Constraining *CP*-violating Higgs sectors at the LHC using gluon fusion

Matthew J. Dolan,[1] Philip Harris,[2] Martin Jankowiak,[3] and Michael Spannowsky[4]

[1]*Theory Group, SLAC National Accelerator Laboratory, Menlo Park, California 94025, USA*
[2]*CERN, CH-1211 Geneva 23, Switzerland*
[3]*Institut für Theoretische Physik, Universität Heidelberg, Philosophenweg 16,*
Heidelberg D-69120, Germany
[4]*Institute for Particle Physics Phenomenology, Department of Physics, Durham University,*
Durham DH1 3LE, United Kingdom

We investigate the constraints that the LHC can set on a 126 GeV Higgs boson that is an admixture of *CP* eigenstates. Traditional analyses rely on Higgs couplings to massive vector bosons, which are suppressed for *CP*-odd couplings, so that these analyses have limited sensitivity. Instead we focus on Higgs production in gluon fusion, which occurs at the same order in α_S for both *CP*-even and -odd Higgs couplings to top quarks. We study the Higgs plus two jet final state followed by Higgs decay into a pair of tau leptons. We show that using the 8 TeV data set it is possible to rule out the pure *CP*-odd hypothesis in this channel alone at nearly 95% C.L, assuming that the Higgs is *CP*-even. We also provide projected limits for the 14 TeV LHC run.

PACS numbers: 14.80.Bn, 12.60.Fr

I. INTRODUCTION

The discovery of the Higgs boson at the Large Hadron Collider (LHC) [1,2] marks the beginning of a long and detailed experimental program to measure and constrain the couplings and quantum numbers of the new resonance. In particular, efforts are underway to measure whether the new particle is even or odd under the *CP* transformation, with current results apparently disfavoring the *CP*-odd hypothesis by nearly 3σ [3–5].

However, there are numerous examples of extensions of the Standard Model Higgs sector where *CP* is violated and is not a good quantum number of the Higgs-like state (see Ref. [6] for a review of a large number of such scenarios). In these models, and indeed in general, one is interested in constraining the properties of the admixture and the extent to which the Higgs is *CP* even or odd, rather than asking whether it is 100% one or the other. The discovery that the Higgs has a nontrivial *CP* coupling structure would be direct evidence for beyond the Standard Model (BSM) physics with many important implications, for instance in baryogenesis [7]. While it is known that such studies are difficult, given the continually advancing nature of reconstruction and statistical techniques, it is worth investigating the prospects for constraining a mixed-*CP* Higgs at the LHC.

Measuring the *CP* eigenvalue of the Higgs (assuming that *CP* is conserved) is a subject with a long pedigree and extensive literature. Many of the searches and variables proposed to constrain the *CP* properties of the Higgs rely on its couplings to massive vector bosons. Constraints can be set either by exploiting angular correlations between the leptons from the $ZZ^* \to 4l$ or $2l2j$ decays [3,8–14] or through angular correlations in the tagging jets in the weak boson fusion (WBF) production mechanism [15–20].

In either case, these methods rely on the existence of unsuppressed (tree-level) couplings between the Higgs and the massive vector bosons.

While this is the case for the *CP*-even component of the Higgs, which couples to the massive vector bosons $V = (W, Z)$ through the $hV^\mu V_\mu$ operator, the *CP*-odd coupling enters at dimension 5 through the $hV^{\mu\nu}\tilde{V}_{\mu\nu}$ operator, where $V^{\mu\nu}$ is the field strength operator for V^μ. Accordingly, *CP*-odd effects in $h \to ZZ^*$ decays and WBF are suppressed by $\mathcal{O}(\alpha_{\rm EW})$, so that these methods effectively project out the *CP*-odd part of the Higgs (although see Ref. [21] for a study which incorporates loop effects and Ref. [22] for a discussion of $h \to VV$ decays in some specific BSM models).

Such studies often assume that BSM physics enters at a low enough scale such that the dimension-5 operator contributes at the same order of magnitude as the tree-level *CP*-even contribution. However, the existence of light electroweakly interacting states necessary for such a large enhancement of the *CP*-odd couplings to massive vector bosons is now being directly probed by LHC searches for BSM physics, where no signals inconsistent with the SM have been observed. Furthermore, such states would likely lead to large deviations from Standard Model (SM) phenomenology in Higgs boson decays to electroweak gauge bosons, which are also in good agreement with the Standard Model. Therefore, we expect that contributions due to *CP*-violating couplings to massive vector bosons should be negligible in our study.

Instead it is more promising to study the possible *CP*-odd admixture of the 126 GeV resonance via interactions where the *CP*-even and *CP*-odd couplings are induced at the same order. At tree level this includes the couplings to quarks and leptons and at loop level the

couplings to gluons and photons. One gluon-induced production process where it is known that sensitivity to Higgs *CP* properties is preserved is $pp \rightarrow h + 2j$ [23]. As in the WBF channel, the main sensitivity is expected to come from angular correlations between the two tagging jets [24–28], correlations which can also be exploited in diffractive Higgs production at the LHC [29]. Unlike WBF, in this case the *CP*-even and $\mathcal{CP}$-odd contributions are of the same order with the relevant operators being $hG^{\mu\nu}G_{\mu\nu}$ and $hG^{\mu\nu}\tilde{G}_{\mu\nu}$, respectively. Other processes relying on fermionic couplings which have recently been studied in the context of setting constraints on the Higgs *CP* properties include $t\bar{t}h$ and Higgs production in association with a single top [30]. It is also possible that new high-scale physics could induce a nonzero contribution to the $hG^{\mu\nu}\tilde{G}_{\mu\nu}$ operator without a change in the $ht\bar{t}$ coupling. While we do not have such a scenario in mind in this article, our method can be straightforwardly extended to this situation. In principle it may be possible to disentangle such contributions using boosted Higgs transverse momenta [31–33], for instance.

While both the Higgs decay mode $h \rightarrow \gamma\gamma$ followed by conversion of both photons to e^+e^- pairs [34,35] and $h \rightarrow Z\gamma \rightarrow l^+l^-\gamma$ [36] have recently been suggested as a possible final state for probing Higgs *CP* properties, we instead elect to utilize the $h \rightarrow \tau\tau$ decay mode. The majority of previous studies on *CP* in $h \rightarrow \tau\tau$ focus on methods for measuring the polarization properties of the Higgs decay products [37–41]. This requires knowledge of the impact parameter or rest frame of the τs, both of which are difficult quantities to reconstruct in a hadron collider environment (although see Refs. [42,43]).

Any collider study of Higgs *CP* properties must be compared with measurements from other sources. Particularly relevant are measurements of electric dipole moments (EDMs) [44,45], which lead to very strong constraints on mixing between *CP*-even and *CP*-odd Higgs components. These constraints, however, rely on the existence of SM-strength interactions of the Higgs to electrons, an assumption that cannot be put to the test at the LHC. Constraints from EDM experiments are therefore complementary to the analysis strategy followed here. Conceivably, we might discover evidence for *CP* violation in gluon fusion (GF), which, together with a null signal from EDM experiments, would reveal invaluable information about Higgs couplings to the first-generation leptons.

We find using a set of cuts modeled on the current CMS $h \rightarrow \tau\tau$ analysis [46] that data from the 8 TeV run of the LHC is already sufficient to exclude a *CP*-odd Higgs boson at nearly 95% C.L.. This can be compared with current bounds presented in Refs. [47–49], which reinterpret current data to set limits on Higgs *CP* properties using measured rates for Higgs production and find constraints at a similar level. Note, however, that arguments based upon rates alone will always have a flat direction due to possible rescalings of the couplings and Higgs width, and so a

differential analysis strategy such as ours should be more robust.

The rest of this paper is organized as follows. In Sec. II we introduce the parametrization of *CP*-violating effects which we will study: the Standard Model Lagrangian augmented with *CP*-violating terms and higher-dimensional operators encoding the effects of particles running in loops. In Sec. III we discuss our methodology and simulations. In Sec. IV we present our results for the expected limits from current LHC data and projections for the limits that can be set with the 14 TeV data set, before presenting our conclusions and possible directions for future research in Sec. V.

II. MODEL

There is a wide variety of models in the literature that lead to *CP* violation in the Higgs sector, such as generalized two-Higgs doublet models, the *CP*-violating minimal supersymmetric Standard Model (often studied in the CPX [50] scenario), and other supersymmetric models that involve R-parity violation [6]. Such scenarios involve a rich UV spectrum of states that is the subject of various LHC searches. In this article we wish to be as model independent as possible and so keep only the 126 GeV Higgs as part of the spectrum, assuming that other BSM states are either out of direct reach of the LHC or that their effects are subdominant for this analysis.

Our model consists of the Standard Model but with the Lagrangian augmented in the following way to include *CP*-violating couplings. Following Ref. [26] we include couplings between Standard Model fermions and the resonance h which we associate with the Higgs boson:

$$\mathcal{L}_{h\bar{f}f} = \cos\alpha\, y_f \bar{\psi}_f \psi_f h + \sin\alpha\, \tilde{y}_f \bar{\psi}_f i\gamma_5 \psi_f h. \quad (1)$$

We have introduced a mixing angle α such that $\cos\alpha = 1$ (equivalently $\alpha = 0$) corresponds to a Standard Model-like *CP*-even Higgs, while $\sin\alpha = 1$ (equivalently $\alpha = \pi/2$) corresponds to a *CP*-odd pseudoscalar. This allows us to study the *CP* properties of the resonance h as a continuous function of the mixing angle α. We will also assume that $y_f = \tilde{y}_f = m_f/v$. Having fixed the interactions with fermions allows us to derive the dimension-5 operators that govern the interaction of h with massless vector bosons, obtaining [51,52]

$$\mathcal{L}_{hgg} = \cos\alpha \frac{\alpha_S}{12\pi v} h G^a_{\mu\nu} G^{a,\mu\nu} + \sin\alpha \frac{\alpha_S}{4\pi v} h G^a_{\mu\nu} \tilde{G}^{a,\mu\nu} \quad (2)$$

for the gluonic interactions, where v is the vacuum expectation value of the SM Higgs, and $\tilde{G}_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}G^{\rho\sigma}$ is the dual field-strength tensor. Note that, when generating events for our analysis, we do not integrate out the top quark, keeping its full mass dependence throughout, so that the effective operators in Eq. (2) should be understood as convenient shorthand.

The leading-order contribution to the interactions of the Higgs with the massive vector bosons is given by

$$\mathcal{L}_{hVV} \supset \cos\alpha \frac{2m_W^2}{v} h W_\mu W^\mu + \cos\alpha \frac{2m_Z^2}{v} h Z_\mu Z^\mu. \quad (3)$$

We neglect higher-order terms, which are loop suppressed by $\mathcal{O}(\alpha_{EW})$ relative to this, although see Ref. [22] for a discussion of how large these terms can become in some BSM models. Note that, while the SM matter fields also induce dimension-5 operators that lead to the decay $h \to \gamma\gamma$, they do not play a role in this article.

III. METHOD

A. Event generation

We generate signal events at leading order using VBFNLO 2.6.3 [23,53–56] including both the vector boson fusion and gluon fusion production mechanisms, before showering the resulting Les Houches event [57] files using PYTHIA 6 [58] with the Z2 tune [59]. Events are generated at $\sqrt{s} = 8$ and $\sqrt{s} = 14$ TeV with the cteq6ll PDF set [60]. The mixing angle ranges from $\alpha = 0$ to $\alpha = 1.5$ in steps of $\Delta\alpha = 0.3$ for $m_H = 126$ GeV. For each value of α and for each initial state $\mathcal{O}(1M - 4M)$, events are generated. At $\sqrt{s} = 8$ TeV generator level cuts are $|\eta(H)| < 2.5$, $|\eta(j)| < 5.0$ for the two required jets, $p_{Tj} > 20$ GeV, $\Delta R_{jj} > 0.6$, $m_{jj} > 200$ GeV, and $p_{T,H} > 70$ GeV. At $\sqrt{s} = 14$ TeV the cuts are identical, except the m_{jj} cut is instead raised to $m_{jj} > 400$ GeV. No cuts are made on $\Delta\eta_{jj,min}$ or $\eta_{j_1} \times \eta_{j_2}$ at the generator level. For the gluon fusion process, the full top mass dependence is retained in the loop, while the bottom quark contribution is neglected. In the Higgs decay to $\tau\tau$ the Higgs is treated as a *CP*-even scalar, since in this study τ polarization plays no role. This prescription also effects τ kinematics, but only at a negligible level suppressed by $\mathcal{O}(m_\tau/p_T)$.

As demonstrated in the experimental papers [46,61] the dominant backgrounds for $h + 2j$ production followed by $h \to \tau\tau$ are Zjj, $W+$ jets, and to a lesser extent $t\bar{t}$. We generate events for these processes at 8 and 14 TeV using SHERPA 2.0.0 [62] with a similar series of selection cuts ($|\eta(\tau)| < 2.5$, $p_{Tj} > 20$ GeV and $\Delta R_{jj} > 0.6$) to those

described for the signal above. We consider the electroweak and QCD production of Zjj separately. We do not take into account backgrounds arising from $h \to WW$ production, which only lead to small changes in the $e\mu$ channel in our study. We do not generate any QCD multijet backgrounds, which are important for jets faking taus when both taus decay hadronically (see below).

The gluon fusion signal is computed at next-to-leading order (NLO) with differential distributions, which is the state of the art. The WBF signal is only computed at leading order, but as we have checked explicitly using VBFNLO high-order corrections are small, as is well known (see, e.g., Ref. [17]). Finally, all backgrounds are computed at leading order (LO) with parton shower/matrix element merging and corrected with global NLO K factors.

We show in Table I the cross sections at parton level for the signal as a function of the mixing angle α for both the gluon fusion and vector boson fusion channels at 8 (left) and 14 TeV (right). We observe that the WBF contribution decreases with increasing mixing angle α as expected, while the contribution from the gluon fusion component increases.

B. Simulation details

We select four different final states with which to perform our analysis, classified by the τ decay channel. There is the fully hadronic di-τ_h case and the semileptonic and leptonic cases $e\tau_h$, $\mu\tau_h$, and $e\mu$. The initial selection cuts we apply to these final states are shown in Table II. The selection is intended to closely mimic both the CMS and ATLAS di-τ analysis. The one missing background from the simulation is the QCD multijet background where a jet imitates a lepton or fake τ_h. This background is particularly important in the di-τ_h final state. We assume that the QCD contribution is flat and uniformly covers the full phase space of the selected region. This is consistent with the results of Ref. [46]. We set the normalization by considering the differential m_{jj} cross section from QCD, extrapolating this to the Z mass, and multiplying by the fake rate for a jet to fake a tau at 50 GeV. Following the selection, using the 8 TeV samples the yields are found to be comparable to both existing CMS and ATLAS results at the 10% level.

TABLE I. The gluon fusion and weak boson fusion signal cross sections at the generator level before event selection and Higgs decay for 8 (left) and 14 TeV (right).

α	8 TeV GF cross section (fb)	8 TeV WBF cross section (fb)	14 TeV GF cross section (fb)	14 TeV WBF cross section (fb)
0.00	250	467	1141	1481
0.30	278	426	1268	1351
0.60	352	318	1606	1009
0.90	447	181	2038	572
1.20	529	61	2411	194

TABLE II.   Kinematic selection and jet selection for the four different channels ($\tau_h\tau_h$, $\mu\tau_h$, $e\tau_h$ and $e\mu$) used for our di-τ analysis. The di-jet selection includes both exclusive loose and tight categories for all the channels apart from the $\tau_h\tau_h$ channel.

	$\tau_h\tau_h$	$\mu\tau_h$	$e\tau_h$	$e\mu$
Lepton selection	$p_T^\tau > 45$ GeV	$p_T^\mu > 20$ GeV $p_T^\tau > 30$ GeV	$p_T^e > 25$ GeV $p_T^\tau > 30$ GeV	$p_T^{\text{lead}} > 20$ GeV $p_T^{\text{trail}} > 10$ GeV
Kinematic selection	$p_T^H > 100$ GeV	$m_T^\mu < 30$ GeV	$m_T^e < 30$ GeV	b-tag veto with $p_T^b > 20$ GeV
Loose jet selection	$m_{jj} > 500$ GeV $|\Delta\eta_{jj}| > 3.5$	$m_{jj} > 500$ GeV $|\Delta\eta_{jj}| > 3.5$	$m_{jj} > 500$ GeV $|\Delta\eta_{jj}| > 3.5$	$m_{jj} > 500$ GeV $|\Delta\eta_{jj}| > 3.5$
Tight jet selection		$m_{jj} > 700$ GeV $|\Delta\eta_{jj}| > 4.5$ $p_T^H > 100$ GeV	$m_{jj} > 700$ GeV $|\Delta\eta_{jj}| > 4.5$ $p_T^H > 100$ GeV	$m_{jj} > 700$ GeV $|\Delta\eta_{jj}| > 4.5$ $p_T^H > 100$ GeV

To emulate the performance of the detectors all reconstructed physics objects are smeared by a standard set of resolution functions. For the muons, electrons, and τ_h, resolution functions with widths of 2 GeV, 3 GeV, and 4 GeV are used. For the jets a series of resolution functions binned in η is used. The parametrizations for these are taken from Refs. [63,64].

The smearing is parametrized in the unclustered p_T and smeared separately for the parallel and perpendicular components of the unclustered energy with respect to the Higgs p_T. To simulate the instance of fake τ_h being produced from a jet, the jet having the smallest energy in an annulus about the jet axis of $0.1 < \Delta R < 0.4$ and having a $p_T > 20$ GeV is selected and deemed to be the fake τ_h. Provided a fake τ_h candidate exists, the event is then reweighted as a function of the p_T of the τ using the fake rates reported by CMS [65]. Finally, for events in the $e\tau_h$ channel a non-negligable fake background results from the instance in which one electron is reconstructed as a fake τ_h candidate. To simulate this number we take the fake rate to be roughly consistent with the tight working points for both the ATLAS and CMS antielectron vetoes [46,61]. For the lepton efficiencies a flat efficiency corresponding to 90% is taken for the muons, 80% for the electrons, and 60% for the taus. These numbers take into account both the expected trigger and identification efficiencies for the leptons after a typical $e/\mu/\tau_h$ selection. For the τ_h efficiency in the case where an antielectron veto is applied the corresponding τ_h efficiency is scaled down by an additional 10%. For the $e\mu$ channel the efficiencies are scaled up by 5% in electrons and muons, corresponding to the improved trigger efficiency in these cases.

The uncertainties applied in the extraction of the signal closely resemble the current LHC analyses. For each background, an uncorrelated normalization uncertainty of 10%–50% is applied in each category. The variation in the uncertainty is dependent on whether a real or fake τ_h is present. Additional correlated normalization uncertainties of 1%–5% are also applied reflecting the effects of lepton

efficiencies, jet scale, and luminosity. No lepton energy scale uncertainty is applied, since this is well constrained from other categories in the LHC analyses. Regarding the theoretical uncertainties on the Higgs yields, the cross section uncertainties from the Higgs working group are applied [66], along with an additional uncorrelated 25% uncertainty on the overall gluon fusion yield to reflect the current knowledge of dijet production in gluon fusion. For projections with an integrated luminosity > 20 fb^{-1}, this uncertainty is reduced to 10%, reflecting expected theoretical improvements in the signal yield calculation. As with the current LHC analyses, these systematic uncertainties are added to the signal extraction, separately floating each uncertainty under a Gaussian prior for which the width is specified by the systematic uncertainty.

As with the CMS and ATLAS analyses, extraction of the signal relies on exploiting full knowledge of the τ decays to improve the mass separation of the signal from the largest background, $Z \to \tau_h\tau_h$. Such a scenario benefits greatly from incorporating knowledge of the τ decay matrix elements into the kinematic reconstruction of the di-τ mass. To perform this mass reconstruction, we developed a di-τ_h mass reconstruction that computes a weighted likelihood of the di-τ_h mass on an event-by-event basis by randomly sampling the allowed neutrino kinematics from the leading-order matrix elements and weighting each event by the consistency with the observed missing transverse energy (MET), using the full MET resolution covariance matrix. This mass reconstruction is very similar to the MCT approach used by ATLAS and the SVFit mass approach used by CMS [46,67]. As a final cross-check, our simulation was checked against the current CMS di-τ analysis and gave yields, shapes, and results similar to those reported in their paper [46].

C. Observable distributions

We show in Fig. 1 starting from the top left and working clockwise the distributions for the invariant mass m_{jj} of the two tagging jets, the azimuthal angle difference $\Delta\phi_{jj}$
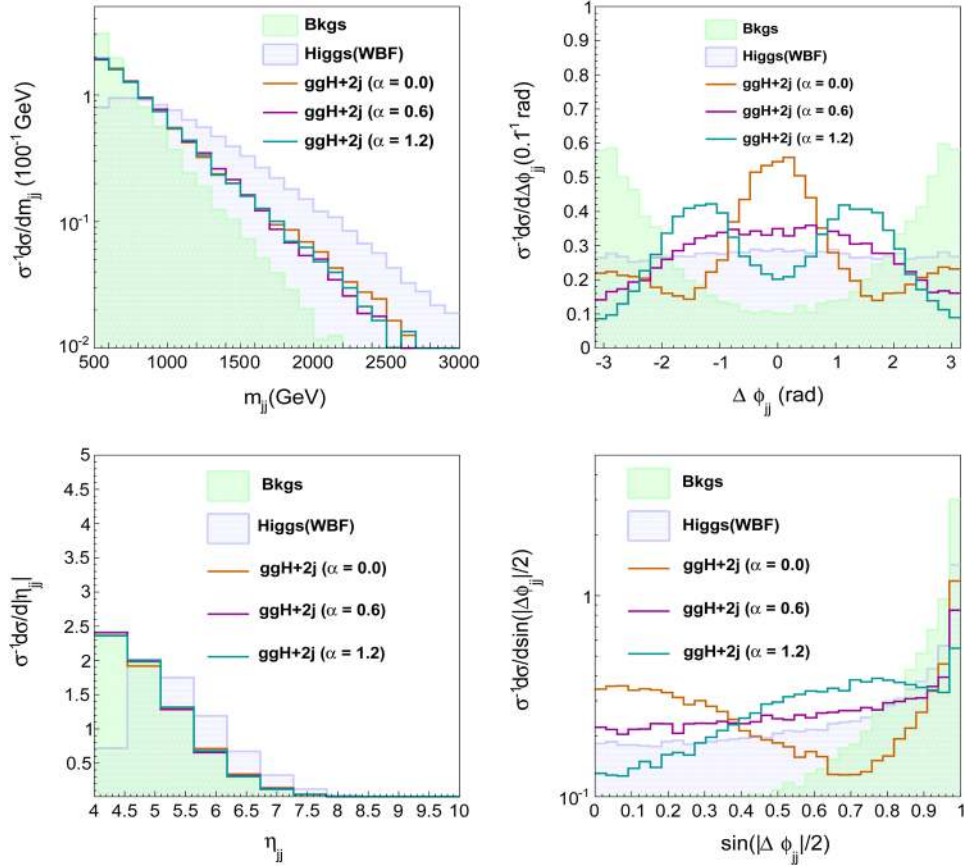
FIG. 1 (color online). Observable distributions for the signal and background. From the top left and proceeding clockwise: m_{jj}, $\Delta\phi_{jj}$, $\sin(|\Delta\phi_{jj}|/2)$, and $\Delta\eta_{jj}$. For each figure the yields are normalized to the expected yields at 8 TeV for the gluon fusion channel at 20 fb^{-1} with $\alpha = 0$. Samples have been passed through the detector pseudosimulation and subjected to the full selection on all channels. The loose WBF selection and the additional category selections are applied in all cases.

between the tagging jets, the rapidity difference $\Delta\eta_{jj}$ between the jets, and finally the discriminating variable $\sin(|\Delta\phi_{jj}|/2)$. Each figure shows the total background contribution, along with that from WBF Higgs production for $\alpha = 0$ and the GF signal component for $\alpha = 0, 0.6$, and

1.2. The individual contributions are normalized to the expected yields at 8 TeV for 20 fb^{-1} for $\alpha = 0$. The variable showing the largest sensitivity to the mixing angle is the azimuthal angle between the two tagging jets, $\Delta\phi_{jj} = \phi_{y>0} - \phi_{y<0}$, which has long been known to provide a good
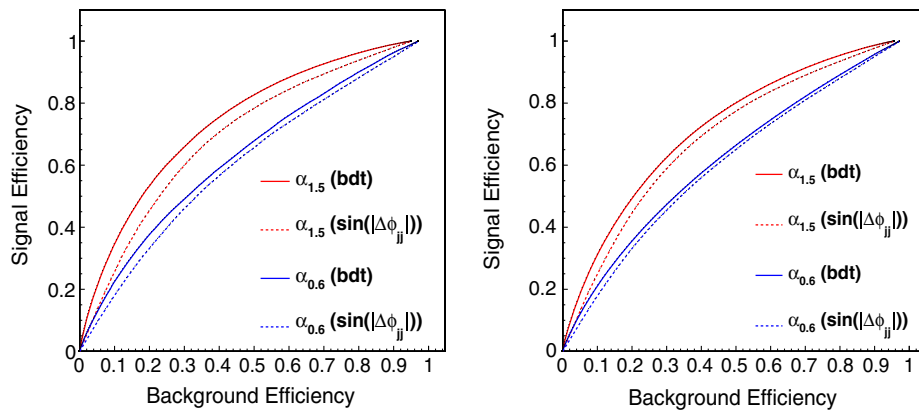

FIG. 2 (color online). Efficiency curves for our boosted decision tree for 8 (left) and 14 TeV (right). The red curves are for $\alpha = 1.5$ and the blue curves for $\alpha = 0.6$. The dashed curves show the results only including the $\sin(|\Delta\phi_{jj}|)$ variable and the solid curves those for the full BDT with all 18 observables included, as described in the text.

handle for discriminating Higgs *CP* properties [25,26]. In addition some small dependence on α can also be observed at large values of the dijet invariant mass m_{jj}. All the distributions we show are for events that have been showered and smeared using our detector pseudosimulation. We have also investigated the p_T distribution of the leading jet, which shows some limited sensitivity to α near the peak of the distribution.

As a cross-check of the possible performance gain that can be had by utilizing other observables we have applied a multivariate analysis (MVA), specifically a boosted decision tree (BDT), that was trained to discriminate a fully simulated gluon fusion sample with $\alpha = 1.2$ from one with $\alpha = 0$. To train this decision tree, we used 18 observables obtained from the pseudosimulation. These include the two leading jet η's and p_T's; the 3-vectors for the visible components of the τ decays; the kinematically fitted mass $m_{\tau\tau}$; the Higgs p_T constructed from the MET and the visible decay products; the MET; the transverse mass of either lepton combined with the MET; and the m_{jj}, $\Delta\eta_{jj}$, and $\Delta\phi_{jj}$ variables. The training was performed separately for each channel, so as to improve the individual performance of each observable. The performance gain of these variables with respect to $\sin(|\Delta\phi_{jj}|/2)$ is shown in Fig. 2 for both 8 and 14 TeV.

As part of the optimization studies for the WBF selection, a BDT was used to train both the WBF and gluon fusion signals against a weighted sum of all the backgrounds using the same variables as described in the previous paragraph. After the optimization, only marginal gains were found beyond the addition of four main variables, m_{jj}, $|\Delta\eta_{jj}|$, the di-τ mass $m_{\tau\tau}$, and $\Delta\phi_{jj}$. The addition of $\Delta\phi_{jj}$, in particular, brought a performance improvement of 20% in the WBF sensitivity. In both CMS and ATLAS, this variable had been used minimally, so as to avoid spin sensitivity and to avoid complications resulting from theoretical modeling of the second jet in gluon fusion. Once $\Delta\phi_{jj}$ was added, it was further found that a category-based analysis binning in mass, $\Delta\phi_{jj}$, m_{jj}, and $\Delta\eta_{jj}$ performed as well as a BDT trained on the full set of observables.

D. $\Delta\phi_{jj}$ analytics

We now briefly discuss the $\Delta\phi_{jj}$ dependence of the two different production mechanisms. To begin, consider the gluon fusion process (specifically $gg \rightarrow Hgg$) in the $m_t \rightarrow \infty$ limit. Apart from the strength of the coupling constants in Eq. (2), the only difference between $\alpha = 0$ and $\alpha = \pi/2$ is to be found in the form of the helicity-conserving amplitudes [68–71]. We note that the helicity-violating amplitudes do exhibit $\Delta\phi_{jj}$ dependence, but the resulting terms are independent of the mixing angle α (apart from the strength of the coupling constants). Consider a final state configuration in which the Higgs is central ($y_H = 0$) and the two jets have opposite rapidities ($y_j \equiv y_{j1} = -y_{j2}$).

Given our selection cuts described above, this represents a typical final state. As a further simplification let the lab frame and the c.m. frame be identical so that the initial state gluons have equal and opposite 3-momenta, $\vec{p} = \pm\frac{1}{2}E_{CM}\hat{z}$, where $\hat{z}$ is the direction along the beam axis. In the limit where the final state jets have large rapidities, one finds that the helicity-conserving squared matrix element for *CP*-even (+ sign) and *CP*-odd (− sign) is given by (omitting coupling constants and other numerical prefactors)

$$|\mathcal{M}|^2_{GF\pm} \sim \exp(4y_j)\{A \pm B\cos(2\Delta\phi_{jj})\}, \qquad (4)$$

where

$$A = \xi^4 + \xi^{-4} + \frac{1}{2}(\xi^5 + \xi^{-3}) \quad \text{and} \quad B = 2 + \xi^2 \quad \text{with}$$

$$\xi \equiv \frac{E_{CM}}{E_{CM} - m_h} \qquad (5)$$

so that in the limit where $m_h \ll E_{CM}$ we have

$$|\mathcal{M}|^2_{GF\pm} \sim \exp(4y_j)\{3 \pm 3\cos(2\Delta\phi_{jj})\}. \qquad (6)$$

In the case of the WBF production mechanism Ref. [15] argued that the matrix element squared in the limit of forward jets is approximately given by

$$|\mathcal{M}|^2_{WBF} \sim \hat{s}m^2_{jj}, \qquad (7)$$

where $\hat{s}$ is the partonic center-of-mass energy, which results in an essentially flat distribution in $\Delta\phi_{jj}$. Note, however, that the dimension-5 operators that we assume to be negligible and omit ($hV^{\mu\nu}V_{\mu\nu}$ and $hV^{\mu\nu}\tilde{V}_{\mu\nu}$) lead to a nontrivial $\Delta\phi_{jj}$ dependence.

The dependence of the gluon fusion and WBF production mechanisms on $\Delta\phi_{jj}$ can be seen explicitly in Fig. 1. The form of the distributions follows the expectations from Eqs. (4) and (7). The fact that the approximations leading to these matrix elements are quite good makes it clear why $\Delta\phi_{jj}$ by itself is nearly optimal as a discriminating observable between the *CP*-even and *CP*-odd cases. In principle, we can include in Eq. (4) the next term in the series in $\exp(y_j)$. Doing so upsets the factorized form $|\mathcal{M}|^2_{GF\pm} \sim f(y_j)g(\Delta\phi_{jj})$. In particular, the next term in the series, which is proportional to $\exp(2y_j)$, includes subterms with $\cos(n\Delta\phi_{jj})$ for $n = 1, 2, 3$ and breaks the degeneracy $|\mathcal{M}|^2_{GF\pm}(\Delta\phi_{jj}) = |\mathcal{M}|^2_{GF\pm}(\Delta\phi_{jj} \pm \pi)$ in Eq. (4), reflecting the observed behavior in Fig. 1. This clarifies why the BDT has an edge in discriminatory power. However, since the correction due to the next term in $\exp(y_j)$ is small for the phase space region of interest, the BDT exhibits only marginally better discriminatory power than $\Delta\phi_{jj}$ by itself (as is evidenced in Fig. 2).

The importance of the WBF-like cuts (in particular the cut on m_{jj} which pushes the tagging jets to large rapidities)
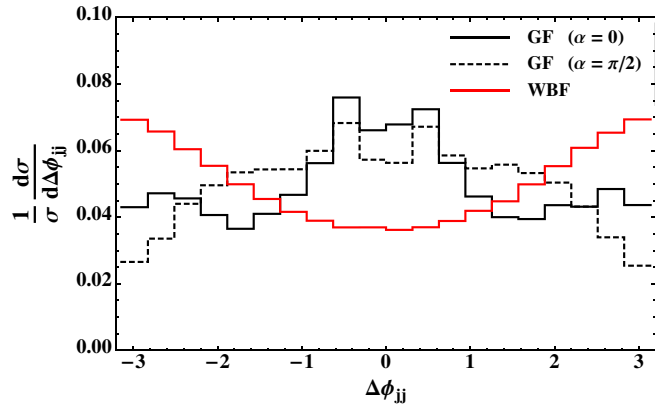
FIG. 3 (color online). Signal distributions for $\Delta\phi_{jj}$ with inclusive (non-WBF-like) cuts at $\sqrt{s} = 8$ TeV: $p_{Tj} > 30$ GeV, $\Delta R_{jj} > 0.4$. In particular there are no cuts on m_{jj} or $p_{T,H}$. These are generator-level distributions, with each channel separately normalized by its cross section, and the whole plot is normalized to $1/\pi$. Note that by assumption the shape of the WBF distribution does not vary with α.

for emphasizing the *CP*-dependent $\pm \cos(2\Delta\phi_{jj})$ terms in the gluon fusion channel can be seen in Fig. 3. A much more inclusive set of selection cuts results in a signal distribution for $\Delta\phi_{jj}$ that is still sensitive to *CP* effects but at a substantially reduced level.

IV. ESTIMATED LIMITS ON *CP* PROPERTIES

We now discuss our results. In Fig. 4 we show the significances that can be achieved using the 20 fb^{-1} of data from the 8 TeV run and projected limits for 50 fb^{-1} of data at 14 TeV, corresponding to around two years of running. Those results from the analyses marked with "loose" were performed using the loose analysis cuts from Table II, while those marked "tight" were performed with the tight analysis

cuts from the same table, which forms a subset of the loose category.

The dashed curves show the estimated significance of the total signal over the Standard Model backgrounds. The dark yellow dashed curve shows the results obtained doing a standard WBF-style analysis with loose cuts, achieving a significance of barely 2σ over the background. The purple dashed curve shows our best approximation to the current CMS analysis [46] with tighter cuts, while the upper three dashed curves either utilize the 18 variables described above (blue dashed) or use a simultaneous fit to both the ditau invariant mass $m_{\tau\tau}$ and the discriminating variable $\sin(|\Delta\phi_{jj}|/2)$ (green and maroon dashed, with loose and tight cuts, respectively).

The solid curves show the exclusion significance computed using the CL$_s$ method [72] relative to the $\alpha = 0$ case. The maroon curve again shows the results using the loose event selection and ditau invariant mass and $\sin(|\Delta\phi_{jj}/2|)$, while the blue and green curves utilize the tight selection and (in the green case) the BDT. We observe from the left-hand figure that with our best analysis a pure *CP*-odd Higgs corresponding to $\alpha = \pi/2$ is already nearly ruled out at 95% C.L. With 20 fb^{-1} of luminosity at 14 TeV this should improve to $\alpha \leq 0.9$, while with 50 fb^{-1} of luminosity it should improve further to $\alpha \leq 0.7$.

To further elucidate how the constraints on *CP* mixing will improve, in the left-hand plot of Fig. 5 we show the expected exclusion limit on the mixing angle α as a function of the integrated luminosity at 14 TeV. This shows that the limit should improve to $\alpha \leq 0.3$ with 500 fb^{-1}. In the right-hand plot we increase the theoretical uncertainty from 10% to 25%, in case theoretical advances do no keep up with experimental ones. We find that the two curves are within errors of each other, since the 25% uncertainty on the theory prediction only starts to affect things at 4σ level. As can be seen from the figure precision measurements of
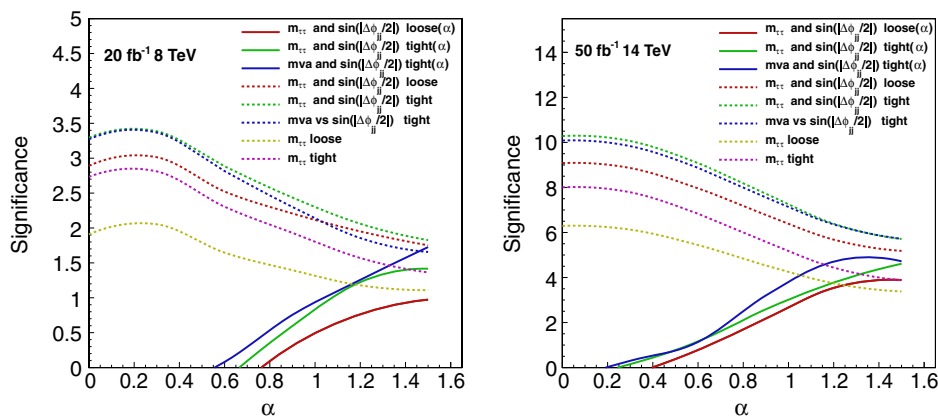

FIG. 4 (color online). Expected limits that can be achieved with our analysis using the 20 fb^{-1} 8 TeV data set (left) and using a 50 fb^{-1} data set at 14 TeV (right). The dashed curves show the estimated significance of the total signal over the Standard Model backgrounds, and the solid curves show the exclusion significance computed using the CL$_s$ method relative to the $\alpha = 0$ case. See the text for details.
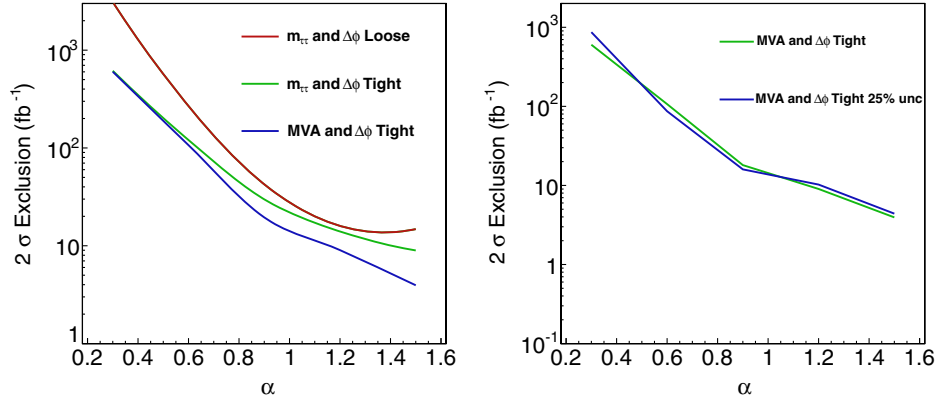
FIG. 5 (color online). Left: The projected 95% exclusion confidence limit on the mixing angle α that can be set as a function of the integrated luminosity at the 14 TeV run of the LHC (assuming $\alpha = 0$ as the null hypothesis). Right: As in the left-hand plot, but comparing with an inflated theoretical error of 25%.

Higgs CP properties will benefit greatly from a high luminosity LHC run.

We note that the limits we have set can in principle be improved upon by including other techniques which are sensitive to the CP properties of the Higgs, such as including detailed information about the τ decay products as in Refs. [37,38]. Further discriminatory power between the gluon fusion and weak boson fusion production mechanisms could also be gained by using likelihood methods as in Ref. [73]. We are thus hopeful that it may be possible to improve upon our projections. With a similar analysis it may even be possible to extract information from the $h \to b\bar{b}$ decay.

V. CONCLUSIONS

Measuring the detailed CP properties of the Higgs is one of the most important aspects of the precision Higgs program in the upcoming 14 TeV run at the LHC. Previous theoretical and current experimental analyses have focused on exploiting the Higgs couplings to massive vector bosons. However, the CP-odd couplings to W and Z are suppressed, so that analyses based on these couplings project out much of the physics of interest. Instead, we focused on Higgs interactions that have the same parametric strength for the CP-even and -odd Higgs components. This led us to consider Higgs production in association with two jets. We interpreted our results in a phenomenological parametrization with CP-violating $ht\bar{t}$ couplings. Our analysis is also sensitive to dimension-5 $hG^{\mu\nu}\tilde{G}_{\mu\nu}$ operators induced by new high scale physics and

is unlikely to be sensitive to CP-violating couplings to massive vector bosons since their effects should be small taking into account constraints from orthogonal Higgs-property measurements.

We then focused on Higgs decay into a pair of τ leptons. Our analysis exploits the jet correlations in Higgs production and is thus relatively independent of the CP nature of the $h\tau\tau$ coupling. Changes in the $h\tau\tau$ coupling will change the statistics but not affect in any fundamental way our ability to set a limit on the CP mixing in this channel.

We have carried out a detailed simulation of the signal and backgrounds, taking detector effects such as acceptances and fake rates into account, and used a multivariate analysis to achieve excellent discriminating power in the mixing angle α. We have presented estimates of the constraints that can be set using the current 8 TeV data set, as well as 20 and 50 fb^{-1} of data at 14 TeV, corresponding to approximately one and two years of running. We find that the 8 TeV data set should be able to achieve nearly 95% C.L. exclusion of a CP-odd Higgs relative to a CP-even one. This should improve even further with the 14 TeV run such that $\alpha \geq 0.7$ could be excluded with 50 fb^{-1} and $\alpha \geq 0.3$ with 500 fb^{-1}. By including other Higgs decay modes, e.g., $H \to \gamma\gamma$, the exclusion reach can be extended even further.

ACKNOWLEDGMENTS

[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] S. Chatrchyan *et al.* (CMS Collaboration), Phys. Rev. Lett. **110**, 081803 (2013).

[4] G. Aad *et al.* (ATLAS Collaboration), Phys. Lett. B **726**, 88 (2013).

[5] S. Chatrchyan *et al.* (CMS Collaboration), Phys. Rev. D **89**, 092007 (2014).

[6] E. Accomando *et al.*, arXiv:hep-ph/0608079.

[7] W. Bernreuther, Lect. Notes Phys. **591**, 237 (2002).

[8] C. Buszello, I. Fleck, P. Marquard, and J. van der Bij, Eur. Phys. J. C **32**, 209 (2004).

[9] S. Choi, D. Miller, M. Muhlleitner, and P. Zerwas, Phys. Lett. B **553**, 61 (2003).

[10] R. M. Godbole, D. Miller, and M. M. Muhlleitner, J. High Energy Phys. 12 (2007) 031.

[11] C. Hackstein and M. Spannowsky, Phys. Rev. D **82**, 113012 (2010).

[12] C. Englert, C. Hackstein, and M. Spannowsky, Phys. Rev. D **82**, 114024 (2010).

[13] S. Bolognesi, Y. Gao, A. V. Gritsan, K. Melnikov, M. Schulze, N. V. Tran, and A. Whitbeck, Phys. Rev. D **86**, 095031 (2012).

[14] I. Anderson *et al.*, Phys. Rev. D **89**, 035007 (2014).

[15] T. Plehn, D. L. Rainwater, and D. Zeppenfeld, Phys. Rev. Lett. **88**, 051801 (2002).

[16] K. Odagiri, J. High Energy Phys. 03 (2003) 009.

[17] T. Figy and D. Zeppenfeld, Phys. Lett. B **591**, 297 (2004).

[18] C. Englert, D. Goncalves-Netto, K. Mawatari, and T. Plehn, J. High Energy Phys. 01 (2013) 148.

[19] P. Artoisenet *et al.*, J. High Energy Phys. 11 (2013) 043.

[20] F. Maltoni, K. Mawatari, and M. Zaro, Eur. Phys. J. **74**, 2710 (2014).

[21] Y. Chen, R. Harnik, and R. Vega-Morales, arXiv:1404.1336.

[22] W. Bernreuther, P. Gonzalez, and M. Wiebusch, Eur. Phys. J. C **69**, 31 (2010).

[23] V. Del Duca, W. Kilgore, C. Oleari, C. Schmidt, and D. Zeppenfeld, Nucl. Phys. **B616**, 367 (2001).

[24] V. Hankele, G. Klamke, and D. Zeppenfeld, arXiv:hep-ph/0605117.

[25] V. Del Duca, G. Klämke, M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, A. D. Polosa, and D. Zeppenfeld, J. High Energy Phys. 10 (2006) 016.

[26] G. Klamke and D. Zeppenfeld, J. High Energy Phys. 04 (2007) 052.

[27] F. Campanario, M. Kubocz, and D. Zeppenfeld, Phys. Rev. D **84**, 095025 (2011).

[28] F. Campanario and M. Kubocz, arXiv:1402.1154.

[29] V. Khoze, A. Martin, and M. Ryskin, Eur. Phys. J. C **34**, 327 (2004).

[30] J. Ellis, D. S. Hwang, K. Sakurai, and M. Takeuchi, J. High Energy Phys. 04 (2014) 004.

[31] C. Grojean, E. Salvioni, M. Schlaffer, and A. Weiler, J. High Energy Phys. 05 (2014) 022.

[32] M. Schlaffer, M. Spannowsky, M. Takeuchi, A. Weiler, and C. Wymant, arXiv:1405.4295.

[33] M. Buschmann, C. Englert, D. Goncalves, T. Plehn, and M. Spannowsky, Phys. Rev. D **90**, 013010 (2014).

[34] M. Voloshin, Phys. Rev. D **86**, 093016 (2012).

[35] F. Bishara, Y. Grossman, R. Harnik, D. J. Robinson, J. Shu, and J. Zupan, J. High Energy Phys. 04 (2014) 084.

[36] Y. Chen, A. Falkowski, I. Low, and R. Vega-Morales, arXiv:1405.6723.

[37] G. Bower, T. Pierzchala, Z. Was, and M. Worek, Phys. Lett. B **543**, 227 (2002).

[38] S. Berge, W. Bernreuther, and J. Ziethe, Phys. Rev. Lett. **100**, 171605 (2008).

[39] S. Berge, W. Bernreuther, B. Niepelt, and H. Spiesberger, Phys. Rev. D **84**, 116003 (2011).

[40] S. Berge, W. Bernreuther, and H. Spiesberger, arXiv:1208.1507.

[41] S. Berge, W. Bernreuther, and H. Spiesberger, Phys. Lett. B **727**, 488 (2013).

[42] C. Englert, M. Spannowsky, and M. Takeuchi, J. High Energy Phys. 06 (2012) 108.

[43] R. Harnik, A. Martin, T. Okui, R. Primulando, and F. Yu, Phys. Rev. D **88**, 076009 (2013).

[44] D. McKeen, M. Pospelov, and A. Ritz, Phys. Rev. D **86**, 113004 (2012).

[45] J. Brod, U. Haisch, and J. Zupan, J. High Energy Phys. 11 (2013) 180.

[46] S. Chatrchyan *et al.* (CMS Collaboration), J. High Energy Phys. 05 (2014) 104.

[47] A. Freitas and P. Schwaller, Phys. Rev. D **87**, 055014 (2013).

[48] K. Cheung, J. S. Lee, and P.-Y. Tseng, J. High Energy Phys. 05 (2013) 134.

[49] A. Djouadi and G. Moreau, Eur. Phys. J. C **73**, 2512 (2013).

[50] M. S. Carena, J. R. Ellis, A. Pilaftsis, and C. Wagner, Phys. Lett. B **495**, 155 (2000).

[51] R. P. Kauffman and W. Schaffer, Phys. Rev. D **49**, 551 (1994).

[52] M. Spira, A. Djouadi, D. Graudenz, and P. Zerwas, Nucl. Phys. **B453**, 17 (1995).

[53] K. Arnold *et al.*, Comput. Phys. Commun. **180**, 1661 (2009).

[54] K. Arnold *et al.*, arXiv:1107.4038.

[55] T. Figy, C. Oleari, and D. Zeppenfeld, Phys. Rev. D **68**, 073005 (2003).

[56] V. Del Duca, W. Kilgore, C. Oleari, C. Schmidt, and D. Zeppenfeld, Phys. Rev. Lett. **87**, 122001 (2001).

[57] J. Alwall *et al.*, Comput. Phys. Commun. **176**, 300 (2007).

[58] T. Sjostrand, S. Mrenna, and P. Z. Skands, J. High Energy Phys. 05 (2006) 026.

[59] P. Z. Skands, Phys. Rev. D **82**, 074018 (2010).

[60] J. Pumplin, D. Robert Stump, J. Huston, H.-L. Lai, P. Nadolsky, and W.-K. Tung, J. High Energy Phys. 07 (2002) 012.

[61] ATLAS Collaboration, ATLAS-CONF-2013-108.

[62] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, F. Siegert, and J. Winter, J. High Energy Phys. 02 (2009) 007.

[63] CMS Collaboration, CMS-PAS-JME-10-014.

[64] G. Aad *et al.* (ATLAS Collaboration), Eur. Phys. J. C **73**, 2306 (2013).

[65] CMS Collaboration, CMS-PAS-HIG-12-051.

[66] S. Heinemeyer *et al.* (LHC Higgs Cross Section Working Group), arXiv:1307.1347.

[67] A. Elagin, P. Murat, A. Pranko, and A. Safonov, Nucl. Instrum. Methods Phys. Res., Sect. A **654**, 481 (2011).

[68] S. Dawson and R. Kauffman, Phys. Rev. Lett. **68**, 2273 (1992).

[69] R. P. Kauffman, S. V. Desai, and D. Risal, Phys. Rev. D **55**, 4005 (1997).

[70] R. P. Kauffman and S. V. Desai, Phys. Rev. D **59**, 057504 (1999).

[71] R. P. Kauffman, S. V. Desai, and D. Risal, arXiv:hep-ph/9903330.

[72] A. L. Read (unpublished).

[73] J. R. Andersen, C. Englert, and M. Spannowsky, Phys. Rev. D **87**, 015019 (2013).