13 TeV has been analysed in Ref. [45]. The study adopted the template overlap method [46,47] (see also [48] for the ATLAS implementation of the method). For our study we will use the signal efficiency and background-rejection rates of the "Cuts 5" scenario in Ref. [45] and a cut on the fat jet containing the Higgs (or its $b\bar{b}$ daughter products), $p_T(h) > 350$ GeV. Given the above requirements, the Wh signal has an efficiency of 22% while the $t\bar{t}$ and $Wb\bar{b}$ backgrounds have a fake rate of only 1.3% and 5.1%, respectively (see Table III in Ref. [45]). We will assume that these jet-substructure efficiencies do not change from 13 to 100 TeV.

To make use of these jet-substructure results for our 100 TeV study we follow their analysis and simulate signal and background for both 13 and 100 TeV applying the same basic cuts. The main requirement is the presence of two b-tagged jets inside the fat jet and a few basic cuts the most relevant of which are $p_T(W), p_T(\text{fat jet}) > 350$ GeV and $0.4 < \Delta R_{bb} < 0.8$ (see Eq. (12) and (13) in Ref. [45]). Their simulation of the signal Wh, and the backgrounds $Wb\bar{b}$, $t\bar{t}$ includes matching to parton shower and next-to-leading-order (NLO) k-factors from MCFM 6.3 [49]. We include these NLO effects by rescaling our LO parton-level simulation at 13 TeV to their results and applying the same rescaling factors to the 100 TeV results. In a similar way, we also include in our study the two-lepton sample, namely Zh production with leptonically decaying Z and the dominant corresponding backgrounds $Zb\bar{b}$ and leptonic $t\bar{t}$. We use the same rescaling factors as for the Wh sample. For a charm-Yukawa measurement it is necessary to include the $Wc\bar{c}$ and $Zc\bar{c}$ backgrounds, because they can be relevant when we employ c-tagging with a large tagging efficiency for charm quarks. The rescaling factor from $Wb\bar{b}$ is used for both of them to rescale their 100 TeV cross sections. Finally, we note that the main $t\bar{t}$ background in the one-lepton analysis originates from a fat jet consisting of a bottom quark and a mistagged charm quark from the associated hadronically decaying W [45]. Such a configuration is absent in the two-lepton, Zh, sample, which has thus a reduced $t\bar{t}$ background.

Having simulated the signal and the dominant backgrounds at 100 TeV and rescaled with the 13 TeV k-factors we find the expected signal and background events and multiply with the corresponding efficiencies depending on whether b-tagging or c-tagging I, II, III is applied. The rest of the analysis is analogous to the 14 TeV study. Also here we combine b-tagging with one c-tagging scenario and profile the resulting distributions to obtain the future sensitivity in signal strengths and Yukawa-coupling modifications. Here, we present results assuming a total luminosity of 3000 fb^{-1} and 2×3000 fb^{-1}. In Fig. 4 we show the result of combining b- and c-tagging II for the uncorrelated scenario using 2×3000 fb^{-1} of data. The expected signal-strength (Yukawa-coupling) regions of sensitivity are plotted in the left (right) panel. The 95% CL region has no overlap with the shaded grey region that is unphysical if SM-Higgs production is assumed. Therefore, modifications in the Higgs production are small and we can safely neglect them. In Table III we list the projected 1-σ uncertainties for μ_b, μ_c and κ_b, κ_c for various scenarios. For the Yukawa coupling modification we also present the 95% CL region after profiling.

Regarding μ_b, we find that the expected precision is better in the correlated scenario than in the uncorrelated one, see Table III. The reason is that, as discussed, the main background in this boosted regime is $t\bar{t}$ with a mistagged c quark. The separation of these background events with the correlated prescription of Eq. (5) assigns most of them to category (ii). This results in an increased S/B ratio in category (i) and leads to a better expected precision in μ_b than in the corresponding uncorrelated case. As far as μ_c is concerned, we find moderate improvements in sensitivity with respect to HL-LHC, compare the results in Table II with those in Table III. For instance for 2×3000 fb^{-1} using c-tagging II the improvement in the sensitivity for μ_c is approximately 24%. The reason for this is that, even though the jet-substructure cuts remove a lot of background, a lot of

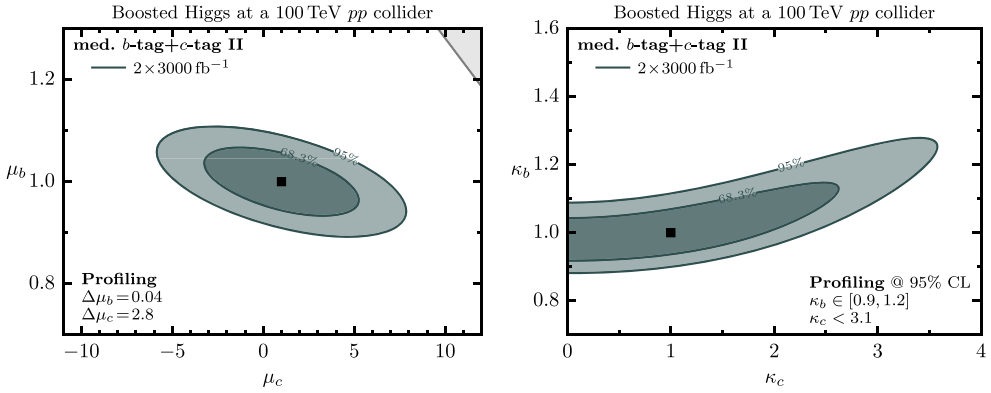

FIG. 4. Prospects for probing μ_b, μ_c (left panel), and κ_b, κ_c (right panel) at a 100 TeV collider, with $h \to b\bar{b}$ and $h \to c\bar{c}$ based on b- and c-tagging II for the uncorrelated scenario using boosted Higgses and 2×3000 fb^{-1} of data. The profiled likelihood ratio [41] is used for the respective reach of $\Delta\mu_b$, $\Delta\mu_c$, κ_b and κ_c.

$h \to c\bar{c}$ signal is also lost in this boosted regime. We, therefore, look into the orthogonal direction of "unboosting" the Higgs.

2. "Unboosted"-Higgs analysis

Our "unboosted" 100 TeV analysis is conceptually not much different than the 14 TeV projection analysis of ATLAS [35]. Unlike the previous analysis, here we also include less energetic events in which the Higgs is not necessarily boosted. Relaxing the requirement for a large boost increases the otherwise statistically challenged $h \to c\bar{c}$ signal. We therefore expect an improved sensitivity for μ_c without affecting much the μ_b sensitivity.

We will use three bins of inclusive H_T,

$$\begin{aligned} H_T &< 340 \text{ GeV}, \\ 340 \text{ GeV} < H_T &< 500 \text{ GeV}, \\ 500 \text{ GeV} < H_T. & \end{aligned} \tag{7}$$

In the lower H_T bins the $t\bar{t}$ background is reduced. The main basic cuts that we apply are: $p_T(W/Z) > 100$ GeV, $p_T(j_1) > 60$ GeV, $p_T(j_2) > 40$ GeV, $\Delta R(j_1, j_2) > 0.4$ and 100 GeV $< m_{j_1 j_2} <$ 140 GeV, where j_1 and j_2 is the leading and next-to-leading in p_T jet, respectively. Following Ref. [35] we demand for the one-lepton sample $E_T^{\text{miss}} > 40$ GeV and for the two-lepton channel $E_T^{\text{miss}} <$ 60 GeV. We simulate the same background processes as for the boosted 100 TeV analysis and rely on LO parton-level simulation supplemented with the same inclusive k-factors. A difference with respect to the boosted analysis is that the $t\bar{t}$ background in this case is not completely dominated by the mistagged c-quark, i.e. the background from two b-quarks is at least as equally important. The ratio of these two backgrounds depends on the jet-tagging employed, so in our study, we keep them as separate backgrounds that depend differently on ϵ_b and ϵ_c.

Having simulated the Wh and Zh signal processes and the corresponding background processes for the three H_T bins we find the projected sensitivities by combining all bins using b-tagging together with one c-tagging scenario. The details are analogous to the boosted analysis. In Table IV we list the projected 1-σ uncertainties for signal strengths and Yukawa-coupling modifications as well as

TABLE IV. 1-σ uncertainties after profiling [41] for the signal strengths, μ_b and μ_c, and the Yukawa-coupling modifications, κ_b and κ_c, for nonboosted Higgses at a pp collider with $\sqrt{s} = 100$ TeV. The results for different scenarios (uncorrelated and correlated) and different c-tagging are shown for a total luminosity of 3000 fb^{-1} and 2×3000 fb^{-1}. The 95% CL regions for κ_b and κ_c are given in the last two columns.

$\mathcal{L}$	unboosted $\sqrt{s} = 100$ TeV	$\Delta\mu_b$	$\Delta\mu_c$	$\Delta\kappa_b$	$\Delta\kappa_c$	κ_b @ 95% CL	$\lvert\kappa_c\rvert$ @ 95% CL
3000 fb^{-1}	correlated c-tagging I	±0.047	±3.2	$+0.086$	$+1.1$	[0.91, 1.20]	<3.0
	uncorrelated c-tagging I	±0.047	±3.2	$+0.086$	$+1.1$	[0.91, 1.20]	<3.0
	uncorrelated c-tagging II	±0.044	±2.2	$+0.057$	$+0.80$	[0.92, 1.12]	<2.4
	uncorrelated c-tagging III	±0.041	±1.1	$+0.043$	$+0.46$	[0.92, 1.09]	<1.8
2×3000 fb^{-1}	correlated c-tagging I	±0.034	±2.3	$+0.058$	$+0.83$	[0.93, 1.13]	<2.5
	uncorrelated c-tagging I	±0.034	±2.3	$+0.058$	$+0.83$	[0.93, 1.13]	<2.5
	uncorrelated c-tagging II	±0.031	±1.5	$+0.039$	$+0.59$	[0.94, 1.08]	<2.1
	uncorrelated c-tagging III	±0.029	±0.8	$+0.030$	$+0.34$	[0.95, 1.06]	<1.6

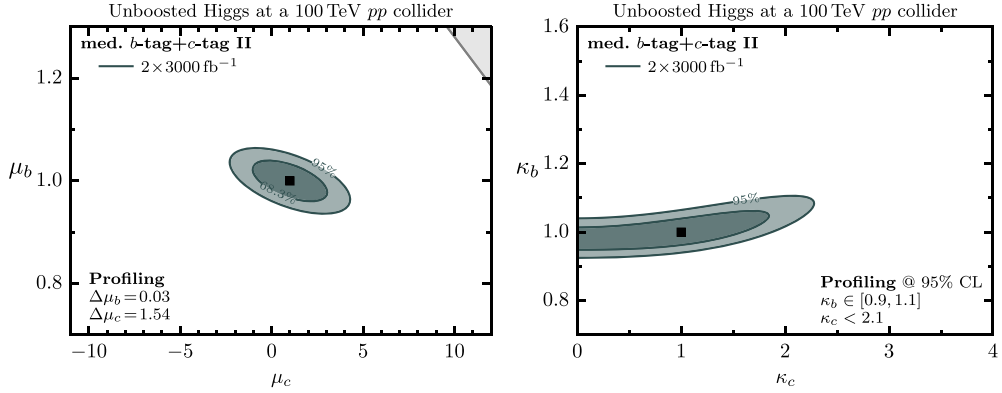

FIG. 5. Prospects for probing μ_b, μ_c (left panel), and κ_b, κ_c (right panel) at a 100TeV collider, with $h \to b\bar{b}$ and $h \to c\bar{c}$ based on b- and c-tagging II for the uncorrelated scenario using nonboosted Higgses and $2 \times 3000\text{fb}^{-1}$ of data. The profiled likelihood ratio [41] is used for the respective reach of $\Delta\mu_b$, $\Delta\mu_c$, κ_b and κ_c.

the 95% CL region for the latter assuming 3000 fb^{-1} and 2×3000 fb^{-1}. In Fig. 5 we present the expected signal-strength and Yukawa-coupling sensitivity regions when c-tagging II is employed and 2×3000 fb^{-1} of data are assumed. In this case, we find for the signal strength μ_c a significant improvement of approximately 60% in the projected uncertainty with respect to the analogous 14 TeV expectation. We also find the promising result that, if c-tagging III is possible, a factor of ~2 modification in the charm-quark Yukawa will be probed by a 100 TeV collider with 95% CL.

C. e^+e^- colliders

Since the future e^+e^- colliders will be experiments aiming at Higgs precision, we expect significant improvement in measuring Higgs couplings. Here, we show the summary of their prospects. There are two types of e^+e^- colliders proposed, linear type, such as ILC, and circular type, such as TLEP. The advantage of the ILC is that it relies on established technology and its collision energy can be potentially increased up to 1 TeV, while in the TLEP one expects much larger luminosity by an order of magnitude and so many Higgs events will be collected.

The technical design report of ILC [50] presents the expected precision in different channels based on dedicated analyses,

$$\Delta\mu_b = 1.1\%(0.47\%), \qquad \Delta\mu_c = 7.4\%(6.7\%),$$
$$\Delta\mu_\tau = 4.2\%(3.5\%), \qquad \Delta\mu_\mu = 100\%(32\%) \tag{8}$$

at $\sqrt{s} = 250$ GeV (1 TeV) with 250 fb^{-1} (1 ab^{-1}). Also the TLEP presents a preliminary analysis [51] of the expected precisions

$$\Delta\mu_b = 0.2\%, \qquad \Delta\mu_c = 1.2\%,$$
$$\Delta\mu_\tau = 0.7\%, \qquad \Delta\mu_\mu = 13\% \tag{9}$$

at $\sqrt{s} = 240$ GeV with 10 ab^{-1} ($\Delta\mu_c$ is based on an extrapolation of the ILC [50]).

Furthermore, there are ongoing discussions on whether it would be possible to also run precisely on the Higgs resonance and being able to measure the electron Yukawa [52]. The above information is based on the inclusive approach to particle identification. As one cannot apply u, d, s-jet-tagging with reasonable efficiencies, no direct information can be extracted on the Higgs coupling to these light-quark states.

III. EXCLUSIVE HIGGS DECAYS

Recently, the ATLAS collaboration provided the first upper bound on the rate for exclusive Higgs decays in the $h \to J/\psi\gamma$ mode, $\sigma_h \text{BR}_{J/\psi\gamma} < 33$ fb [32]. This result is interesting not only because it can be interpreted as a bound on the Higgs couplings, in particular on the charm Yukawa [22], but also because other exclusive decay modes are potentially subject to similar backgrounds. In Ref. [32] it is stated that the main background is inclusive quarkonium production where a jet in the event is reconstructed as a photon. This knowledge allows us to estimate the reach of future searches for exclusive Higgs decays in the $J/\psi\gamma$ channel as well as in other modes such as $\phi\gamma$ by using a PYTHIA simulation of the backgrounds and rescaling to the ATLAS data. We also note that we find a sizeable contribution to the background from a real photon and QCD J/ψ or ϕ production in addition to the jet-conversion background.

The Higgs signal-strength measurements at the LHC are sensitive only to the product of production cross section times branching ratio to a specific final state. The dependence on both the production and the total width cancels to good approximation in the ratio between the rates of two processes with similar production but different final states. In particular, we choose to normalize the exclusive decay signal strength by $h \to ZZ^* \to 4\ell$ [22],

$$\mathcal{R}_{M\gamma,Z} \equiv \frac{\mu_{M\gamma}}{\mu_{ZZ^*}} \frac{\mathrm{BR}^{\mathrm{SM}}_{M\gamma}}{\mathrm{BR}^{\mathrm{SM}}_{ZZ^*\to 4\ell}} \simeq \frac{\Gamma_{M\gamma}}{\Gamma_{ZZ^*\to 4\ell}} = \begin{cases} 2.8\times 10^{-2}(\kappa_\gamma - 8.7\times 10^{-2}\kappa_c)^2/\kappa_V^2 & \text{for } M = J/\psi \\ 2.4\times 10^{-2}(\kappa_\gamma - 2.6\times 10^{-3}\kappa_s)^2/\kappa_V^2 & \text{for } M = \phi \end{cases}, \tag{10}$$

where $\mu_{M\gamma} = \sigma_h \mathrm{BR}_{M\gamma}/\sigma_h^{\mathrm{SM}}\mathrm{BR}^{\mathrm{SM}}_{M\gamma}$ and $\kappa_X \equiv y_X/y_X^{\mathrm{SM}}$ and $V = Z, W$. Here, we assumed a perfect cancellation of the production cross sections and that the Higgs decay width to a Z and two leptons (e.g. $h \to Z\gamma^* \to 4\ell$) is close to its SM value. The theoretical predictions for $h \to J/\psi\gamma$ and $h \to \phi\gamma$ are taken from Ref. [53] and [24], respectively, using the SM predictions $\mathrm{BR}^{\mathrm{SM}}_{J/\psi\gamma} = 2.9\times 10^{-6}$ [53] and $\mathrm{BR}^{\mathrm{SM}}_{\phi\gamma} = 3.0\times 10^{-6}$ [24]. Ref. [34] gives $\mathrm{BR}^{\mathrm{SM}}_{ZZ^*\to 4\ell} = 1.25\times 10^{-4}$.

We are now in a position to study the prospects of the exclusive modes in the next phases of the LHC, HL-LHC, and a future 100 TeV pp collider. We first define the inequality,

$$\mathcal{R}_{M\gamma,Z} < \frac{\mu^{95}_{M\gamma,E}}{\mu_{ZZ^*}} \frac{\mathrm{BR}^{\mathrm{SM}}_{M\gamma}}{\mathrm{BR}^{\mathrm{SM}}_{ZZ^*\to 4\ell}}, \tag{11}$$

where $\mu^{95}_{M\gamma,E}$ is a 95% CL upper bound for the $h \to M\gamma$ channel at the energy of $E = 8, 14, 100$ TeV. We neglect the uncertainty in μ_{ZZ^*}, because it is expected to be smaller than 10% [54,55]. The inequality in Eq. (11) together with Eq. (10) leads to the following bound for the charm and strange Yukawa couplings,

$$11\kappa_\gamma - 10\kappa_V \left(\frac{\mu^{95}_{J/\psi\gamma,E}}{\mu_{ZZ^*}}\right)^{1/2} < \kappa_c < 11\kappa_\gamma + 10\kappa_V \left(\frac{\mu^{95}_{J/\psi\gamma,E}}{\mu_{ZZ^*}}\right)^{1/2}, \tag{12}$$

$$3.8\kappa_\gamma - 3.8\kappa_V \left(\frac{\mu^{95}_{\phi\gamma,E}}{\mu_{ZZ^*}}\right)^{1/2} < \frac{\kappa_s}{100} < 3.8\kappa_\gamma + 3.8\kappa_V \left(\frac{\mu^{95}_{\phi\gamma,E}}{\mu_{ZZ^*}}\right)^{1/2}. \tag{13}$$

If the upper bounds on the $J/\psi\gamma$ and $\phi\gamma$ signal strengths are similar, the resulting bound on κ_s is weaker than the bound on κ_c by a factor of $\mathcal{O}[(m_c/m_s)\times(m_{J/\psi}/m_\phi)]$.

We note in passing that during the last preparation stage of this paper, Ref. [56] appeared. The authors of Ref. [56] presented an interpretation of the same ATLAS exclusive Higgs decay result, and obtained a weaker bound than the one found in Ref. [22]. The reason for this is threefold: (i) we normalized the signal strength of the exclusive channels by μ_{ZZ^*} to reduce the dependence on κ_γ (which is more sensitive to new-physics contributions) while [56] chose to normalise it by $\mu_{\gamma\gamma}$, which leads to weakening the bound by 10%. This happens because the observed central value of $\mu_{\gamma\gamma}$ [57] is smaller than that of the ATLAS μ_{ZZ^*} result. (ii) we did not include the order 10% theoretical uncertainty in the bound. (iii) most importantly, Ref. [56] has provided an improved and more precise calculation of the central value of the relevant matrix element that leads to a significant 40% reduction in the dependence of κ_c (and a slight increase of the theoretical uncertainties), which translates to a 40% increase in the bound.

Next, we move to provide a rough estimation of the future bound on the $pp \to h \to J/\psi\gamma$ rate given the current ATLAS upper bound [32]. We denote by S_E^{95} the 95% CL upper bound on the number of signal events and by B_E the expected number of background events at the center-of-mass energy E. Based on the available 8 TeV result (S_8^{95}), we estimate the future sensitivity by assuming that

$$\frac{S_E^{95}}{\sqrt{B_E}} \approx \frac{S_8^{95}}{\sqrt{B_8}}. \tag{14}$$

Using this, we find the following scaling,

$$\mu^{95}_{J/\psi\gamma,E} = \frac{S_E^{95}}{S_E^{\mathrm{SM}}} \approx \left(\frac{B_E}{B_8}\frac{S_8^{\mathrm{SM}}}{S_E^{\mathrm{SM}}}\right)^{1/2}\left(\frac{S_8^{\mathrm{SM}}}{S_E^{\mathrm{SM}}}\right)^{1/2}\frac{S_8^{95}}{S_8^{\mathrm{SM}}} = \frac{1}{R_E^{1/2}}\left(\frac{\sigma^{\mathrm{SM}}_{h,8}\mathcal{L}_8}{\sigma^{\mathrm{SM}}_{h,E}\mathcal{L}_E}\right)^{1/2}\mu^{95}_{J/\psi\gamma,8}, \tag{15}$$

where

$$R_E \equiv \frac{S_E^{\mathrm{SM}}/B_E}{S_8^{\mathrm{SM}}/B_8} \quad \text{and} \quad \mu^{95}_{J/\psi\gamma,8} = \frac{S_8^{95}}{S_8^{\mathrm{SM}}}. \tag{16}$$

S_E^{SM} is the number of signal events as expected in the SM, $\sigma^{\mathrm{SM}}_{h,E}$ is the SM Higgs-production cross section and $\mathcal{L}_E$ is the integrated luminosity. We have implicitly assumed above that the signal and background efficiencies are equal across the different runs. If future findings indicate that the efficiencies differ from each other, then the corresponding modification to Eq. (15) can be absorbed by an appropriate rescaling of R_E. The rate for Higgs production is characterized by a harder physical scale than the one of the corresponding QCD background. Consequently, colliders with larger center of mass are expected to have a larger signal to background ratio, i.e., $R_E \gtrsim 1$.

The expected upper bound on the signal strength in Eq. (15) can be easily interpreted as a bound on the Higgs couplings using Eq. (12). For pp colliders with a center-of-mass energy of 14 TeV and 100 TeV, assuming $\mu_{ZZ^*} = \kappa_\gamma = \kappa_V = 1$ and SM Higgs production, we find that the expected reach at 95% CL is

$$11 - 80\left(\frac{1}{R_{14}}\frac{2\times 300\ \mathrm{fb}^{-1}}{\mathcal{L}_{14}}\right)^{1/4} < \kappa_c < 11 + 80\left(\frac{1}{R_{14}}\frac{2\times 300\ \mathrm{fb}^{-1}}{\mathcal{L}_{14}}\right)^{1/4}, \tag{17}$$

$$11-45\left(\frac{1}{R_{14}}\frac{2\times 3000\ \text{fb}^{-1}}{\mathcal{L}_{14}}\right)^{1/4} < \kappa_c < 11+45\left(\frac{1}{R_{14}}\frac{2\times 3000\ \text{fb}^{-1}}{\mathcal{L}_{14}}\right)^{1/4}, \quad (18)$$

$$11-40\left(\frac{1}{R_{100}}\frac{2\times 300\ \text{fb}^{-1}}{\mathcal{L}_{100}}\right)^{1/4} < \kappa_c < 11+40\left(\frac{1}{R_{100}}\frac{2\times 300\ \text{fb}^{-1}}{\mathcal{L}_{100}}\right)^{1/4}, \quad (19)$$

$$11-22\left(\frac{1}{R_{100}}\frac{2\times 3000\ \text{fb}^{-1}}{\mathcal{L}_{100}}\right)^{1/4} < \kappa_c < 11+22\left(\frac{1}{R_{100}}\frac{2\times 3000\ \text{fb}^{-1}}{\mathcal{L}_{100}}\right)^{1/4}. \quad (20)$$

Here, we used $\sigma^{\text{SM}}_{h,(8,14,100)} = 22.3,\ 57.2,\ 897$ pb [34], $\mathcal{L}_8 = 19.2\ \text{fb}^{-1}$ and $\mu^{95}_{J/\psi,8} = 515$ [32]. These bounds may be compared to the current bound of $\kappa_c \lesssim 220$ [22]. We see that the projected bounds depend only weakly on the integrated luminosity and on R_E. The corresponding expected upper bound on the branching ratio is $\text{BR}_{J/\psi,14(100)} < 2.4(0.60)\times 10^{-4}$, where we assume SM production and $\mathcal{L}_{14(100)} = 300\ \text{fb}^{-1}$.

The different exclusive channels are expected to be subject to analogous backgrounds, namely QCD production and an associated fake jet or a real photon. The ATLAS result for $h \to J/\psi\gamma$ [32] can be, thus, used to estimate the future reach in the different channels. In particular, we focus on the case of $h \to \phi\gamma$ decay, but the generalization of our analysis to other final states, such as ρ or ω is straightforward. However, as our results are very pessimistic we do not expect good results for the other analyses.

To make the following discussion more transparent, we supplement our previously used symbols for signal and background (S, B) with a subscript $J/\psi\gamma$ ($S_{J/\psi,E}$ and $B_{J/\psi,E}$). Symbols regarding $h \to \phi\gamma$ will contain a $\phi\gamma$ subscript.

In order to estimate the upper bound on the $h \to \phi\gamma$ signal strength, we use an approximation,

$$\frac{S^{95}_{\phi\gamma,E}}{\sqrt{B_{\phi\gamma,E}}} \approx \frac{S^{95}_{J/\psi\gamma,E}}{\sqrt{B_{J/\psi\gamma,E}}}. \quad (21)$$

We then estimate $S_{\phi\gamma,E}$ and $B_{\phi\gamma,E}$ in the following way. The ratio between the number of signal events in each channel is given by

$$\frac{S_{\phi\gamma,E}}{S_{J/\psi\gamma,E}} = \frac{\sigma_{h,E}\text{BR}(h\to\phi\gamma)\mathcal{L}_E}{\sigma_{h,E}\text{BR}(h\to J/\psi\gamma)\mathcal{L}_E}\frac{\text{BR}(\phi\to K^+K^-)}{\text{BR}(J/\psi\to\mu^+\mu^-)}\frac{\epsilon_\phi}{\epsilon_{J/\psi}} \quad (22)$$

where $\epsilon_{J/\psi(\phi)}$ stands for the triggering and reconstruction efficiency (including the isolation and various kinematical cuts following Ref. [32]). The J/ψ is observed via its rather clean $J/\psi \to \mu^+\mu^-$ leptonic decay mode, while the ϕ is assumed to decay to $\phi \to K^+K^-$, which is a much more challenging final state for triggering, identification and background rejection. Nevertheless, we focus on this final state because it has a large branching ratio. In that sense, the bound below is rather conservative, given that we ignore these challenges when we rescale the ATLAS J/ψ result. The ratio of backgrounds for the two different exclusive final states can be written as

$$\frac{B_{\phi\gamma,E}}{B_{J/\psi,E}} = \frac{\sigma_E(pp\to\phi\text{“}\gamma\text{”})}{\sigma_E(pp\to J/\psi\text{“}\gamma\text{”})}\frac{\text{BR}(\phi\to K^+K^-)}{\text{BR}(J/\psi\to\mu^+\mu^-)}\frac{\epsilon_\phi}{\epsilon_{J/\psi}}, \quad (23)$$

where by “γ” we refer to a photon candidate, namely an object that has passed the (ATLAS) tight-photon selection cuts, i.e., it is either a genuine photon or a jet faking a photon. The corresponding rate reads

$$\sigma_E(pp\to\phi,J/\psi\text{“}\gamma\text{”}) = \sigma_E(pp\to\phi,J/\psi j)P(j\to\gamma) + \sigma_E(pp\to\phi,J/\psi\gamma),$$

where $P(j\to\gamma)\sim 2\times 10^{-4}$ stands for the rate that a jet is misidentified as a photon under the tight photon selection [58].

Combining Eq. (20) with Eqs. (21)–(22) leads to an extrapolation of the upper bound for the $h \to \phi\gamma$ signal strength,

$$\begin{aligned}\mu^{95}_{\phi\gamma,E} &= \frac{S^{95}_{\phi\gamma,E}}{S^{\text{SM}}_{\phi\gamma,E}} \approx \left(\frac{B_{\phi\gamma,E}}{B_{J\gamma,E}}\right)^{1/2}\frac{S^{\text{SM}}_{J\psi\gamma,E}}{S^{\text{SM}}_{\phi\gamma,E}}\frac{S^{95}_{J\psi\gamma,E}}{S^{\text{SM}}_{J\psi\gamma,E}} \\ &= \left(\frac{\sigma_E(pp\to\phi\text{“}\gamma\text{”})}{\sigma_E(pp\to J/\psi\text{“}\gamma\text{”})}\frac{\text{BR}(J/\psi\to\mu^+\mu^-)}{\text{BR}(\phi\to K^+K^-)}\frac{\epsilon_{J/\psi}}{\epsilon_\phi}\right)^{1/2} \\ &\quad\times\frac{\text{BR}^{\text{SM}}_{J/\psi\gamma}}{\text{BR}^{\text{SM}}_{\phi\gamma}}\mu^{95}_{J/\psi\gamma,E} \\ &= 0.34\left(\frac{\sigma_E(pp\to\phi\text{“}\gamma\text{”})}{\sigma_E(pp\to J/\psi\text{“}\gamma\text{”})}\frac{\epsilon_{J/\psi}}{\epsilon_\phi}\right)^{1/2}\mu^{95}_{J/\psi,E} \end{aligned} \quad (24)$$

where $\text{BR}(\phi\to K^+K^-) = 48.9\%$ and $\text{BR}(J/\psi\to\mu^+\mu^-) = 5.93\%$ [40]. As the J/ψ is reconstructed from a dimuon pair while the ϕ from a K^+K^- pair, we expect that $\epsilon_{J/\psi} > \epsilon_\phi$. Moreover, we expect that $\sigma_E(pp\to\phi j) > \sigma_E(pp\to J/\psi j)$ because J/ψ's are more rarely produced than ϕ in the QCD process. Therefore, we expect the upper bound for $h \to \phi\gamma$ to be weaker than that for $h \to J/\psi\gamma$, more precisely $\mu^{95}_\phi \gg 0.34\mu^{95}_{J/\psi}$.

The estimation of the ratio of $\sigma_E(pp\to\phi\text{“}\gamma\text{”})/\sigma_E(pp\to J/\psi\text{“}\gamma\text{”})$ was performed in two steps. Using PYTHIA8.2 [59,60] we generated two samples. The first of events with a photon and a jet and the second one with di-jets at $\sqrt{s} = 8$ TeV. Prior to showering and hadronization, we

required that the two objects have $p_T > 20$ GeV at the parton level. Following Ref. [32], we then selected events that contain $J/\psi(\phi)$ with $p_T > 36$ GeV, a photon or an anti-k_T jet (of cone 0.4) with $p_T > 36$ GeV, $|\eta| < 2.37$ and $\Delta\phi(J/\psi(\phi), \text{“}\gamma\text{”}) > 0.5$, where “$\gamma$” stands either for the second jet multiplied by $P(j \to \gamma) \sim 2 \times 10^{-4}$ [58] or the actual photon. Following the ATLAS analysis, we also required an isolation cut for the J/ψ (ϕ). We evaluated the sum of the energy of the extra hadrons (or photons from π^0 decay) that are within a cone of 0.2 away form the J/ψ (ϕ) and required it to be less than 10% of the energy of the J/ψ (ϕ). We found that the number of events that pass these cuts for the fake jet sample is very close to that of the real photon, and thus retained both samples. The resulting ratio from our simulation is

$$\left.\frac{\sigma_8(pp \to \phi\text{“}\gamma\text{”})}{\sigma_8(pp \to J/\psi\text{“}\gamma\text{”})}\right|_{\text{PYTHIA}} \sim 9. \tag{25}$$

We have verified that for the jj sample a similar ratio has been obtained for the $\sqrt{s} = 14$ TeV case. As a sanity check we compare our total simulated rate of $\sigma(pp \to J/\psi\text{“}\gamma\text{”})$ with the one reported by ATLAS and to agree within ~50%.

All the above information can be combined to constrain κ_s. Using Eqs. (13), (24), and (25) we find,

$$3.8 - 29\left(\frac{1}{R_{14}}\frac{2\times 300\ \text{fb}^{-1}}{\mathcal{L}_{14}}\right)^{1/4} < \frac{\kappa_s}{100} < 3.8 + 29\left(\frac{1}{R_{14}}\frac{2\times 300\ \text{fb}^{-1}}{\mathcal{L}_{14}}\right)^{1/4}, \tag{26}$$

$$3.8 - 16\left(\frac{1}{R_{14}}\frac{2\times 3000\ \text{fb}^{-1}}{\mathcal{L}_{14}}\right)^{1/4} < \frac{\kappa_s}{100} < 3.8 + 16\left(\frac{1}{R_{14}}\frac{2\times 3000\ \text{fb}^{-1}}{\mathcal{L}_{14}}\right)^{1/4}, \tag{27}$$

$$3.8 - 14\left(\frac{1}{R_{100}}\frac{2\times 300\ \text{fb}^{-1}}{\mathcal{L}_{100}}\right)^{1/4} < \frac{\kappa_s}{100} < 3.8 + 14\left(\frac{1}{R_{100}}\frac{2\times 300\ \text{fb}^{-1}}{\mathcal{L}_{100}}\right)^{1/4}, \tag{28}$$

$$3.8 - 8.2\left(\frac{1}{R_{100}}\frac{2\times 3000\ \text{fb}^{-1}}{\mathcal{L}_{100}}\right)^{1/4} < \frac{\kappa_s}{100} < 3.8 + 8.2\left(\frac{1}{R_{100}}\frac{2\times 3000\ \text{fb}^{-1}}{\mathcal{L}_{100}}\right)^{1/4}, \tag{29}$$

where we assumed $\mu_{ZZ^*} = \kappa_\gamma = \kappa_V = 1$. Since the only known possibility to probe the strange Yukawa in hadron machines is via $h \to \phi\gamma$, the resulting reach is still weak, $\mathcal{O}(\kappa_s) \lesssim 2000$ even at HL-LHC. However, this disappointing situation may be improved by the development of new methods that would take advantage of the relatively quiet QCD environment of a Higgs event. For instance, one direction is to consider jet-substructure techniques.

The first-generation Yukawas may be probed via the $h \to (\rho, \omega)\gamma$ decays [24]. However, since the ρ and ω mesons are lighter than ϕ, we expect a larger QCD background in hadron colliders, resulting in weaker sensitivity. Nevertheless, due to the large $\gamma - \rho$ mixing, the process $h \to \rho\gamma$ has a branching ratio, $\text{BR}^{\text{SM}}_{\rho\gamma} = 1.9 \times 10^{-5}$, larger than in the other modes, and may be probed in the clean environment of future e^+e^- colliders.

It is interesting to compare the projected sensitivity reach on κ_c between the inclusive rate with c-tagging and the exclusive decays to $h \to J/\psi\gamma$. From Tables II, III, IV and Eqs. (17), (18), (19), (20), we learn that the prospects of the inclusive analysis to probe the charm Yukawa are much better than of the exclusive analysis. For example, with 300 fb^{-1} the projected reach in the inclusive analysis is stronger than in the exclusive analysis by roughly a factor of 4, and with 3000 fb^{-1} in the high-luminosity stage it is stronger by a factor of 10. So given the current background understanding for $h \to J/\psi\gamma$, we expect that the inclusive c-tagging method will be more powerful in probing modifications of the charm Yukawa.

IV. CONCLUSIONS

In this work, we have presented the projections for probing light-quark Yukawas within the LHC, its high-luminosity stage (HL-LHC) and a future 100 TeV hadron collider. Using charm tagging we find that the HL-LHC can probe the charm Yukawa with a sensitivity of a few times the SM value. With an improved tagger at a 100 TeV machine a sensitivity close to the one necessary for probing the SM value appears feasible.

We have also provided a preliminary study of the sensitivity of the various exclusive decay modes. These channels are particularly important as they provide a unique opportunity to probe the Higgs couplings to the three lightest quarks. Our study shows that the reach of the exclusive modes, however, is rather limited, as follows. ATLAS recently provided the first measurement of the background relevant to charmonia and a photon final state; a final state that can potentially also probe the Higgs–charm coupling. ATLAS observes a large continuous background due to QCD production of charmonia and a jet converted into a photon. We also find, using a leading-order simulation, a sizeable contribution from charmonia plus a photon production. Given the small signal and the large background we find that the reach based on the $J/\psi\gamma$ mode is more than an order of magnitude weaker than that expected in the inclusive approach.

Focusing on analogous backgrounds, we further study the sensitivity reach of the search for the $\phi\gamma$ final state, which can probe the Higgs to strange coupling. In this case, the resulting sensitivity is poorer, allowing the HL-LHC to only prove a strange Yukawa of order 10^3 times the SM

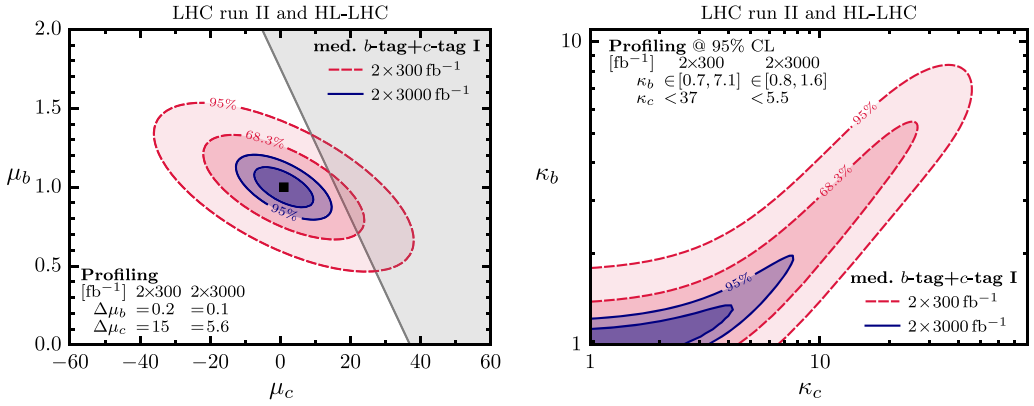

FIG. 6. 300 fb^{-1} and 3000 fb^{-1} prospects for the signal strengths and couplings of $h \to b\bar{b}$ and $h \to c\bar{c}$ at the LHC based on b- and c-tagging I for the correlated scenario. Future sensitivity of μ_b and μ_c is in the left panel. The grey shaded region is unphysical unless Higgs production is modified with respect to the SM case. Future sensitivity κ_b and κ_c is in the right panel. All other Higgs couplings are assumed to be like in the SM. The profiled likelihood ratio [41] is used for the respective reach of $\Delta\mu_b$, $\Delta\mu_c$, κ_b, and κ_c.

value or more. The current analysis strategy for the exclusive modes is subject to large backgrounds. These consist of pure QCD production (with a jet faking a photon) and QCD plus associated-photon production that limit the reach of the analysis. However, these backgrounds are not irreducible. The situation may be improved as follows. One can modify the search, limiting the kinematics such that the Higgs is boosted and captured by a fat jet. In this case, the energy deposition for the signal and background inside the fat jet should be very different leading to a better background rejection and an improved sensitivity to the Yukawa couplings.

ACKNOWLEDGMENTS

We thank David Kosower for useful discussions. We acknowledge help from the ATLAS collaboration for providing us with details of the analysis of Ref. [35]. The work of K. T. is supported in part by the Grant-in-Aid for JSPS Fellows, the work of G. P. is supported by ERC, IRG and ISF.

APPENDIX: SUPPLEMENT FOR INCLUSIVE ANALYSIS

1. Correlation between b- and c-tagging

When b- and c-taggers are simultaneously imposed, we study two scenarios with respect to correlations. The actual situation is expected to be something between the two scenarios. In the main text, we showed only the uncorrelated scenarios, and here we show in Fig. 6 the projection using b-tagging and c-tagging I in the correlated scenario. Comparing to Fig. 2 (left) and Fig. 3 (left), we find the results in the two scenarios to be very similar. This is for the following reason. The main difference of the two scenarios is whether the two b-tagged category (i) contains c-tagged events or not. However, this statistical difference does not change the overall number in category (i), and hence it does not affect the main significance of the category (i) that determines μ_b. In addition, the two c-tagged category (iii), which is the most important to measure μ_c is exactly the same in both scenarios. Therefore, the two scenarios give almost the same sensitivity.

For c-tagging II and III, we can only assume the uncorrelated scenario because the assumption for the correlated scenario leads to an inequality,

$$\epsilon_x^{(c\text{-tag})} \leq \epsilon_x^{(b\text{-tag})} \quad \text{for all } x = b, c, l. \tag{A1}$$

c-tagging II or III with medium b-tagging does not satisfy this inequality. However, the c-tagger is expected to be less correlated with b-tagging after its possible improvement, so it is reasonable to expect the uncorrelated scenario to be realistic.

We emphasize that the correlation of the two taggers are not an issue for the actual experiments because they have the information for the respective jet. We studied the two different scenarios simply because the information is not available for us to take into account the correlation.

2. Non-Standard Model Vh production

For large $\kappa_c \sim \mathcal{O}(10\text{–}100)$, new contributions to the Vh final states, shown in Fig. 7, become important and the

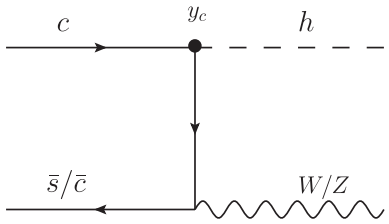

FIG. 7. Example diagram that modifies Vh production when the charm-quark Yukawa is enhanced.

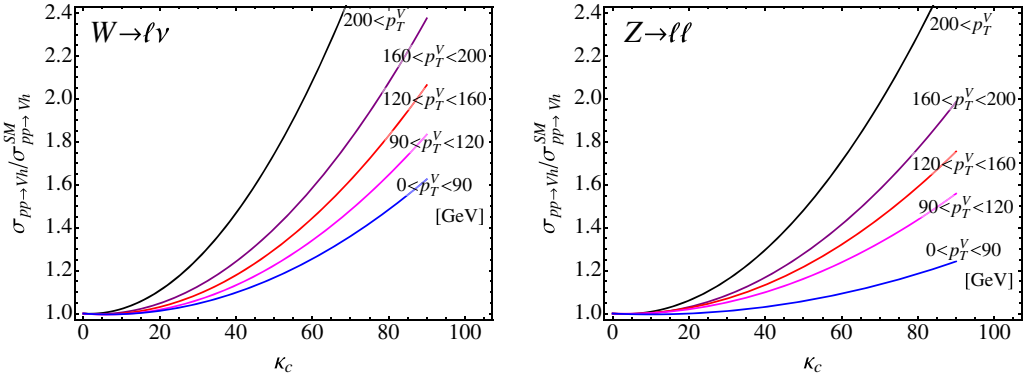

FIG. 8. Vh enhancement with κ_c from the new production mechanism, using the selection cuts ATLAS.

signal strength is modified not only by the branching ratio but also by the production cross section. The contributions to the Vh production cross section at $\sqrt{s} = 14$ TeV as a function of κ_c are presented in Fig. 8 and are roughly given by

$$\frac{\sigma_{pp\to Vh}}{\sigma^{\mathrm{SM}}_{pp\to Vh}} \simeq 1 + \left(\frac{\kappa_c}{60\text{–}180}\right)^2 \tag{A2}$$

for large κ_c. Here, the Higgs coupling to the W/Z is assumed to be SM like, i.e. $\kappa_V = 1$. We obtained these results using MadGraph 5.2 [43] at the parton level and leading order, applying the ATLAS selection cuts for the LHC 14 TeV run [35].

For 100 TeV pp collider, the effect of the new production mechanism is not important because the sensitivity in κ_c is $\mathcal{O}(1)$. Therefore, we neglect this effect.

[1] G. Aad *et al.* (ATLAS Collaboration), Phys. Lett. B **716**, 1 (2012).
[2] S. Chatrchyan *et al.* (CMS Collaboration), Phys. Lett. B **716**, 30 (2012).
[3] C. Delaunay, C. Grojean, and G. Perez, J. High Energy Phys. 09 (2013) 090.
[4] C. Delaunay, T. Flacke, J. Gonzalez-Fraile, S. J. Lee, G. Panico, and G. Perez, J. High Energy Phys. 02 (2014) 055.
[5] M. Blanke, G. F. Giudice, P. Paradisi, G. Perez, and J. Zupan, J. High Energy Phys. 06 (2013) 022.
[6] R. Mahbubani, M. Papucci, G. Perez, J. T. Ruderman, and A. Weiler, Phys. Rev. Lett. **110**, 151804 (2013).
[7] A. L. Kagan, G. Perez, T. Volansky, and J. Zupan, Phys. Rev. D **80**, 076002 (2009).
[8] A. Dery, A. Efrati, G. Hiller, Y. Hochberg, and Y. Nir, J. High Energy Phys. 08 (2013) 006.
[9] G. F. Giudice and O. Lebedev, Phys. Lett. B **665**, 79 (2008).
[10] L. Da Rold, C. Delaunay, C. Grojean, and G. Perez, J. High Energy Phys. 02 (2013) 149.
[11] A. Dery, A. Efrati, Y. Nir, Y. Soreq, and V. Susic, Phys. Rev. D **90**, 115022 (2014).
[12] F. Bishara, J. Brod, P. Uttayarat, and J. Zupan, arXiv: 1504.04022.
[13] C. Delaunay, T. Golling, G. Perez, and Y. Soreq, Phys. Rev. D **89**, 033014 (2014).
[14] The ATLAS Collaboration, Reports No. ATLAS-CONF-2014-011, No. ATLAS-COM-CONF-2014-004, 2014.
[15] G. Aad *et al.* (ATLAS Collaboration), J. High Energy Phys. 01 (2015) 069.
[16] G. Aad *et al.* (ATLAS Collaboration), J. High Energy Phys. 04 (2015) 117.
[17] V. Khachatryan *et al.* (CMS Collaboration), J. High Energy Phys. 09 (2014) 087.
[18] S. Chatrchyan *et al.* (CMS Collaboration), Phys. Rev. D **89**, 012003 (2014).
[19] S. Chatrchyan *et al.* (CMS Collaboration), J. High Energy Phys. 05 (2014) 104.
[20] G. Aad *et al.* (ATLAS Collaboration), Phys. Lett. B **738**, 68 (2014).
[21] V. Khachatryan *et al.* (CMS Collaboration), Phys. Lett. B **744**, 184 (2015).
[22] G. Perez, Y. Soreq, E. Stamou, and K. Tobioka, Phys. Rev. D **92**, 033016 (2015).
[23] G. T. Bodwin, F. Petriello, S. Stoynev, and M. Velasco, Phys. Rev. D **88**, 053003 (2013).
[24] A. L. Kagan, G. Perez, F. Petriello, Y. Soreq, S. Stoynev, and J. Zupan, Phys. Rev. Lett. **114**, 101802 (2015).
[25] G. Isidori, A. V. Manohar, and M. Trott, Phys. Lett. B **728**, 131 (2014).
[26] M. Mangano and T. Melia, Eur. Phys. J. C **75**, 258 (2015).
[27] T.-C. Huang and F. Petriello, Phys. Rev. D **92**, 014007 (2015).
[28] Y. Grossmann, M. König, and M. Neubert, J. High Energy Phys. 04 (2015) 101.

[29] G. Aad *et al.* (ATLAS Collaboration), Phys. Rev. D **90**, 052008 (2014).
[30] G. Aad *et al.* (ATLAS Collaboration), Phys. Rev. Lett. **114**, 161801 (2015).
[31] CERN Technical Report No. ATL-PHYS-PUB-2015-001, 2015.
[32] G. Aad *et al.* (ATLAS Collaboration), Phys. Rev. Lett. **114**, 121801 (2015).
[33] V. Khachatryan *et al.* (CMS), arXiv:1507.03031.
[34] S. Heinemeyer *et al.* (LHC Higgs Cross Section Working Group), arXiv:1307.1347.
[35] The ATLAS Collaboration, Report No. ATL-PHYS-PUB-2014-011, 2014.
[36] M. Capeans, G. Darbo, K. Einsweiller, M. Elsing, T. Flick, M. Garcia-Sciveres, C. Gemme, H. Pernegger, O. Rohne, and R. Vuillermet, CERN Reports No. CERN-LHCC-2010-013, No. ATLAS-TDR-19, 2010.
[37] ATLAS Collaboration, Report No. ATL-PHYS-PUB-2013-010.
[38] S. Chatrchyan *et al.* (CMS Collaboration), JINST **8**, P04013 (2013).
[39] K. Chetyrkin, J. H. Kuhn, and M. Steinhauser, Comput. Phys. Commun. **133**, 43 (2000).
[40] K. Olive *et al.* (Particle Data Group), Chin. Phys. C **38**, 090001 (2014).
[41] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, Eur. Phys. J. C **71**, 1554 (2011).
[42] CERN Report No. ATL-PHYS-PUB-2014-016, 2014.
[43] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, J. High Energy Phys. 06 (2011) 128.
[44] A. Altheimer, A. Arce, L. Asquith, J. Backus Mayes, E. Bergeaas Kuutmann *et al.*, Eur. Phys. J. C **74**, 2792 (2014).
[45] M. Backovic, J. Juknevich, and G. Perez, J. High Energy Phys. 07 (2013) 114.
[46] L. G. Almeida, S. J. Lee, G. Perez, G. Sterman, and I. Sung, Phys. Rev. D **82**, 054034 (2010).
[47] L. G. Almeida, O. Erdogan, J. Juknevich, S. J. Lee, G. Perez, and G. Sterman *et al.*, Phys. Rev. D **85**, 114046 (2012).
[48] G. Aad *et al.* (ATLAS), J. High Energy Phys. 01 (2013) 116.
[49] J. M. Campbell and R. Ellis, Nucl. Phys. B, Proc. Suppl. **205-206**, 10 (2010).
[50] H. Baer, T. Barklow, K. Fujii, Y. Gao, A. Hoang *et al.*, arXiv:1306.6352.
[51] M. Bicer *et al.* (TLEP Design Study Working Group), J. High Energy Phys. 01 (2014) 164.
[52] http://indico.cern.ch/event/337673/session/6/contribution/20/material/slides/0.pdf.
[53] G. T. Bodwin, H. S. Chung, J.-H. Ee, J. Lee, and F. Petriello, Phys. Rev. D **90**, 113010 (2014).
[54] CMS Collaboration, arXiv:1307.7135.
[55] ATLAS Collaboration, Reports No. ATL-PHYS-PUB-2012-004, No. ATL-COM-PHYS-2012-1455, 2012.
[56] M. König and M. Neubert, J. High Energy Phys. 08 (2015) 012.
[57] G. Aad *et al.* (ATLAS Collaboration), Phys. Rev. D **90**, 112015 (2014).
[58] ATLAS Collaboration, Repors No. ATL-PHYS-PUB-2011-007, No. ATL-COM-PHYS-2010-1051, 2011.
[59] T. Sjöstrand, S. Mrenna, and P. Z. Skands, J. High Energy Phys. 05 (2006) 026.
[60] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, Comput. Phys. Commun. **191**, 159 (2015).