

Probing charm Yukawa coupling through ch associated production at the hadron colliders

Nuoyu Dong^{*,†}, Hongsheng Hou[†], Zhuoni Qian[‡], Bowen Wang[§], Pei Xu^{||} and Qingjun Xu[¶]
School of Physics, Hangzhou Normal University, Hangzhou, Zhejiang 311121, China



(Received 19 December 2024; accepted 26 February 2025; published 13 March 2025)

At present, the study of the charm-quark Yukawa coupling at the Large Hadron Collider mainly focuses on the Higgs decay processes. Such signal suffers from overwhelming QCD background and derives its sensitivity primarily from the Vh associated production channel. In addition, sensitivity to a possible CP phase in charm-quark Yukawa at the hadron collider is not discussed. We investigate the charm-Higgs associated production signal that contains a potentially detectable interference term between cch Yukawa coupling mediated diagrams and ggh coupling mediated diagram. Such an interference term is sensitive to the relative CP phase between contributing diagrams. High dimensional kinematic information is exploited by machine learning techniques to separate the different contribution, and sensitivity on the coupling is derived. Assuming a real κ_c modification framework, the 1σ bound of $-5.6 < \kappa_c < 5.6$ (HL-LHC) and $-1.51 < \kappa_c < 1.62$ (FCC) are achieved. When allowing for CP phase in the charm Yukawa, a combined 1σ bound of $0.32 < |\kappa_c| < 1.69$, $-77^\circ < \alpha < 77^\circ$ (HL-LHC) and $0.70 < |\kappa_c| < 1.29$, $-55^\circ < \alpha < 55^\circ$ (FCC) can be achieved on the magnitude and CP phase of the coupling, respectively.

DOI: [10.1103/PhysRevD.111.053003](https://doi.org/10.1103/PhysRevD.111.053003)

I. INTRODUCTION

Since the discovery of the Higgs boson by ATLAS [1] and CMS [2], determining the properties of the Higgs boson has been a crucial task at the Large Hadron Collider (LHC). Current constraints from Run 2 LHC data show that coupling strengths between all third-generation fermions and the Higgs boson deviate by no more than 20% from the Standard Model (SM) predictions [3–8]. At the High Luminosity (HL-)LHC, it is expected that the precision of the third-generation Yukawa couplings will be measured to within 5% [9–11]. It is worth noting that in general a complex CP phase can be introduced in Yukawa couplings in the presence of new physics contribution. Currently, LHC data put bounds on the top and tau Yukawa coupling CP phase of $(11^{+52}_{-73})^\circ$, $(9^{+16}_{-16})^\circ$ [12], respectively. At the HL-LHC, constraints are expected to reach $|\alpha_t| < 36^\circ$ [13], $|\alpha_b| < 23.4^\circ$ [10] for top and bottom, and $|\alpha_\tau| < 8^\circ$ [11,14]

for tau at 2σ confidence level. Measurements of Yukawa couplings in the second generation remain challenging, among which the muon Yukawa observation is the most promising. The current observed upper limit on $h \rightarrow \mu\mu$ decay rate is 2.2 times the SM prediction [15], signifying an $\mathcal{O}(1)$ muon Yukawa coupling measurement in the near future. The allowed range of the charm Yukawa coupling is measured to be 8.5 times that of the SM at 95% confidence level from ATLAS [16] or $1.1 < \kappa_c < 5.5$ from CMS [17].

With the accumulation of data at the HL-LHC, extensive exploration on the charm Yukawa coupling are expected. Following the κ framework, constraints on possible rescaling of SM couplings are described through $\kappa_i = g_i/g_i^{\text{SM}}$. We compile here a representative list of proposals and prospected bounds on charm Yukawa coupling in Table I under the HL-LHC dataset. Among them, the Higgs rare decay channel $H \rightarrow J/\psi\gamma$ offers a small but clean signal in detecting the charm Yukawa coupling at the LHC [18–20]. A current constraint of $\kappa_c/\kappa_\gamma \in (-133, 175)$ [21,22] is already performed using this signal with 139 fb^{-1} data at $\sqrt{s} = 13 \text{ TeV}$. A future reach of $0.32 < \kappa_c/\kappa_\gamma < 1.53$ assuming a 10% precision on the decay branching is targeted for example at FCC [20]. Similarly, the $h \rightarrow c\bar{c} + J/\psi$ decay provides another exclusive hadron decay channel to further constrain the coupling [23]. Precision measurement of the differential distribution of the Higgs from inclusive Higgs production are claimed to constrain the light quark Yukawa coupling as well [24,25]. With current Run 2 data, $\kappa_c \in (-4.46, 4.81)$ is achieved [26,27],

*Contact author: 2022111030003@stu.hznu.edu.cn

†Contact author: hshou@hznu.edu.cn

‡Contact author: zhuoniqian@hznu.edu.cn

§Contact author: bowenw@hznu.edu.cn

||Contact author: 2022111030014@stu.hznu.edu.cn

¶Contact author: xuqingjun@hznu.edu.cn

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

TABLE I. The table summarizes representative proposals and prospective constraints on the charm Yukawa coupling at the HL-LHC and FCC-hh, while the last three rows include constraints from related studies on future hadron-electron collider, electron-positron collider, and muon collider.

Channel	Machine	Bound
Higgs differential	HL-LHC	$-0.6 < \kappa_c < 3.0(2\sigma)$ [24]
$h \rightarrow c\bar{c}\gamma$	HL-LHC	$ \kappa_c < 6.3(2\sigma)$ [28]
$hh \rightarrow c\bar{c}\gamma\gamma$	HL-LHC	$-4.8 < \kappa_c < 4.6(2\sigma)$ [35]
$pp \rightarrow qq\bar{h}\gamma(h \rightarrow c\bar{c})$	HL-LHC	$ \kappa_c < 13(2\sigma)$ [36]
$pp \rightarrow VVc\bar{j}$	HL-LHC	$-2.4 < \kappa_c < 1.77(2\sigma)$ [29]
$h \rightarrow J/\psi + c\bar{c}$	HL-LHC	$ \kappa_c < 2.4(2\sigma)$ [23]
$pp \rightarrow ch$	HL-LHC	$ \kappa_c < 2.6(2\sigma)$ [40]
$pp \rightarrow Zh(h \rightarrow c\bar{c})$	HL-LHC	$ \kappa_c < 2.5(2\sigma)$ [34]
$h \rightarrow J/\psi + \gamma$	FCC-hh	$0.32 < \kappa_c/\kappa_\gamma < 1.53(2\sigma)$ [20]
$e^-p \rightarrow \nu_e h j(h \rightarrow c\bar{c})$	LHeC	$ \kappa_c < 1.18(2\sigma)$ [37]
$e^-e^+ \rightarrow Zh(h \rightarrow c\bar{c})$	CEPC	$0.98 < \kappa_c < 1.02(2\sigma)$ [38]
$\mu^+\mu^- \rightarrow \mu^+\mu^-(\nu_\mu\bar{\nu}_\mu)h(h \rightarrow c\bar{c})$	Muon collider	$0.94 < \kappa_c < 1.06(1\sigma)$ [39]

and an order one reach is expected at the HL-LHC. Other proposals include the $h \rightarrow c\bar{c}\gamma$ decay where the electro-weak loop contribution is non-negligible [28] and associated production of charm quark with diboson processes $pp \rightarrow VVcX$ [29]. All these could offer contributions to probing the charm Yukawa coupling, awaiting real simulations to be performed and realistic bounds to be achieved. Note that the $h \rightarrow J/\psi + \gamma$ and the Higgs differential signal give asymmetric bounds on a real κ_c modification, which come from sizable interference term that could be extrapolated to meaningful bounds on possible CP phase of the coupling at HL-LHC and FCC. The Vh ($h \rightarrow c\bar{c}$) channel allows for a direct probe of the Higgs decay to a pair of charm jets [17,30–33]. A prospects of $|\kappa_c| < 2.5$ is projected from current Run 2 data and offers a realistic estimate for HL-LHC reach [34]. Other processes involving the Higgs decay to a pair of charm jet also contribute, such as the Higgs pair production $hh \rightarrow c\bar{c}\gamma\gamma$ [35], and the vector boson fusion $pp \rightarrow qq\bar{h}\gamma(h \rightarrow c\bar{c})$ [36]. All the production modes collectively improve bounds on the decay branching and therefore improve bounds on magnitude of charm Yukawa coupling. Much more precise determination of the branching ratio are expected to be obtained at future facilities such as the Large Hadron electron Collider (LHeC) [37], future lepton collider such as the Circular electron positron collider (CEPC) [38], or a high energy muon collider [39]. The prospected bounds are all summarized as in Table I.

The ch associated production process probes the coupling from the production side and tags the charm jet in the final state exclusively. An early analysis from Ref. [40] estimates a sensitivity to the charm Yukawa coupling from the production rate as a function of the coupling. It however lacks a study in Higgs decay nor simulation with realistic effects. It also neglects contribution from the interference term which could become detectable with future data. To achieve an

realistic sensitivity and understanding of the channel motivates the work of this study. Building upon the foundation of the literature [40], we further conduct event simulation and analysis, provide a more realistic conclusions. Moreover, we consider a possible CP -odd component of the coupling and its probe through relevant collider signals.

The remaining content of the article proceeds as follows. In Sec. II, we introduce the chosen signal and all relevant backgrounds, provide the details for data simulation, and define a set of observable. In Sec. III, we present event distributions of signal and background. We then analyze and categorize the simulated data with Boosted Decision Tree (BDT) and interpret the optimization with shapley values. In Sec. IV, we translate the categorization (confusion matrix) to constraints on the charm Yukawa coupling and possible CP phase at the HL-LHC and FCC-hh. We combine complementary collider constraints as well as bounds electron dipole moment (EDM) measurement. Finally, in Sec. V, we conclude. A derivation and the numerical dependence of the modified ggh and $\gamma\gamma h$ effective coupling as functions of the modified charm Yukawa coupling is given in the Appendix.

II. SIGNAL AND BACKGROUND

Treating ggh as an effective vertex $G_{\mu\nu}G^{\mu\nu}\Phi$, the leading order (LO) contribution to the signal comes mainly from two types of diagrams. The first type involves the charm Yukawa coupling, depicted in Figs. 1(a) and 1(b), whose amplitude is proportional to y_c . The second type involves the ggh effective coupling, shown in Fig. 1(c). With a dominant top loop contribution, such diagram involves minimal dependence on light quark Yukawa, which is summarized in the Appendix and can be considered almost backgroundlike. Moreover, these two types of Feynman diagrams exhibit interference effects. Within the total cross section, we designate the $|M1 + M2|^2$ contribution from

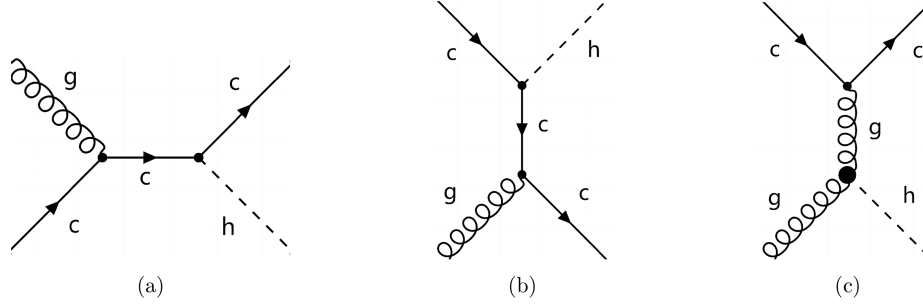


FIG. 1. The leading-order contribution Feynman diagrams for $pp \rightarrow hc$. Diagrams (a) and (b) depict typical diagrams involving the cch Yukawa coupling, while diagram (c) illustrates contribution mediated by the ggh effective coupling. The corresponding $\bar{c}$ diagrams are implied and not explicitly shown.

the first two diagrams as cch , which is proportional to y_c^2 . The $|M3|^2$ term is designated as ggh , which is mostly independent of y_c . The interference term, $2\text{Re}[(M1 + M2)^*M3]$ is designated as int and is proportional to y_c . In the SM, the predicted contribution from ggh diagram is significantly larger, by about an order of magnitude, compared to cch and the int term. This implies that while the signal ch has a sizable total cross section, its sensitivity to y_c is suppressed. Additionally, if the coupling contains a CP phase angle, it shows in the interference part and affects the cross section as well.

As a first attempt for a realistic simulation and analysis of the ch associated production signal, we chose the diphoton decay channel of the Higgs. The observed final states are $c\gamma\gamma$, where c is a c -tagged jet that originates from a c or $\bar{c}$ quark.¹ With this final state, the dominant irreducible background contribution is from the QCD-QED processes. Additionally, there are fake background such as $b\gamma\gamma$, $j\gamma\gamma$ ($j = u, d, s, g$) where the jet is mistagged as a c jet. Following the c -tagging strategy ($c \rightarrow 41\%$, $b \rightarrow 50\%$, $j \rightarrow 3.3\%$) from Refs. [41,42], the cross sections of $b\gamma\gamma$ and $j\gamma\gamma$ contribute less than 25% compared to that from $c\gamma\gamma$. Therefore, we do not include them in the full simulation of background events. When we define our signal to be inclusive with additional jets in the final states, contribution from the process $pp \rightarrow Wh \rightarrow cj\gamma\gamma$ become non-negligible. The main distinction is that this electroweak background contains two relatively hard jets with their invariant mass m_{cj} peaked around the W boson mass, whereas for the signal a second jet is mainly from higher order radiation which is soft. In the detailed simulation below, we include both the dominant irreducible QCD-QED background, and this large electroweak background Wh .

A. Simulation and kinematic cuts

The leading-order events for both signal and background are generated using MG5@MCNLO v3.5.3 [43], and the

¹The alternative $\bar{c}$ quark involving processes are always implied and included, which is not to be explicitly repeated in the following.

parton-level events are showered and hadronized using Pythia8 [44]. Detector simulation is performed using Delphes [45]. We use NNPDF31_lo_as_0118 [46] for simulations at both the HL-LHC with a center-of-mass energy of 14 TeV and the FCC-hh with a center-of-mass energy of 100 TeV. For the input parameters, the mass of the Higgs boson are set at $M_h = 125$ GeV, and the pole mass of the charm quark as $m_c = 1.3$ GeV. The charm Yukawa coupling at the Higgs mass scale is set according to the running mass $\bar{m}_c(M_h) = 0.81$ GeV. The renormalization and factorization scales are set at M_h as in Ref. [47].

As performed in the default Delphes setting with HL-LHC and FCC-hh Delphes parameter card, Jets are reconstructed using the anti-kt algorithm with $\Delta R = 0.4$. Then the simulated jet flavor is assigned by identifying with the heaviest flavor parton in the vicinity $\Delta R < 0.5$ of the center of the jet. Afterwards we estimate the realistic tagging efficiency with an overall rescaling following the tagging strategy discussed above. Note that this universally performed jet flavor definition can become problematic when going beyond the current LO + PS simulation because of infrared singularities induced by the emission of flavored quark-anti-quark pairs of negligible energy. Recent discussion and efforts on extensions of the algorithm with IRC safe jet flavor association can be found in Refs. [48–52]. In the final event selection, at least one charm-tagged jet is required to satisfy $|\eta_c| < 2.5$, $|p_{T,c}| > 20$ GeV. We also require two photons in the final states according to detector configuration defined with corresponding HL-LHC and FCC Delphes card. Additionally, to improve the efficiency of event generation, we require the following: $115 \text{ GeV} < m_{\gamma\gamma} < 135 \text{ GeV}$ (HL-LHC), $118 \text{ GeV} < m_{\gamma\gamma} < 132 \text{ GeV}$ (FCC-hh). We leave additional signal and background classification to BDT analysis with input observables to be defined in the next section. We summarize the cross section and corresponding event numbers after these basic selection cuts in Table II. Notably, the cross section contribution from interference of the cch and ggh diagrams is negative.

We would also like to add a note on correction from next leading order (NLO) contributions. The NLO introduces a

TABLE II. The total cross sections and corresponding number of events for the two types of background $c\gamma\gamma$, Wh and the three types of signal ggh , cch , int contributions. Full detector simulation, flavor tagging strategy, and basic selection cuts are applied at the HL-LHC and the FCC-hh, respectively.

Channel	14 TeV	6 ab ⁻¹ #	100 TeV	30 ab ⁻¹ #
	σ (fb)		σ (fb)	
$c\gamma\gamma$	73.2	439,387	897.59	269,279,58
Wh	0.052	314	0.31	9,258
ggh	0.046	276	0.94	28,218
cch	0.0033	20	0.067	2,023
int	-0.00033	-2	-0.0074	-223

new set of diagrams involving gg, cq, cc , and $c\bar{c}$ initial states, leading to a significant enhancement to the LO cross section. So far, the NLO simulation is available for the cch mediated diagrams at MG5@MCNLO. The ggh NLO simulation can be obtained using available calculation of the $h + j$ at NLO process from MCFM [53,54]. For the interference term, the calculation is done in Ref. [47], with the code not publicly available. In our simulation, we thus stay with the LO calculation at parton level using MG5@MCNLO and include only the QCD real radiation shape effects with parton shower. While we differ such involved study including NLO correction to future work, we would expect a better sensitivity on the charm Yukawa coupling with improved statistics after including the full NLO correction. The methodology and overall conclusion we obtain here with LO simulation should however remain the same.

B. Observables

To perform interpretable and full analysis on the simulated events, we need to construct a complete set of physical observables. First considering the degrees of freedom involving a three particle final state, we have six observables assuming momentum conservation. E_{cm} denotes the invariant mass of the $c\gamma\gamma$ final state system. η_c and ϕ_c are the rapidity and the azimuthal angle of the outgoing charm jet. $\cos(\theta_{\gamma 1})$ is the polar angle of the outgoing photon with the leading transverse momentum. The angle between the $cg \rightarrow ch$ interaction plane and the Higgs decay plane in the center of mass frame is denoted as $\Delta\phi$. $m_{\gamma\gamma}$ is the invariant mass of the two leading p_T photons. Additionally, the momentum along the beam direction of the $c\gamma\gamma$ system p_{z-all} encodes the PDF asymmetry of the incoming partons. The transverse momentum p_{T-all} and azimuthal angle ϕ_{all} of the $c\gamma\gamma$ system, as well as the scalar sum of transverse momenta H_T of all observed final states, are defined to account for showering effects. To effectively reduce the Wh background contribution, we also include an observable m_{cj} , the invariant mass of the charm jet and the hardest p_T jet identified without a flavor tag. In all, we choose an overcomplete set of observables for a final state of $c\gamma\gamma + X$, inclusive with additional jets. By calculating

the correlation matrix among the observations, we find that most correlations are negligible, with only E_{cm} , H_T , p_{T-all} , and m_{cj} showing high levels of correlation at the level of about 30%. This is reasonable to expect, as they all relate to the overall energy scale of the collision event, yet provide marginally additional information for distinguishing between signal and background processes.

III. $pp \rightarrow ch$ AT FUTURE COLLIDERS

As mentioned earlier, we categorize contribution to the total $c\gamma\gamma$ cross section into three types of ch signals and two types of non-Higgs backgrounds, an overall five categories. The physics goal is to evaluate the sensitivity to both the magnitude and a possible CP phase of the charm Yukawa, we aim to distinguish the cch and int contribution from all. Thus we must address the challenge of optimizing a multi-class discrimination based on multiobservable inputs. We use the BDT algorithm implemented in XGBoost [55] in our analysis. We include the overcomplete set of observables as inputs to the BDT training, to ensure all physics information known at the simulation level are exploited. In recent years, interpretable machine learning has gained significant popularity in high-energy physics and offered additional insights and understanding in collider signal analysis. Therefore, we use shapley values [56] in the analysis to dissect the optimization results quantitatively. The efficiency of this approach are demonstrated in Ref. [10]. After training the BDT model, we calculate the shapley values based on the BDT network, and then interpret the averaged absolute shapley values for each observable ($|\bar{S}_v|$) as their importance in making the optimized categorization. The larger the $|\bar{S}_v|$ value of a certain observable, the greater its contribution to distinguish different signal and background channel contributions. We show the main analysis results at the two collider settings in the following.

A. HL-LHC

The importance ranking of the observables is shown in the first panel of Fig. 2. In order to understand the importance ranking obtained from the BDT analysis, we also include the differential cross section distributions of the top most important observables in the rest of Fig. 2. In the importance ranking plot, we show the mean of absolute shapley value $|\bar{S}_v|$ representing different channels with different colors. The length quantify the contributing importance of that observable from distinguishing that specific channel.

Taking $m_{\gamma\gamma}$ as an example, since the diphoton from the three signal channels as well as the Wh background are all from the Higgs decay, their invariant mass distribution has the same shape peaking at the Higgs mass, while the background $c\gamma\gamma$ exhibits a relatively flat distribution. This observable naturally plays the most important role in distinguishing these two types of contributions. In the $|\bar{S}_v|$

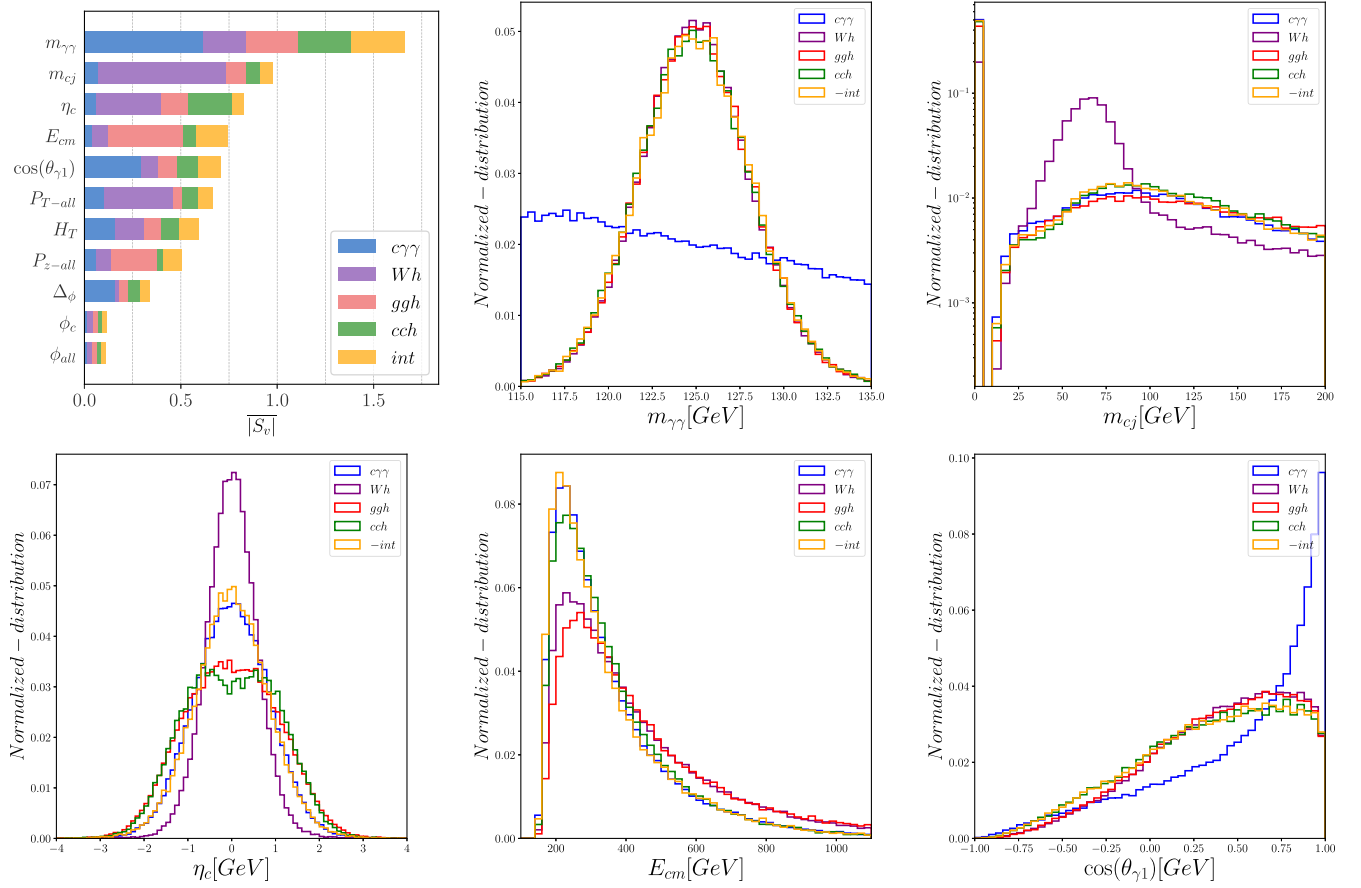


FIG. 2. Based on simulated events at the HL-LHC and the BDT categorization analysis, shapley importance ranking and the differential distributions of the top most important kinematic variables for each channel are shown.

shapley importance ranking, the score of $m_{\gamma\gamma}$ from $c\gamma\gamma$ is the highest, while the scores for the other three channels are similar. The m_{c_j} distribution from Wh channel has an expected wide peak below the W boson mass and distinguishes the channel from the others effectively. As expected, m_{c_j} importance ranking has dominant contribution from the Wh channel. For the observable η_c , we observe that it has distinctive shape for the cch and Wh channels from the others. The cch contribution exhibits a relative plateau and small dent at small absolute values, while the others all peak at the center. Additionally, the sharpness of the peak differentiates Wh from the rest of the channels. It thus offers high distinction power for cch and Wh in the importance ranking. Similar features can be observed and understood in the importance ranking for the other observables as well.

After understanding the importance of observables in differentiating different contributing channels, we proceed to use the trained BDT model to predict and classify the simulated data. Since this analysis involves five channels, we present the predictions of the BDT using a 5×5 confusion matrix. The element at position (i, j) in the matrix represents the number of occurrences of the “real

(simulated)” i th event type being predicted within j th event category. The confusion matrix for HL-LHC is presented in Table III. We choose a commonly used significance formula ($\sigma = N_S / \sqrt{N_S + N_B}$) to represent the corresponding “signal” strength within each predicted category. In the confusion matrix, we define it as

TABLE III. Trained BDT classification (confusion matrix) of the five channel contributions at HL-LHC with an integrated luminosity of 6 ab^{-1} (ATLAS + CMS). The right-most column provides the total number of events expected from each true channel in the SM. The last row shows the expected significance for the corresponding channel.

Predicted no. of events at HL-LHC						
Channel	$c\gamma\gamma$	wh	ggh	cch	int	Total
$c\gamma\gamma$	389,531	22,373	10,194	9582	7707	439,387
wh	91	193	13	7	10	314
ggh	131	31	77	25	12	276
cch	12	2	2	3	1	20
int	-2	0	0	0	0	-2
σ_j	623.93	1.28	0.76	0.031	0	

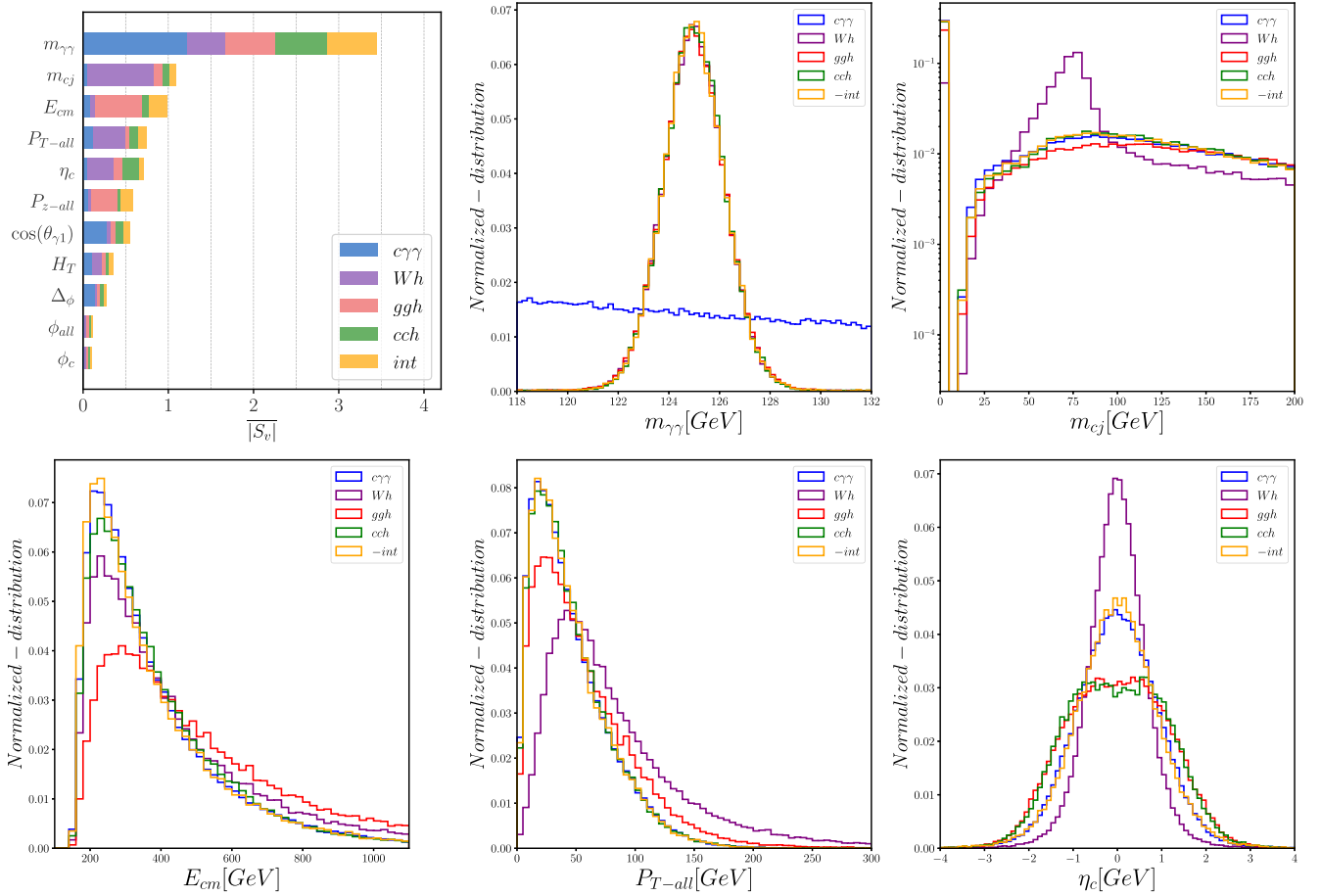


FIG. 3. Based on simulated events at the FCC-hh and the BDT categorization analysis, shapley importance ranking and the differential distributions of the top most important kinematic variables for each channel are shown.

$$\sigma_i = \frac{|N_{ii}|}{\sqrt{\sum_j N_{ij}}}. \quad (3.1)$$

At HL-LHC, the predicted number of events for the *cch* and *int* channel is small given the limited cross section. The significance obtained for the *cch* channel is 0.03σ . For the *int* channel we lose sensitivity for the phase angle. Based on this, we can still provide a constraint on a real modification of charm Yukawa κ_c , which is calculated in the next section.

B. FCC-hh

As in the case for the HL-LHC, we conduct the same event simulations and analysis scheme for the future collider scenario FCC-hh, considering 100 TeV pp collision with 30 ab^{-1} integrated luminosity. At the FCC-hh, the *cch* and *int* channel have many enhanced statistics, not only due to the increase in accumulated luminosity but also because of the increase in the cross section. From Table II, it can be observed that the cross sections of *cγγ*, *cch* have increased by approximately 18, *ggh* by approximately 20,

and *int* by approximately 11 times compared to the HL-LHC. After training the BDT model, we generated the $|\tilde{S}_v|$ importance ranking and event distribution plots for the five most important observables, which are shown in Fig. 3. The basic features remain similar to the HL-LHC where the top two most important observable remain $m_{\gamma\gamma}$ and m_{cj} in helping to distinguish the dominant background process *cγγ* and the electroweak background *Wh*. Notably at the FCC-hh, better prospected detector resolution reconstructs a narrower Higgs mass peak from the diphoton invariant mass, making $m_{\gamma\gamma}$ much more effective in distinguishing between signal and background and much larger importance ranking score. The resulting confusion matrix from BDT optimized classification and the observation significance achieved for each channel are summarized in Table IV. Compared to the results in HL-LHC, the number of events in all categories increased significantly, especially with the events in the *int* category beginning to be discernible. The significance of the *cch* channel reaches 0.5σ , and the *int* channel reaches 0.06σ . Although their significance levels do not reach 1σ , we derive constraints on the modified space of charm Yukawa coupling.

TABLE IV. Trained BDT classification (confusion matrix) of the five channel contributions at FCC-hh with a luminosity of 30 ab^{-1} . The right-most column provides the total number of events expected from each true channel in the SM.

Channel	Predicted no. of events at FCC-hh					total
	$c\gamma\gamma$	wh	ggh	cch	int	
$c\gamma\gamma$	22,748,500	1,056,348	781,197	1,005,789	1,336,122	26,927,958
wh	1282	6061	526	482	871	9258
ggh	5994	2949	11 478	4338	3459	28 218
cch	529	221	311	516	446	2,023
int	-64	-30	-23	-42	-64	-223
σ_j	4768	5.87	12.88	0.51	0.055	

IV. CONSTRAINTS ON THE YUKAWA COUPLING

Given the analysis above, we utilize the event counts of the two signal categories from the confusion matrix obtained to constrain the Charm-Yukawa coupling. To investigate a possible CP -violating term in the Yukawa sector, we use κ_c to describe the CP -even part and $\tilde{\kappa}_c$ the CP -odd part. Alternatively, we can use $|\kappa_c|$ and the phase angle α to describe the complex Yukawa coupling. Note that the sensitivity we get on the CP phase depends on sensitivity achievable to the interference term between cch and ggh with the latter assumed to be SM.

In addition to our probe of CP phase through the interference term, direct probe of the CP structure in the fermion-Higgs coupling by constructing CP -sensitive observable is done in Refs. [57–68]. Such observable construction are so far limited to probe the top and tau Yukawa couplings where the polarization information could be retrieved through their decay kinematics. With the accumulation of LHC data and the expectation of increased event rate, the CP -structure information for bottom and charm Yukawa couplings may be probed with the help of jet substructure observable. This however is beyond the scope of current study and we leave for future exploration.

A. Constraints on a real charm Yukawa

First we assume that charm Yukawa coupling remains real and the SM Lagrangian only changes in the charm Yukawa sector with a κ_c rescaling:

$$\mathcal{L} \supset -\kappa_c \frac{m_c}{v} \bar{c}ch. \quad (4.1)$$

Here, κ_c is a real number representing the deviation of the modified charm Yukawa from the SM value, and $v = 246 \text{ GeV}$ is the SM vacuum expectation value. The Lagrangian corresponds to the SM when $\kappa_c = 1$. Under our assumption, the correction for cch is proportional to κ_c^2 , while the correction for int is proportional to κ_c . The correction for ggh has a very minor dependence on κ_c as

well, whose numerical form is derived and included in the Appendix.² The respective χ -square deviation from the SM by probing the three different signal channel can thus be calculated from the confusion matrix as

$$\chi_j^2 = \frac{1}{\sum_i N_{ij}} \left(N_{4j} \kappa_c^2 + N_{5j} \kappa_c (1.01 - 0.01 \kappa_c) - N_{4j} - N_{5j} \right)^2. \quad (4.2)$$

Here N_{ij} are the confusion matrix elements from the SM prediction. When j takes the value of 4 and 5, it represents the number of events in the cch and int columns of the confusion matrix, respectively. We plot the significance curves as a function of κ_c in Fig. 4. Notably, the dominant signal contribution is from κ_c^2 , hence the mostly symmetric shape of the significance. Given limited sensitivity to the interference part, the probe gives bound region of $-5.6 < \kappa_c < 5.6$ (HL-LHC) and $-1.51 < \kappa_c < 1.62$ (FCC-hh, slightly asymmetric) with 1σ significance.

B. Constraints on a CP -complex charm Yukawa coupling

When including possible CP phase in the charm Yukawa coupling, we consider a modified SM Lagrangian as follows:

$$\mathcal{L} \supset -\frac{m_c}{v} |\kappa_c| \bar{c}(\cos \alpha + i\gamma_5 \sin \alpha)ch. \quad (4.3)$$

When $|\kappa_c| = 1$ and $\alpha = 0$, it returns to the SM. Here $|\kappa_c|$ denotes the rescaling of the overall magnitude of a CP -complex charm Yukawa coupling, and α the CP phase. Alternatively, we can also parametrize in terms of the CP real and imaginary part of the coupling with $\kappa_c = |\kappa_c| \cos \alpha$ and $\tilde{\kappa}_c = |\kappa_c| \sin \alpha$. The phase angle α does not affect the cch channel contribution that is always proportional to y_c^2 , but it introduces a factor of $\cos \alpha$ in the y_c dependent int

²For the Higgs decay, we assume a SM decay branching ratio to photons, where a correction from κ_c affects the deviation at most at subpercentage level, see the Appendix.

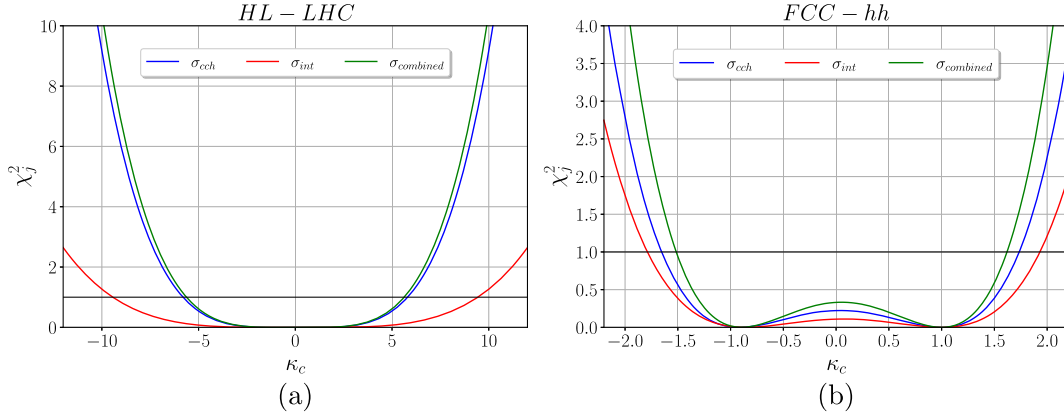


FIG. 4. The expected significance for a deviation from the SM as a function of κ_c by combining significance from the two signal channels as defined in Eq. (4.2). Panel (a) shows the expected significance at the HL-LHC (ATLAS + CMS, 6 ab^{-1}), and panel (b) represents the significance at the FCC-hh (30 ab^{-1}).

channel contribution. Such dependence arises from the relative phase between the modified cch diagram and the ggh diagram, with the latter mostly remain SM. Therefore, the chi-square deviation as a function of the charm Yukawa modification factors are now defined as follows:

$$\chi_j^2 = \frac{1}{\sum_i N_{ij}} (N_{4j} |\kappa_c|^2 + (1.01 |\kappa_c| \cos \alpha - 0.01 |\kappa_c|^2 - 0.001 |\kappa_c|^2 \sin^2 \alpha) N_{5j} - N_{4j} - N_{5j})^2. \quad (4.4)$$

Next we attempt to compare constraints on a CP -complex charm Yukawa coupling from this ch study with other collider signals. As discussed in the introduction, the $Zh, Zh(h \rightarrow c\bar{c})$ channel offers a competitive probe of the charm Yukawa coupling with realistic data and collider simulation. At the HL-LHC assuming 3 ab^{-1} integrated luminosity and 14 TeV center of mass energy, ATLAS prospects an upper limit of $\mu_{Zh(h \rightarrow c\bar{c})} < 6.3$ [34] with 2σ . At FCC-hh assuming 6 ab^{-1} integrated luminosity and 100 TeV center of mass energy, constraints of $|\kappa_c| < 2.1$ at 2σ are expected from $Vh(c\bar{c})$ analysis in Ref. [31]. To compare with our study, we rescale the two bounds above to $|\kappa_c| < 1.69$ for a 6 ab^{-1} luminosity at the HL-LHC and $0.47 < |\kappa_c| < 1.33$ for a 30 ab^{-1} luminosity at the FCC-hh at 1σ . Additionally, indirect constraints from the precisely measured Higgs production and decay rate such as $gg \rightarrow h$ and $h \rightarrow \gamma\gamma$ have sensitivity in the charm Yukawa coupling through the quark loop contribution. The charm quark loop contributes at LO mostly through interfering with the dominant top loop (and W loop for $\gamma\gamma h$) diagram. The relative CP phase between diagrams thus affects the size of the interference directly. We include in the Appendix the dependence of $\kappa_{g,\gamma}$ and $\tilde{\kappa}_{g,\gamma}$ on κ_c and $\tilde{\kappa}_c$ after integrating the quark loop contribution as function of the modified Yukawa coupling. The projected constraints on κ_g and κ_γ from the corresponding production and decay rate measurement can be directly mapped to bounds on the charm

Yukawa coupling modification. The projected 1σ sensitivity at the HL-LHC (ATLAS + CMS combined, 6 ab^{-1}) for κ_g (0.8%), κ_γ (1.3%) are derived from the projected bounds on the production cross section and decay branching ratios σ_{ggh} (1.6%), $\text{BR}(h \rightarrow \gamma\gamma)$ (2.6%) from Figs. 28–29 of the HL-LHC projections study [11]. The CP -phase contribution to the total cross section comes in through the dependence on $|\kappa_{g,\gamma}|^2 + |\tilde{\kappa}_{g,\gamma}|^2$. These bounds on $\kappa_{g,\gamma}$ are dominated by inclusive Higgs measurements assumed to have no additional Yukawa coupling or New Physics (NP) dependence other than the κ_c considered. At FCC-hh, the expected sensitivity from a global κ fit to κ_g, κ_γ are about 0.49%, 0.29%, respectively, taken from Table 3 of Ref. [69] including experimental and theory uncertainties.

Taking all these collider signal into consideration, constraints on the complex charm Yukawa are shown in Fig. 5, with HL-LHC on the left and the FCC-hh in the right panel. Reading from the plot, the most stringent constraint on the magnitude of the charm Yukawa comes from the measurement of the $Zh, h \rightarrow c\bar{c}$ (yellow region) process as expected. Being proportional to $|\kappa_c|^2 + |\tilde{\kappa}_c|^2$, it does not allow for the determination of the CP phase α_c . It nevertheless probes the charm Yukawa contribution directly by identifying the charm jet in the final state as in our study. Bounds from our ch analysis (green) are comparatively weak at the HL-LHC, but improved significantly at the FCC-hh since the large center of mass energy at the FCC significantly increased the cross section of the ch process. In the plot, we also show the combined constraint achieved from our ch study and $Zh, h \rightarrow c\bar{c}$ as “charm combined.” Improvement to the bound by including the ch study is about percent level, and slightly shift the bound from the center. The indirect constraint from $gg \rightarrow h$ (purple region) production rate measurement offer sensitive probe to a large CP space. There the charm loop contributes through interfering with the dominant top loop at percent level, which is reachable for the Higgs

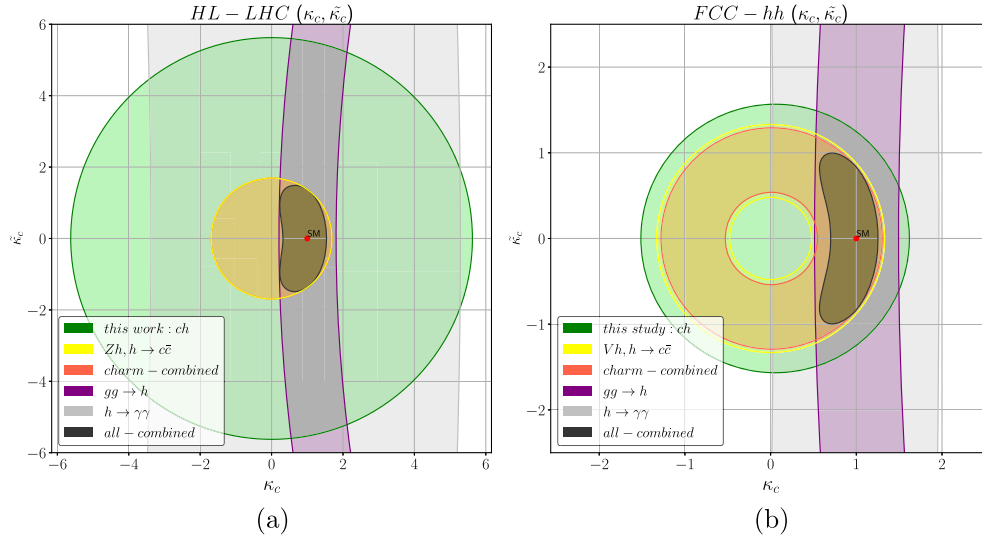


FIG. 5. (a) and (b) show the 1σ sensitivity contours in the $(\kappa_c, \tilde{\kappa}_c)$ space at the HL-LHC (ATLAS + CMS, 6 ab^{-1}) and the FCC-hh (30 ab^{-1}), respectively. All other parameters except for the charm Yukawa coupling are fixed to their SM values. The green region represents constraint from this ch study, the yellow region from $Zh, h \rightarrow c\bar{c}$, while the red region “charm combined” represents the bound from the two signals combined. Additionally, we show the purple and light gray contour regions that are indirect bounds from charm loop contribution to $gg \rightarrow h$ and $h \rightarrow \gamma\gamma$ inclusive Higgs production and decay measurements, respectively.

production through gluon fusion. Similar sensitivity are to be reached for $h \rightarrow \gamma\gamma$ (light gray region) decay rate measurement, whereas the charm loop generates a even minor contributions. The indirect constraints from $gg \rightarrow h$ offers a complementary probe of the CP structure of the charm Yukawa coupling.

Reading off from Fig. 5, the study from the ch channel alone give symmetric bounds of $|\kappa_c| < 5.6$ at the HL-LHC and a slightly shifted bound of $-1.50 < \kappa_c < 1.61$, $-1.57 < \tilde{\kappa}_c < 1.57$ at the FCC. While the bound from $Vh, h \rightarrow c\bar{c}$ signal offers the best sensitivity, it is improved from $0.47 < |\kappa_c| < 1.33$ to $0.52 < |\kappa_c| < 1.28$ at the FCC, by about 12%. After combining with the Higgs inclusive production and decay rate measurement $gg \rightarrow h$ and $h \rightarrow \gamma\gamma$, a marginal constraint of $0.32 < |\kappa_c| < 1.69$, $-77^\circ < \alpha < 77^\circ$ at the HL-LHC and $0.70 < |\kappa_c| < 1.29$, $-55^\circ < \alpha < 55^\circ$ at the FCC-hh can be achieved.

C. Bounds from EDM

EDM measurements at low energies provide constraints on CP -odd components of the quark Yukawa couplings. Recent update from ACME [70] puts an upper bound on electron EDM (eEDM) at $|d_e| < 1.1 \times 10^{-29} e \text{ cm}$ (at 90% CL). Currently, the constraints on the charm Yukawa from eEDM are strongest compared to those from the neutron or other hadronic EDMs [71]. The latest constraints on the charm Yukawa from eEDM are $\tilde{\kappa}_c < 0.18$ (at 1σ) from Eq. (4.42) in Ref. [72]. The complementary constraints imposed by eEDM and the results of this work are shown in Fig. 6. As expected, the EDM results sensitively constrain the CP -odd component of the charm Yukawa,

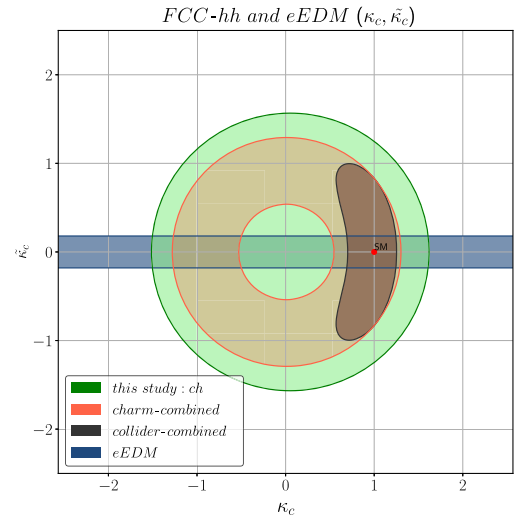


FIG. 6. The 1σ sensitivity contours are plotted in the $(\kappa_c, \tilde{\kappa}_c)$ space while fixing all other parameters to their SM values for FCC-hh (30 ab^{-1}). The green and red contours correspond to interpretations of ch and charm-combined constraints, respectively. The dark purple contour represents the collider-combined constraints. The blue contour represents the constraints from the eEDM measurement.

while the collider results probe the overall magnitude more precisely. It should be stressed though that, despite the seemingly stringent constraints from EDMs upon the CP -odd component, multiple flavor Yukawa couplings could contribute through the loop. Therefore, the probe of CP -odd term at collider processes where the charm flavor is exclusively identified such as in this study plays an

irreplaceable role. In addition, eEDM bounds rely on the assumption of a nonvanishing electron Yukawa coupling with the Higgs, which is hardly possible to test experimentally and may very well not hold under NP models. In this light, hadronic EDM measurements offers weaker yet complementary tests on the CP -odd component of the quark Yukawa couplings.

V. CONCLUSION

In this work, we study the $ch(h \rightarrow \gamma\gamma)$ signal and its probe to the charm Yukawa coupling. We perform full detector simulation and use BDT method with the calculation of Shapley values from an interpretable machine learning framework. We identify and classify the different types of signal and background contribution, along with an understanding of the kinematics spanned by the main collider observables. With the optimized classification results, we achieve an improved bounds on the charm Yukawa coupling as well as the CP phase. By combining with collider signals as well as EDM measurements, we achieve a prospects for probing the CP -complex charm Yukawa coupling space at HL-LHC and FCC-hh. In summary, this work provides the following main conclusions:

- (1) At the HL-LHC, limited by statistics, we reach constraints of $|\kappa_c| < 5.6$ at the 1σ level, and negligible constraints on the phase. After combining with the results from $Zh, h \rightarrow c\bar{c}$ and Higgs production rate ($gg \rightarrow h, h \rightarrow \gamma\gamma$), we obtain $0.32 < |\kappa_c| < 1.69$, $-77^\circ < \alpha < 77^\circ$ at 1σ level.
- (2) At the FCC-hh, the significantly increased event numbers of the ch effectively improved the precision of constraints. At the 1σ level, an asymmetric constraint of $-1.51 < \kappa_c < 1.62$ is achieved for a real κ_c . This improves bound from $Vh, h \rightarrow c\bar{c}$ by about 12%. Combing further with Higgs production and decay rate measurement we arrive a bound of $0.70 < |\kappa_c| < 1.29$, $-55^\circ < \alpha < 55^\circ$.
- (3) EDM measurements offer complementary probe to the charm Yukawa coupling as shown in Fig. 6. Whereas the eEDM currently gives the most stringent bound on the CP -odd component of charm Yukawa coupling $\tilde{\kappa}_c$, the collider signals give the best constraint on the overall magnitude. Notably, collider signals with exclusive charm jet identification along with the Higgs resonance signal such as the $h \rightarrow c\bar{c}$ and our study of the ch provide more direct or indisputable probe into the charm Yukawa coupling space.
- (4) Machine learning methods, when combined with interpretable frameworks, provide valuable insights. We demonstrate in Sec. III the shape distributions of different observables along with their corresponding absolute mean of Shapley values, a well defined

importance measure in the BDT classification procedure. The contribution of qualitative shape difference and different correlation patterns among the complete set of observables can be better understood with the importance measure from the BDT machine learning process.

ACKNOWLEDGMENTS

We would like to thank Professor Kirill Melnikov giving an explanation on available NLO calculations for the $cg \rightarrow ch$ process. Z. Q. is supported by the National Natural Science Foundation of China (Grant No. 1240050404) and the HZNU start-up fund.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available upon publication because it is not technically feasible and/or the cost of preparing, depositing, and hosting the data would be prohibitive within the terms of this research project. The data are available from the authors upon reasonable request.

APPENDIX: EFFECTIVE κ_g AND κ_γ

As described in the text, it is useful to list the numerical dependence of complex ggh and $\gamma\gamma h$ effective coupling on a complex charm Yukawa and its modification. Following the procedure in Ref. [73], the couplings at leading order contribution are defined as

$$\mathcal{L}_{\text{eff}} \supset c_g \frac{\alpha_s}{12\pi} \frac{h}{v} G_{\mu\nu}^a G^{\mu\nu,a} + \tilde{c}_g \frac{\alpha_s}{8\pi} \frac{h}{v} G_{\mu\nu}^a \tilde{G}^{\mu\nu,a}, \quad (\text{A1})$$

$$\mathcal{L}_{\text{eff}} \supset c_\gamma \frac{\alpha}{\pi} \frac{h}{v} F_{\mu\nu} F^{\mu\nu} + \tilde{c}_\gamma \frac{3\alpha}{2\pi} \frac{h}{v} F_{\mu\nu} \tilde{F}^{\mu\nu}. \quad (\text{A2})$$

The CP -even and CP -odd couplings are defined, respectively, as

$$c_g = \sum_{f=t,b,c} \kappa_f A(\tau_f), \quad \tilde{c}_g = \sum_{f=t,b,c} \tilde{\kappa}_f B(\tau_f), \quad (\text{A3})$$

with $\tau_f = 4m_f^2/m_h^2$ and

$$A(\tau) = \frac{3\tau}{2} \left[1 + (1-\tau) \arctan^2 \frac{1}{\sqrt{\tau-1}} \right],$$

$$B(\tau) = \tau \arctan^2 \frac{1}{\sqrt{\tau-1}}. \quad (\text{A4})$$

Hence, κ_g and κ_γ , the rescaling of the effective ggh and $\gamma\gamma h$ couplings, can be written as functions of κ_c and $\tilde{\kappa}_c$ as

$$\kappa_g = \frac{c_g}{c_g^{\text{SM}}} = \frac{A(\tau_t) + A(\tau_b) + \kappa_c A(\tau_c)}{A(\tau_t) + A(\tau_b) + A(\tau_c)}, \quad (\text{A5})$$

$$\tilde{\kappa}_g = \frac{3}{2} \frac{\tilde{c}_g}{c_g^{\text{SM}}} = \frac{3}{2} \frac{\tilde{\kappa}_c B(\tau_c)}{A(\tau_t) + A(\tau_b) + A(\tau_c)}, \quad (\text{A6})$$

such that the modifier for inclusive gluon fusion rate is $\mu_g = |\kappa_g|^2 + |\tilde{\kappa}_g|^2$:

$$\kappa_g \sim (-0.01 + 0.01i)\kappa_c + (1.01 - 0.01i), \quad (\text{A7})$$

$$\tilde{\kappa}_g \sim (-0.011 + 0.01i)\tilde{\kappa}_c. \quad (\text{A8})$$

Similarly, the $h\gamma\gamma$ effective couplings can be expressed in terms of κ_c . In addition to the linear dependence on the Yukawa coupling, there is also a large constant term in the CP -even coupling from the leading W and subleading top contributions at one loop in the SM. We omit here the well-known analytical forms and show the numerical dependence that matters for this analysis:

$$\kappa_\gamma = (0.997 + 0.0024i) + (0.003 - 0.0024i)\kappa_c, \quad (\text{A9})$$

$$\tilde{\kappa}_\gamma = (0.003 - 0.0024i)\tilde{\kappa}_c. \quad (\text{A10})$$

-
- [1] 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**, 1 (2012).
 - [2] CMS Collaboration, Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC, *Phys. Lett. B* **716**, 30 (2012).
 - [3] ATLAS Collaboration, Observation of $H \rightarrow b\bar{b}$ decays and VH production with the ATLAS detector, *Phys. Lett. B* **786**, 59 (2018).
 - [4] CMS Collaboration, Observation of Higgs boson decay to bottom quarks, *Phys. Rev. Lett.* **121**, 121801 (2018).
 - [5] ATLAS Collaboration, Cross-section measurements of the Higgs boson decaying into a pair of τ -leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, *Phys. Rev. D* **99**, 072001 (2019).
 - [6] CMS Collaboration, Search for the associated production of the Higgs boson and a vector boson in proton-proton collisions at $\sqrt{s} = 13$ TeV via Higgs boson decays to τ leptons, *J. High Energy Phys.* **06** (2019) 093.
 - [7] CMS Collaboration, Search for the Higgs boson decaying to two muons in proton-proton collisions at $\sqrt{s} = 13$ TeV, *Phys. Rev. Lett.* **122**, 021801 (2019).
 - [8] ATLAS Collaboration, Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector, *Phys. Lett. B* **784**, 173 (2018).
 - [9] R. Mammen Abraham, D. Gonçalves, T. Han, S. C. I. Leung, and H. Qin, Directly probing the Higgs-top coupling at high scales, *Phys. Lett. B* **825**, 136839 (2022).
 - [10] C. Grojean, A. Paul, and Z. Qian, Resurrecting $b\bar{b}h$ with kinematic shapes, *J. High Energy Phys.* **04** (2021) 139.
 - [11] M. Cepeda *et al.*, Report from Working Group 2: Higgs physics at the HL-LHC and HE-LHC, *CERN Yellow Rep. Monogr.* **7**, 221 (2019).
 - [12] ATLAS Collaboration, Characterising the Higgs boson with ATLAS data from run 2 of the LHC, *Phys. Rep.* **11**, 001 (2024).
 - [13] D. Gonçalves, J. H. Kim, K. Kong, and Y. Wu, Direct Higgs-top CP -phase measurement with $t\bar{t}h$ at the 14 TeV LHC and 100 TeV FCC, *J. High Energy Phys.* **01** (2022) 158.
 - [14] R. Harnik, A. Martin, T. Okui, R. Primulando, and F. Yu, Measuring CP violation in $h \rightarrow \tau^+\tau^-$ at colliders, *Phys. Rev. D* **88**, 076009 (2013).
 - [15] ATLAS Collaboration, A search for the dimuon decay of the Standard Model Higgs boson with the ATLAS detector, *Phys. Lett. B* **812**, 135980 (2021).
 - [16] ATLAS Collaboration, Direct constraint on the Higgs-charm coupling from a search for Higgs boson decays into charm quarks with the ATLAS detector, *Eur. Phys. J. C* **82**, 717 (2022).
 - [17] CMS Collaboration, Search for Higgs boson decay to a charm quark-antiquark pair in proton-proton collisions at $s = 13$ TeV, *Phys. Rev. Lett.* **131**, 061801 (2023).
 - [18] G. T. Bodwin, F. Petriello, S. Stoynev, and M. Velasco, Higgs boson decays to quarkonia and the $H\bar{c}c$ coupling, *Phys. Rev. D* **88**, 053003 (2013).
 - [19] A. L. Kagan, G. Perez, F. Petriello, Y. Soreq, S. Stoynev, and J. Zupan, Exclusive window onto Higgs Yukawa couplings, *Phys. Rev. Lett.* **114**, 101802 (2015).
 - [20] M. König and M. Neubert, Exclusive radiative Higgs decays as probes of light-quark Yukawa couplings, *J. High Energy Phys.* **08** (2015) 012.
 - [21] CMS Collaboration, Search for Higgs boson decays into Z and J/ψ and for Higgs and Z boson decays into J/ψ or Y pairs in pp collisions at $s = 13$ TeV, *Phys. Lett. B* **842**, 137534 (2023).
 - [22] ATLAS Collaboration, Searches for exclusive Higgs and Z boson decays into a vector quarkonium state and a photon using 139 fb $^{-1}$ of ATLAS $\sqrt{s} = 13$ TeV proton-proton collision data, *Eur. Phys. J. C* **83**, 781 (2023).
 - [23] T. Han, A. K. Leibovich, Y. Ma, and X.-Z. Tan, Higgs decay to charmonia and the charm-quark Yukawa coupling, *Proc. Sci. ICHEP2022* (2022) 517 [arXiv:2211.10727].
 - [24] F. Bishara, U. Haisch, P. F. Monni, and E. Re, Constraining light-quark Yukawa couplings from Higgs distributions, *Phys. Rev. Lett.* **118**, 121801 (2017).
 - [25] Y. Soreq, H. X. Zhu, and J. Zupan, Light quark Yukawa couplings from Higgs kinematics, *J. High Energy Phys.* **12** (2016) 045.

- [26] ATLAS Collaboration, Measurement of the total and differential Higgs boson production cross-sections at $\sqrt{s} = 13$ TeV with the ATLAS detector by combining the $H \rightarrow ZZ^* \rightarrow 4\ell$ and $H \rightarrow \gamma\gamma$ decay channels, *J. High Energy Phys.* **05** (2023) 028.
- [27] CMS Collaboration, Combination and interpretation of fiducial differential Higgs boson production cross sections at $\sqrt{s} = 13$ TeV, CERN Report No. CMS-PAS-HIG-23-013, 2024, <http://cds.cern.ch/record/2905139>.
- [28] T. Han and X. Wang, Radiative decays of the Higgs boson to a pair of fermions, *J. High Energy Phys.* **10** (2017) 036.
- [29] N. Vignaroli, Off-shell probes of the Higgs Yukawa couplings: Light quarks and charm, *Symmetry* **14**, 1183 (2022).
- [30] G. Perez, Y. Soreq, E. Stamou, and K. Tobioka, Constraining the charm Yukawa and Higgs-quark coupling universality, *Phys. Rev. D* **92**, 033016 (2015).
- [31] G. Perez, Y. Soreq, E. Stamou, and K. Tobioka, Prospects for measuring the Higgs boson coupling to light quarks, *Phys. Rev. D* **93**, 013001 (2016).
- [32] C. Delaunay, T. Golling, G. Perez, and Y. Soreq, Enhanced Higgs boson coupling to charm pairs, *Phys. Rev. D* **89**, 033014 (2014).
- [33] L. M. Carpenter, T. Han, K. Hendricks, Z. Qian, and N. Zhou, Higgs boson decay to light jets at the LHC, *Phys. Rev. D* **95**, 053003 (2017).
- [34] ATLAS Collaboration, Prospects for $H \rightarrow c\bar{c}$ using charm tagging with the ATLAS experiment at the HL-LHC, Technical Report ATL-PHYS-PUB-2018-016, CERN, Geneva, 2018.
- [35] L. Alasfar, R. Corral Lopez, and R. Gröber, Probing Higgs couplings to light quarks via Higgs pair production, *J. High Energy Phys.* **11** (2019) 088.
- [36] B. Carlson, T. Han, and S. C. I. Leung, Higgs boson to charm quark decay in vector boson fusion plus a photon, *Phys. Rev. D* **104**, 073006 (2021).
- [37] R. Li, B.-W. Wang, K. Wang, X. Zhang, and Z. Zhou, Probing the charm Yukawa coupling at future e^-p and e^+e^- colliders, *Phys. Rev. D* **100**, 053008 (2019).
- [38] F. An *et al.*, Precision Higgs physics at the CEPC, *Chin. Phys. C* **43**, 043002 (2019).
- [39] M. Forslund and P. Meade, High precision Higgs from high energy muon colliders, *J. High Energy Phys.* **08** (2022) 185.
- [40] I. Brivio, F. Goertz, and G. Isidori, Probing the charm quark Yukawa coupling in Higgs + charm production, *Phys. Rev. Lett.* **115**, 211801 (2015).
- [41] T. Han, B. Nachman, and X. Wang, Charm-quark Yukawa coupling in $h \rightarrow c\bar{c}\gamma$ at LHC, *Phys. Lett. B* **793**, 90 (2019).
- [42] ATLAS Collaboration, Search for the decay of the Higgs boson to charm quarks with the ATLAS experiment, *Phys. Rev. Lett.* **120**, 211802 (2018).
- [43] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations, *J. High Energy Phys.* **07** (2014) 079.
- [44] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, An introduction to Pythia8.2, *Comput. Phys. Commun.* **191** (2015) 159.
- [45] DELPHES 3 Collaboration, Delphes3, A modular framework for fast simulation of a generic collider experiment, *J. High Energy Phys.* **02** (2014) 057.
- [46] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht, M. Schönherr, and G. Watt, LHAPDF6: parton density access in the LHC precision era, *Eur. Phys. J. C* **75**, 132 (2015).
- [47] W. Bizon, K. Melnikov, and J. Quarroz, On the interference of ggH and $c\bar{c}H$ Higgs production mechanisms and the determination of charm Yukawa coupling at the LHC, *J. High Energy Phys.* **06** (2021) 107.
- [48] A. Banfi, G. P. Salam, and G. Zanderighi, Infrared safe definition of jet flavor, *Eur. Phys. J. C* **47**, 113 (2006).
- [49] S. Caletti, A. J. Larkoski, S. Marzani, and D. Reichelt, Practical jet flavour through NNLO, *Eur. Phys. J. C* **82**, 632 (2022).
- [50] M. Czakon, A. Mitov, and R. Poncelet, Infrared-safe flavoured anti- k_T jets, *J. High Energy Phys.* **04** (2023) 138.
- [51] R. Gauld, A. Huss, and G. Stagnitto, Flavor identification of reconstructed hadronic jets, *Phys. Rev. Lett.* **130**, 161901 (2023).
- [52] F. Caola, R. Grabarczyk, M. L. Hutt, G. P. Salam, L. Scyboz, and J. Thaler, Flavored jets with exact anti- k_T kinematics and tests of infrared and collinear safety, *Phys. Rev. D* **108**, 094010 (2023).
- [53] V. Ravindran, J. Smith, and W. L. Van Neerven, Next-to-leading order QCD corrections to differential distributions of Higgs boson production in hadron hadron collisions, *Nucl. Phys. B* **634**, 247 (2002).
- [54] C. R. Schmidt, $H \rightarrow ggg(gqq)$ at two loops in the large M_t limit, *Phys. Lett. B* **413**, 391 (1997).
- [55] T. Chen and C. Guestrin, XGBoost: A scalable tree boosting system, *KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2016), arXiv:1603.02754.
- [56] L. S. Shapley, *Notes on the N -person game II: The value of an N -person game*, (RAND Corporation, Santa Monica, CA, 1951), 10.7249/RM0670.
- [57] M. E. Cassidy, Z. Dong, K. Kong, I. M. Lewis, Y. Zhang, and Y.-J. Zheng, Probing the CP structure of the top quark Yukawa at the future muon collider, *J. High Energy Phys.* **05** (2024) 176.
- [58] R. K. Barman *et al.*, Directly Probing the CP -structure of the Higgs-top Yukawa at HL-LHC and future colliders, in Snowmass 2021 (2022), 3, arXiv:2203.08127.
- [59] CMS Collaboration, Measurement of the CP structure of the Higgs-tau Yukawa coupling, *SciPost Phys. Proc.* **8**, 008 (2022).
- [60] F. Boudjema, R. M. Godbole, D. Guadagnoli, and K. A. Mohan, Lab-frame observables for probing the top-Higgs interaction, *Phys. Rev. D* **92**, 015019 (2015).
- [61] D. Gonçalves, K. Kong, and J. H. Kim, Probing the top-Higgs Yukawa CP structure in dileptonic $t\bar{t}h$ with M_2 -assisted reconstruction, *J. High Energy Phys.* **06** (2018) 079.
- [62] S. Biswas, E. Gabrielli, and B. Mele, Single top and Higgs associated production as a probe of the Htt coupling sign at the LHC, *J. High Energy Phys.* **01** (2013) 088.
- [63] S. Biswas, E. Gabrielli, F. Margaroli, and B. Mele, Direct constraints on the top-Higgs coupling from the 8 TeV LHC data, *J. High Energy Phys.* **07** (2013) 073.

- [64] J. Ellis, D. S. Hwang, K. Sakurai, and M. Takeuchi, Disentangling Higgs-top couplings in associated production, *J. High Energy Phys.* **04** (2014) 004.
- [65] J. Chang, K. Cheung, J. S. Lee, and C.-T. Lu, Probing the top-Yukawa coupling in associated Higgs production with a single top quark, *J. High Energy Phys.* **05** (2014) 062.
- [66] CMS Collaboration, Search for associated production of a single top quark and a Higgs boson in events where the Higgs boson decays to two photons at $\sqrt{s} = 8$ TeV, Technical Report CMS-PAS-HIG-14-001, CERN, Geneva (2014).
- [67] CMS Collaboration, Search for H to $b\bar{b}$ in association with a single top quark as a test of Higgs boson couplings at 13 TeV, Technical Report CMS-PAS-HIG-16-019, CERN, Geneva, 2016.
- [68] ATLAS Collaboration, Search for $H \rightarrow \gamma\gamma$ produced in association with top quarks and constraints on the Yukawa coupling between the top quark and the Higgs boson using data taken at 7 TeV and 8 TeV with the ATLAS detector, *Phys. Lett. B* **740**, 222 (2015).
- [69] J. de Blas *et al.*, Higgs boson studies at future particle colliders, *J. High Energy Phys.* **01** (2020) 139.
- [70] ACME Collaboration, Improved limit on the electric dipole moment of the electron, *Nature (London)* **562**, 355 (2018).
- [71] J. Brod and E. Stamou, Electric dipole moment constraints on cp -violating heavy-quark yukawas at next-to-leading order, *J. High Energy Phys.* **07** (2021) 080.
- [72] J. Brod, Z. Polonsky, and E. Stamou, A precise electron EDM constraint on CP -odd heavy-quark Yukawas, *J. High Energy Phys.* **06** (2024) 091.
- [73] J. Brod, U. Haisch, and J. Zupan, Constraints on CP -violating Higgs couplings to the third generation, *J. High Energy Phys.* **11** (2013) 180.