FERMILAB-PUB-10-019-T

Higgs look-alikes at the LHC

A. De Rújula[a,b], Joseph Lykken[c], Maurizio Pierini[d], Christopher Rogan[e], and Maria Spiropulu[d,e]

[a] *Instituto de Física Teórica, Univ. Autónoma de Madrid, Madrid, and CIEMAT, Madrid, Spain,*
[b] *Physics Dept., Boston University, Boston, MA 02215,*
[c] *Fermi National Accelerator Laboratory, P.O. Box 500, Batavia, IL 60510,*
[d] *Physics Department, CERN, CH 1211 Geneva 23, Switzerland,*
[e] *Lauritsen Laboratory of Physics, California Institute of Technology, Pasadena, CA 91125*

The discovery of a Higgs particle is possible in a variety of search channels at the LHC. However, the true identity of any putative Higgs boson will, at first, remain ambiguous until one has experimentally excluded other possible assignments of quantum numbers and couplings. We quantify the degree to which one can discriminate a Standard Model Higgs boson from "look-alikes" at, or close to, the moment of discovery at the LHC. We focus on the fully-reconstructible "golden" decay mode to a pair of Z bosons and a four-lepton final state. Considering both on-shell and off-shell Z's, we show how to utilize the full decay information from the events, including the distributions and correlations of the five relevant angular variables. We demonstrate how the finite phase space acceptance of any LHC detector sculpts the decay distributions, a feature neglected in previous studies. We use likelihood ratios to discriminate a Standard Model Higgs from look-alikes with other spins or nonstandard parity, CP, or form factors. For a benchmark resonance mass of 200 GeV/c^2, we achieve a median expected discrimination significance of $3\,\sigma$ with as few as 19 events, and even better discrimination for the off-shell decays of a 145 GeV/c^2 resonance.

I. INTRODUCTION

The CDF and DØ experiments [1] at the Fermilab Tevatron are continuously improving their Higgs mass limits, and the ATLAS and CMS detectors at the CERN LHC are designed to discover [2, 3] the standard Higgs in all of the unexplored mass range, up to the high masses at which its raison d'être is lost. While an un-discovery would be momentous, we focus here on the possibility that evidence resembling the standard expectation is found.

Because the idea is so venerable, one may have grown insensitive to how special a Higgs boson would be. Its quantum numbers must be those of the vacuum, which its field permeates. Its couplings to the electroweak gauge bosons $W^{\pm}$ and Z are proportional to their masses, as are its couplings to quarks and leptons. Any deviation from the predicted quantum numbers or couplings of a putative Higgs boson would have deep ramifications for particle physics. An experimental program for Higgs physics must be focused on the rigorous determination of these fundamental quantities.

A Higgs boson discovery at the LHC will arise from excesses or resonances observed in one or more final states. Since the couplings and partial widths of a SM Higgs boson are predicted as a function of its mass, the size of any excess, the width of a reconstructed resonance, or a comparison of different channels may immediately give clues as to whether the putative new particle is consistent with a SM Higgs boson. Nevertheless, the true identity of the new particle will at first remain ambiguous, until one has experimentally excluded other possible assignments of quantum numbers and couplings. We shall refer to these other possibilities as Higgs look-alikes (HLLs).

The purpose of this paper is to quantify to what degree one can discriminate a Standard Model Higgs boson from HLLs at, or close to, the moment of discovery at the LHC. In spite of much preparation for Higgs searches at hadron colliders, no previous published study of this question exists. There is a vast literature about determining Higgs properties from signals in a variety of final states (for a review, see [4]), but this research mostly addresses only the related question of whether it is possible *at all* to determine Higgs quantum numbers and couplings at a hadron collider. The current situation in this respect is similar to the LHC experimental program for supersymmetry, where only recently are there quantitative studies of the potential to discriminate supersymmetry look-alikes at the moment of discovery [5]-[9].

Our study focuses on the so-called "golden channel" for Higgs discovery, namely the Higgs decay $H \to ZZ^* \to \ell_1^+\ell_1^-\ell_2^+\ell_2^-$, where $\ell_{1,2}^{\pm}$ denotes an electron or a muon. Obviously, this channel has the advantage that the kinematics of the Higgs and its decay products are fully reconstructible from a completely leptonic final state. Approximately half of the events will be $\mu^+\mu^-e^+e^-$, where all four leptons are easily distinguishable, and even in the 4μ and $4e$ final states all four leptons can be distinguished by the requirement that one or both Z bosons are reconstructed within an on-shell mass window. A well-measured, four-body, closed kinematic final state provides many independent observables for determining properties of the observed resonance; thus this channel provides more information than, e.g., the fully-reconstructible Higgs decay into two photons, where the photon polarizations are not measured.

A drawback of the golden mode is the small branching fraction; example values for a SM Higgs $\to ZZ \to 4\ell$ are 0.0011 for $m_H = 200$ GeV/c^2, 0.0014 for $m_H = 350$ GeV/c^2, and 0.00036 for $m_H = 145$ GeV/c^2 [10]. Even

for favorable Higgs masses, this branching fraction is 130 times smaller than that for semileptonic $H \to W^+W^- \to \ell\nu jj$, a channel which, though hampered by large backgrounds, is also fully reconstructible up to a two-fold ambiguity in the determination of the longitudinal neutrino momentum [11, 12]. The golden mode branching fraction is also much smaller than that for the fully leptonic SM Higgs decay $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$. Nevertheless, for a wide range of SM Higgs masses, this mode is a promising discovery channel and would, in any event, be populated at or around the time of a putative discovery in a different channel.

We factorize the HLL problem into observables related to production and observables related to decay. In this paper we perform a systematic analysis including all of the information from the putative Higgs decays, leaving the analysis of Higgs versus HLL production to later work. While this factorization of production and decay is not completely clean, we show that the resulting model-dependent uncertainty introduced into the decay analysis is small. A full analysis will include production information and could produce stronger results than those presented here, since large cross section differences are expected between SM Higgs production and the production of many Higgs look-alikes. However, including Higgs production also introduces new theoretical and measurement uncertainties involving associated hadronic jets and the parton distribution functions (PDFs) that describe the initial state.

One advantage of focusing only on Higgs decay in the four lepton final state is that we can perform a realistic study without resorting to full simulation of a detector. This is demonstrated in Section IV, where we parametrize the relevant efficiencies, resolutions and acceptances for an LHC detector. Because both CMS and ATLAS in general measure muons and electrons with exquisite precision, we show that uncertainties about detector performance have no impact on our results.

This is not to say that detector effects are not important. We will show that the finite phase space acceptance of any LHC detector has strong effects on the HLL analysis, causing a detector-induced sculpting of the angular distributions used for HLL discrimination. We demonstrate that these effects must be accounted for in order to avoid serious errors in the characterization of a Higgs signal.

Our analysis depends on five distinct angles that describe the $H \to ZZ \to 4\ell$ decay process. In the case where one of the Z bosons is strongly off-shell, the SM Higgs versus HLL decays also differ in their dependence on the reconstructed Z^* invariant mass. Because we are interested in HLL discrimination with small data samples, at or near the moment of discovery, we need to use all of the decay information in the events, including not just the distributions but also the correlations between all five (or six) of the relevant observables.

In the same spirit, we disentangle the Standard Model ZZ background from the putative Higgs signal using the ${}_sPlots$ technique [13]. This produces an effectively background-subtracted data sample where, instead of making stringent requirements that reduce signal, we reweight the selected events according to how likely each event is considered to be signal by the fit, keeping the normalization to the signal yield found in the search.

Previous analyses of the Higgs golden mode decay properties have examined the dependence on some of the relevant angular distributions [14]-[19] and have shown the potential for LHC measurements to discriminate a SM Higgs from look-alikes with different spin and parity assignments or CP properties [4],[17]-[32]. However, none of these studies utilized all of the decay information in the events and all of them have ignored the effects of detector phase space sculpting of distributions.

In our analysis we compare a SM Higgs signal to a variety of Higgs look-alikes. We consider the most general couplings, containing up to two derivatives, of a massive, spinless boson to ZZ or ZZ^*; this corresponds to gauge-invariant couplings up to dimension six. Some of the corresponding HLLs can be considered as small modifications of the SM Higgs properties via P or CP violation or Higgs compositeness. Another spin 0 HLL corresponds to a new massive pseudoscalar, a particle occurring in models with extended Higgs sectors such as supersymmetry.

Our HLL analysis also includes the most general couplings, containing up to two derivatives, of a massive neutral spin 1 boson to ZZ or ZZ^*. The off-shell case has not been presented before, to our knowledge. Of course a spin 1 HLL is a special case of what is usually denoted as a Z' vector boson. The spin 1 part of our results is thus also part of a Z' look-alike analysis, which is interesting in its own right [33].

We also include as one of our HLLs a massive spin 2 resonance with gravity-like couplings to ZZ. Although massive gravitons from existing models are already experimentally excluded in the relevant mass range [34], general spin 2 HLLs offer new challenges for our spin discriminations.

In outline, Section II defines our notation for the observables of the four-lepton final state. Section III gives the general gauge and Lorentz invariant couplings of an HLL to ZZ or ZZ^*, with a discussion of other symmetry properties. We describe in Section IV event generation, detector simulation, and the construction of effectively background-subtracted samples using ${}_sPlots$; here also we show the sculpting of the angular distributions and correlations by the finite phase space acceptance of the detector. In Section V we describe our statistical approach to HLL discrimination using hypothesis testing with likelihood ratios. We demonstrate in Section VI the consistency of our methods by applying them to the discrimination of signal from SM ZZ background. In Section VII we detail many examples quantifying our ability to discriminate a SM Higgs from a variety of HLLs, showing in each case the expected discrimination significance as a function of the number of signal events; we use

benchmark Higgs masses of 145, 200, and 350 GeV/c^2. We summarize, in Section VIII, our results and outlook for further improvements. Here we explicitly quantify to what extent our expected discrimination significance would be degraded by using a less complete or less rigorous analysis.

II. THE GOLDEN CHANNEL

We are interested in the case of a SM Higgs boson, or a Higgs look-alike, decaying via ZZ or ZZ^* into a four-lepton final state. We will denote the putative Higgs and its mass by H and m_H, regardless of whether it is a SM Higgs or a look-alike. This notation is also used to describe background events, where the four-lepton object is treated as a Higgs or HLL in the sense that m_H stands for $M[4\ell]$. Since the events are fully reconstructible the lab frame kinematics of the candidate H particles are known: its transverse momentum p_T, pseudorapidity η, and azimuthal angle. These three variables define the direction and boost from the lab frame to the H rest frame. All other observables can then be defined with respect to the H rest frame.

In the approximation that the final state leptons are massless, 12 observables are measured per event. Since all 12 are well-measured there is no experimental reason not to re-express these in terms of whatever combinations most naturally capture the underlying physics. Thus we choose four observables to be m_H and the three production observables just described that define the H rest frame. The remaining eight observables are taken to be the two reconstructed masses of the Z bosons together with six decay angles defined with respect to the H rest frame.

In the H rest frame the reconstructed Z bosons are back-to-back. We label these bosons as Z_1, Z_2 and take the direction of Z_2 as defining the positive z-axis. Because of Bose symmetry, the labeling is arbitrary; in the case of an $e^+e^-\mu^+\mu^-$ final state we will follow the literature [29] and choose Z_2 to be the Z boson that decayed to muons. We then adopt the additional convention that the transverse direction of the μ^- lies along the positive y-axis; thus the Z_2 decay leptons lie in the y-z plane.

With the above choices, the reconstructed Z boson masses m_1 and m_2 also define the longitudinal boosts from the H rest frame to the rest frames of the decaying Z_1 and Z_2 bosons. The boost parameters are given by

$$\gamma_1 = \frac{m_H}{2m_1}\left(1+\frac{m_1^2-m_2^2}{m_H^2}\right) , \quad (1)$$

$$\gamma_2 = \frac{m_H}{2m_2}\left(1-\frac{m_1^2-m_2^2}{m_H^2}\right) . \quad (2)$$

We let θ_1, φ_1 denote the ℓ_1^- decay angles in the Z_1 rest frame, while θ_2, φ_2 denote the ℓ_2^- decay angles in the Z_2 rest frame.

There are two additional angles Θ, Φ defining the direction of the initial state partons as reconstructed in the H rest frame. For a gluon-gluon initial state these angles measure a rotation from the z-axis defined above to the direction of the initial state gluon with positive z-component of momentum. For quark-antiquark ($q\bar{q}$) initiated production of an HLL we have the problem that we do not know event-by-event which proton contributed the antiquark; this is resolved by symmetrizing the expected angular distributions under the replacement $\cos\Theta \to -\cos\Theta$.

As expected, one combination of the three azimuthal angles Φ, φ_1 and φ_2 is physically redundant. We take advantage of this fact to make the replacements $\varphi_1 \to \Phi+\phi$, $\varphi_2 \to \Phi$. Thus ϕ then represents the azimuthal rotation between the Z_2 and Z_1 decay planes.

In summary, the 4-momenta of the process $gg \to H \to Z_1Z_2 \to \ell_1^-\ell_1^+\ell_2^-\ell_2^+$ are explicitly parametrized in the H rest frame as

$$\begin{aligned}
p_{g_2} &= \frac{m_H}{2}\,(\,1, S\cos\Phi, S\sin\Phi,\; C) \\
p_{g_1} &= \frac{m_H}{2}\,(\,1, -S\cos\Phi, -S\sin\Phi,\; -C) \\
k &= m_H\,(\,1,\;0,\;0,\;0)\,, \\
p_2 &= m_2\,(\gamma_2,\;0,\;0, \beta_2\gamma_2)\,, \\
p_1 &= m_1\,(\gamma_1,\;0,\;0, -\beta_1\gamma_1)\,, \\
p_{\ell_2^-} &= \frac{m_2}{2}\,(\gamma_2(1+\beta_2 c_2),\;0,\;s_2, \gamma_2(\beta_2+c_2))\,, \\
p_{\ell_2^+} &= \frac{m_2}{2}\,(\gamma_2(1-\beta_2 c_2),\;0,\;-s_2, \gamma_2(\beta_2-c_2))\,, \\
p_{\ell_1^-} &= \frac{m_1}{2}\,(\gamma_1(1+\beta_1 c_1),\;-s\;s_1,\;-c\;s_1, -\gamma_1(\beta_1+c_1))\,, \\
p_{\ell_1^+} &= \frac{m_1}{2}\,(\gamma_1(1-\beta_1 c_1),\;s\;s_1,\;c\;s_1, -\gamma_1(\beta_1-c_1))\,.
\end{aligned} \quad (3)$$

Here k denotes the 4-momentum of H, while p_1,p_2 are the 4-momenta of Z_1,Z_2. We have used the condensed notation $C, S{=}\cos\Theta,\sin\Theta$, $c, s{=}\cos\phi,\sin\phi$, $c_1, s_1{=}\cos\theta_1,\sin\theta_1$, and $c_2, s_2{=}\cos\theta_2,\sin\theta_2$.

Of the five relevant angles, Θ and Φ are Z-pair production angles, while the remaining three are 4ℓ production angles. We will use the notation

$$\vec{\omega} = \{\phi,\, \cos\theta_1,\, \cos\theta_2\}\,, \quad (4)$$

$$\vec{\Omega} = \{\Phi,\, \cos\Theta\}\,. \quad (5)$$

For a SM Higgs, the distributions in Θ and Φ are flat if we ignore the phase space acceptance effects. In previous studies these two angles have been either ignored or integrated over.

Although we have tried to conform to the literature in our parametrization of the decay angles, we note that the literature itself is divided over the choice of which decay plane orientation corresponds to $\phi = 0$ rather than $\phi = \pi$. We conform to the convention of Buszello et al [29], which is opposite to that of Djouadi [4] and Bredenstein et al [35].

The decay amplitudes defined in the next section depend on two combinations of the boost parameters γ_1

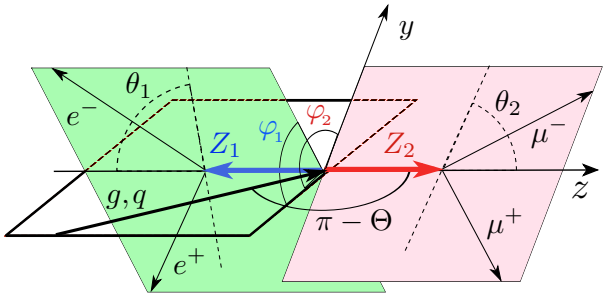

FIG. 1: The Cabibbo-Maksymowicz angles [36] in the $H \to ZZ$ decays.

and γ_2, defined by

$$\gamma_a = \gamma_1\gamma_2(1+\beta_1\beta_2)\,, \tag{6}$$
$$\gamma_b = \gamma_1\gamma_2(\beta_1+\beta_2)\,, \tag{7}$$

which are in fact just the cosh and sinh of the rapidity difference of Z_2 and Z_1, such that

$$\gamma_a^2 - \gamma_b^2 = 1\,. \tag{8}$$

More explicitly, we have

$$\gamma_a = \frac{1}{2m_1m_2}\left(m_H^2 - (m_1^2+m_2^2)\right)\,. \tag{9}$$

III. COUPLINGS AND ANGULAR DISTRIBUTIONS

A. General couplings to ZZ*

The vertex Feynman rules for the most general coupling of a spinless particle to the polarization vectors ϵ_1^μ and ϵ_2^α of two Zs of four-momenta p_1 and p_2 are given by the expression:

$$L_{\mu\alpha} = X\, g_{\mu\alpha} - (Y + i\,Z)\,\frac{k_\alpha k_\mu}{M_Z^2} + (P + i\,Q)\,\epsilon_{\mu\alpha}\frac{p_1p_2}{M_Z^2}\,, \tag{10}$$

where we have suppressed repeated indices in the contraction of the four-index ϵ tensor, $k = p_1 + p_2$ and only Lorentz-invariance has been assumed. The dimensionless form factors X to Q are functions of k^2 and $p_1\cdot p_2$ which, with no loss of generality, can be taken to be real (but for their absorptive parts, expected to be perturbatively small). The rescalings by $1/M_Z^2$ are just for definiteness, since the true mass scale of the underlying operators is as yet unspecified. In practice we also remove an overall factor of $igM_Z/\cos\theta_W$, so that $X = 1$ corresponds to the tree level coupling of a SM Higgs boson.

Similarly, the most general vertex describing the coupling of a spin $J = 1$ particle to two Z-polarizations (indices μ and α, momenta p_1 and p_2, respectively) and to its own polarization (index ρ) is:

$$L^{\rho\mu\alpha} = X\,(g^{\rho\mu}\,p_1^\alpha + g^{\rho\alpha}\,p_2^\mu) + (P + i\,Q)\,\epsilon^{\rho\mu\alpha}(p_1 - p_2), \tag{11}$$

again with X, P and Q real.

The most general parity-conserving vertex describing the coupling of a $J = 2^+$ particle of polarization tensor $\epsilon^{\rho\sigma}$ to our two vector bosons is:

$$\begin{aligned} L^{\rho\sigma\mu\alpha} = \quad & X_0\, m_H^2\, g^{\mu\rho}\, g^{\alpha\sigma} \\ & +(X_1 + i\,Y_1)\,(p_1^\alpha\, p_2^\rho\, g^{\sigma\mu} + p_1^\rho\, p_2^\mu\, g^{\sigma\alpha}) \\ & +(X_2 + i\,Y_2)\, p_1^\rho\, p_2^\sigma\, g^{\mu\alpha}, \end{aligned} \tag{12}$$

again with all coefficients real. The special case of tree level gravity-like couplings corresponds to

$$X_0 = -\frac{1}{2}\kappa\,, \quad X_1 = \kappa\,, \quad X_2 = -\kappa\,, \tag{13}$$

with all other coefficients vanishing and κ an overall coupling strength.

These general couplings, with naive mass dimensions $d = 3$, 4, and 5, can arise from $SU(2)_L \times U(1)_Y$ invariant operators of dimension 5, 6, or higher. Since, for HLLs with non-vanishing weak charges, this parentage introduces model dependence, we relegate it to a brief discussion in Appendix A.

B. 'Pure' cases of specified J^{PC}

We specify in this section the results for four cases (scalar, pseudoscalar, vector and axial vector) that would be 'pure' in the sense of having a single dominant term in their HZZ couplings, which we use to define their spin and parity. This allows one to illustrate the mass and angular dependences of the predictions, setting the stage for the later discussion of the impure cases for which P and/or CP are not symmetries of the theory, and to establish comparisons with the existing literature (but for the ZZ^* case for $J = 1$, which we have not found elsewhere).

The general expressions for the angular correlations in the ZZ^* case (which includes ZZ when the two Z masses are fixed at M_Z) are given in an appendix. The notation is as defined in the previous section; we also need to introduce the quantity

$$\eta \equiv \frac{2\,c_v\,v_a}{(c_v^2 + c_a^2)} \approx 0.15, \tag{14}$$

arising from the SM couplings of the Z bosons to the final state leptons.

1. *The standard Higgs, $J^{PC} = 0^{++}$*

The tree level SM coupling of the Higgs to two Z's of polarisation ϵ_1 and ϵ_2 is $\propto \epsilon_1\epsilon_2$, see Eq. (10). The angular

distribution of the leptons in $H \to ZZ \to 4\,l$ decay, for on or off-shell Z's of mass m_1 and m_2, is:

$$\frac{d\Gamma[0^+]}{dc_1\,dc_2\,d\phi} \propto m_1^2\,m_2^2\,m_H^4\big[1 + c_1^2c_2^2 + (\gamma_b^2 + c^2)s_1^2s_2^2 + 2\gamma_a\,c\,s_1s_2\,c_1c_2 + 2\eta^2(c_1c_2 + \gamma_a\,c\,s_1s_2)\big]\,. \tag{15}$$

2. A pure pseudoscalar, $J^{PC} = 0^{-+}$

The coupling of a $J^{PC} = 0^{-+}$ pseudoscalar to two Z's of polarisation ϵ_1 and ϵ_2 and four-momenta p_1 and p_2 is $\propto \epsilon[\epsilon_1, \epsilon_2, p_1, p_2]$, see Eq. (10). The angular distribution of the leptons in its $ZZ \to 4\,l$ decay is:

$$\frac{d\Gamma[0^-]}{dc_1\,dc_2\,d\phi} \propto m_1^4\,m_2^4\,\gamma_b^2 \big(1 + c_1^2c_2^2 - c^2s_1^2s_2^2 + 2\,\eta^2\,c_1c_2\big)\,. \tag{16}$$

3. A pure vector, $J^{PC} = 1^{--}$

The coupling of a $J^{PC} = 1^{--}$ vector particle of polarization ϵ_H to two Z's of polarisation ϵ_1 and ϵ_2 and four-momenta p_1 and p_2 is $\propto \epsilon_H{\cdot}\epsilon_1\;\epsilon_2{\cdot}p_1 + \epsilon_H{\cdot}\epsilon_2\;\epsilon_1{\cdot}p_2$, see Eq. (11). Unlike for the scalar cases, the fully differential decay amplitude depends nontrivially on the angles Θ and Φ, representing correlations between the helicities of the initial and final state particles. Assuming a quark-antiquark initial state this, in principle, introduces two new parameters: the vector and axial couplings of the (massless) quarks to the spin 1 HLL. However, once we symmetrize over $\cos\Theta \leftrightarrow -\cos\Theta$, reflecting our ignorance of which colliding proton contributes the antiquark of the hard scattering, the dependence on these new couplings disappears except for an overall factor. After this symmetrization, the angular distribution of the leptons in $H \to ZZ^* \to 4\,l$ decay is:

$$\begin{aligned}\frac{d\Gamma[1^-]}{dC\,dc_1\,dc_2\,d\Phi\,d\phi} \propto\; & 4m_1^2m_2^2\gamma_b^2\Big[S^2s_1^2s_2^2\left(2m_d^4 - m_H^2\big[m_1^2\cos(2(\Phi+\phi)) + m_2^2\cos(2\Phi)\big]\right) \\ & + m_H^2(1+C^2)\big[2m_2^2s_1^2 + 2m_1^2s_2^2 - (m_1^2+m_2^2)s_1^2s_2^2\big] + 4m_Hm_d^2\,C\,S\left[m_1c_1\,s_1s_2^2\sin(\Phi+\phi) - m_2c_2\,s_2s_1^2\sin\Phi\right] \\ & - 2m_H^2m_1m_2s_1s_2\left((1+C^2)(c_1c_2-\eta^2)c + S^2(c_1c_2+\eta^2)\cos(2\Phi+\phi)\right)\Big]\,,\end{aligned} \tag{17}$$

where we have introduced the notation

$$m_d^2 \equiv m_1^2 - m_2^2\,. \tag{18}$$

4. A pure axial vector, $J^{PC} = 1^{++}$

The coupling of a $J^{PC} = 1^{++}$ axial-vector particle of polarization ϵ_H to two Z's of polarisation ϵ_1 and ϵ_2 and four-momenta p_1 and p_2 is $\propto \epsilon[\epsilon_H, \epsilon_1, \epsilon_2, p_1 - p_2]$, see Eq. (11). After the same symmetrization in $\cos\Theta$ described above, and introducing the notation

$$\begin{aligned} M_1^2 &\equiv m_H^2 - 3m_1^2 - m_2^2\,, \\ M_2^2 &\equiv m_H^2 - m_1^2 - 3m_2^2\,, \end{aligned} \tag{19}$$

the angular distribution of the final state leptons is given by:

$$\begin{aligned}\frac{d\Gamma[1^+]}{dC\,dc_1\,dc_2\,d\Phi\,d\phi} \propto\; & m_H^2S^2s_1^2s_2^2\left[M_2^4m_1^2\cos(2(\Phi+\phi)) + M_1^4m_2^2\cos(2\Phi)\right] + 8m_1^2m_2^2m_d^4S^2\left[\,c_1^2 + c_2^2 + s_1^2s_2^2s^2 + 2\eta^2c_1c_2\right] \\ & + m_H^2(1+C^2)\big[2M_1^4m_2^2s_1^2 + 2M_2^4m_1^2s_2^2 - (M_2^4m_1^2 + M_1^4m_2^2)s_1^2s_2^2\big] \\ & - 8m_Hm_d^2m_1m_2C\,S\left[M_2^2m_1s_2\left(c_2s_1^2c\sin(\Phi+\phi) + c_1(c_1c_2+\eta^2)\sin\Phi\right)\right. \\ & \qquad \left. - M_1^2m_2s_1\left(c_1s_2^2c\sin\Phi + c_2(c_1c_2+\eta^2)\sin(\Phi+\phi)\right)\right] \\ & + 2m_H^2M_1^2M_2^2m_1m_2s_1s_2\big[(1+C^2)(c_1c_2-\eta^2)c - S^2(c_1c_2+\eta^2)\cos(2\Phi+\phi)\big]\,,\end{aligned} \tag{20}$$

5. A pure massive graviton, $J^{PC} = 2^{++}$

Since the general analysis of spin 2 coupling to off-shell Z's is quite cumbersome, we will only quote results for a simpler example: a positive parity spin 2 with gravity-like couplings produced by gluon fusion and decaying to

two on-shell Z's. Defining the on-shell ratio

$$x \equiv \frac{m_H}{M_Z} \,, \tag{21}$$

and using the massive graviton formalism of [37], we obtain the tree level angular distribution of the final state leptons as:

$$\begin{aligned}\frac{d\Gamma[gg \to \text{graviton} \to ZZ]}{dC\, dc_1\, dc_2\, d\Phi\, d\phi} \propto\; & 16x^4C^2 + 2(x^4+16)S^4 + s_1^2 s_2^2[(x^4+16)S^4 - 4x^2(x^2+4)S^2 + 4x^4] \\ & +8x^2S^2\Big[[2+S^2+(2-3S^2)c_2^2]s_1^2\cos(\Phi+\phi)^2 + [2+S^2+(2-3S^2)c_1^2]s_2^2\cos^2\Phi\Big] \\ & +S^4 s_1^2 s_2^2[x^4\cos(2\Phi+\phi)^2 + 16\,c^2] - (s_1^2+s_2^2)[(x^2+4)^2C^4 + 2(3x^4-16)C^2 + (x^2-4)^2] \\ & +2S^2 c_1\, c_2\, s_1\, s_2\left[x^2\,[2(x^2+4) - (x^2+12)S^2]\cos(2\Phi+\phi) + 4\,[4x^2 - (3x^2+4)S^2]c\right] . \end{aligned} \tag{22}$$

Note the $\cos^4\Theta$ dependence characteristic of a spin 2 resonance.

C. Tests of symmetries

Now we discuss the behaviour of the HZZ couplings under various symmetries, including CP and Bose-Einstein statistics. The discussion attempts to clarify the literature on these issues.

Consider the $J=0$ case. The most general coupling of a spinless particle to the polarization vectors ϵ_1 and ϵ_2 of two Z's is that of Eq. (10). In computing the ensuing $H \to ZZ^* \to 4\ell$ process one finds that the XP interference term is of the form:

$$\begin{aligned}&\frac{d\Gamma[0,\text{Todd}]}{dc_1\, dc_2\, d\phi} \propto \\ &2\, m_1^3\, m_2^3\, m_H^2\, \gamma_b\, s_1\, s_2\, s\left[s_1\, s_2\, c + \gamma_a\,(c_1 c_2 + \eta^2)\right] ,\end{aligned} \tag{23}$$

where the term $\sin\theta_1 \sin\theta_2 \sin\phi \propto \vec{p}_{e^+} \cdot \vec{p}_{\mu^-} \times \vec{p}_{\mu^+}$. By definition, this observable is $\tilde{T}$-odd: it changes sign as all three-momentae are reversed (the tilde in "$\tilde{T}$-odd" emphasizes that past and future are not being interchanged).

The Born approximation is, by definition, the result of squaring the amplitude dictated by the Lagrangian to lowest order in its couplings: a quadratic result, in our case, in any pair of the quantities X to Q in Eq. (10). To this order, a $\tilde{T}$-odd observable must vanish if CP is a symmetry, as shown in [38]. Thus, a non-vanishing $\tilde{T}$-odd observable such as that of Eq. (23) can only arise if CP-invariance is also violated.

The XQ interference term resulting from Eq. (10) is:

$$\begin{aligned}&\frac{d\Gamma[0,\text{Codd}]}{dc_1\, dc_2\, d\phi} \propto \\ &-2\,\eta\, m_1^3\, m_2^3\, m_H^2\, \gamma_b\,[c_1 + c_2]\,(1 + c_1 c_2 + \gamma_a\, s_1 s_2\, c) .\end{aligned} \tag{24}$$

This term is CP odd and $\tilde{T}$-even, a combination not addressed by the theorem quoted above. It is a C-odd observable, in that it changes sign under the interchange of $p_{e^+} \leftrightarrow p_{e^-}$ and $p_{\mu^+} \leftrightarrow p_{\mu^-}$, tantamount to $\cos\theta_i \leftrightarrow -\cos\theta_i$ in our chosen notation.

1. Bose-Einstein statistics

The general coupling, up to two derivatives, of a $J=1$ particle to two Z's is that of Eq. (11). This is true whether or not the Z's are on-shell, which seems to be a point of confusion in the literature. Thus for example [23] contains extra "off-shell" couplings, such as $g^{\mu\alpha}(p_1 - p_2)^\rho$ and $\epsilon^{\rho\mu\alpha}(p_1+p_2)$, that violate Bose symmetry and vanish for two on-shell Z's. However, Bose symmetry is a property manifest at the Lagrangian level, and thus independently of any on- or off-shell considerations. The two Z's in an $H \to ZZ^*$ decay are described by the same bosonic Z field, whether or not they are on-shell, and they do not obey the laxer rules that different particles ($Z \neq Z'$) would.

D. Tests of compositeness.

If the couplings of a HLL conserve P and CP, but the object is not point-like, there will be deviations from the standard $g_{\mu\nu}$ coupling to Z's. To lowest order in the dimensions of the corresponding effective operators, these will be of two types. The first is a non-vanishing Y in Eq. (10), and the second is a nontrivial form for X. Barring large effects –quite conceivable in a model with multiple SM Higgs-like fields– deviations in X are much harder to limit or measure than a non-zero Y/X which is governed by the shapes of angular distributions.

It is useful to introduce the notation $\tan\xi \equiv Y/X$. In this notation, the "composite" HLL angular distribution is of the form:

$$d\Gamma_C = \cos^2\xi d\Gamma_{XX} + \cos\xi\sin\xi\, d\Gamma_{XY} + \sin^2\xi\, d\Gamma_{YY} \,, \tag{25}$$

where $d\Gamma_{XX}$ is the standard result of Eq. (15). The in-

terference term is:

$$\frac{d\Gamma_{XY}}{dc_1\,dc_2\,d\phi} \propto -2\,m_1^3\,m_2^3\,m_H^2\,\gamma_b^2\,s_1\,s_2\,(c_1c_2c + \gamma_a\,s_1s_2 + \eta^2\,c)\,, \quad (26)$$

and the last term is:

$$\frac{d\Gamma_{YY}}{dc_1\,dc_2\,d\phi} \propto m_1^4\,m_2^4\,\gamma_b^4\,s_1^2s_2^2\,. \quad (27)$$

Contrary to all of the other cases we study, the interference term in this instance is between two operators whose P and C are identical: the HLL is not point-like, but it is 'pure' 0^{++}. As a consequence, the angular distribution of the interference term is not very different from that of the XX and YY terms and the interference can, for certain values of Y/X, be very destructive. This can be seen even at the level of the $H \to ZZ$ branching fraction, the integral of Eq. (25) over $\cos\theta_1$, $\cos\theta_2$, and ϕ:

$$\Gamma_C \propto m_1^2m_2^2\,[2\cos^2\xi + (\gamma_a\cos\xi - m_1m_2\gamma_b^2\sin\xi)^2]\,. \quad (28)$$

If ξ has a value close to the (mass-dependent) point of maximal interference, the golden mode channel can be suppressed by a large factor. For this to happen X and Y ought to be of the same order of magnitude, signifying a low dynamical scale for a composite Higgs.

IV. ANALYSIS

In this section we describe the modeling of the detector effects and the analysis strategy to extract an effectively pure sample of signal events. We describe the Monte Carlo (MC) event generation and the simulation of the detector response. We use parameterized reconstruction resolutions and efficiencies based on the published CMS performance results [39]. A similar study can be performed with parameterizations based on the ATLAS detector. We focus on the four-muon (4μ) final state, but the results can be generalized to include final states with electrons. Since a four-lepton final state is relatively "clean" in the LHC environment, we apply a loose event selection and use a maximum likelihood (ML) fit technique to separate the signal from the background. This maximizes the statistical power and the possibility of characterizing the nature of the discovered particle through the study of the multi-dimensional angular distribution of the four leptons in the resonance rest frame.

A. Event generation

The knowledge of the four-momenta of the leptons fully specifies the information needed in this analysis. We generate the four-momenta of the four leptons from the five-dimensional (six-dimensional) probability density functions (*pdfs*) of

$$\vec{X} = \{\Phi, \cos\Theta, \cos\theta_1, \cos\theta_2, \phi\} \text{ for } ZZ\,, \quad (29)$$
$$\vec{X} = \{\Phi, \cos\Theta, \cos\theta_1, \cos\theta_2, \phi, M_{Z^*}\} \text{ for } ZZ^*\,.$$

The $\vec{X}$ quantities are generated in the rest frame of the decaying resonance. The muons are then boosted to the laboratory frame and the detector effects (acceptance, efficiency and resolution) are applied to the boosted momenta. We use the azimuthal symmetry of the LHC detectors to reduce the remaining kinematic degrees of freedom to the knowledge of the p_T, η and the invariant mass $m_{4\mu}$ of the 4μ system. The p_T, η for the signal is taken from a two-dimensional *pdf* generated using `MC@NLO` [40]. We consider proton-proton collisions at $\sqrt{s} = 10$ TeV and we model the parton probability density functions (PDFs) using CTEQ5L [41].

In this analysis we don't assume a specific signal production mechanism and cross section; Rather we rely exclusively on the discrimination provided by the angular distributions of the leptons in the final state. We show the p_T vs. η *pdfs* for a spin-0 and a spin-1 neutral HLL in Fig. 2. As discussed in Section I, for all the signal generation we use the p_T vs. η *pdfs* of the scalar. For the

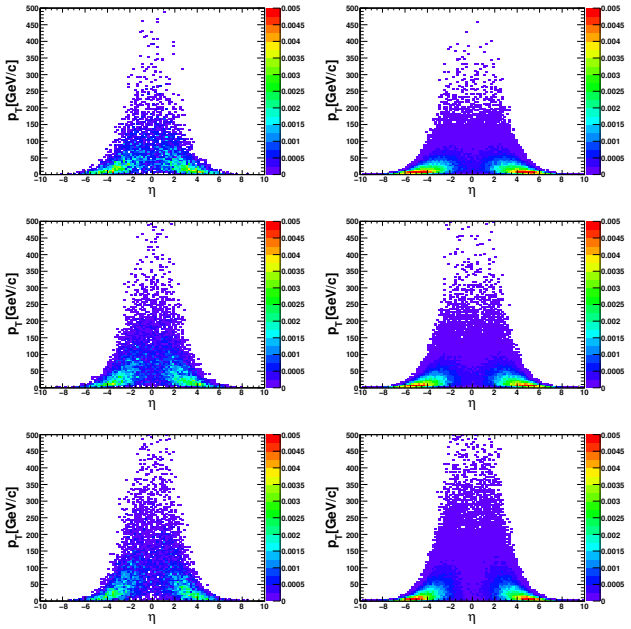

FIG. 2: 2D p_T-η *pdf* of a 0^+ HLL resonance (left) and a 1^- one (right) for $\sqrt{s} = 10$ TeV collisions, obtained using `PYTHIA` and the CTEQL5 parton density functions and for m_H =145, 200, 350 GeV/c^2 (top, middle and bottom).

SM ZZ background the p_T, η and $m_{4\mu}$ are taken from a three-dimensional *pdf* generated using the `PYTHIA` [42] leading-order MC generator. The momenta of the four muons in the rest frame of the $ZZ^{(*)}$ system as a function of $m_{4\mu}$ are generated according to the theoretical distributions.

B. Detector emulation and event selection

Muon reconstruction efficiency and resolution are parameterized as a function of the muon p_T and η according to [39], where the muon reconstruction efficiency is close to 100% for muons with $p_T \geq 10$ GeV/c and $|\eta| \leq 2.3$, corresponding to the event selection in this analysis. The reconstruction efficiency is applied through a hit-or-miss technique. For muon candidates accepted by the efficiency filter, the reconstructed momentum is determined by applying Gaussian smearing functions to the true p_T, η and ϕ with p_T- and η-dependent resolutions. We verified the goodness of our *very-fast* muon simulation by comparing the parameters of the fit of the Z invariant mass distribution obtained in our analysis and shown in Fig. 3 with the corresponding ones from a published full-simulation analysis [43].

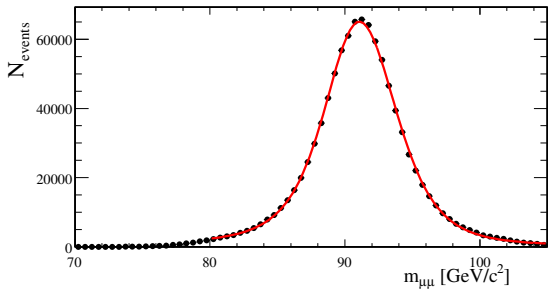

FIG. 3: Distribution of the dimuon invariant mass for a sample of signal $H \to ZZ$ events, generated using our very-fast muon simulation. The parameters of the superimposed fit are extracted from [43].

A number of detector related effects can modify the $\vec{X}$ observables' *pdfs*. The resolution of the observables used in the analysis is shown in Fig. 4 and is found to be small independent of the HLL resonance mass and quantum numbers. The systematic bias in the reconstruction of the same variables is shown in Fig. 5 and is found to be also negligible. This suggests that the sculpting of the observables' *pdfs* is not a result of reconstruction resolution or bias. Rather, it depends on the simulated kinematics of the HLL resonance, including its mass, and on the specific model considered (0^+, 0^- etc). Specifically, the overall phase space acceptance, implemented in the signal selection by means of the p_T and η requirements produces the largest effects on the observables. This is shown in Fig. 6 for a resonance of mass 145 GeV/c^2 generated with no explicit angular correlations. Adding the angular correlations in the matrix element can enhance or reduce the overall selection efficiency depending on the details of the multidimensional *pdf*. Our selection is 60% (74%) efficient for a 0^+ resonance of mass 200 GeV/c^2 (350 GeV/c^2) as shown in Fig. 7. The same figure demonstrates that the efficiency has a non-trivial dependence on the nature of the spin correlations. Specifically, for a 0^- resonance of 200 GeV/c^2 (350 GeV/c^2) the

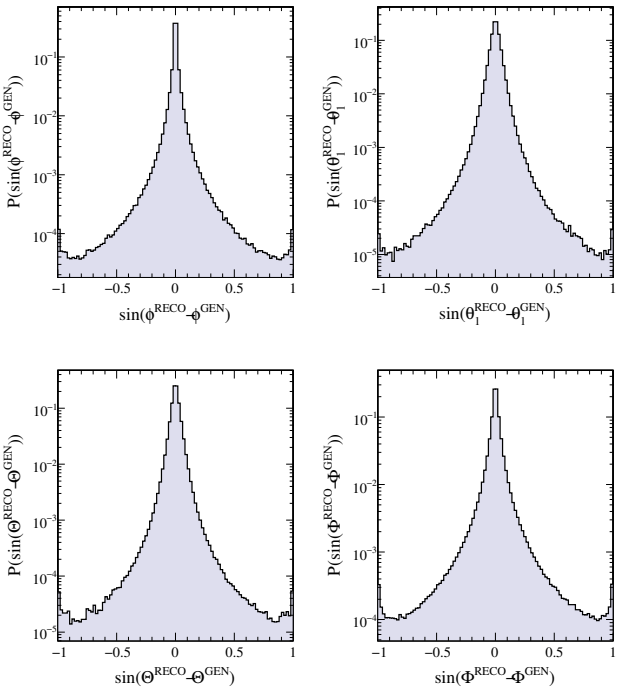

FIG. 4: Reconstruction resolution for the angular variables of $\vec{X}$ shown here for a resonance with mass 145 GeV/c^2. The $\cos\theta_2$ and $\cos\theta_1$ distributions are very similar in this case. Only events surviving the signal selection are included. All distributions are normalized to unit integral.

efficiency is 60% (69%). With an absence of explicit spin correlations the efficiency for a 350 GeV/c^2 resonance is 71%.

We find that changes in the $\vec{X}$ distributions are strongly correlated with the kinematics of the off-shell Z, e.g for $\cos\theta_2$, the largest inefficiencies corresponding to the kinematic configurations where at least one of the muons is soft. While the angles $\vec{\Omega}$ can be neglected for $J = 0$ resonances we find that this is not the case when correlations between the variables $\vec{\omega}$ and $\vec{\Omega}$ appear explicitly in the differential cross-sections. For example, for $J = 1^-$, the ϕ and Φ theoretical distributions are highly correlated. The consequences on model discrimination are discussed in Sec. VII B

Since the shapes of the reconstructed $\vec{\omega}$ and $\vec{\Omega}$ distributions depend on the phase space acceptance, we find that the distributions are nearly identical when applying the same p_T and η acceptance requirements to electrons in $H \to ZZ \to 2e2\mu$ or $4e$ final states with the p_T cuts considered in this study. This observation is illustrated in Fig. 8, demonstrating the excellent agreement between kinematic distributions reconstructed in muon and electron final states. As a consequence, results concerning model discrimination, as a function of the number of observed signal events, will be nearly identical when the additional final states are included ($2e2\mu$, $4e$), especially

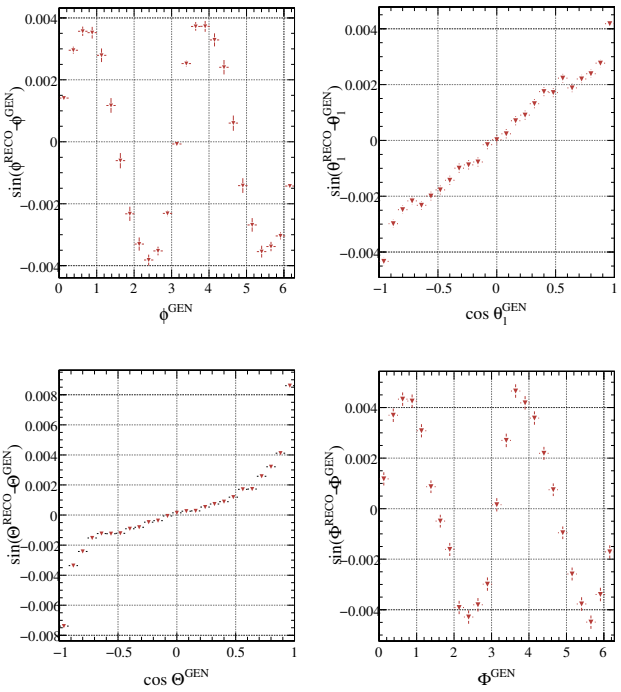

FIG. 5: Distributions showing systematic biases for a subset of the reconstructed variables $\vec{X}$ for a resonance with mass 145 GeV/c^2. Only events that survive the signal selection are included. All biases are negligible.

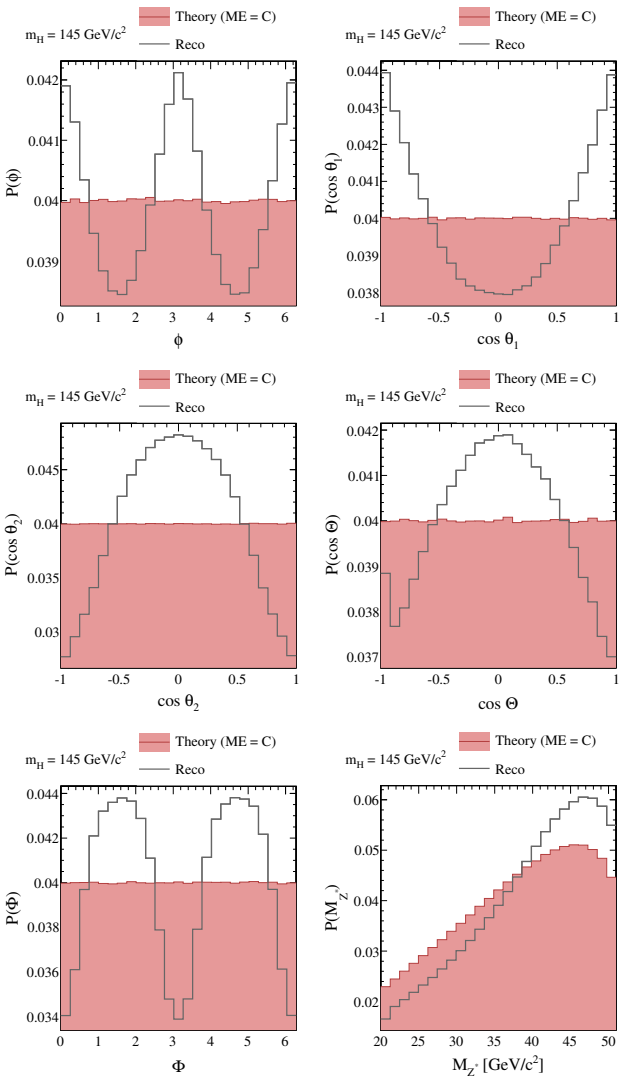

FIG. 6: The variables $\vec{X}$ used in this analysis for a 145 GeV/c^2 resonance. The off-shell M_{Z^*} is required to lie a window between 20 and 50 GeV/c^2. The shaded histograms are the 1D distributions using a constant matrix element (*i.e.* no angular correlations included). The overlaid histograms show the same distributions for reconstructed events passing the p_T and η signal selection after the detector parameterization. All distributions are normalized to unit integral.

when the off-shell Z mass is not used as an observable. This is not necessarily the case for results concerning the discovery of a resonance in these final states with respect to the background-only hypothesis, since different backgrounds need to be considered for electron and muon final states.

C. Fit definition and signal extraction

The $H \to ZZ$ signal events can be discriminated from SM backgrounds using an extended and unbinned ML fit. Since there is no resonant 4μ background in the SM, the fit can use as a discriminating variable the 4μ mass distribution. In the presence of a sizable background due to fake Z candidates (such as top decays) the 2μ mass distributions can be included in the likelihood. Since this is not a conceptually different situation, we ignore this possibility and assume for simplicity that the only relevant background is given by events with two real Z candidates. We write the likelihood function as:

$$\mathcal{L} = \frac{1}{N!}\exp\left(-\sum_j N_j\right) \prod_{i=1}^{N}\left(N_S P_S[m^i_{4\mu}] + N_B P_B[m^i_{4\mu}]\right), \quad (30)$$

where N_j $(j = S, B)$ represents the yield of each component; $m^i_{4\mu}$ is the 4μ candidate mass for the event i; and $P_S[m]$ ($P_B[m]$) is the signal (background) distribution for the variable m. The *pdfs* for the signal and background components are described using the template distributions from the simulation, as shown in Fig. 9 for $m_H = 250$ GeV/c^2. This fit configuration is appropriate for the HLL characterization.

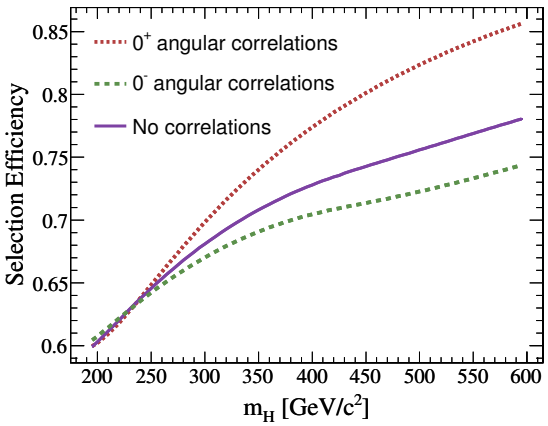

FIG. 7: The analysis efficiency for 0^+, 0^- and no correlations as a function of the resonance mass.

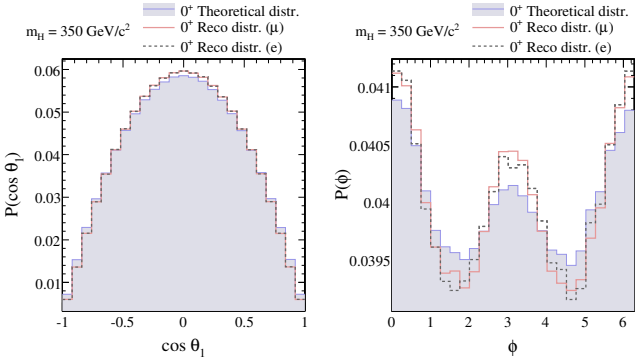

FIG. 8: Kinematic distributions for the variables $\cos\theta_1$ (left) and ϕ (right) for a 0^+ resonance with mass 350 GeV/c^2. The shaded histograms show the 1D projections of the variables as described by the analytic *pdfs*. The overlaid histograms (blue, red) show the same 1D projections for reconstructed events passing the p_T and η signal selection after the detector parameterization for 4μ and $4e$ final states. All distributions are normalized to unit integral.

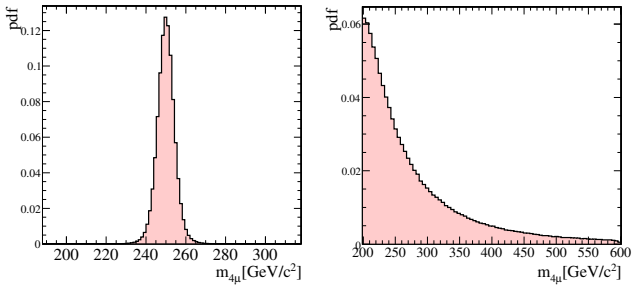

FIG. 9: Distribution of the 4μ invariant mass for a sample of signal with $m_H = 250$ (left), and background (right) ZZ events.

D. Background subtraction

In order to establish if a newly-discovered resonance is indeed the Higgs boson or not, a hypothesis test is performed (see Sec. VII). In this context, a tool to disentangle signal and background events from the selected dataset is an important prerequisite. We use the ${}_sWeight$ [13] technique and reweight the selected dataset according to *how likely* each event is considered to be signal by the fit. The ${}_sWeight$ technique is statistically optimal when the discriminating variable ($m_{4\mu}$ in our case) in the fit is uncorrelated with the subsequently used variables ($\vec{X}$ in our case). The narrowness of the $m_{4\mu}$ peak in this example guarantees that the signal events are, to an excellent approximation, uncorrelated with the angular variables. On the upper plot of Fig. 10, the 4μ invariant mass distribution is shown for a sample of $N_S = 70$

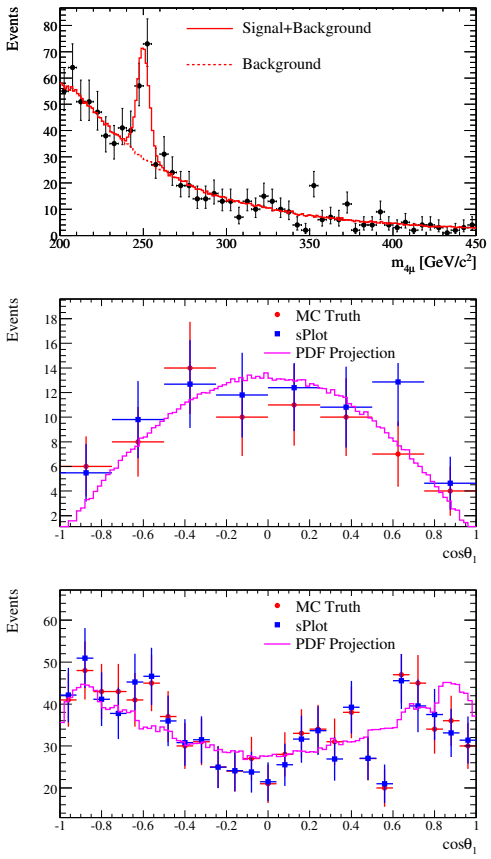

FIG. 10: The 4μ invariant mass distribution for a sample of $N_S = 70$ $H \to ZZ$ events with $m_H = 250$ GeV/c^2 and $N_B = 1000$ ZZ background events. The superimposed curves represent the likelihood function returned by an ML fit, with N_S, N_B, and $m_{4\mu}$ as free parameters (top). Comparison of the signal-only MC distribution of $\cos\theta_1$, with the background-subtracted distribution obtained with the ${}_sWeight$ technique (middle). Comparison of the background-only MC distribution of $\cos\theta_1$, with the signal-subtracted distribution obtained with the ${}_sWeight$ technique (bottom).

$H \to ZZ$ events (with $m_H = 250$ GeV/c^2) on top of $N_B = 1000$ continuum ZZ background events, corresponding to a $\sim 5\sigma$ deviation from the background-only hypothesis. The superimposed curves represent the likelihood function returned by an ML fit (with N_S, N_B, and m as free parameters). The middle plot shows the signal $sWeighted$ $\cos\theta_1$ distribution. Similarly, the bottom plot shows the background $sWeighted$ $\cos\theta_1$ distribution. The comparison of the two sets of points shows how the background (signal) subtraction allows one to recover the signal (background) distribution for the considered variable in the given sample, the deviation from the expected *pdf* being due to statistical fluctuations already present at the MC level.

V. STATISTICAL APPROACH

In this section we discuss the statistical formulation we use to address comparisons between different hypotheses as well as relevant measurements on the search and discovery and characterization of an HLL resonance. We focus on four statistical approaches:

- (1) Search analysis of a signal in the presence of backgrounds.
- (2a) Comparisons between two "pure" spin-parity hypotheses (such as 0^+ vs. 1^-).
- (2b) Comparisons between two spin-parity hypotheses, with at least one of the two being an "impure" admixture of two pure HLL states (e.g. 0^+ vs. a combination of 1^+ and 1^-). This case is similar to (2a), except for the presence of one or more nuisance parameters.
- (3) The measurement of mixing parameters in the case of impure Higgs look-alikes.

In case (1) we consider two hypotheses. $\mathbb{H}_1$ is the "standard Higgs signal plus background", and $\mathbb{H}_0$ is the null, "background only" hypothesis.

Cases (2) and (3) establish the nature of a newly discovered particle. Guided by our results on ${}_sPlots$, we contend that it is a very good approximation to confront two different "signal" hypotheses in the absence of background – which has been statistically subtracted. This assumes that the discovery of a new resonance has already been established.

The case (2) hypotheses refer to an m_H peak with two different J^P interpretations. In the (2a) case the two hypotheses under consideration are simple *i.e.* the corresponding likelihoods are fully specified once the values $\vec{X}$ are measured. In the (2b) case the unknown mixing angles, referred to as $\vec{\alpha}$ (and including e.g. the various ξ and δ), for the impure hypothesis are treated as nuisance parameters. The analysis in case (3) is a traditional parameter estimate, based on the ML fit for which we obtain a confidence interval by using the Feldman-Cousins approach [44]. We discuss the three cases starting from the last.

A. Coupling admixtures

Consider the example of a one-parameter mixture of two types of HZZ coupling, such as the composite case discussed in Sec. III D. For a fixed value of the resonance mass m_H and the mixing angle α, Eq. (25) is the theoretical probability-distribution of the events as a function of the variables $\vec{X}$ for ZZ and ZZ^* final states. The experimental *pdf* is a numerical representation of the result of sieving –through a specific detector and its resolution, trigger and analysis requirements– a very large number of events, generated with the theoretical *pdf* of Eq. (25). This experimental *pdf*, referred to as P, is a function $P = P_M(\alpha, \vec{X})$ of $M = m_H$, (which is kept fixed through this exercise), α, and $\vec{X}$. The dependence on $\vec{\Omega} \equiv \{\cos\Theta, \Phi\}$ is, in this example, exclusively a phase space acceptance effect.

Many experiments with a fixed number of events, N_S, are simulated, assuming the same detector response. The probability of each event, evaluated with the experimental *pdf*, is P_i. The likelihood of a given experiment is $\mathcal{L}(\alpha) = \prod_{i=1}^{N_S} P_i$. The experimentally measured value of the α parameter, $\hat{\alpha}$ corresponds to the value that maximizes $\mathcal{L}(\alpha)$. The simulation is repeated many times, as a function of the true value of the mixing angle α. Running many experiments one can derive the confidence interval, i.e. the range *covering* the true value of α for some confidence level and some measured value $\hat{\alpha}$ [44].

It is customary to estimate the error (or the number n of standard deviations σ) in the measured α from the expression $\mathcal{L}(\alpha_{max} \pm n\,\sigma) = \mathcal{L}(\alpha_{max}) - n^2/2$. While this method is accurate for large samples with Gaussian errors, it is not the one used to draw the σ contours in Fig. 11 (where $\alpha = \xi_{XQ}$) and in the similar figures of Sec. VII. Instead, the CL is evaluated measuring the frequency of a given result in the set of generated pseudo-experiments.

B. Confronting a J^P hypotheses

Consider two hypotheses, $\mathbb{H}_{0,1}$, for the spin-parity assignment of a signal candidate sample, detected via its ZZ mass peak and background subtracted using the ${}_sPlot$ method. Large numbers of events are generated assuming each hypothesis and used to construct two unbinned experimental *pdfs*: $P_{\mathbb{H}_{0,1}}(\hat{=}P_M(\vec{x}\,|\,\mathbb{H}_{0,1})$. For our pure spin-parity cases, the simple nature of the hypotheses considered guarantees through the Neyman-Pearson (NePe) lemma [45] that the hypothesis test is *universally most powerful*. Next, we explicitly identifying one hypothesis as $\mathbb{H}_0$ and the other as $\mathbb{H}_1$. Additionally, we specify the test *statistic* Λ which we define as the log-

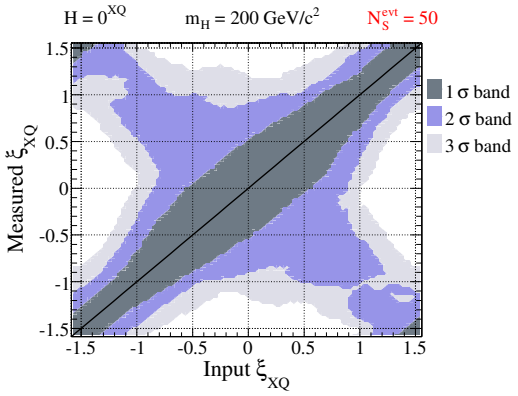

FIG. 11: Confidence intervals for measured values of ξ_{XQ} for a C-violating $J = 0$ resonance with a mass 200 GeV/c^2.

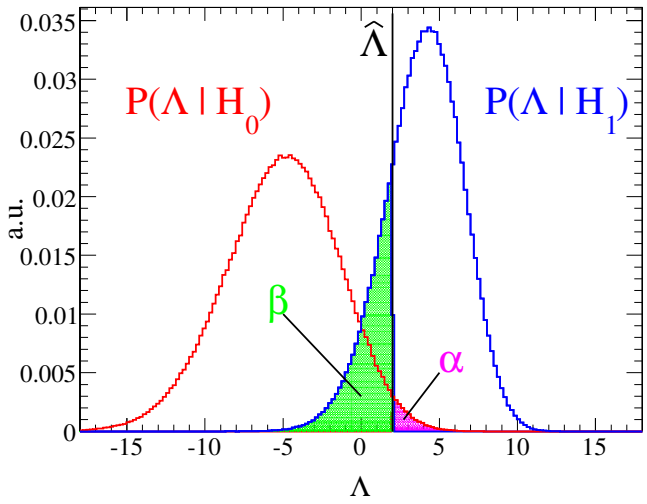

FIG. 12: Distribution of Λ for $m_H = 200$ GeV/c^2 and $N_S = 23$, constructed with $\sim 10^9$ pseudo-experiments. The hypothesis being confronted are $\mathbb{H}_0 = 0^+$ and $\mathbb{H}_1 = 0^-$.

likelihood ratio $\log[\mathcal{L}(\mathbb{H}_1)/\mathcal{L}(\mathbb{H}_0)]$. Finally, we must a priori choose the acceptable probability level α, of rejecting $\mathbb{H}_0$ in favor of $\mathbb{H}_1$, even though $\mathbb{H}_0$ is true (Type I error). We generate a series of pseudo-experiments with a fixed number of events, N_S, to construct the *pdf* of Λ for the two hypotheses. A typical result is illustrated in Fig. 12. We first generate pseudo-experiments considering $\mathbb{H}_0$ as true. For each experiment we construct two likelihoods $\mathcal{L}(\mathbb{H}_0) \equiv \prod_{i=1}^{N_S} P_{\mathbb{H}_0}(\vec{X}_i)$ for the correct interpretation of the true theory, and $\mathcal{L}(\mathbb{H}_1) \equiv \prod_{i=1}^{N_S} P_{\mathbb{H}_1}(\vec{X}_i)$ for its incorrect interpretation. With the ensemble of experiments one constructs the distribution $P(\Lambda \mid \mathbb{H}_0)$ with $\Lambda \equiv \log[\mathcal{L}(\mathbb{H}_1)/\mathcal{L}(\mathbb{H}_0)]$. The result is the leftmost (red) curve in Fig. 12. The exercise is repeated with the pseudo-experiments generated considering $\mathbb{H}_1$ as true and the result is the rightmost (blue) curve in the figure.

An a priori chosen value of α implicitly defines a value $\hat{\Lambda}(\alpha)$ via

$$\alpha = \int_{\hat{\Lambda}(\alpha)}^{\infty} P(\Lambda \,|\, \mathbb{H}_0)\, d\Lambda \ . \tag{31}$$

This fixed value $\hat{\Lambda}(\alpha)$ implies that

$$\beta(\alpha) = \int_{-\infty}^{\hat{\Lambda}(\alpha)} P(\Lambda \,|\, \mathbb{H}_1)\, d\Lambda \tag{32}$$

is the probability of accepting $\mathbb{H}_0$ even if $\mathbb{H}_1$ is correct (Type II error). The value $1 - \beta$ is called the *power of the test*. When the real experiment is performed, a specific value $\Lambda_{\rm exp}$, is obtained for $\Lambda = \mathcal{L}(\mathbb{H}_1)/\mathcal{L}(\mathbb{H}_0)$. The associated *p-value* $= \int_{\Lambda_{\rm exp}}^{\infty} P(\Lambda \,|\, \mathbb{H}_0)\, d\Lambda$, is compared to α to determine if the measurement favors one hypothesis versus the other.

Instead of the α and β values, the σ (significance) is commonly used. To convert to an equivalent number of σ's using Fig. 12 we calculate the same α-area in a Gaussian distribution centered at 0 with σ =1. The number n of α-equivalent standard deviations is obtained by inverting

$$\alpha = \frac{1}{\sqrt{2\pi}} \int_n^{\infty} dx\, e^{-x^2/2} \ . \tag{33}$$

The a priori (subjective) choice of α (and subsequently β and corresponding significances) is heavily discussed in the literature. *The Physical Review*, for example, requires a 5σ (3σ) significance to the claim of a discovery (evidence). The caveat is, of course, that when one minimizes as much as possible the probability of an error of Type I (wrongly claiming a discovery) one risks making an error of Type II (and e.g. delaying the claim of a discovery to the next luminosity upgrade).

The pure vs. impure Higgs look-alike hypothesis test has an additional complication due to the dependence of the likelihood function on the value of the $\vec{\alpha}$ in at least one of the two hypothesis. In this case, we are testing the simple (i.e. $\vec{\alpha}$ independent) hypothesis against a class of alternative hypotheses, connected by the variation of a continuous unknown parameter(s). The test is performed by comparing the simple hypothesis to the impure hypothesis with values $\vec{alpha}$ that best fit the data. The impure vs. impure Higgs look-alike test is technically identical to the pure vs. impure. Here, we try to exclude some value of the $\vec{\alpha}$ parameter for one of the two composite hypotheses in favor of the alternative impure hypothesis, where the mixing angles are treated as nuisance parameters. With fixed $\vec{\alpha}$, one impure look-alike becomes a simple hypothesis (like a pure one) tested against an impure hypothesis.

C. Higgs searches

When searching for a new particle two hypotheses are tested against each other: the background-only, $\mathbb{H}_0$, and signal plus ackground, $\mathbb{H}_1$.

Assuming that the event distributions for signal and background are fully specified (an unrealistic situation in that the value of the Higgs mass and width are not known a priori), one still has to determine the signal and background yields. Hence, the likelihood function has a parametric dependence on at least one nuisance parameter. There is no guarantee that the Neyman-Pearson construction is, in this case, the optimal hypothesis test one could perform. We are, however, not concerned with what the optimal statistical test is, but rather on the physics content of the likelihood function. Our aim is to illustrate how different analyses that fully or partially exploit a matrix-element approach compare with each other. For this purpose it is sufficient to use a consistent statistic among the various cases and discuss their relative merits. We still perform a hypothesis test based on the likelihood ratio. The dependence on the nuisance parameters is removed through a maximization (profiling) of $\mathcal{L}(\mathbb{H}_{0,1})$ relative to the nuisance parameter(s), prior to the construction of the likelihood ratio, as done for the case of impure hypothesis-testing discussed in Sec. V B.

Given a specific analysis setting (i.e. a set of variables defining the likelihood function) we evaluate its discovery power by evaluating the expected significance (the number of σ's) as a function of the signal yield and for different values of signal over background yields ratio. We define an expected value for the signal to background ratio, $\langle N_S/N_B\rangle$, between the signal events constituting the $M(ZZ)$ peak and the integral of the background distribution in the same variable in the range 190 GeV/c^2 to 600 GeV/c^2. To address the uncertainties, we compare the two hypotheses for various pre-selected values of $\langle N_S/N_B\rangle$, in a large range including and bracketing the central current expectation. The likelihood for $\mathbb{H}_0$ is then that of Eq. (31), expressed as a function of the angular variables at fixed m_H, as opposed to a function of only m_H.

When adding the $\vec{X}$ variables to the likelihood, one should consider the event-by-event dependence of their *pdf* on the value of m_H. This can done using a different $\vec{X}$ *pdf* for each bin of the template function of Fig. 9. This step is straight-forward when performing the real analysis, but CPU intensive when performing hundreds of billions of pseudo-experiments. We simplify the exercise by fixing the $\vec{X}$ *pdf* to the peak value of m_H.

In our search results we compare the significance, as given by an m_H-based peak search, with the corresponding quantity following from the whole angular-distribution analysis. In the case of a discovery test, the p-value of any toy experiment is compared to the equivalent of a $\geq 5\,\sigma$ significant p-value, in order to establish if a discovery could be claimed for that experiment. By repeating the exercise many times, we can associate a probability to the discovery potential. The $5\,\sigma$ convention fixes the value of α for the hypothesis test, as well as the value of β for a given likelihood function.

VI. SIGNAL SIGNIFICANCE USING THE ANGULAR INFORMATION

As described in the two previous sections, discrimination of Higgs look-alikes first requires an event sample constituting a putative Higgs discovery. We can be agnostic about the details of the discovery analysis, although we have made the reasonable assumption that, after all selections, the dominant residual background component is from SM ZZ production. As noted already, the search analysis could be model independent, relying only on the reconstruction of a resonant excess over non-resonant backgrounds. In this case a discovery is completely factorized from its characterization.

In order to discriminate the look-alikes based on angular variables, we require the background-subtracted event sample provided by the ${}_sPlot$ technique described in Section IV D. The ${}_sPlot$ fit provides not only this ${}_sWeighted$ dataset but also an estimate of the mass m_H of the putative Higgs, as well as the fit values of N_S and N_B, the number of signal and background events in the sample. In turn this provides an estimate of the expected signal to background ratio $< N_S/N_B >$, with an error that should be dominated by the Poissonian fluctuations. Given a hypothesis of the identity of the putative Higgs, the $< N_S/N_B >$ can be independently estimated by theory, with a nontrivial systematic error. A significant discrepancy between these two estimates could signify either a problem with the experimental analysis, theory uncertainties from parton fluxes and radiative corrections, or the presence of a Higgs imposter whose production cross section differs from that of a SM Higgs.

Despite the natural factorization between discovery, HLL discrimination based on production, and HLL discrimination based on decay, it is important to check the consistency of the entire chain of analysis. This is especially true for the small datasets considered in this work, where we demonstrate HLL discrimination with datasets not much larger than, or identical to, the original discovery sample.

A powerful check is to fix the extracted (or assumed) values of m_H and $< N_S/N_B >$, and then compare the signal significance of two nominal analyses:

- An "$m(ZZ)$ only" fit, for which the discrimination between signal and background is given only by the ZZ invariant-mass peak. This is an example of a model-independent discovery analysis (although not necessarily the actual discovery analysis used in the experiment).

- An "$m(ZZ)$+$\vec{X}$" fit, in which the *pdf* for the angular variables $\vec{X}$ is also included. Thus here we are using the angular information to improve the

discrimination of the signal from the background, rather than discriminate SM Higgs from Higgs look-alikes.

We compare the signal significance of the two analyses, corresponding to different physics content for the likelihood function. A common statistical framework is used, since we are interested to compare the physics performance rather that determining the optimal statistical approach. The overall normalization is obtained by assuming $\sqrt{s} = 10$ TeV with a corresponding SM Higgs production cross section [39].

The two fit configurations answer different questions. The "$m(ZZ)$ only" fit does not assume a priori any hypothesis on the nature of the resonance (neglecting finite width effects), whereas the "$m(ZZ)$+$\vec{X}$" fit requires as input a hypothesis for the signal *pdf*. On the other hand, it is natural to expect that the "m(ZZ)+$\vec{X}$" fit offers better signal significance, since more information is added to the fit.

A direct comparison of the two analyses in a common framework is a way to quantify the price to pay in order to run a completely model-independent search. At the same time, it is a consistency check on the HLL discrimination analysis, since the background events are themselves Higgs imposters. If, as we claim, HLL discrimination is possible with datasets not much larger than, or identical to, the original discovery sample, then we should also find that the "m(ZZ)+$\vec{X}$" fit offers comparable improvements in signal significance over the "$m(ZZ)$ only" fit, for similarly small datasets.

To make the likelihood comparison meaningful, a common fit setting is used. For the ZZ invariant mass, we consider the range $190 < m_H < 600$ GeV, whereas for the signal m_H is fixed at some nominal value. The fit configuration is specified by the nominal expected signal-over-background yield ratio $< N_S/N_B >$ and by the nominal number of signal events N_S. We consider different scenarios by fixing different values of $< N_S/N_B >$ and perform the study as a function of N_S.

For each fit configuration we perform a set of toy Monte Carlo experiments. The actual number of background events are generated according to a Poisson distribution around the nominal value, and the event-by-event values of the variables used in the fit (m_H and, if used, $\vec{X}$) are randomly generated according to the signal and background *pdfs*. The fit is then performed for each toy sample, maximizing the likelihood as a function of the signal and the background. The sets of fits provide a distribution for the expected statistical signal significance obtained in a particular experiment, and in particular a 68% probability range for the spread in this significance.

This information is summarized in Figures 13- 15. The two bands in the figures correspond to the expected spread (at 68% confidence level) for the signal significance achieved in a single experiment, as a function of the signal yield N_S, for the "$m(ZZ)$ only" fit (light band) and the "$m(ZZ)$+$\vec{X}$" fit (dark band). The horizontal lines show the 3σ and 5σ thresholds (evidence and discovery, in the usual convention). The intersection with each band provides a corresponding range for the needed signal yield, the spread being due to statistical fluctuations. For a correct interpretation of the separation between the two bands, one should consider that the statistical fluctuations in the two fits are strongly correlated since they both depend on the invariant mass observable and background fluctuations for this mass distribution will be the same for both.

Fig. 13 corresponds to the case of a $m_H = 200$ GeV/c^2 SM Higgs boson, while Fig. 14 provides similar results for a $m_H = 350$ GeV/c^2 SM Higgs boson. For each mass, different values for $< N_S/N_B >$ are considered; we show here the results for $< N_S/N_B >$=1/5, 1/10 for 200 and $< N_S/N_B >$=1/10, 1/20 for 350. We note that better discrimination between the signal and background in the higher mass case (compared to the lower mass) especially in the invariant mass observable, and despite the lower cross section, results in higher significance for the higher mass case and for the same luminosity. Similarly,

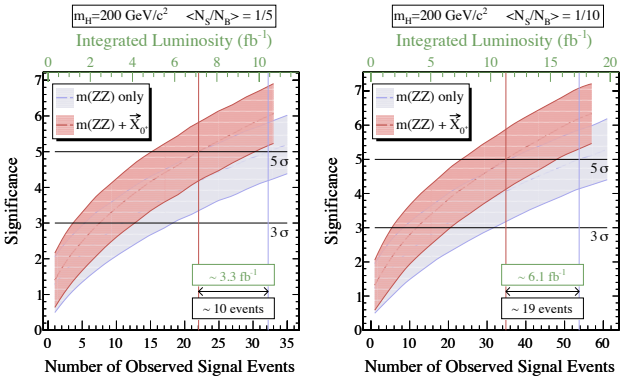

FIG. 13: Distribution of signal significance for a 200 GeV/c^2 SM Higgs boson decaying in the $H \to ZZ \to 4\mu$ channel for pp collisions with $\sqrt{s} = 10$ TeV. The mean signal to background ratios used are $< N_S/N_B > = 1/5$ (top) and 1/10 (bottom).

Fig. 15 shows the corresponding results for a $m_H = 200$ GeV/c^2 and $m_H = 350$ GeV/c^2 pseudoscalar Higgs impostor. Here the input parameters (such as the cross section) are assumed to be those of a SM Higgs boson; only the shape of the *pdfs* defining the likelihood (and in particular the correlations between the angles) are different from the SM case. The angular distributions and correlations for a pseudoscalar resonance are similar to the ZZ background, resulting in a much smaller improvement in the signal significance over the "$m(ZZ)$ only" fit, and thus a smaller distance between the two bands in the plots.

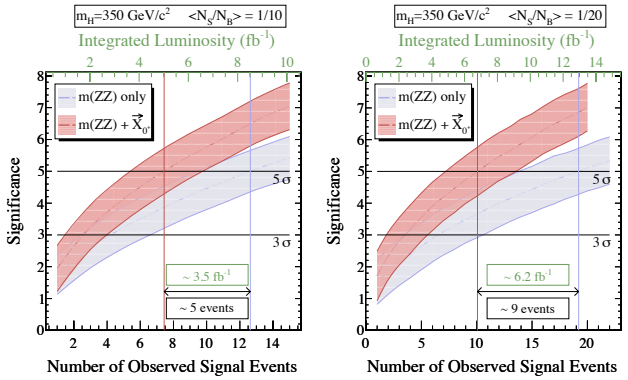

FIG. 14: Distribution of signal significance for a 350 GeV/c^2 SM Higgs boson decaying in the $H \to ZZ \to 4\mu$ channel for pp collisions with $\sqrt{s} = 10$ TeV. The mean signal to background ratios used are $< N_S/N_B > = 1/10$ (top) and 1/20 (bottom).

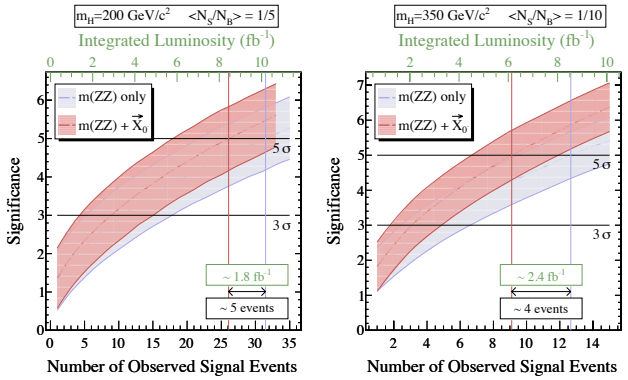

FIG. 15: Distribution of signal significance for a 200 (top) GeV/c^2 and 350 (bottom) pseudoscalar particle in the $0^- \to ZZ \to 4\mu$ channel for pp collisions with $\sqrt{s} = 10$ TeV. The mean signal to background ratios used are $< N_S/N_B > = 1/5$ (top) and 1/10 (bottom), and assumes NLO SM Higgs production cross section for the pseudoscalar.

VII. RESULTS

We present results for three HLL masses: $m_H = 145$, 200, and 350 GeV/c^2.

A. 0^+ vs. 0^-

We consider here two different "pure" scalar hypotheses: 0^+, corresponding to a SM Higgs, and 0^-, a psuedoscalar. Neither of these possibilities have explicit dependence on the angles $\vec{\Omega}$ in their differential cross-section, meaning that only the variables $\vec{\omega}$ (and the off shell Z mass, $m_2 = M_{Z^*}$, in $m_H < 2M_Z$ cases) are used to discriminate between the two hypotheses.

In Fig. 16 we show the distributions in ϕ and $\cos\theta_1$ at $m_H = 350$ GeV/c^2 for $J^P = 0^+$ and 0^-. These angular variables (along with $\cos\theta_2$, whose distribution is identical to that of $\cos\theta_1$ except when Z_2 is off-shell) provide the discrimination between these two hypotheses at all masses m_H. For masses m_H below the $2\,M_Z$ threshold,

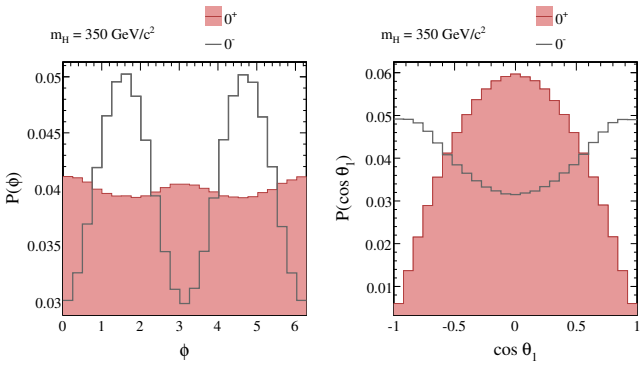

FIG. 16: Distributions of the variables ϕ (left) and $\cos\theta_1$ (right) for 0^+ and 0^- resonances with $m_H = 350$ GeV/c^2. All distributions are normalized to a unit integral.

the kinematic factors in Eqs. 15,16 result in the differential cross section dependences on the off-shell Z mass M_{Z^*} that differ for the 0^+ and 0^- cases. This is illustated in Fig. 17(left) for $m_H = 145$ GeV/c^2. For all the discriminating variables we consider, the ability to distinguish between two hypotheses is degraded when their correlations are neglected. This is shown in Fig. 17(right) where we present the results of the NePe hypothesis test between 0^+ and 0^- for likelihoods built using different subsets of variables and correlations thereof. Specifically $P(M_{Z^*}, \vec{\omega})$ denotes the use of the full set of variables while in $P(\vec{\omega})$ the probability distribution of M_{Z^*} is ignored. The product of all one-dimensional probabilities, ignoring correlations, is $\prod_i P(X_i)$. As expected, the likelihood including all discriminating variables and their correlations is optimal. The other two definitions give similar results. We note that, regardless of the results, the use of $\prod_i P(X_i)$ is an improper approximation, since the X_i variables are far from being uncorrelated.

The expected significance for discriminating between the 0^+ and 0^- hypotheses (assuming one of the other to be correct), as a function of N_S, where N_S is the number of observed signal events,is shown in Fig. 18 for $m_H = 145$, 200, and 350 GeV/c^2. In all cases, results correspond to the case where $\mathbb{H}_1$ is the true hypothesis (see Sec. V). The model discrimination is based on an NePe test between these simple hypotheses with test statistic $\log(L[0^+]/L[0^-])$. The variables $\vec{\omega}$ (and M_{Z^*}, when applicable), along with their correlations, are used in the likelihood construction. The median expected significance for rejecting one hypothesis in favor of the other at the time of $5\,\sigma$ excess (see Sec. VI) is better than $3\,\sigma$ for $m_H = 200$ and 350 GeV/c^2 while a $5\,\sigma$ discrimination can be achieved with double the observed signal events

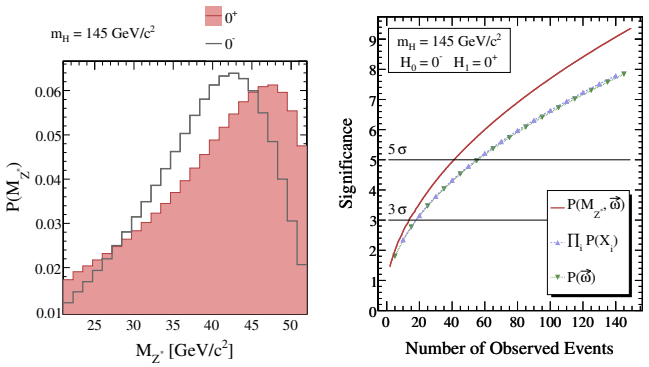

FIG. 17: Left: Distribution of M_{Z^*} for 0^+ and 0^- $H \to ZZ^*$ decays at $m_H = 145$ GeV/c^2, normalized to a unit integral. Right: Median expected significance for rejecting 0^- in favour of 0^+, assumed to be correct, as a function of N_S. The different likelihood constructions are specified in the text.

(less than ∼40 events in both mass cases presented here).

B. 0^+ vs. 1^- and 1^+

We consider here two different "pure" $J = 1$ models specified by their HZZ couplings: "vector" ($J = 1^-$) and "axial vector" ($J = 1^+$). Unlike in the 0^+ case, the differential cross sections have non-trivial dependences on the Z-production angles $\vec{\Omega}$ that provide additional discrimination between 0^+ and $J = 1$. In Fig. 19 we show the distributions for some of these variables.

We note that when a $J = 1$ resonance decays in ZZ^*, the shapes of the $c_1 \equiv \cos\theta_1$ and $c_2 \equiv \cos\theta_2$ are not any longer qualitatively similar, as illustrated in Fig. 20 (in striking contrast to the $J = 0$ cases). In the $J = 1^-$ case, this asymmetric effect arises from the configurations in which the object, in its rest system, is polarized along the direction of motion of one of its Z-decay products. These helicity configurations result in an addend proportional to $m_2^2 s_1^2 c_2^2 + m_1^2 s_2^2 c_1^2$ in the *pdf*, which can be rewritten as $2M_Z^2(s_1^2 + s_2^2 - s_1^2 s_2^2) - m_d^2 s_1^2(2 - s_2^2)$, with $m_d^2 \equiv M_Z^2 - m_2^2$. The second term is $1 \leftrightarrow 2$ asymmetric at fixed m_d and induces the difference between the c_1 and c_2 one-dimensional distributions. In the $J = 1^+$ case the asymmetric *pdf* term is, in the notation of Appendix D, $2M_1^4 m_2^2 s_1^2 + 2M_2^4 m_1^2 s_2^2 - (M_2^4 m_1^2 + M_1^4 m_2^2) s_1^2 s_2^2$ and its origin is similar. These asymmetric effects significantly enable the discrimination between $J = 1$ and $J = 0$ models when $m_H < 2M_Z$.

In Fig. 21 we compare the discrimination between the 0^+ and 1^- hypotheses for likelihood definitions that exploit different variables. The obvious qualitative conclusion is that likelihoods defined in terms of *pdfs* containing the most information are the most performant. The figure shows the relative discriminating power of the different choices: $P(a_1, \cdots, a_N)$ denotes N-dimensional *pdfs* in

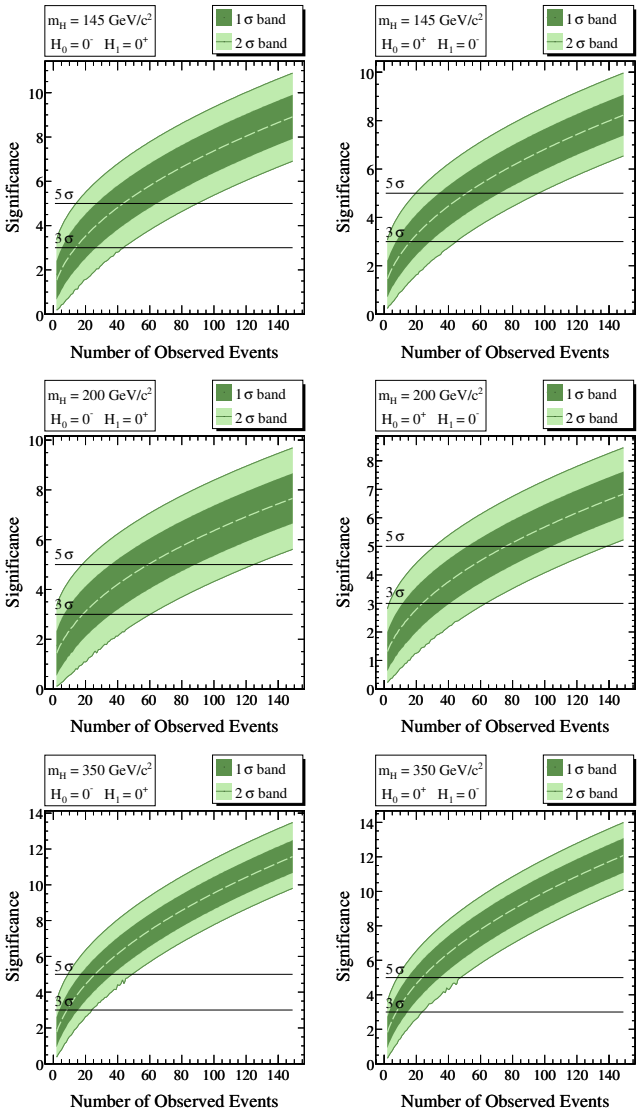

FIG. 18: Expected significance for rejecting 0^- in favor of 0^+, assuming 0^+ is true (left), and vice-versa, $0^+ \leftrightarrow 0^-$ (right), for $m_H = 145$, 200 and 350 GeV/c^2 (top, middle and bottom). The dashed central line is the median significance. The 1 and $2\,\sigma$ bands correspond to 68% and 95% confidence intervals, centered on the median.

the correlated variables $\{a_1, \cdots, a_N\}$. $\prod_i P(X_i)$ is constructed from one-dimensional *pdfs* for all variables, ignoring (erroneously) their correlations. $P(\vec{\omega}\,|\langle\vec{\Omega}\rangle_{\rm TH})$ are *pdfs* including the variables $\vec{\omega}$ and their correlations, but with the hypothesis 1^+ represented by a *pdf* in which dependence on the variables $\vec{\Omega} = \{\Phi, \cos\Theta\}$ has been integrated out of the analytic differential cross-section. The likelihood $P(\vec{\omega}\,|\langle\vec{\Omega}\rangle_{\rm TH})$ performs badly relative to $P(\vec{\omega})$, where the two differ only in that the first construction implicitly assumes a uniform $4\,\pi$ coverage of the observed leptons, as if the muon p_T and η analysis requirements did not depend on the $\vec{\Omega}$ angular variables. The pri-

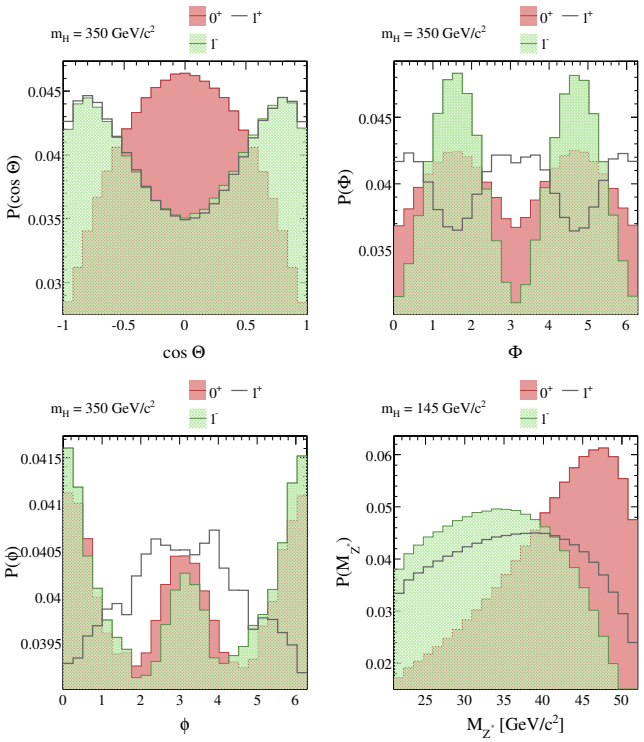

FIG. 19: Distributions of the variables $\cos\Theta$ (upper left), Φ (upper right), ϕ (lower left) and M_{Z^*} (lower right, $m_H = 145$ GeV/c^2) for 0^+, 1^- and 1^+ resonances. All distributions are normalized to a unit integral. The angular disctributions are shown for $m_H = 350$ GeV/c^2.

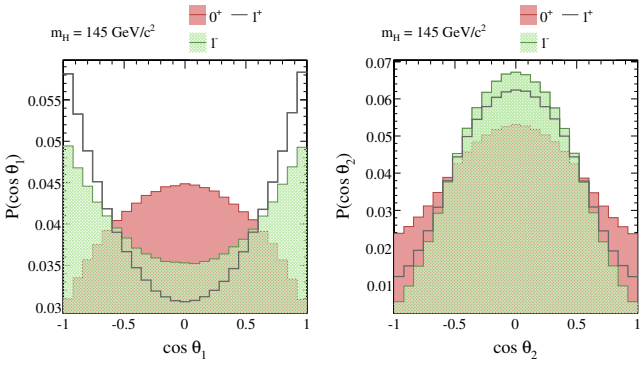

FIG. 20: Distributions in $\cos\theta_1$ (left) and $\cos\theta_2$ (right), for 0^+, 1^- and 1^+ resonances with mass 145 GeV/c^2, normalized to a unit integral.

mary reason for this difference is the strong correlation between the variables Φ and ϕ in the $J = 1$ *pdfs*, which causes phase space acceptance sculpting of the Φ distribution that in turn alters the ϕ distribution, as discussed in detail in Sec.IV.

The expected significance for discriminating between the 0^+ and 1^- (1^+) hypotheses, as a function of N_S, is summarized in Fig. 22 (Fig. 23). The full correlated set

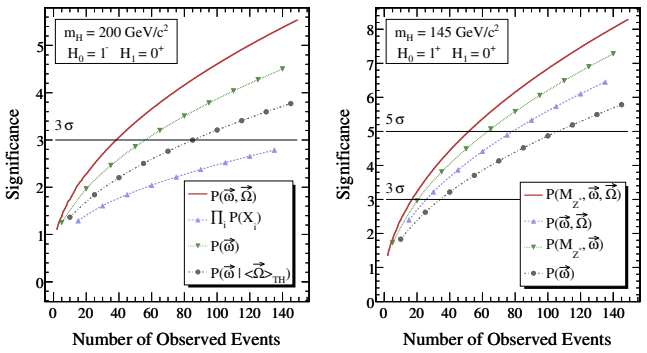

FIG. 21: Median expected significance for rejecting 1^- in favour of 0^+ (assuming 0^+ is true), for different likelihood constructions used in the likelihood ratio test statistic.

of variables $\vec{\Omega}$, $\vec{\omega}$, and M_{Z^*} (when applicable) is used in the likelihood construction. The discriminations are based on the NePe tests between simple hypotheses with statistic $\log(L[0^+]/L[1^-])$ ($\log(L[0^+]/L[1^+])$). The discrimination between 0^+ and 1^- or 1^+ is similar.

C. 0^+ vs. 2^+

We consider one "pure" $J = 2$ model: a $J = 2^+$ heavy-graviton-like resonance. A $J = 2$ object has *pdfs* with non-trivial dependence on the angles $\vec{\Omega}$ up to quartic order in $\cos\Theta$. In Fig. 24 we show the corresponding distributions in the $\vec{\Omega}$ variables and for $m_H = 200$ and 350 GeV/c^2. The ability to discriminate between the 0^+ and $J = 2$ hypotheses improves with increasing resonance mass. Despite the presence of quartic terms in $\cos\Theta$ in the 2^+ *pdf* and the absence of this variable in the 0^+ *pdf*, their corresponding one-dimensional *pdfs* are similar for the 0^+ and 2^+ resonances for values of m_H close to $2\,M_Z$, as shown in Fig. 24. Similar behavior is observed in the distributions of $\cos\theta_1$ and $\cos\theta_2$, as illustrated in Fig. 25. Nevertheless, the inclusion of all angular variables and their correlations improves the discrimination power between the 2^+ and 0^+ hypothesis as shown in Fig. 26.

The expected significance for discriminating between 0^+ and 2^+ as a function of N_S, is summarized in Fig. 27 for $m_H = 200$ and 350 GeV/c^2. For these tests the variables $\vec{\Omega}$ and $\vec{\omega}$ and their correlations were used in the likelihood. Model discrimination is based on the NePe test between simple hypotheses with test statistic $\log(L[0^+]/L[2^+])$ and $\log(L[0^+]/L[2^-])$.

D. Other pure J^{PC} comparisons

If a resonance discovered in the 4ℓ final state does not have the quantum numbers of the SM Higgs boson, it is likely that 0^+ will be rejected in favor of other pure-J^{PC} hypotheses. The issue of abandoning a particular J^{PC} in favor of others becomes a combinatoric exercise, where the compatibility of the data is assessed against

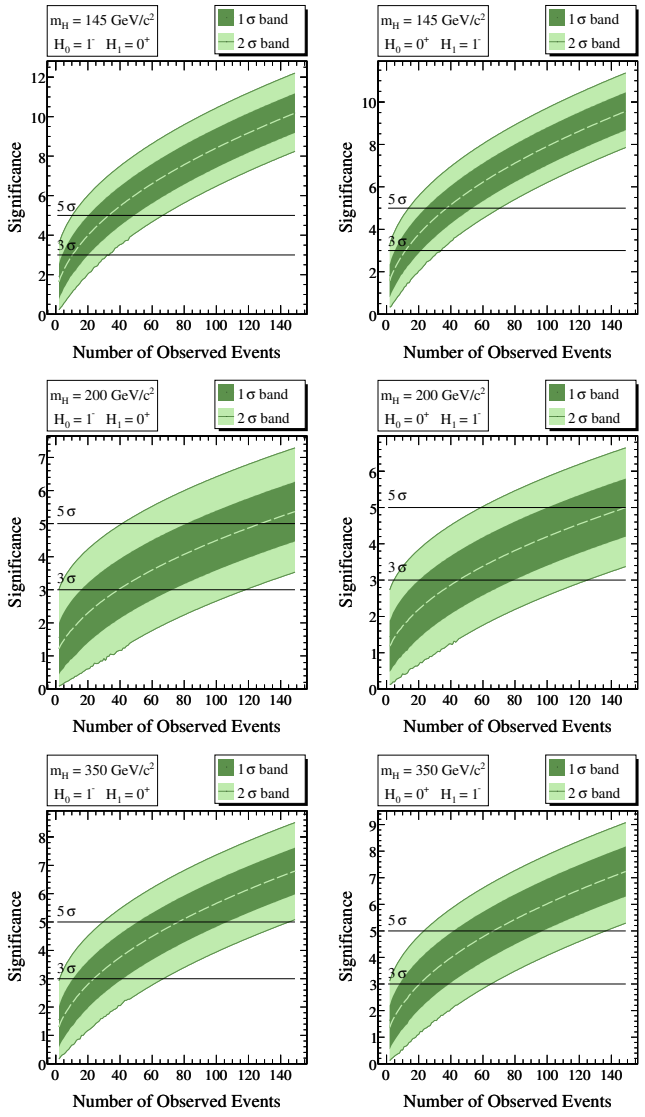

FIG. 22: Expected significance for rejecting 1^- in favour of 0^+, assuming 0^+ is true (left), or vice-versa ($0^+ \leftrightarrow 1^-$, right), for $m_H = 145$, 200 and 350 GeV/c^2 (top, middle and bottom).

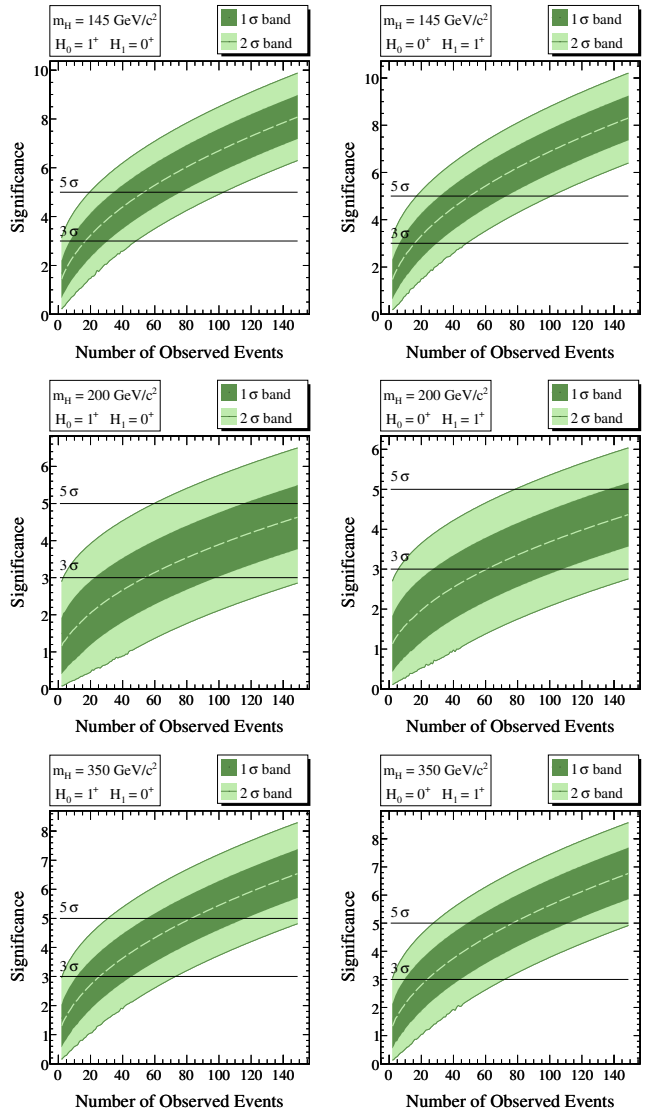

FIG. 23: Expected significance for rejecting 1^+ in favour of 0^+, assuming 0^+ is true (left), or vice-versa ($0^+ \leftrightarrow 1^+$, right), for $m_H = 145$, 200 and 350 GeV/c^2 (top, middle and bottom).

each possible pair of hypotheses in a simple NePe test, in view of selecting the optimal assumption. In this section we present the expected results for these comparison tests, as a function of the observed number of events N_S. Following the results of the previous section, we alawys use the full set of angular variables, plus, when appropriate, M_{Z^*}, corresponding to the optimal statistic for model discrimination.

The discrimination between the 0^- hypothesis and the pure $J = 1$ ones is very similar to the case of distinguishing the latter from 0^+, described in Sec. VII B. The *pdf* for 0^- has also no explicit dependence on the angles $\vec{\Omega}$. Differences in the *pdfs* of these variables provide discrimination between 0^- and $J = 1$ states, as Fig. 28 illustrates. The one-dimensional M_{Z^*} *pdfs* are similar for 0^- and 0^+, as well as for 1^- and 1^+, while the differences between the two J-values are maximal. The $\cos\theta_{1,2}$ distributions for $J = 1$ have qualitatively different behavior when $m_H < 2\,M_Z$, as discussed in Sec. VII B. This results in the $J = 1$ $\cos\theta_1$ ($\cos\theta_2$) distribution being more "0^--like" ("0^+-like"), resulting in similar levels of discrimination between $J = 0$ and $J = 1$.

The expected significance to distinguish the 0^- and 1^- (1^+) hypotheses, as functions of N_S, is shown in Fig. 29 (30). The $m_H = 145$ GeV/c^2 results and the ones for 0^+ vs. $J = 1$ (Figs. 22 and 23) are nearly identical. A similar comparison of 0^- vs. $J = 1$ (Figs. 29 and 30) with 0^+ vs. $J = 1$ (Figs. 22 and 23) for $m_H = 200$ GeV/c^2 reveals that it is more difficult to discriminate between 0^+ and

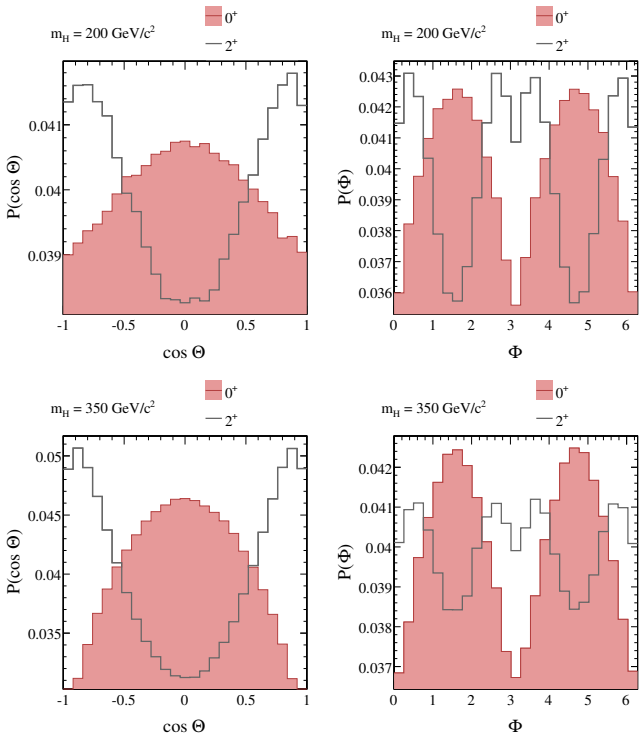

FIG. 24: Distributions of the variables $\cos\Theta$ (left) and Φ (right) for 0^+, 2^+ resonances with masses of 200 and 350 GeV/c^2 (top, bottom). All distributions are normalized to a unit integral.

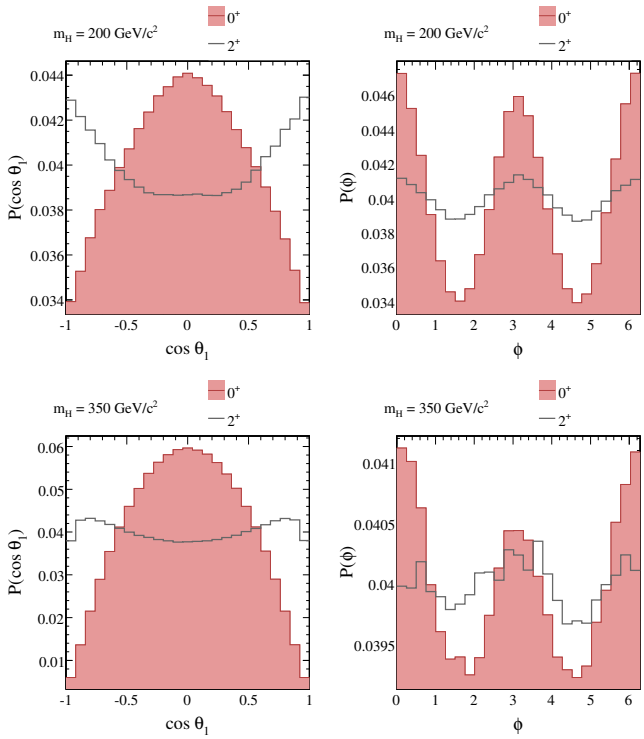

FIG. 25: Distributions of the variables $\cos\theta_1$ (left) and ϕ (right) for 0^+, 2^+ resonances with masses of 200 and 350 GeV/c^2 (top, bottom). All distributions are normalized a unit integral.

$J = 1$ at this mass. This is predominantly due to the *pdfs* for the angles $\cos\theta_{1,2}$ (which are similar between 0^+ and $J = 1$ for $m_H > 2\,M_Z$).

The distributions for the variables $\vec{\Omega}$ and $\vec{\omega}$ for all the pure J^{PC} hypotheses considered in our analysis are shown in Fig. 31, for $m_H = 200$ GeV/c^2. The $\vec{\Omega}$ distributions are nearly identical for the two $J = 0$ cases, since they are only induced by detector limitations.

The potential to distinguish between 0^- and 2^+ resonances is shown in Fig. 32 for $m_H = 200$ and 350 GeV/c^2. If both of the $J = 0$ cases are excluded in favour of $J = 1$ or $J = 2$, one needs to discriminate between the latter. Relative to $J = 0$ case, the two pure $J = 1$ resonances have the most similar *pdfs*, as we saw in Sec. VII B while comparing them to the 0^+ case. The comparison to the $J = 2$ reflects the same limitation, as shown in Fig. 34 (35) for 1^- vs. 2^+ (1^+ vs. 2^+).

The hardest differentiation is between 1^- and 1^+. Figures 28, 31 and 33, show that the one-dimensional $\cos\Theta$, $\cos\theta_1$, $\cos\theta_2$ and M_{Z^*} *pdfs* are similar. While the Φ and ϕ *pdfs* provide some discrimination, the phase space acceptance tends to sculpt the Φ distributions (and ϕ distribution through correlations) in ways that render the two cases very similar. The expected significance for distinguishing between the two $J = 1$ cases is shown in Fig. 36. We conclude that the discriminating potential

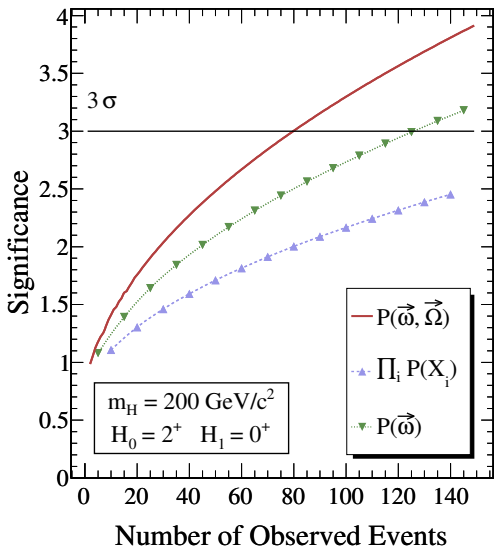

FIG. 26: Median expected significance for rejecting 2^+ in favour of 0^+, assuming 0^+ is true, for the different likelihood constructions discussed in the text.

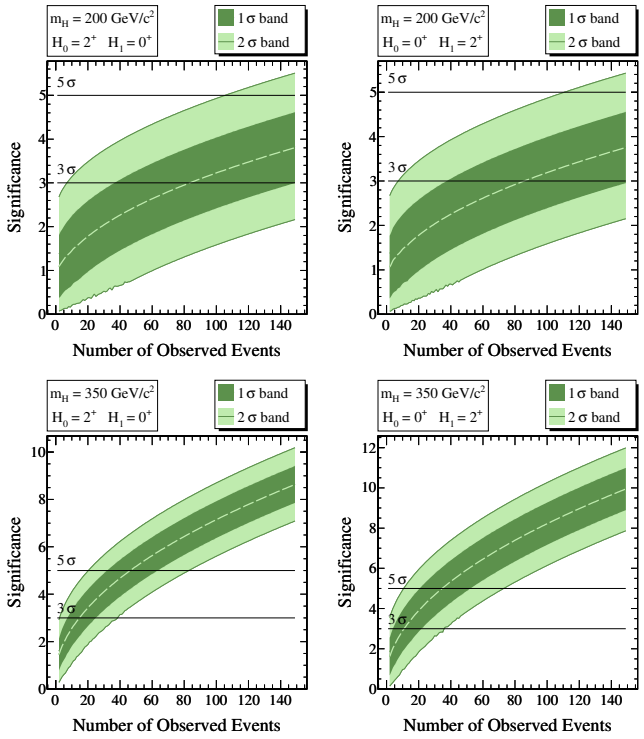

FIG. 27: Expected significance for rejecting 2^+ in favour of 0^+, assuming 0^+ is true (left) or vice-versa ($0^+ \leftrightarrow 2^+$, right), for $m_H = 200$ and 350 GeV/c^2 (top, bottom).

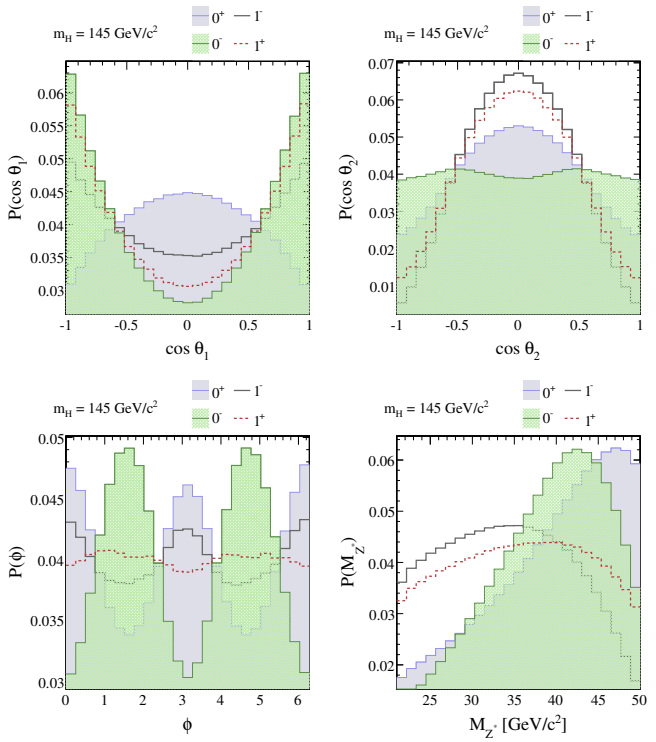

FIG. 28: Distributions, normalized to a unit integral, of the variables $\cos\theta_1$ (top left), $\cos\theta_2$ (top right), ϕ (bottom left) and M_{Z^*} (bottom right) for 0^+, 0^-, 1^+ and 1^- resonances with $m_H = 145$ GeV/c^2.

is weakest for 1^+ vs. 1^-, for all m_H. We revisit this result in Sec. VII G in the context of measuring mixing parameters in a general $J = 1$ Lagrangian.

E. 0^+ vs. mixed scalar states

Consider the "vertex Feynman rules" of Eq.10 for the most general Lorentz-covariant coupling $\mathcal{L}_{\mu\alpha}$ of a spinless object to a Z pair. Rather than studying the general case, for which any of the quantities X to Q can be nonzero, we investigate three cases, each with only two non-vanishing types of coupling, resulting in one free mixing "angle" and an overall normalization (which we ignore):

- $X \neq 0,\ P \neq 0$: A scalar whose ZZ coupling has a CP-violating phase, ξ_{XP}:

$$\mathcal{L}_{\mu\alpha} \propto \cos(\xi_{XP})\, g_{\mu\alpha} + \sin(\xi_{XP})\, \epsilon_{\mu\alpha} p_1 p_2 / M_Z^2 \qquad (34)$$

- $X \neq 0,\ Q \neq 0$: A scalar whose ZZ coupling violates C, described in terms of an angle as:

$$\mathcal{L}_{\mu\alpha} \propto \cos(\xi_{XQ})\, g_{\mu\alpha} + i\sin(\xi_{XQ})\, \epsilon_{\mu\alpha} p_1 p_2 / M_Z^2 \qquad (35)$$

- $X \neq 0,\ Y \neq 0$: A composite 0^+, parameterized in terms of an angle as:

$$\mathcal{L}_{\mu\alpha} \propto \cos(\xi_{XY})\, g_{\mu\alpha} - \sin(\xi_{XY})\, k_\alpha k_\mu / M_Z^2 \qquad (36)$$

For each of the above three cases we attempt to answer two different questions, both depending on N_S, the number of signal events (and the resulting significance):

- (a) If the resonance is SM-like (0^+, with all mixing angles vanishing), which values of the angles can be excluded in favour of a pure 0^+?
- (b) If the resonance corresponds to one of the three mixed cases discussed above, can a pure 0^+ be excluded in favour of a non-trivial mixture?

We first explain how tests to answer these two questions are constructed in the example of a CP-violating HZZ coupling with $m_H = 350$ GeV/c^2.

Question (a) corresponds to a series of simple hypothesis tests of the type we considered earlier for distinguishing between pure J^{PC} states. Specifically, for a given number of observed signal events at a fixed value of m_H, we perform a NePe test between two simple hypotheses: that the resonance is 0^+ (denoted hypothesis $\mathbb{H}_1$) or that the resonance is $J = 0$ with ξ_{XP} *fixed to a specific nonzero value* (denoted hypothesis $\mathbb{H}_0$). The test statistic we use is $\log[L^{XP}(\xi_{XP})/L(0^+)]$, where $L(0^+)$

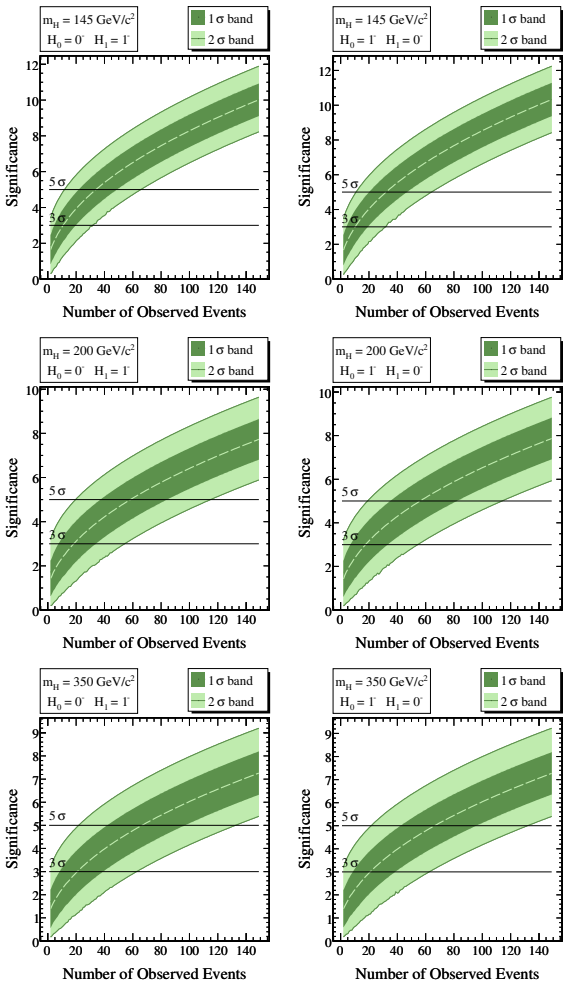

FIG. 29: Expected significance for rejecting 0^- in favour of 1^-, assuming 1^- is true (left) or vice-versa ($0^- \leftrightarrow 1^-$, right) for $m_H = 145$, 200 and 350 GeV/c^2 (top, middle and bottom).

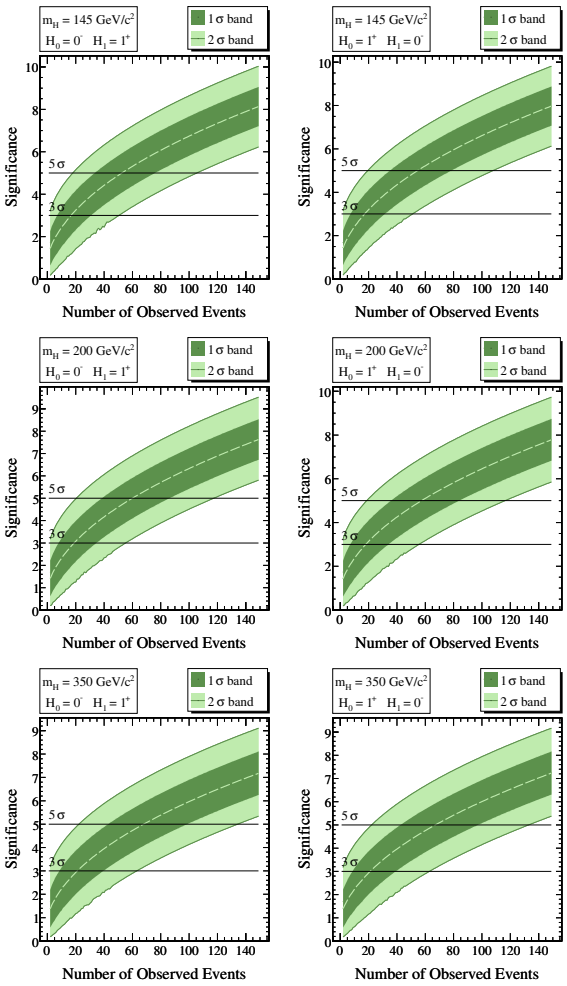

FIG. 30: Expected significance for rejecting 0^- in favour of 1^+, assuming 1^+ is true (left) or vice-versa ($0^- \leftrightarrow 1^+$, right) for $m_H = 145$, 200 and 350 GeV/c^2 (top, middle and bottom).

and $L^{XP}(\xi_{XP})$ denote the likelihoods for a set of events agreeing with the hypotheses $\mathbb{H}_1$ and $\mathbb{H}_0$, respectively. Obviously, the test cannot be performed for $\xi_{XP} = 0$, since in this case the $\mathbb{H}_1$ CP-violating hypothesis we want to test reduces to the alternative $\mathbb{H}_0$ hypothesis (the CP-conserving SM Higgs).

The result of this test is the significance with which hypothesis $\mathbb{H}_0$ can be rejected in favor of the hypothesis $\mathbb{H}_1$, or similarly, the significance with which a particular value of ξ_{XP} can be excluded in favour of the 0^+ hypothesis. This test is then repeated with different fixed values of ξ_{XP}, *i.e.* different NePe tests with different hypotheses $\mathbb{H}_0$. The results for a large ensemble of such tests are shown in Fig. 37. Here, $\mathbb{H}_0 = 0^{XP}$ denotes the simple $J = 0$ CP-violating hypothesis with ξ_{XP} fixed at values chosen on the x-axis.

In this example we see that, for $N_S = 50$, the median expected significance for excluding a CP-violating coupling exceeds $3\,\sigma$ for $|\xi_{XP}| > 0.5$ and $5\,\sigma$ for $|\xi_{XP}| > 0.9$.

Question (b) is complementary to (a). Here, nature has chosen a value of $\xi_{XP} \neq 0$ and we would like to know the confidence with which one can exclude the $\xi_{XP} = 0$ hypothesis. This time, extra care needs to be taken in choosing the test statistic, since we do not know ξ_{XP} a priori . Consequently, we cannot construct a simple NePe test between 0^+ and a fixed-ξ_{XP} hypothesis. Instead, we treat ξ_{XP} as a nuisance parameter and choose a value, $\hat{\xi}_{XP}$, that maximizes the CP-violating likelihood for the given set of observed events. Specifically, we fix ξ_{XP} at a particular value (the "true" value) to generate events and perform NePe tests comparing $\xi_{XP} = 0$ (denoted hypothesis $\mathbb{H}_0$) and $\xi_{XP} \neq 0$ ($\mathbb{H}_1$). This test is repeated for many different values of the fixed "input" ξ_{XP}.

An example of results from an ensemble of these tests is shown in Fig. 38. Because of the addition of a nuisance parameter, the figure's interpretation is not simply related to the interpretation of Fig. 37, which answered

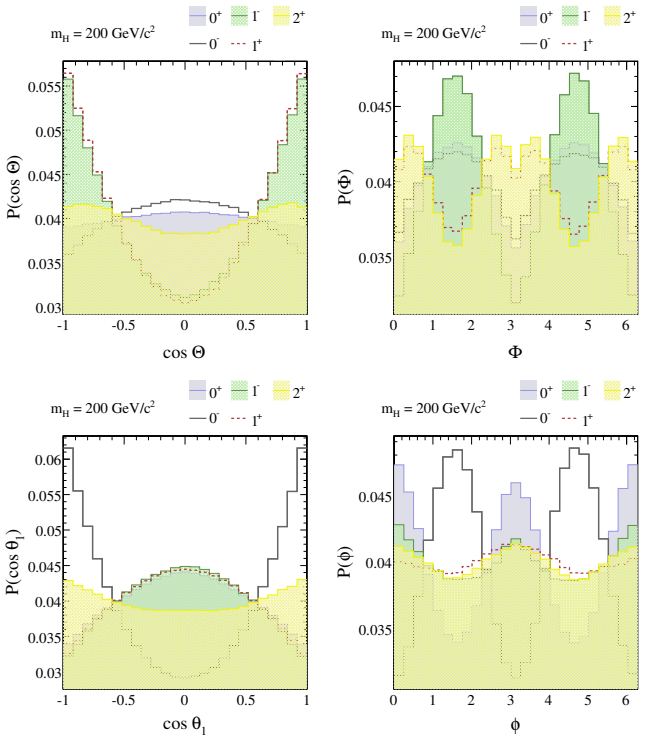

FIG. 31: Distributions of $\cos\Theta$ (top left), Φ (top right), $\cos\theta_1$ (bottom left) and ϕ (bottom right) for all the pure J^{PC} choices we study, for $m_H = 200$ GeV/c^2. All distributions are normalized to a unit integral.

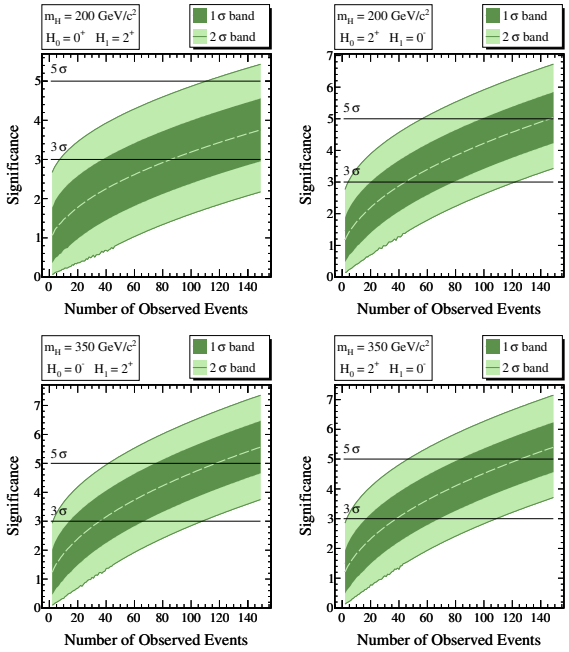

FIG. 32: Expected significance for rejecting 0^- in favour of 2^+, assuming 0^- is true (left) or vice-versa ($0^- \leftrightarrow 2^+$, right) for $m_H = 200$ and 350 GeV/c^2 (top, bottom).

question (a). What Fig. 38 shows is the expected significance with which one can exclude the SM hypothesis in favour of the CP-violating hypothesis with $\xi_{XP} \neq 0$, as a function of the true value of ξ_{XP} (given on the x-axis). No a priori knowledge of the actual value of ξ_{XP} is required to perform this test.

From Figs. 37 and 38 we observe that the expected significances are symmetric around $\xi_{XP} = 0$. This is due to the *pdfs* of the "pure 0^+" and "pure 0^-" terms being even under $\xi_{XP} \to -\xi_{XP}$, while the $\tilde{T}$-odd interference term vanishes under the integration of $\cos\theta_1$, $\cos\theta_2$ or ϕ. We shall see that there are exceptions to this trivial statement. Comparing these two figures we observe a remarkable similarity of the significances of the two tests. Since two different statistics are used, this is somewhat of a coincidence. To explain it, consider the example with $\xi_{XP} = \pi/5$, which corresponds to vertical slices of Figs. 37 and 38. We denote the two different test statistics $\Lambda^{\rm fix} = \log[L^{XP}(\xi_{XP})/L(0^+)]$, with ξ_{XP} fixed at its true value, corresponding to a simple hypothesis test and $\Lambda^{\rm max} = \log[\max L^{XP}(\hat{\xi}_{XP})/L(0^+)]$, *marginalized* to the value $\hat{\xi}_{XP}$ at which it peaks. The distributions of $\Lambda^{\rm fix}$ and $\Lambda^{\rm max}$ are shown in Fig. 39.

In the top figure the bell-shaped curves $P(\Lambda^{\rm fix}|0^+)$ and $P(\Lambda^{\rm fix}|0^{XP})$ are characteristic of a simple hypothesis test. The distributions of $\Lambda^{\rm max}$ have a sharp cut-off

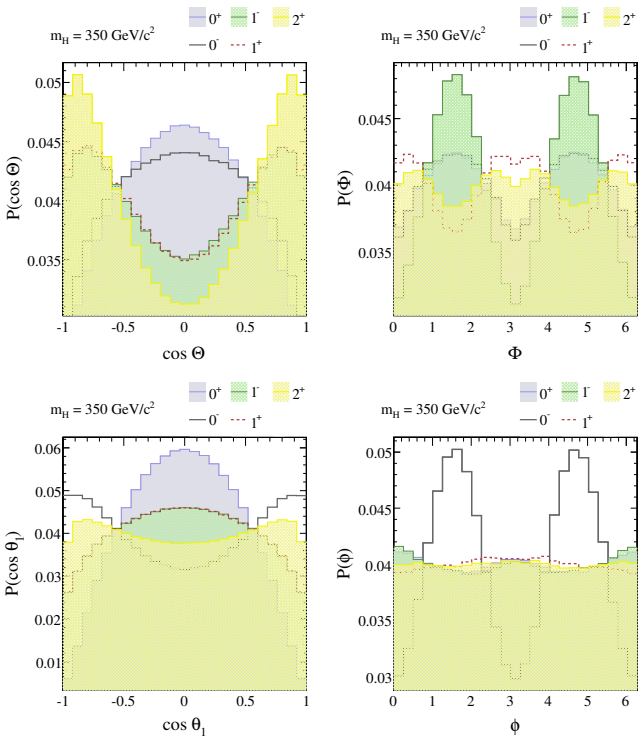

FIG. 33: Distributions of $\cos\Theta$ (top left), Φ (top right), $\cos\theta_1$ (bottom left) and ϕ (bottom right) for all the pure J^{PC} choices we study, for $m_H = 350$ GeV/c^2. All distributions are normalized to a unit integral.

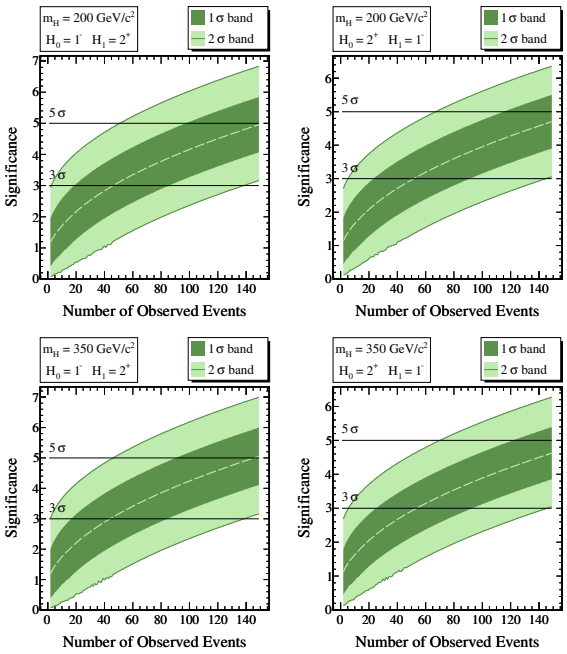

FIG. 34: Expected significance for rejecting 1^- in favour of 2^+, assuming 2^+ is true (left) or vice-versa ($1^- \leftrightarrow 2^+$, right) for $m_H = 200$ and 350 GeV/c^2 (top, bottom).

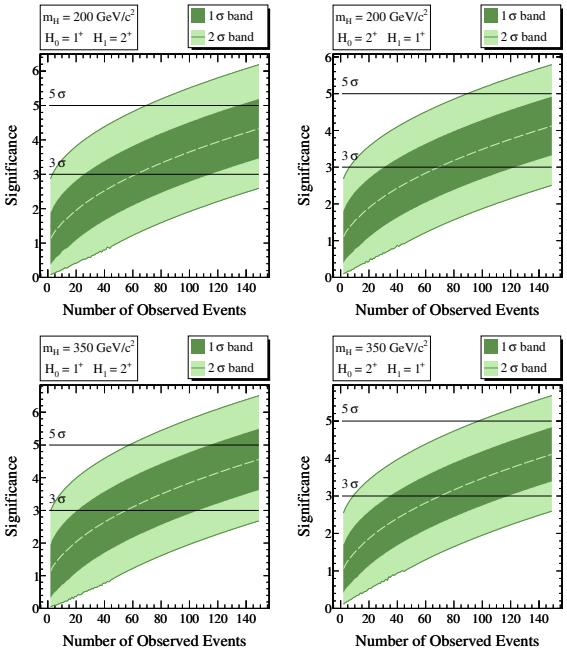

FIG. 35: Expected significance for rejecting 1^+ in favour of 2^+, assumed to be correct (left) or vice-versa ($1^+ \leftrightarrow 2^+$, right) for $m_H = 200$ and 350 GeV/c^2 (top, bottom).

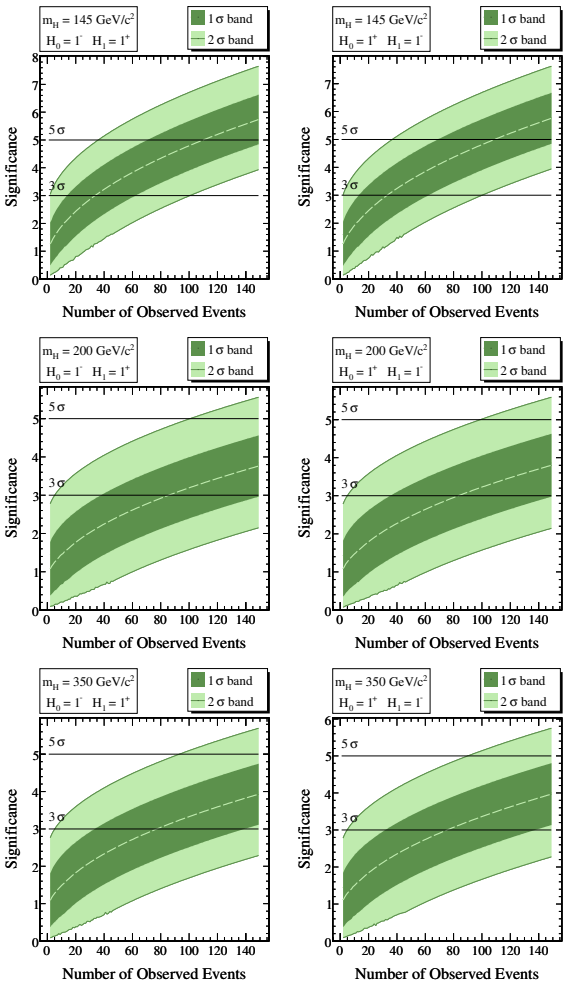

FIG. 36: Expected significance for rejecting 1^- in favour of 1^+, assuming 1^+ is true (left) or vice-versa ($1^- \leftrightarrow 1^+$, right) for $m_H = 145$, 200 and 350 GeV/c^2 (top, middle and bottom).

at $\Lambda^{\max} = 0$, since the 0^+ model is a member of the 0^{XP} family with $\xi_{XP} = 0$, and $\max L^{XP}(\hat{\xi}_{XP})/L(0^+) \geq 1$, which are also features characteristic of this type of test.

The reason for two very different hypothesis tests to end up in the similar-looking results of Figs. 37 and 38 is that the statistically-significant features of the different-looking $P(\Lambda)$ distributions shown in Fig. 39 are actually very similar. $P(\Lambda^{\text{fix}}|0^{XP})$ and $P(\Lambda^{\max}|0^{XP})$ differ, but the distributions of ξ_{XP} close to the maxima are localized around the true input value, their median values and 68% and 95% confidence intervals are nearly identical (try to tell apart the two vertical dotted lines in the lower half of Fig. 39, at $\Lambda \sim 7$). Also, the tails of one-minus-cumulative distributions for $P(\Lambda^{\text{fix}}|0^+)$ and $P(\Lambda^{\max}|0^+)$ coalesce for p-values exceeding $2\,\sigma$ significance, despite large differences in the distributions themselves.

In Fig. 40 we show the results for the distinction between pure 0^+ and CP-violating $J = 0$ hypotheses for

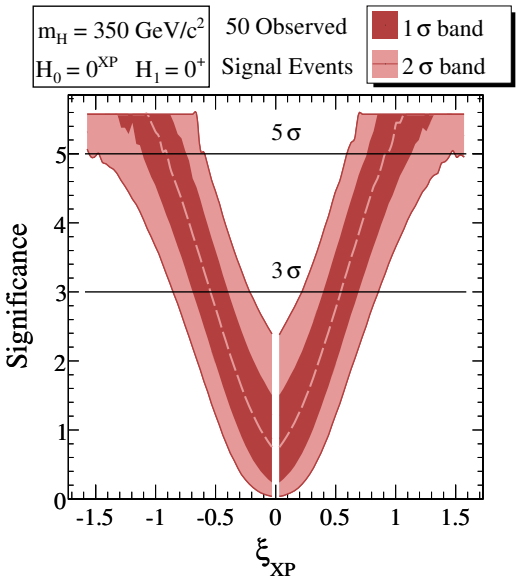

FIG. 37: Expected significance for excluding values of ξ_{XP} in the CP-violating $J = 0$ hypothesis in favour of the 0^+ one, assumed to be correct, for $m_H = 350$ GeV/c^2 and $N_S = 50$. The dashed line corresponds to the median expected significance. The 1 and 2σ bands correspond to 68% and 95% confidence intervals centered on the median value.

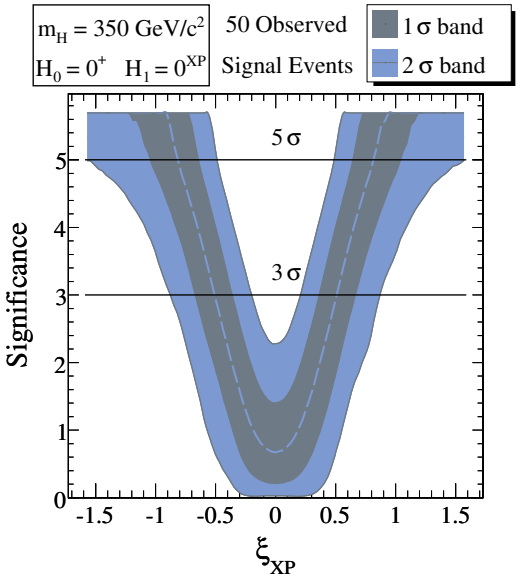

FIG. 38: The expected significance for excluding a pure 0^+ in favour of a CP-violating HZZ coupling ($\xi_{XP} \neq 0$), assuming the latter to be correct, with ξ_{XP} given by its x-axis values. Example for $N_S = 50$, $m_H = 350$ GeV/c^2. Dashed line and bands as in Fig. 37.

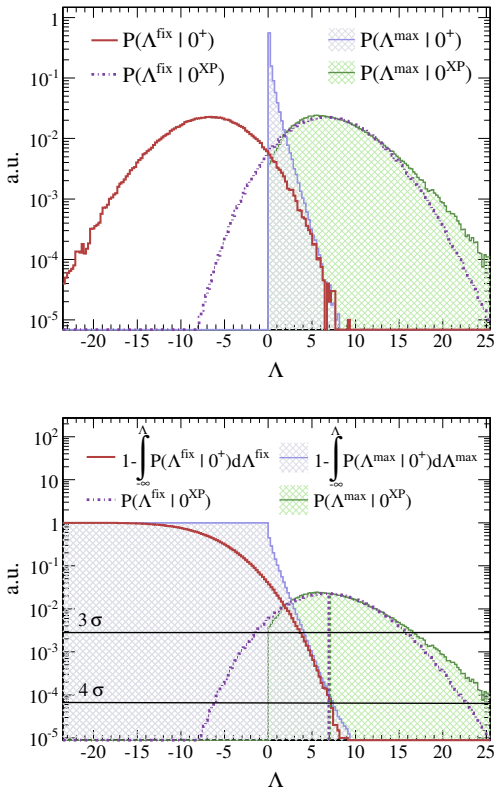

FIG. 39: Distributions of the two statistics Λ, defined in the text, for $m_H = 350$ GeV/c^2 and $N_S = 50$. The hypotheses are $\mathbb{H}_0 = 0^+$, and $\mathbb{H}_1 = 0^{XP}$ with the CP-phase ξ_{XP} fixed at $\pi/5$. (Top) Probability distributions P($\Lambda|\mathbb{H}$). (Bottom) The same with the 0^+ results traded for 1 minus their cumulative values. The two nearly indistinguishable vertical dotted lines correspond to the median values of the P($\Lambda|\mathbb{H}_1$) distributions.

$m_H = 145$ and 200 GeV/c^2.For $m_H = 145$ GeV/c^2, the "flat" behavior around $\xi_{XP} = 0$ is due to the coupling strength of the 0^+ part relative to 0^-, an order of magnitude larger for $m_H = 145$ GeV/c ^{2} and closer to unity for the higher m_H values. The corresponding results at $m_H = 350$ GeV/c^2 are those of Figs. 37 and 38.

The next mixed $J = 0$ case that we consider is that of a C-violating scalar, with mixing angle ξ_{XQ}. This scenario is very similar to that of the CP-violating scalar: only the interference term between the 0^+ and 0^- amplitudes is different (C-odd, instead of T-odd).

The expected results of hypothesis tests distinguishing between a C-violating scalar and a 0^+ state are shown in Fig. 41. Comparing this figure with Figs. 37, 38 and 40, we observe identical behavior in all the results. This shows that the relative strength between the 0^+ and 0^- parts of the matrix element squared, rather than the nature of the interference term, is the most relevant factor in resolving the values of ξ_{XP} and ξ_{XQ}.

If a pure 0^+ hypothesis is rejected in favour of both $\xi_{XP} \neq 0$ and $\xi_{XQ} \neq 0$, the next question would be

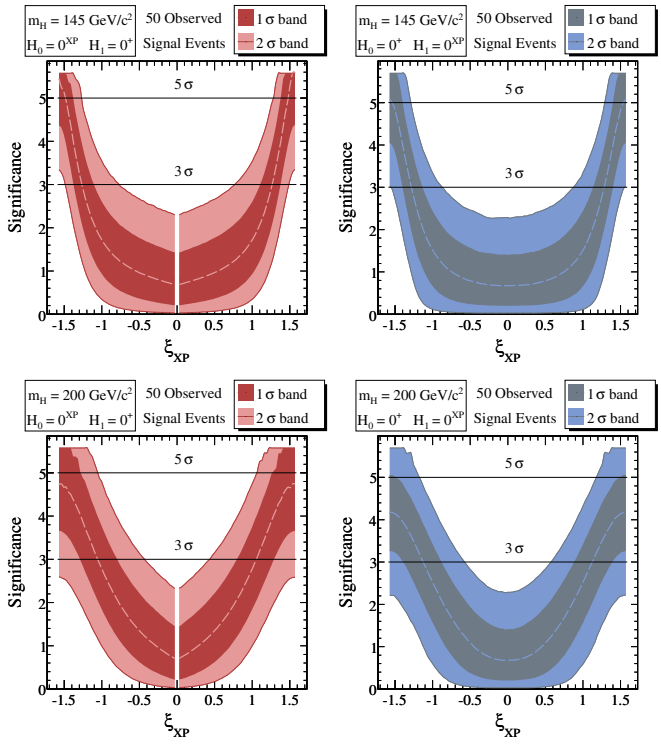

FIG. 40: Left: Expected significance for the exclusion of values of a CP-violating $\xi_{XP} \neq 0$ in favour of 0^+ ($\xi_{XP} = 0$), assumed to be correct. Right: Expected significance for excluding a pure 0^+ in favour of $\xi_{XP} \neq 0$, assumed correct with ξ_{XP} given by its x-axis values. Results for $m_H = 145$, 200 GeV/c^2 (top, bottom) and $N_S = 50$.

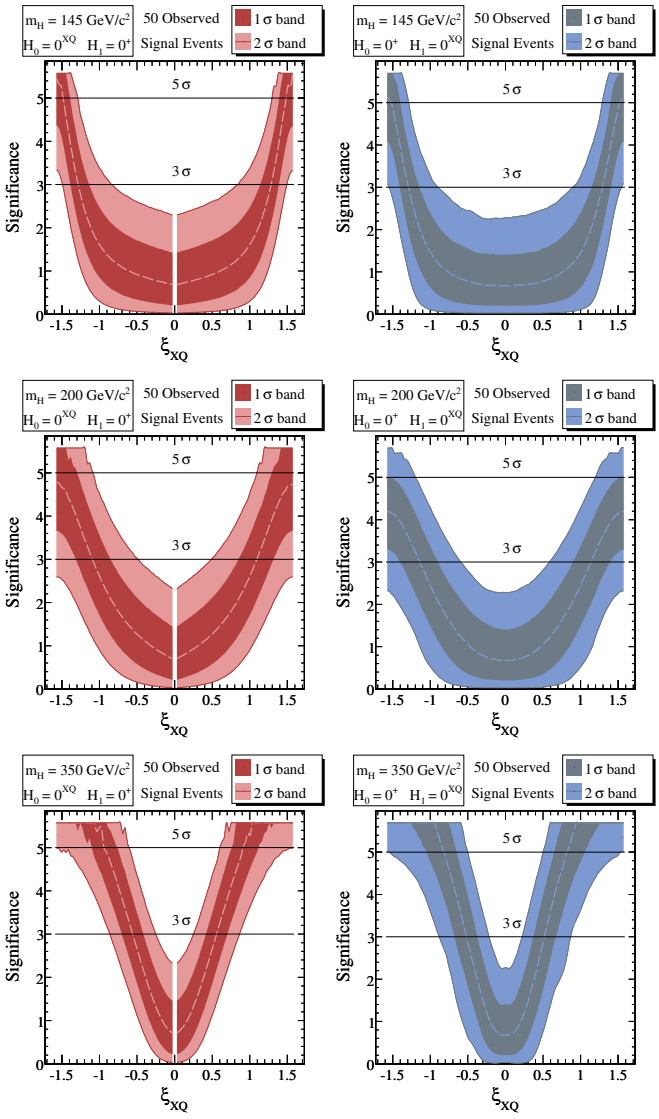

FIG. 41: Left: Expected significance for excluding values of a C-violating $\xi_{XQ} \neq 0$ in favour of 0^+ ($\xi_{XQ} = 0$), assumed to be correct. Right: Expected significance for excluding a pure 0^+ in favour of $\xi_{XQ} \neq 0$, assumed correct for the ξ_{XQ}-values on the x-axis. Hypothesis tests are for $m_H = 145$, 200 and 350 GeV/c^2 (top, middle and bottom), for $N_S = 50$.

whether it is possible to distinguish between these two cases. To address this question, we perform a series of hypothesis tests similar to the one described to answer type (b) questions. Specifically, we first assume a given CP-violating $\xi_{XP} \neq 0$ as "true". We then assess the expected significance with which particular values of ξ_{XQ} can be excluded in favour of the true hypothesis. Hence, for each fixed value of ξ_{XP} we perform a test against the C-violating case using a fixed ξ_{XQ}. The test statistic is $\Lambda = \log[\max L^{XP}(\hat{\xi}_{XP})/L(\xi_{XQ})]$, where the 0^{XQ} hypothesis is simple (fixed ξ_{XQ}) and $L(\xi_{XP})$ is marginalized "experiment-by-experiment". The test is repeated over a matrix of values for ξ_{XP} and ξ_{XQ}. Next, we switch the roles of the hypotheses to assess the significance for excluding given values of ξ_{XP} in favour of $\xi_{XQ} \neq 0$. The results are shown in Fig. 42. The color-coded z-"axis" is the median expected significance for ruling out the hypothesis $\mathbb{H}_0$ with the value of $\xi_{\mathbb{H}_0}$ given on the y-axis in favour of the $\mathbb{H}_1$ hypothesis with $\xi_{\mathbb{H}_1} \neq 0$, assumed to be correct for $\xi_{\mathbb{H}_1}$-values chosen on the x-axis.

The similarities between the C- and CP- mixed scalars are reflected in the $y \leftrightarrow x$ symmetries of Figs. 42. Moreover, switching the roles of the two hypotheses (comparing the figures on the left with those on the right) one only sees small changes. Still, the fact that the diagonals ($|\xi_{XP}| = |\xi_{XQ}|$) are not all at the same significance shows that the tests are sensitive to the differences between the $\tilde{T}$- and C-odd interference terms, but it would require an order of magnitude larger N_S to draw $5\,\sigma$-level conclusions over most of the (ξ_{XP}, ξ_{XQ}) plane. For example, we show in Fig. 43 the expected significance with which one can distinguish between the two cases, as a function of the number of observed events, for $\xi_{XY,XQ} = \pi/4$ and $m_H = 200$ GeV/c^2. The ambiguity between ξ_{XP}^{meas}, $-\xi_{XP}^{meas}$, $\xi_{XQ} = \xi_{XP}^{meas}$ and $\xi_{XQ} = -\xi_{XP}^{meas}$ would be very hard to lift.

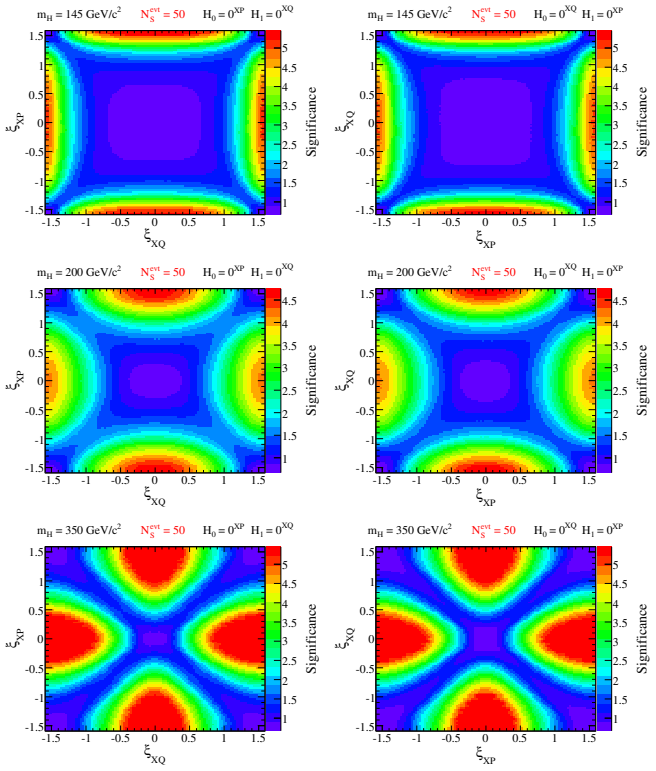

FIG. 42: The median expected significance (coloured z-"axis") for excluding values of $\xi_{\mathbb{H}_0}$ (y-axis) in favour of the $\xi_{\mathbb{H}_1} \neq 0$ hypothesis assuming as correct the values $\xi_{\mathbb{H}_1}$ of the x-axis. The tests are performed for $\mathbb{H}_1 = 0^{XP}$, $\mathbb{H}_0 = 0^{XQ}$ (left) and $\mathbb{H}_1 = 0^{XQ}$, $\mathbb{H}_0 = 0^{XP}$ (right); $m_H = 145$, 200 and 350 GeV/c^2 (top, middle and bottom), for $N_S = 50$.

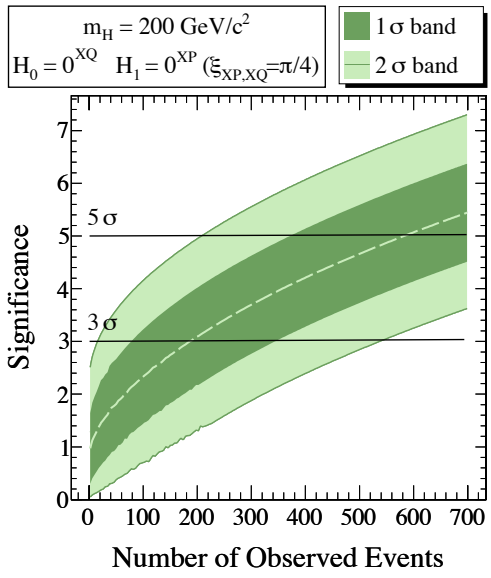

FIG. 43: The expected significance for excluding the C-violating $J = 0$ hypothesis in favour of a CP-violating case, assuming the latter to be correct, with $\xi_{XP,XQ} = \pi/4$. Example for $m_H = 200$ GeV/c^2.

The last $J = 0$ mixed case that we consider has unique features; this is the "composite Higgs" in which a term $\propto k_\mu k_\nu$ is present in the HZZ coupling. This case is different from the previous ones in that a composite scalar has well defined $J^{PC} = 0^{++}$, regardless of the value of the angle ξ_{XY} characterizing the mixing between its pointlike and derivative couplings. As a consequence, the angular integrals of their interference term do not vanish, and there is no symmetry around $\xi_{XY} = 0$. All the terms in the *pdf* having the same discrete symmetries and similar angular dependences, there happen to be large cancellations in the *pdf* for a 'critical' m_H-dependent value of ξ_{XY}, as in the example shown in Fig. 44 for the fully angular-integrated result.

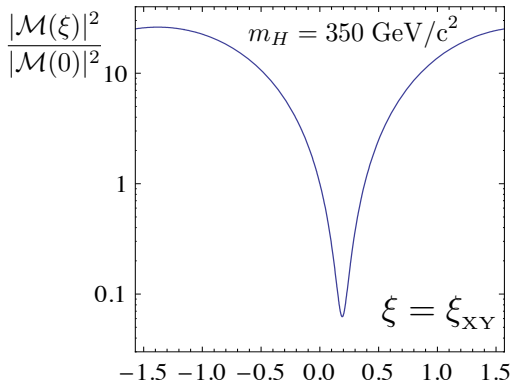

FIG. 44: The fully angularly-integrated matrix element squared for a "composite" 0^+, showing a strong destructive interference at a given ξ_{XY}. The result, shown here for $m_H = 350$ GeV/c^2, is normalized to $\xi_{XY} = 0$.

The appearance of an order of magnitude enhancement of the squared matrix element in Fig. 44 for $O(1)$ values of ξ_{XY} can be regarded as an artifact of our choosing a rather low mass scale (M_Z) in the definition of the dimensionless coupling Y in Eq. 10; thus if e.g. we instead chose the compositeness scale $= m_H = 350$ GeV/c^2, this enhancement would be much smaller. Nevertheless the possible enhancement from a nonzero Y coupling, and the possible suppression from XY interference, flags a dramatic scenario: it is possible to discover an HLL that is in fact a 0^{++} resonance, and is produced by exactly the same pp production processes as a SM Higgs, but for which the cross section times branching fraction to ZZ is several times higher *or* several times lower than Standard Model expectation.

We evaluate the expected significance with which one can distinguish between a pointlike and a composite 0^+ using the same hypothesis-test approach described earlier for the CP-violating scalar case. The results are shown in Fig. 45. We observe a non-trivial behavior of the significance values at and around the critical ξ_{XY}. Interestingly, the qualitative nature of these cancellations also changes with mass. For $m_H = 145$ GeV/c^2 and $m_H = 200$ GeV/c^2, the composite scalar with ξ_{XY} near

the critical point is 0^+-like, relative to nearby values of ξ_{XY}. For $m_H = 350$ GeV/c^2, it is very difficult to distinguish between the composite and elementary hypotheses, except if ξ_{XY} is close to critical. Near this critical value the significance is greatly improved, because after the large cancellations the angular distributions of the pure 0^+ and the mixed case no longer resemble each other.

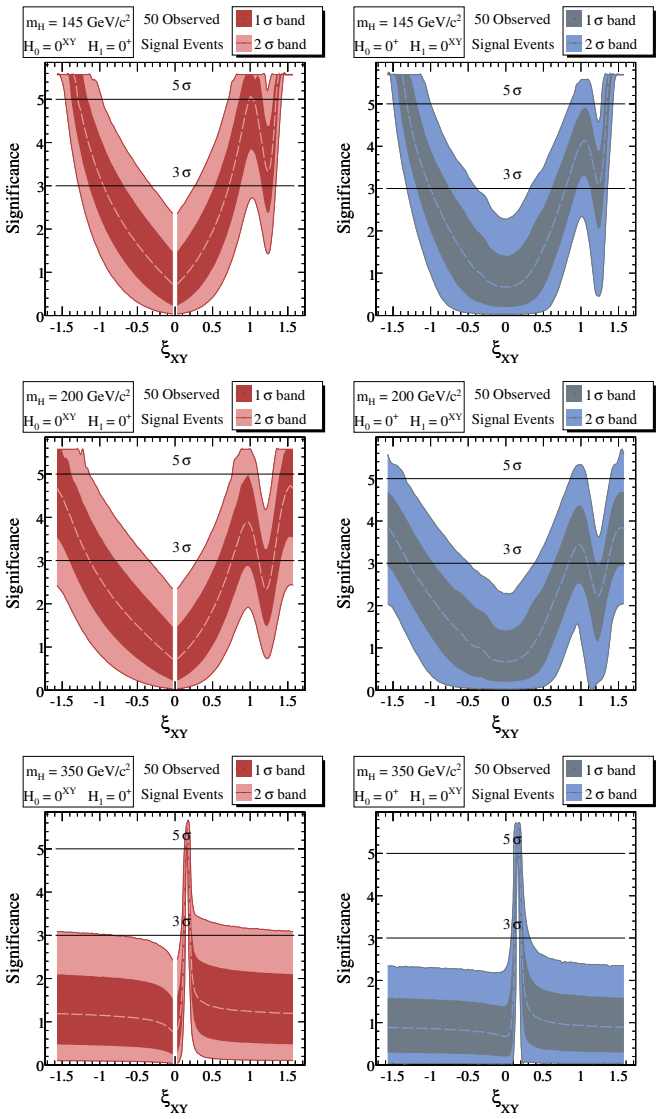

FIG. 45: Left: Expected significance for excluding values of ξ_{XY} in favour of a pointlike 0^+ ($\xi_{XY} = 0$), assumed to be correct. Right: Expected significance for excluding a pointlike 0^+ in favour of a "composite" one ($\xi_{XY} \neq 0$), assumed correct for the ξ_{XY} values on the x-axis, for $m_H = 145$, 200 and 350 GeV/c^2 (top, middle and bottom) and $N_S = 50$.

As we discussed for the C- and CP- violating cases, an additional question is whether one can distinguish a composite scalar from other mixed scalars. We find that, compared to the composite case, the two other mixed cases are nearly identical. The results for the distiction between the CP-violating and composite cases are shown in Fig. 46. For large values of ξ_{XY} and ξ_{XP}, it is possible to distinguish between the two hypotheses at a large significance with a mere $N_S = 50$. For $m_H = 350$ GeV/c^2, the composite scalar is very similar to the pointlike 0^+ –and cannot be distinguished from it– except if ξ_{XY} is near its critical point.

Replacing the CP-violating scalar with the C-violating one yields results nearly identical to the ones in Fig. 46.

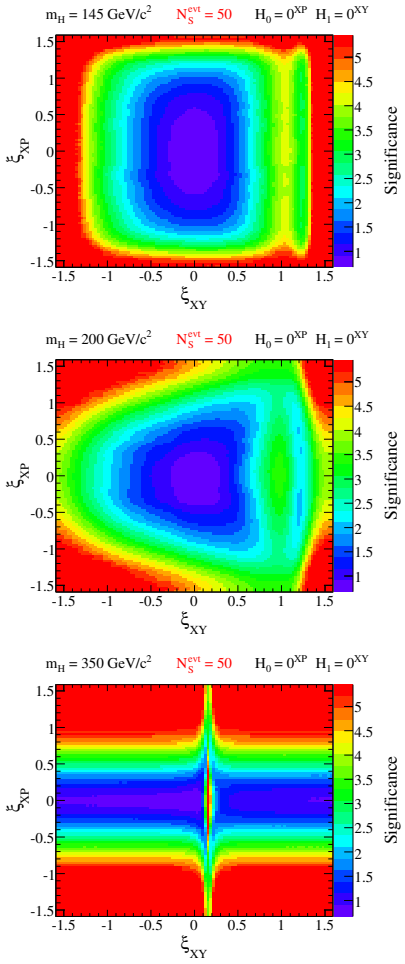

FIG. 46: The median expected significance (colored-labeled z-"axis") for excluding values of ξ_{XP} (y-axis) in favour of the composite scalar assuming it to be correct with the ξ_{XY} values of the x-axis, for $m_H = 145$, 200 and 350 GeV/c^2 (top, middle and bottom) and $N_S = 50$.

F. 0^+ vs. general $J = 1$

In Sec. VII B we discussed the prospects for distinguishing a 0^+ from the two pure J^{PC} spin-one objects, vector and axial-vector. Here, we address a more general question: how well can one distinguish between 0^+ and the general family of $J = 1$ states?

The most general vertex describing the coupling of a $J = 1$ particle a Z pair, see Eq. D can be parametrized, for non-vanishing X, P and Q, as:

$$\mathcal{L}^{\rho\mu\alpha} \propto \cos\xi\,(g^{\rho\mu}p_1^{\alpha} + g^{\rho\alpha}p_2^{\mu}) + e^{i\delta}\sin\xi\,\epsilon^{\rho\mu\alpha}(p_1 - p_2), \tag{37}$$

in terms of two mixing angles ξ and δ.

The mixing between the pure vector and axial couplings is described by ξ, while δ parametrizes the mixing between the CP- and C-violating parts of the interference term in the matrix element squared. In order to quantify the significance at which one can distinguish between the 0^+ hypothesis and the general $J = 1$ case, we consider two different types of tests, which answer two similar questions.

The first question is assuming a 0^+ resonance to be the correct choice, at what significance can we exclude values of ξ and δ for a $J = 1$ hypothesis? To answer, we perform a series of simple hypothesis tests, for each set of fixed values ξ and δ, between the two hypotheses: the test statistic is $\Lambda = \log[L(0^+)/\mathcal{L}(\xi,\delta)]$. The results, as a function of ξ for $\delta = \pi/2$ and $m_H = 350$ GeV/c^2, are shown in Fig. 47. The points $\xi = 0$ and $|\xi| = \pi/2$ correspond to the pure vector and pure axial-vector limits, respectively, and are consistent with Figs. 22 and 23 on these pure cases.

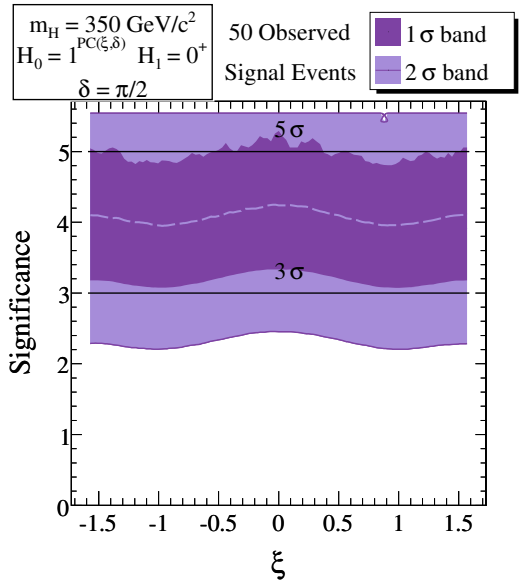

FIG. 47: Expected significance for excluding values of ξ, for $\delta = \pi/2$, in the general $J = 1$ hypothesis [dubbed $1^{\mathrm{PC}(\xi\delta)}$] in favour of the 0^+ one, assumed to be correct. Results for $m_H = 350$ GeV/c^2 and $N_S = 50$. The dashed line is the median significance. The 1 and 2 σ bands correspond to 68% and 95% median-centered confidence intervals.

The second question is if a $J = 1$ resonance with given ξ and δ is the correct choice, at what significance can we exclude the 0^+ case in favour of $J = 1$? Again, we have to treat ξ and δ as nuisance parameters, since we are considering the *general* $J = 1$ case. The statistic is $\log[\max L(\hat{\xi},\hat{\delta})/L(0^+)]$. The results, as functions of ξ for $\delta = \pi/4$ and $m_H = 200$ GeV/c^2, are given in Fig. 48, which shows that one can potentially exclude the 0^+ hypothesis without knowing the actual values of ξ and δ. Prospects for measuring these angles are discussed in Sec. VII G.

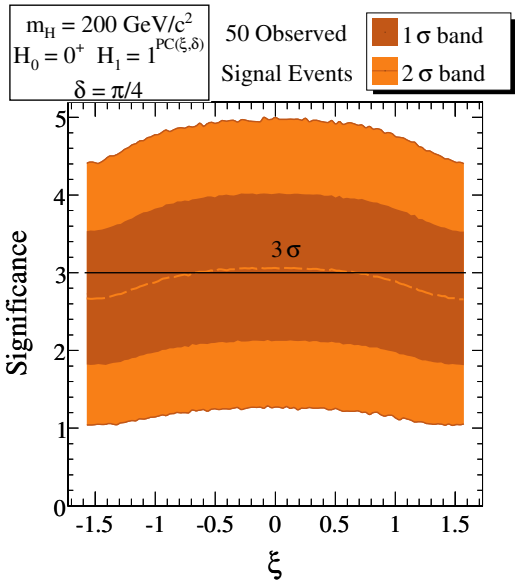

FIG. 48: Expected significance for excluding the 0^+ hypothesis in favour of the general $J = 1$ case [dubbed $1^{\mathrm{PC}(\xi\delta)}$], assumed correct for ξ as in the x-axis and $\delta = \pi/4$. Results for $m_H = 200$ GeV/c^2 and $N_S = 50$. The dashed line and bands are as in Fig. 47.

In Fig 49 we show the expected significance for the distinction between the 0^+ and the general $J = 1$ cases, as a function of ξ and δ, for $m_H = 145$, 200, and 350 GeV/c^2. Notice that the significance levels colour-coded as a z-"axis" range over a small interval. This means that the entire $J = 1$ family is almost "equally dissimilar" to 0^+. In general, one's ability to exclude $J = 1$ relative to 0^+ is greater than its opposite, due to the required treatment of ξ and δ as nuisance parameters, although the differences are relatively small in magnitude and in ξ- and δ-dependence.

The fact that the significance plane as a function of ξ and δ is relatively flat means that, with some m_H-dependent amount of observed events, one shall be able to unambiguously exclude the general $J = 1$ hypothesis in favour of the 0^+ case (assuming it to be correct) or vice-versa, regardless of the values of ξ and δ. Using the pure J^{PC} hypothesis test as a guide, we conclude that the median expectation for differentiating between 0^+ and $J = 1$ should exceed $5\,\sigma$ with $N_S \sim (60, 200, 85)$ events for $m_H = (145, 200, 350)$ GeV/c^2, respectively.

Additionally, based on our results concerning the distinction between 0^- and the two pure $J = 1$ states, and

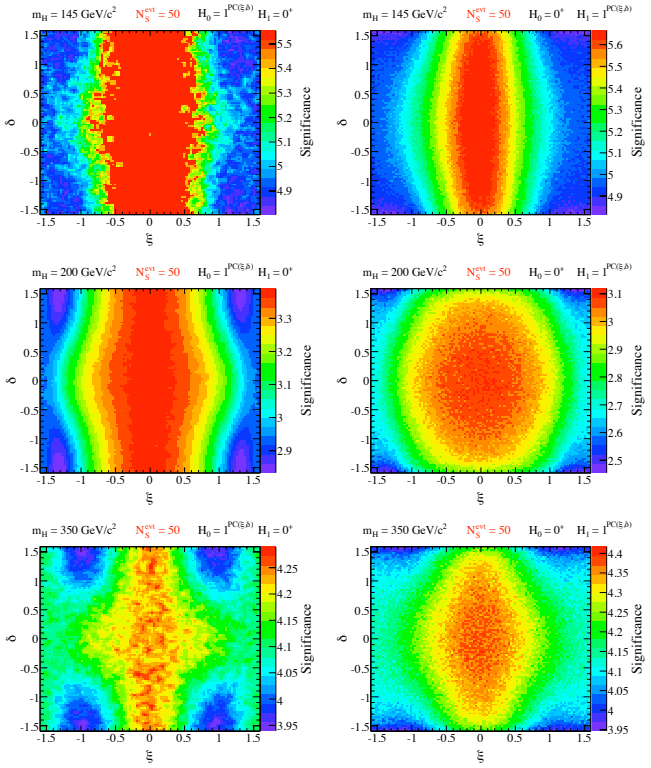

FIG. 49: Left: Median expected significance (coloured z-"axis") for excluding values of ξ and δ corresponding to a $J = 1$ hypothesis [dubbed $1^{\mathrm{PC}(\xi\delta)}$] in favour of 0^+, if the latter is correct. Right: vice-versa, with values of ξ and δ indicated on the axes. Results for $m_H = 145$, 200 and 350 GeV/c^2 (top, middle and bottom), for $N_S = 50$.

the results on the mixed $J = 0$ hypotheses, we conclude that it is equally easy, or even easier, to distinguish between $J = 1$ and a $J = 0$ state other than 0^+. Hence, with the numbers of events listed above, it is likely that one will be able to unambiguously exclude the $J = 1$ family of hypotheses in favour of a general $J = 0$ hypothesis, or vice-versa, if the resonance is either one or the other.

G. Parameter estimation in mixed J = 0 and J = 1 cases

Were one to find out from real data and the hypothesis tests discussed in the previous section that a mixed $J = 0$ or $J = 1$ state is the preferred description, the next item in the context of this analysis would be the measurement of its mixing parameters (in a larger context one would include at this stage the measurement of decay branching ratios).

We have seen in Secs. VII E and VII F that our hypothesis tests can demonstrate –if correct– and with computable significance, that a standard 0^+ particle is disfavoured relative to a mixed scalar or vector with unspecified HZZ coupling ratios (or mixing angles). In these tests, the angles were treated as nuisance parameters. Their measurement proceeds along the same line –the preferred value is simply that which maximizes the likelihood– but the treatment of confidence intervals need be different.

More specifically, each mixed hypothesis family is characterized by mixing angles $\vec{\rho}$. For each "experiment", N events are simulated, each one characterized by a vector $\vec{x}_e = \{\vec{\omega}, \vec{\Omega}, M_{Z^*}\}|_e$. The likelihood for a particular family of hypotheses is $\mathcal{L}(\vec{\rho}) = \prod_{e=1}^{N} P_e(\vec{x}_e, \vec{\rho})$. The measured values of the mixing angles, $\vec{\rho}_{\mathrm{meas}}$, are chosen to be those that maximize the likelihood.

To assign confidence intervals to these measurements we use a fully frequentist approach. An ensemble of "experiments" is performed with fixed input values $\vec{\rho} = \vec{\rho}_{\mathrm{input}}$. For each experiment, the measured values of $\vec{\rho}$ are taken from the maximization of the likelihood. This procedure is repeated for a fine matrix of input values, covering the allowed parameter space. From the probability distribution functions $P(\vec{\rho}_{\mathrm{meas}}|\vec{\rho}_{\mathrm{input}})$, estimated using this ensemble of experiments, the Feldman-Cousins unified approach [44] is used to choose which elements of probability are included in confidence intervals.

As an example, consider the CP-violating scalar case, discussed in Sec. VII E. The confidence intervals for measured values of ξ_{XP} (the mixing parameter that characterizes this hypothesis) are shown in Fig. 50 for different values of m_H. The way to interpret these figures is as follows: For a particular set of data –one experiment, which in this case includes $N_S = 50$ observed events– an input value of ξ_{XP} (to be read on the x-axis) results in a measured value to be read (with its error bands) on the y axis. The confidence intervals are obtained by drawing a horizontal line passing through the measured ξ_{XP}. The overlap of this line with the $n\,\sigma$ bands dictates which values of "input ξ_{XP}" should be included in the $n\,\sigma$ confidence intervals. For example, for $m_H = 200$ GeV/c^2 (middle of Fig. 50) we see that, if $\xi_{XP}^{\mathrm{meas}} = 0$, the $3\,\sigma$ confidence interval is approximately $\xi_{XP} \in [-1, 1]$.

The $1\,\sigma$ bands in Fig. 50 are centered on the diagonal $\xi_{XP}^{\mathrm{meas}} = \xi_{XP}^{\mathrm{input}}$, implying that there is no significant bias in the measurement. In addition to this, the $2\,\sigma$ and $3\,\sigma$ bands also cover most of the diagonal $\xi_{XP}^{\mathrm{meas}} = -\xi_{XP}^{\mathrm{input}}$. This confirms our observation from Sec. VII F that our ability to pin down this parameter comes predominantly from measuring the relative strengths of the 0^+ and 0^- parts of the *pdf* rather than the nature ($\tilde{T}$-odd) of its interference term. An increased number of observed events is needed to fully resolve this sign ambiguity.

In Fig. 50 we see that for $m_H = 145$ GeV/c^2 (but not for $m_H = 200$ GeV/c^2) the size of the confidence intervals for ξ_{XP} decreases with increasing $|\xi_{XP}|$. This is due to the effective coupling strengths of the 0^+ and 0^- parts of the *pdf* differing by a factor of ~ 10 at $m_H = 145$ GeV/c^2 but not at the other masses. Hence, at the lowest mass, only at $\tan^2(\xi_{XP}) \sim 10$ does the *pdf* exhibit 0^+- and 0^--like behaviours of similar magnitude.

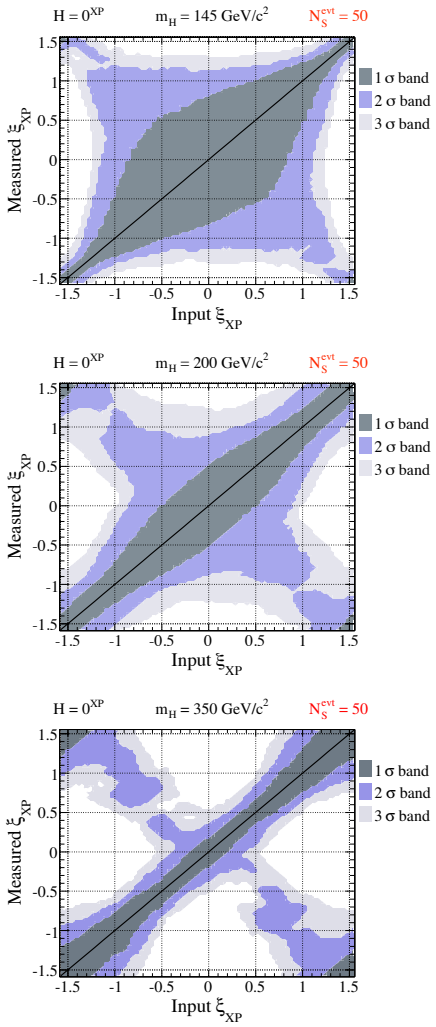

FIG. 50: Confidence intervals for measured values of ξ_{XP} for a CP-violating $J = 0$ resonance, for $m_H = 145$, 200 and 350 GeV/c^2 (top, middle and bottom), all for $N_S = 50$. For measured values of ξ_{XP} on the y-axis, confidence intervals should be read horizontally, see text.

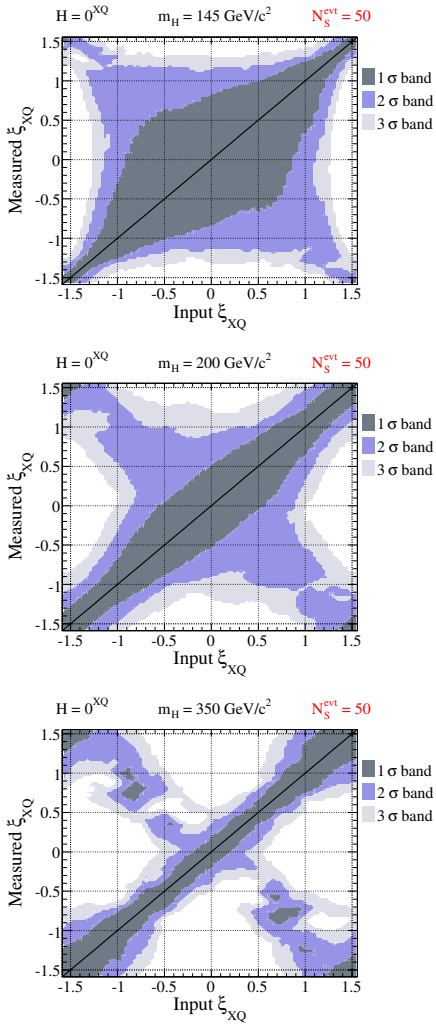

FIG. 51: Confidence intervals for measured values of ξ_{XQ} for a C-violating $J = 0$ resonance for $m_H = 145$, 200 and 350 GeV/c^2 (top, middle and bottom), all for $N_S = 50$. For measured values of ξ_{XP} on the y-axis, confidence intervals should be read horizontally, see text.

Confidence intervals for measurements of the parameter ξ_{XQ} for a scalar with C-violating HLL couplings are shown in Fig. 51, very similar to Fig. 50. Both hypotheses involve mixtures of 0^+ and 0^- couplings, only their interference is different, and we already learned that its effects are small. For the C-odd case, the sign ambiguity of $\xi_{XQ}^{\rm meas}$ is slightly worse than for the $\tilde{T}$-odd one, see the $1\,\sigma$ confidence bands appearing on the $\xi_{XQ}^{\rm meas} = -\xi_{XQ}^{\rm input}$ diagonal for $m_H = 350$ GeV/c^2. This is also expected, since the C-odd interference term is proportional to the relatively small number $\eta \approx 0.15$, see Eq. 24.

One's ability to distinguish between $J = 0$ C- and $\tilde{T}$-odd admixtures relies on the resolution of the interference terms. With a factor of 10 more statistics ($N_S \sim 500$), one would be able to resolve the sign ambiguity in ξ_{XP} and ξ_{XQ} and to distinguish between the two cases.

The confidence intervals associated with measurements of ξ_{XY} for a composite scalar are shown in Fig. 52. We observe that, for $m_H = 145$ and 200 GeV/c^2, the $1\,\sigma$ intervals are centered on the diagonal $\xi_{XY}^{meas} = \xi_{XY}^{input}$. There are no bands along $\xi_{XY}^{meas} = -\xi_{XY}^{input}$, since the interference term is of a different nature than that of the discrete-symmetry violating cases. The extensions of the 2 and $3\,\sigma$ bands along almost horizontal and vertical lines around $\xi_{XY} \sim 1.3$ result from large cancellations in the *pdf*, discussed in Sec. VII E.

The figure for $m_H = 350$ GeV/c^2 is hard to decipher. With a magnifier one sees that at the critical value of ξ_{XY} the confidence intervals are tiny. Everywhere else, the intervals essentially include all possible values *except* the critical one. This is tantamount to saying that at this mass we cannot tell, on the basis of our analysis, a com-

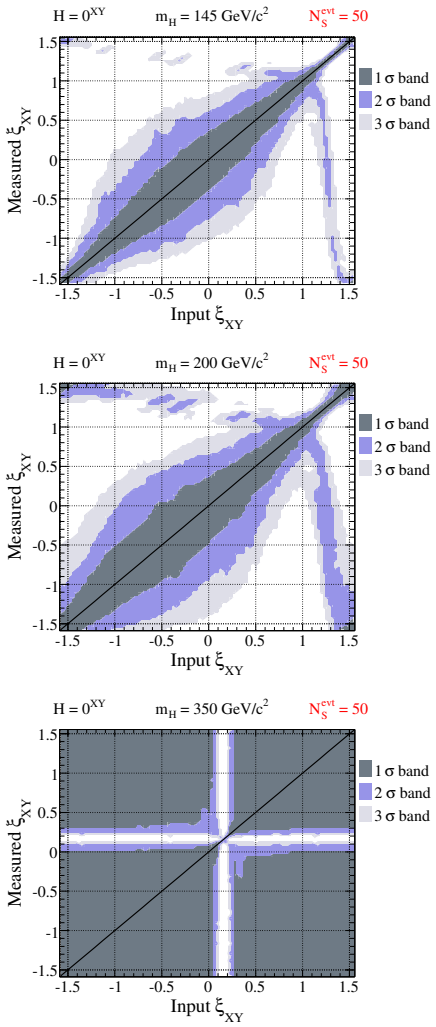

FIG. 52: Confidence intervals for measured values of ξ_{XY} for a "composite" $J = 0$ resonance, for $m_H = 145$, 200 and 350 GeV/c^2 (top, middle and bottom), all for $N_S = 50$. For measured values of ξ_{XY} on the y-axis, confidence intervals should be read horizontally.

posite from a pointlike scalar *unless* is has a particular value of ξ_{XY}, a fact made clearer by Fig. 45.

The other mixed case we study is that of a general $J = 1$ resonance, parameterized by angles ξ and δ as described in Sec. VII F. We saw in Sec. VII D, that most difficult distinction is the one between the two pure J^{PC} spin-one resonances, indicating that these two cases are very similar. This is what we find again when exploring the potential for measuring ξ and δ.

In Figs. 53 we show as an example the confidence intervals for measurements of ξ and δ at $m_H = 145$ GeV/c^2 The ability to resolve the value of the P-mixing angle ξ is modest. The measurement of the CP-mixing angle δ is still harder. Specifically, we see a large sign ambiguity in the measured δ, indicating that, with $N_S \sim 50$, it is difficult to resolve the nature of the interference term, as was the case for $J = 0$.

Overall, we find that a precise measurement of ξ and δ for a $J = 1$ resonance is very difficult. The conclusion of this section and Sec. VII F is that, if a new $J = 1$ boson is found, a modest number of events will suffice to exclude $J = 0, 2$ alternatives with high significance. Before many more events are gathered, and with only the tools we have studied, it is hard to make precise statements about the nature of a $J = 1$ resonance, other than its spin.

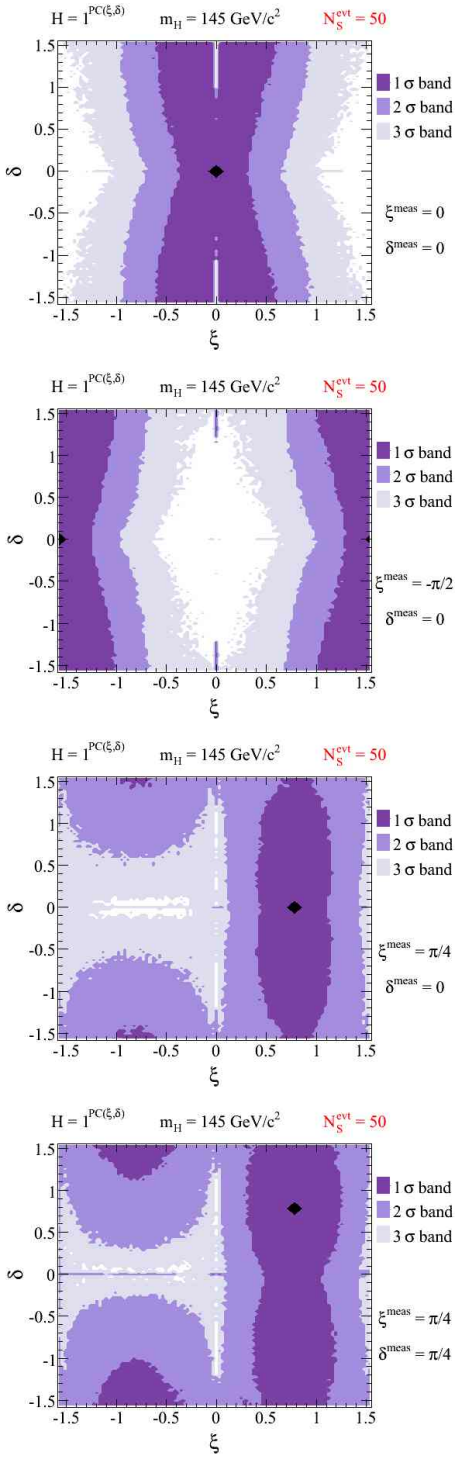

FIG. 53: Confidence intervals for measured values of ξ and δ for a $J = 1$ resonance with a mass 145 GeV/c^2 and $N_S = 50$ events. The input values, indicated by diamonds, are reported alongside the figures.

VIII. CONCLUSIONS, CAVEATS, AND OUTLOOK

It is no surprise that using all of the decay information in a data sample provides better discrimination of the identity of a new heavy resonance than examining a single angular distribution or asymmetry. Nevertheless, one might be tempted, given a small data set constituting an initial discovery, to settle for a stripped-down analysis. Our study quantifies the cost, in units of integrated LHC luminosity, of pursuing such sub-optimal analysis strategies, as illustrated in Fig. 54 for the benchmark $m_H = 200$ GeV/c^2.

In this figure we compare the discrimination between the 0^+ and 1^- hypotheses for likelihood definitions that exploit different sets of variables, with the notation that $P(a_1, \cdots, a_N)$ denotes N-dimensional *pdfs* in the correlated variables $\{a_1, \cdots, a_N\}$. Here $\prod_i P(X_i)$ is constructed from one-dimensional *pdfs* for all variables, ignoring (erroneously) their correlations. $P(\vec{\omega}\,|\langle\vec{\Omega}\rangle_{\rm TH})$ are *pdfs* including the variables $\vec{\omega}$ and their correlations, but with the hypothesis 1^- represented by a *pdf* in which the variables $\vec{\Omega} = \{\Phi, \cos\Theta\}$ have been integrated out.

The likelihood $P(\vec{\omega}\,|\langle\vec{\Omega}\rangle_{\rm TH})$ performs badly even relative to $P(\vec{\omega})$, which uses fewer angular variables. The two differ only in that the first construction implicitly assumes a uniform $4\,\pi$ coverage of the observed leptons (an assumption customary in the literature) as if the muon p_T and η analysis requirements did not depend on the $\vec{\Omega}$ angular variables. The differing results arise from the strong correlation between the variables Φ and ϕ in the $J = 1$ *pdfs*, which causes phase space acceptance sculpting of the Φ distribution that, in turn, alters the ϕ distribution, as discussed in Sec.IV and VII B.

Additionally we find that treating the correlated angular variables as uncorrelated, as in the $\prod_i P(X_i)$ example of Figure 54, not only degrades the discrimination significance but also produces a real chance of falsely labeling the quantum numbers of the new resonance.

Overall, we have demonstrated that small signal samples in the $ZZ \to 4\ell$ or $ZZ^* \to 4\ell$ decay channels, as might be available at the moment of discovery, could be sufficient to characterize a putative Higgs particle. Below we summarize these results in more detail.

A. Summary of pure case discrimination

Amongst the many comparisons considered in our analysis, the ones between simple hypotheses are the most readily summarized. This we do in Tables I,II for $m_H = 145$ GeV/c^2 for all pure-case comparisons between $J = 0, 1$ parent particles, and in Tables III,IV (V,VI) for $m_H = 200$ (350) GeV/c^2, for all pure-case comparisons between $J = 0, 1, 2$ parent particles.

Overall, the discrimination power of the hypothesis tests is very impressive. The $m_H = 200$ GeV/c^2 bench-

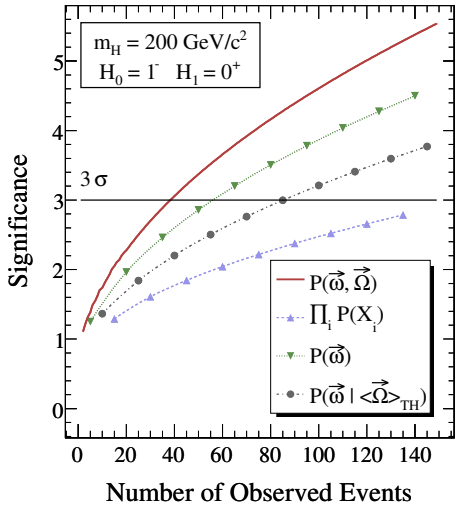

FIG. 54: Median expected significance for rejecting 1^- in favour of 0^+, for different likelihood constructions used in the likelihood ratio test statistic. H_0 is always the considered true hypothesis.

$\mathbb{H}_0 \Downarrow \mathbb{H}_1 \Rightarrow$	0^+	0^-	1^-	1^+
0^+	–	17	12	16
0^-	14	–	11	17
1^-	11	11	–	35
1^+	17	18	34	–

TABLE I: Minimum number of observed events such that the median significance for rejecting $\mathbb{H}_0$ in favour of the hypothesis $\mathbb{H}_1$ (assuming $\mathbb{H}_1$ is right) exceed $3\,\sigma$ with $m_H = 145$ GeV/c^2

$\mathbb{H}_0 \Downarrow \mathbb{H}_1 \Rightarrow$	0^+	0^-	1^-	1^+
0^+	–	52	37	50
0^-	44	–	34	54
1^-	33	32	–	112
1^+	54	55	109	–

TABLE II: Same as Table I, but requiring that the median significance exceed $5\,\sigma$.

mark example is the one requiring the largest statistics to reach a given discrimination at a given level of confidence. Compared with the $m_H = 350$ GeV/c^2 case, this is because various coefficients of the angular dependences vanish at the $m_H = 2\,M_Z$ threshold. The $m_H = 145$ GeV/c^2 example fares better than the 200 GeV/c^2 one for the same reason, amplified by the extra lever-arm supplied by a non-trivial M_{Z^*} distribution.

The Tables also show that the discrimination power between two given hypotheses is approximately symmetric under the interchange of 'right' and 'wrong'. Telling 1^+ from 1^- is always difficult but not impossible, a fact

$\mathbb{H}_0 \Downarrow \mathbb{H}_1 \Rightarrow$	0^+	0^-	1^-	1^+	2^+
0^+	–	24	45	62	86
0^-	19	–	19	19	38
1^-	40	18	–	90	48
1^+	56	19	85	–	66
2^+	86	45	54	70	–

TABLE III: Minimum number of observed events such that the median significance for rejecting $\mathbb{H}_0$ in favour of the hypothesis $\mathbb{H}_1$ (assuming $\mathbb{H}_1$ is right) exceed $3\,\sigma$ with $m_H = 200$ GeV/c^2

$\mathbb{H}_0 \Downarrow \mathbb{H}_1 \Rightarrow$	0^+	0^-	1^-	1^+	2^+
0^+	–	76	146	203	287
0^-	59	–	60	61	123
1^-	130	57	–	297	156
1^+	182	58	278	–	217
2^+	287	146	178	230	–

TABLE IV: Same as Table III, but requiring that the median significance exceed $5\,\sigma$.

$\mathbb{H}_0 \Downarrow \mathbb{H}_1 \Rightarrow$	0^+	0^-	1^-	1^+	2^+
0^+	–	8	21	24	11
0^-	9	–	22	22	36
1^-	24	22	–	81	46
1^+	26	22	80	–	56
2^+	15	39	55	73	–

TABLE V: Minimum number of observed events such that the median significance for rejecting $\mathbb{H}_0$ in favour of the hypothesis $\mathbb{H}_1$ (assuming $\mathbb{H}_1$ is right) exceed $3\,\sigma$ with $m_H = 350$ GeV/c^2

of relevance for a Z' look-alike analysis. The level of significance does not obey a naïve $N(\sigma) \propto \sqrt{N_S}$ law. However we find by inspection that an approximation of the form $N(\sigma) = a + b\sqrt{N_S}$ works well, allowing one to extrapolate to larger numbers of events than presented here.

Other lessons from the Tables are case-by-case specific, reflecting the mass-dependent quantum-mechanical entanglement between the decay variables. Some examples are: distinguishing the 'natural-parity' $J = 0^+$ and 1^-

$\mathbb{H}_0 \Downarrow \mathbb{H}_1 \Rightarrow$	0^+	0^-	1^-	1^+	2^+
0^+	–	25	67	77	35
0^-	26	–	68	68	118
1^-	76	68	–	268	149
1^+	83	68	263	–	184
2^+	46	127	181	240	–

TABLE VI: Same as Table V, but requiring that the median significance exceed $5\,\sigma$.

hypotheses for $m_H = 145$ GeV/c^2 requires only a dozen signal events for $3\,\sigma$ discrimination. For 200 GeV/c^2, discriminating 0^+ from 0^- is relatively easy, but distinguishing 0^+ from 2^+ is difficult. For 350 GeV/c^2, contrariwise, 2^+ is relatively easy to disentangle from 0^+, but not from 0^-.

B. Summary of mixed cases, CP and compositeness discrimination

We find that direct sensitivity to CP odd, parity odd XP interference effects, or to CP odd, parity even XQ interference effects, will require signal samples about an order of magnitude larger than considered here. We have also observed that with much smaller statistics it may be possible to conclude that a mix of X and P (or X and Q) couplings is favored over just the pure X (i.e. 0^+) or pure P (i.e. 0^-) couplings alone. Such a conclusion would be tantamount to demonstrating CP violation in the Higgs sector. However this scenario relies on large CP violation, and even in this favorable case one cannot tell an X and P mixture from an X and Q mixture without more data than what is required to establish discovery.

In the case of a composite Higgs, it may be conceivable that the Higgs is as 'soft' as a pion, in the sense of having an inverse radius and a mass of comparable magnitude. In this scenario we have seen that the angular distributions associated to the X and Y couplings are similar after integrating over the decay angles. As a result there can be strong destructive interference between these contributions. For our lighter mass benchmarks we find good discrimination of pure 0^+ from the mixed composites. For the heavier $m_H = 350$ GeV/c^2 example, discrimination based on decay angles is poor unless the strong interference effects are present; here we also observed that substantial enhancement or suppression of the HLL$\to ZZ$ branching fraction can provide another important discriminator.

For mixed cases, one could worry that certain combinations of exotic couplings might let an HLL successfully masquerade as a 0^+ Higgs, even when all the pure case exotics are excluded. For of spin 1 HLLs we have shown that this does not happen. In fact we find that when we have an SM Higgs, the entire family of mixed coupling spin 1 HLLs can be excluded at approximately the same expected level of significance as for the pure 1^- or 1^+ cases. An even stronger result is that the general spin 0 hypothesis can be conclusively discriminated from the general spin 1 hypothesis, at or close to the moment of discovery.

C. Limitations of our analysis

As mentioned before, our analysis used an approximate factorization between observables related to Higgs (or HLL) production and observables related to decay. We have here focused only on decay information, though it would be straightforward to include at least rudimentary production information about cross sections and the relative weighting of different partonic initial states.

A more complicated issue is how to account for higher order electroweak radiative corrections, which modify the angular distributions expected for the SM Higgs, as well as some of the otherwise pure case look-alikes. On a first pass, ignoring these corrections is not unreasonable for small signal samples, where the HLL discrimination is statistics limited. To include these corrections, we imagine three viable strategies:

- Compute the relevant radiative corrections and include them in the signal pdfs.
- Parametrize the corrections as extra nuisance parameters in the hypothesis tests, or use the hypothesis testing to extract them from the data.
- Treat the signal radiative corrections as part of the background, parametrized in the ${}_sPlots$ fits. This approach is motivated by the observation that the electroweak corrections to the SM Higgs golden channel, in addition to being small in magnitude, also seem to have less angular variation [35].

There are also large QCD radiative corrections which affect both the total cross section and the p_T and η distributions of the Higgs. To first approximation the decay distributions in the Higgs rest frame are unaltered, however since the lab frame p_T and η distributions of the final state leptons are modified, we expect also corrections to the phase space acceptance effects on the angular distributions. This uncertainty is side-stepped by taking the p_T and η distributions from data rather than theory.

It would also be useful to include a more comprehensive treatment of the SM backgrounds to the golden model channel. Here we only considered the dominant ZZ background, and only at leading order.

Our treatment of couplings and HLLs was not exhaustive, since we have ignored gauge invariant operators with dimension > 6, have only examined special cases of spin 2 HLLs, and have not even mentioned the possibility of HLLs with spins > 2. At some point Occam's razor obviates the need for such comparisons: "Raffiniert ist der Herr Gott, aber boshaft ist Er nicht", to quote a known author [46].

The likelihood analyses pursued here are very computing intensive, since $5\,\sigma$ discrimination implies simulating enough pseudo-experiments to fill out the what amounts to the $5\,\sigma$ tails in multidimensional likelihood distributions, where typically the likelihood distributions are highly non-Gaussian. The analysis presented here utilized more than 10^{14} pseudo-experiments in total.

D. Outlook

We have seen that by exploiting the full decay information in the golden channel we should be able to say a lot about the identity of a putative Higgs resonance around the moment of discovery. Our results also show that asymptotically, utilizing the full physics run of the LHC, it should be possible to explore very detailed properties of such a resonance.

It has not escaped our attention that there are many processes other than the ZZ decays of a heavy resonance whose characterization may benefit from an analysis of the kind that we have performed here.

Note added:

While this manuscript was in preparation we received the preprint [47], reporting on an analysis similar to what we have presented here. At first look the two studies appear to be both compatible and complementary.

Acknowledgments

We dedicate this work to the memory of our colleagues Andrew Lange and Juan Antonio Rubio. The authors are grateful to Keith Ellis, Guido Martinelli, Sezen Sekmen, Raman Sundrum, Steven Weinberg, and Jan Winter for useful discussions. JL acknowledges the hospitality of the CERN Theory Department and support from the Aspen Center for Physics. Fermilab is operated by the Fermi Research Alliance LLC under contract DE-AC02-07CH11359 with the U.S. Dept. of Energy. CR and MS are supported in part by the U.S. Dept. of Energy under contract DE-FG02-92-ER40701.

Appendix A: $\mathbf{SU(2)_L \times U(1)_Y}$ gauge-invariant couplings

To write Lagrangians generating the couplings of §III A and respecting the electroweak gauge symmetry one must specify the electroweak charges of the Higgs look-alikes. Consider the example of HLLs that are "neutral", i.e. are weak singlets and have zero hypercharge. For the scalar case, in a conventional notation for isovector and isoscalar gauge fields, the lowest-dimensionality Lagrangian density is:

$$L = \frac{1}{\Lambda} H \left(A_1 \vec{W}_{\mu\alpha} \vec{W}^{\mu\alpha} + A_2 B_{\mu\alpha} B^{\mu\alpha}\right) + \frac{1}{\Lambda} H\, i\, \epsilon^{\mu\alpha\sigma\tau} (A_3 \vec{W}_{\mu\alpha} \vec{W}_{\sigma\tau} + A_4 B_{\mu\alpha} B_{\sigma\tau}), \quad \text{(A1)}$$

with A_i arbitrary constants and Λ a mass parameter. This object generates, amongst others, the couplings of Eq. (10). The "true" dimensionality of the operators in Eq. (10) is that of the ones appearing in Eq. (A1), i.e. $d = 5$.

The form of Eq. (A1) results in a coupling $H Z_{\mu\alpha} Z^{\mu\alpha} \rightarrow 2\, p_1 \cdot p_2\, g_{\mu\alpha} - 2\, k_\mu k_\alpha$, establishing a relation between X and $Y + i\, Z$ in Eq. (10). We do not impose it, for it is not general even at tree level. Consider, for instance, a model with a conventionally-charged but otherwise non-standard HLL, dubbed Φ before the spontaneous symmetry breaking. Call $V_{\mu\nu}$ any of the field tensors in Eq. (A1). The operators in this Lagrangian could be "descendants" of dimension 6 operators of the form $\Phi^\dagger \Phi\, V^2$, with $\Phi \rightarrow H + v$, see. e.g. [22]. In such a case there would be a standard-like $g_{\mu\nu}$ coupling *plus* the one induced by the higher-dimensional operators.

For the case of a spin-1 "neutral" HLL, H_ρ, the lowest-dimension gauge-invariant Lagrangian generating the couplings of Eq. (11) is built of operators of dimension 6:

$$\Lambda^2 L = (\partial^\mu H^\alpha + \partial^\alpha H^\mu)\,(A_1\, \vec{W}^\lambda_\mu\, \vec{W}_{\alpha\lambda} + A_2\, B^\lambda_\mu B_{\alpha\lambda}) + \epsilon^{\mu\nu\alpha\rho}[A_3(\vec{W}^\lambda_\mu \overline{D}_\alpha \vec{W}_{\nu\lambda}) H_\rho + A_4 (B^\lambda_\mu \overline{\partial}_\alpha B_{\nu\lambda}) H_\rho], \quad \text{(A2)}$$

where D_α is the covariant derivative and $(M\overline{D}_\alpha N) \equiv M D_\alpha N - (D_\alpha M) N$.

For a canonical-dimension spin-2 "neutral" HLL, $H_{\mu\nu}$, the lowest-dimension gauge-invariant Lagrangian has couplings of dimension 5:

$$L = \frac{1}{\Lambda} H_{\mu\nu}\,(A_1\, \vec{W}^\mu_\alpha\, \vec{W}^{\nu\alpha} + A_2\, B^\mu_\alpha B^{\nu\alpha}) + \frac{1}{\Lambda} H^{\nu\rho}\, i\, \epsilon_{\mu\nu\alpha\beta} (A_3\, \vec{W}^{\mu\alpha} \vec{W}^{\rho\beta} + A_4\, B^{\mu\alpha} B^{\rho\beta})\,. \quad \text{(A3)}$$

The consideration of gauge-invariant constructions for HLLs with non-trivial electroweak charges would take us well beyond the scope of this paper.

Appendix B: Phase space for ZZ*

In the case in which one of the two Z bosons is off-shell, the dependence on its mass (M_{Z^*}, either m_1 or m_2) is an extra handle in determining the shapes of signal and backgrounds. Let $p_{\rm cms} \equiv |\vec{p}[Z]| = m_1\, \gamma_1\, \beta_1 = m_2\, \gamma_2\, \beta_2$ be the momentum of one or the other Z in the H center-of-mass system:

$$p_{\rm cms} = \frac{1}{2\, m_H}\, \Theta[m_H - (M_Z + M_{Z^*})] \times \sqrt{m_H^2 - (M_Z - M_{Z^*})^2}\, \sqrt{m_H^2 - (M_Z + M_{Z^*})^2} \quad \text{(B1)}$$

Let $\mathcal{M}$ be the matrix element for the process. The expectation for the rate of events, including the dependence on M_{Z^*}, is:

$$\frac{dN}{d\cos\theta_1\, d\cos\theta_2\, d\phi\, d\cos\Theta\, d\Phi\, dM_{Z^*}} \propto |\mathcal{M}|^2\, \frac{M_{Z^*}\, p_{\rm cms}}{\left(M^2_{Z^*} - M^2_Z\right)^2 + M^2_{Z^*}\, \Gamma^2_Z}, \quad \text{(B2)}$$

with $|\mathcal{M}|^2$ an explicit function of c_1, c_2, ϕ, Θ, Φ and M_{Z^*} for each specific case to be discussed.

Appendix C: General results for spin 0 coupled to ZZ*

In Section III we have already written the angular distributions $d\Gamma[0^+]$ and $d\Gamma[0^-]$ for the pure scalar and pseudoscalar cases, see Eqs. 15, 16. We also discussed the T-odd and C-odd interferences between the standard coupling –proportional to X in Eq. (10)– and the P and Q terms of the same equation. Thus we defined $d\Gamma[0, \text{Todd}]$ and $d\Gamma[0, \text{Codd}]$ in Eqs. 23, 24. Similarly we discussed the complete result for the 'composite' case with $X \neq 0$ and $Y \neq 0$, defining $d\Gamma_{XY}$ and $d\Gamma_{YY}$ in Eqs. 26, 27. This allows us to gather the results corresponding to the most general deviations from the SM HIggs couplings:

$$\begin{aligned} d\Gamma[0] \quad &= X^2\, d\Gamma[0^+] + (P^2+Q^2)\, d\Gamma[0^-] \\ &+ X\,P\, d\Gamma[0,\text{ Todd}] + X\,Q\, d\Gamma[0,\text{ Codd}] \\ &+ X\,Y\, d\Gamma_{XY} + (Y^2+Z^2)\, d\Gamma_{YY}\,. \end{aligned} \tag{C1}$$

To obtain the complete spin 0 result one must add to Eq. (C1) the interferences between the non-standard terms themselves:

$$\begin{aligned} \Delta d\Gamma[0] \;&=\; XZ\, d\Gamma_{XZ} + YP\, d\Gamma_{YP} \\ &+\; YQ\, d\Gamma_{YQ} + ZP\, d\Gamma_{ZP} + ZQ\, d\Gamma_{ZQ}\,, \end{aligned} \tag{C2}$$

where

$$d\Gamma_{XZ} = 2\,\eta\, m_1^3\, m_2^3\, m_H^2\, \gamma_b^2\, (c_1+c_2)\, s\, s_1\, s_2\,, \tag{C3}$$

$$d\Gamma_{YP} = d\Gamma_{ZQ} = -2\, m_1^4\, m_2^4 \gamma_b^3\, s\, s_1 s_2 (c_1 c_2 + \eta^2)\,, \tag{C4}$$

$$d\Gamma_{YQ} = -d\Gamma_{ZP} = 2\,\eta\, m_1^4\, m_2^4\, \gamma_b^3\, c\,(c_1+c_2) s_1 s_2\,. \tag{C5}$$

Appendix D: General results for spin 1 coupled to ZZ*

We produce a spin 1 HLL from annihilation of $q\bar{q}$ with quark helicity $\tau/2$, $\tau = \pm 1$. To an excellent approximation the coupling of the HLL to light quarks must conserve helicity, so the antiquark has helicity $-\tau/2$. Then the HLL decays to ZZ (or ZZ^*), with $Z_2 \to \mu^-\mu^+$ with muon helicity $\sigma_2/2$ and $Z_1 \to e^-e^+$ with electron helicity $\sigma_1/2$.

The fully differential cross section is a sum over τ, σ_1, σ_2 of the squared absolute values of the helicity amplitudes. In addition the (unmeasured) helicities λ_1, λ_2 of Z_1, Z_2 are summed over 0, ± 1, before squaring.

We use the following notation to denote the helicity conserving coupling of a Z boson to a massless fermion of helicity $\sigma/2$, $\sigma = \pm 1$:

$$g_\sigma = \frac{1}{2}(c_v - \sigma c_a)\,. \tag{D1}$$

Similarly, we denote the helicity conserving coupling of a vector boson HLL to a massless fermion of helicity $\tau/2$, $\tau = \pm 1$:

$$g_\tau = \frac{1}{2}(g_v - \tau g_a)\,. \tag{D2}$$

In the full matrix element squared, the dependence on these vector-fermion-fermion couplings is

$$\begin{aligned} &\frac{1}{64}\Big[\left(c_v^2+c_a^2\right)^2\left(g_v^2+g_a^2\right) \\ &-2c_v c_a\left(c_v^2+c_a^2\right)\left(g_v^2+g_a^2\right)\,(\sigma_1+\sigma_2) \\ &-2\left(c_v^2+c_a^2\right)^2 g_v g_a\,\tau + 4c_v^2 c_a^2\left(g_v^2+g_a^2\right)\,\sigma_1\sigma_2 \\ &+4c_v c_a\left(c_v^2+c_a^2\right) g_v g_a\,(\sigma_1\tau+\sigma_2\tau) \\ &-8c_v^2 c_a^2 g_v g_a\,\sigma_1\sigma_2\tau\Big]\,, \end{aligned} \tag{D3}$$

from which we derive the shorthand notation

$$\begin{aligned} g_1 \;&\equiv\; \left(c_v^2+c_a^2\right)^2\left(g_v^2+g_a^2\right) \\ g_\sigma \;&\equiv\; -4c_v c_a\left(c_v^2+c_a^2\right)\left(g_v^2+g_a^2\right) \\ g_\tau \;&\equiv\; -2\left(c_v^2+c_a^2\right)^2 g_v g_a \\ g_{\sigma\sigma} \;&\equiv\; 4c_v^2 c_a^2\left(g_v^2+g_a^2\right) \\ g_{\sigma\tau} \;&\equiv\; 8c_v c_a\left(c_v^2+c_a^2\right) g_v g_a \\ g_{\sigma\sigma\tau} \;&\equiv\; -8c_v^2 c_a^2 g_v g_a\,. \end{aligned} \tag{D4}$$

We allow both Z bosons to be off-shell, with invariant masses m_1 and m_2. Some useful mass combinations are

$$\begin{aligned} &m_d^2 \equiv m_1^2 - m_2^2\,, \\ &M_1^2 \equiv m_H^2 - 3m_1^2 - m_2^2\,, \quad M_2^2 \equiv m_H^2 - m_1^2 - 3m_2^2\,, \\ &M_3^2 \equiv m_H^2 - 2(m_1^2+m_2^2)\,, \quad M_4^2 \equiv m_H^2 - (m_1^2+m_2^2)\,. \end{aligned} \tag{D5}$$

One of the advantages of using helicity amplitudes is that we can keep track of which contributions come from the longitudinal polarization of the HLL rather than the transverse polarizations. We use the notation ℓ^2, ℓ_0^2 to flag the parts of the squared matrix element that come from the transverse, longitudinal polarizations of the HLL, and $\ell\ell_0$ to flag contributions from the interference.

We define Θ to be the polar angle of the incoming quark with respect to the z-axis defined by Z_2 in the HLL rest frame. This raises a problem since at a pp collider we cannot distinguish the quark direction from the anitquark direction in a $q\bar{q}$-initiated process. A solution is to symmetrize the cross section between the case where Θ is the polar angle of the quark direction and the case where Θ is the polar angle of the antiquark. In the coupling notation defined in (D4), this symmetrization has the the same effect as setting g_τ, $g_{\sigma\tau}$, and $g_{\sigma\sigma\tau}$ to zero.

The standard convention in the literature has the coordinate axes chosen such that the outgoing muon moves along the y-axis in the rest frame of the HLL (or equivalently of Z_2). Thus the azimuthal angle of the muon is $\pi/2$, while the azimuthal angle of the outgoing electron

is denoted $\phi-\pi/2$. We denote the azimuthal angle of the incoming quark by Φ. This choice of conventions leads to rather awkward expressions for the angular distributions. A better choice is to align the axes such that the quark azimuthal angle $\Phi = 0$. The remaining azimuthal dependence is then denoted by φ_1 and φ_2, such that the substitutions $\varphi_1 \to \Phi + \phi$, $\varphi_2 \to \Phi$ regain the previous convention. We will employ this improved notation in this appendix, which makes the formulae more symmetrical.

After the quark-antiquark symmetrization described above, the XX and PP parts of the full matrix element squared is given by:

$$\begin{aligned}
&4m_1^2m_2^2X^2\gamma_b^2\Big[g_1S^2s_1^2s_2^2\left(2\ell_0^2m_d^4-\ell^2m_H^2\left[m_1^2\cos(2\varphi_1)+m_2^2\cos(2\varphi_2)\right]\right) \\
&+g_1\ell^2m_H^2(1+C^2)\left[2m_2^2s_1^2+2m_1^2s_2^2-(m_1^2+m_2^2)s_1^2s_2^2\right]+4\ell\ell_0g_1m_Hm_d^2\,C\,S\left[m_1c_1s_1s_2^2\sin\varphi_1-m_2c_2s_2s_1^2\sin\varphi_2\right] \\
&\quad -2\ell^2m_H^2m_1m_2s_1s_2\left((1+C^2)(g_1c_1c_2-g_{\sigma\sigma})\cos(\varphi_1-\varphi_2)+S^2(g_1c_1c_2+g_{\sigma\sigma})\cos(\varphi_1+\varphi_2)\right)\Big]\,,
\end{aligned} \tag{D6}$$

$$\begin{aligned}
&P^2\Big[\ell^2g_1m_H^2S^2s_1^2s_2^2\left[M_2^4m_1^2\cos(2\varphi_1)+M_1^4m_2^2\cos(2\varphi_2)\right] \\
&\quad+8\ell_0^2m_1^2m_2^2m_d^4S^2\left[g_1\left(c_1^2+c_2^2+s_1^2s_2^2\sin(\varphi_1-\varphi_2)^2\right)+2g_{\sigma\sigma}c_1c_2\right] \\
&\quad+(1+C^2)\ell^2g_1m_H^2\left[2M_1^4m_2^2s_1^2+2M_2^4m_1^2s_2^2-(M_2^4m_1^2+M_1^4m_2^2)s_1^2s_2^2\right] \\
&\quad-8\ell\ell_0m_Hm_d^2m_1m_2C\,S\left[M_2^2m_1s_2\left(g_1c_2s_1^2\sin\varphi_1\cos(\varphi_1-\varphi_2)+c_1(g_1c_1c_2+g_{\sigma\sigma})\sin\varphi_2\right)\right. \\
&\quad\left.-M_1^2m_2s_1\left(g_1c_1s_2^2\sin\varphi_2\cos(\varphi_1-\varphi_2)+c_2(g_1c_1c_2+g_{\sigma\sigma})\sin\varphi_1\right)\right] \\
&\quad+2\ell^2m_H^2M_1^2M_2^2m_1m_2s_1s_2\left[(1+C^2)(g_1c_1c_2-g_{\sigma\sigma})\cos(\varphi_1-\varphi_2)-S^2(g_1c_1c_2+g_{\sigma\sigma})\cos(\varphi_1+\varphi_2)\right]\Big]\,.
\end{aligned} \tag{D7}$$

The XP interference part is given by

$$\begin{aligned}
&4m_1m_2XP\,\gamma_b\Big[\ell^2g_1m_H^2S^2s_1^2s_2^2(M_1^2m_2^2\sin(2\varphi_2)-M_2^2m_1^2\sin(2\varphi_1)) \\
&\quad+2\ell\ell_0g_1m_Hm_d^2C\,S\left[m_2s_1^2c_2s_2(2m_1^2\sin\varphi_1\sin(\varphi_1-\varphi_2)-M_1^2\cos\varphi_2)\right. \\
&\qquad\left.-m_1s_2^2c_1s_1(2m_2^2\sin\varphi_2\sin(\varphi_1-\varphi_2)+M_2^2\cos\varphi_1)\right] \\
&\quad-2m_1m_2s_1s_2\left[(1+C^2)\ell^2m_H^2M_3^2(g_1c_1c_2-g_{\sigma\sigma})\sin(\varphi_1-\varphi_2)\right. \\
&\qquad\left.+m_d^2s^2(g_1c_1c_2+g_{\sigma\sigma})(\ell^2m_H^2\sin(\varphi_1+\varphi_2)+2\ell_0^2m_d^2\sin(\varphi_1-\varphi_2))\right] \\
&\quad-4\ell\ell_0m_Hm_1m_2m_d^2\,C\,S\left[m_2s_1\left(g_1c_1+g_{\sigma\sigma}c_2\right)\cos\varphi_1+m_1s_2\left(g_1c_2+g_{\sigma\sigma}c_1\right)\cos\varphi_2\right]\Big]\,.
\end{aligned} \tag{D8}$$

The XQ interference part is given by

$$\begin{aligned}
&4m_1m_2XQ\,\gamma_b\Big[\ell\ell_0g_\sigma m_Hm_d^2C\,S\left(m_2s_1^2s_2\left(2m_1^2\cos(\varphi_1-\varphi_2)\sin\varphi_1-M_1^2\sin\varphi_2\right)\right. \\
&\qquad\left.-m_1s_2^2s_1\left(2m_2^2\cos(\varphi_1-\varphi_2)\sin\varphi_2-M_2^2\sin\varphi_1\right)\right)
\end{aligned} \tag{D9}$$

$$+\ell^2 g_\sigma m_H^2(1+c^2)(M_1^2 m_2^2 s_1^2 c_2 + M_2^2 m_1^2 s_2^2 c_1) + m_1 m_2 s_1 s_2 \left[(1+C^2)\ell^2 g_\sigma m_H^2 m_d^2 (c_1-c_2)\cos(\varphi_1-\varphi_2)\right.$$
$$-g_\sigma s^2 (c_1+c_2)(\ell^2 m_H^2 M_3^2 \cos(\varphi_1+\varphi_2) + 2\ell_0^2 m_d^4 \cos(\varphi_1-\varphi_2))\big]$$
$$\left.+2\ell\ell_0 g_\sigma m_H m_d^2 m_1 m_2 C\, S\,(1+c_1c_2)(m_2 s_1 \sin\varphi_1 - m_1 s_2 \sin\varphi_2)\right] .$$

Without the quark-antiquark symmetrization, one adds:

$$8 m_H m_1^2 m_2^2 X^2 \gamma_b^2 g_{\sigma\tau}\Big[\ell^2 m_H C\,(m_2^2 c_2 s_1^2 - m_1^2 c_1 s_2^2 - m_1 m_2 (c_1-c_2) s_1 s_2 \cos(\varphi_1-\varphi_2))$$
$$-\ell\ell_0 m_d^2 S\,(m_2 s_1^2 s_2 \sin\varphi_2 + m_1 s_2^2 s_1 \sin\varphi_1)\Big] \qquad \text{(D10)}$$
$$+2 m_H P^2 g_{\sigma\tau}\Big[\ell^2 m_H C\left(M_1^4 m_2^2 c_2 s_1^2 - M_2^4 m_1^2 c_1 s_2^2 + M_1^2 M_2^2 m_1 m_2 (c_1-c_2) s_1 s_2 \cos(\varphi_1-\varphi_2)\right)$$
$$+2\ell\ell_0 m_1 m_2 m_d^2 S\,((1+c_1c_2)(M_1^2 m_2 s_1 \sin\varphi_1 + M_2^2 m_1 s_2 \sin\varphi_2)$$
$$-\cos(\varphi_1-\varphi_2)(M_2^2 m_1 s_1^2 s_2 \sin\varphi_1 + M_1^2 m_2 s_2^2 s_1 \sin\varphi_2))\Big]$$
$$-4 m_H m_1 m_2 X P\,\gamma_b g_{\sigma\tau}\Big[2\ell^2 m_H M_3^2 m_1 m_2 C\,(c_1-c_2) s_1 s_2 \sin(\varphi_1-\varphi_2)$$
$$+\ell\ell_0 m_d^2 S\left[m_2 s_1^2 s_2 (M_4^2 \cos\varphi_2 - 2m_1^2 \sin\varphi_1 \sin(\varphi_1-\varphi_2)) - m_1 s_2^2 s_1 (M_4^2 \cos\varphi_1 + 2m_2^2 \sin\varphi_2 \sin(\varphi_1-\varphi_2)\right]$$
$$+2 m_1 m_2 (1+c_1c_2)(m_2 s_1 \cos\varphi_1 - m_1 s_2 \cos\varphi_2))\Big]$$
$$4 m_H m_1 m_2 X Q\,\gamma_b\Big[2\ell^2 g_\tau m_H C\,(2M_1^2 m_2^2 s_1^2 - 2M_2^2 m_1^2 s_2^2 + M_4^2 m_d^2 s_1^2 s_2^2)$$
$$-4\ell^2 m_H m_d^2 m_1 m_2 c (g_{\sigma\sigma\tau} - g_\tau c_1 c_2) s_1 s_2 \cos(\varphi_1-\varphi_2)$$
$$+2\ell\ell_0 g_\tau m_d^2 S\left[m_2 c_2 s_1^2 s_2 (2m_1^2 \cos(\varphi_1-\varphi_2)\sin\varphi_1 - M_4^2 \sin\varphi_2) + m_1 c_1 s_2^2 s_1 (2m_2^2 \cos(\varphi_1-\varphi_2)\sin(\varphi_2) - M_4^2 \sin\varphi_1)\right]$$
$$+4\ell\ell_0 m_1 m_2 m_d^2 S\,(m_2 s_1 (g_{\sigma\sigma\tau} c_2 + g_\tau c_1)\sin\varphi_1 + m_1 s_2 (g_{\sigma\sigma\tau} c_1 + g_\tau c_2)\sin\varphi_2)\Big] .$$

In the limit that both Z's are on-shell, $m_1{=}m_2{=}M_Z$, we introduce the notation of Buszello et al: $x = m_H/M_Z$, $y^2 = (x^2-4)/4$. Then we can simplify using $m_d \to 0$, $M_1{=}M_2{=}M_3 \to 4m_H^2 y^2/x^2$, $M_4 \to m_H^2(x^2-2)/x^2$, and $\gamma_b \to xy$. For the full symmetrized matrix element squared the result is:

$$\frac{4}{x^6}\ell^2 m_H^8 y^2\Big[2(x^2X^2 + (x^2-4)P^2)\big[g_1(1+C^2)(1-c_1^2c_2^2) - S^2(g_{\sigma\sigma} + g_1c_1c_2)s_1s_2\cos(\varphi_1+\varphi_2)\big]$$
$$-(x^2X^2 - (x^2-4)P^2)\big[g_1 S^2 s_1^2 s_2^2(\cos(2\varphi_1) + \cos(2\varphi_2)) - 2(1+C^2)(g_{\sigma\sigma} - g_1c_1c_2)s_1s_2\cos(\varphi_1-\varphi_2)\big]$$
$$-4XP\,x\,y\,s_1s_2\big[g_1S^2 s_1 s_2(\sin(2\varphi_1) - \sin(2\varphi_2)) - 2(1+C^2)(g_{\sigma\sigma} - g_1c_1c_2)\sin(\varphi_1-\varphi_2)\big]$$
$$+4XP\,x\,y\,g_\sigma\big[(1+C^2)(c_2s_1^2 + c_1s_2^2) - S^2(c_1+c_2)s_1s_2\cos(\varphi_1+\varphi_2)\big]\Big] . \qquad \text{(D11)}$$

If we simply ignore the polar angle Θ, i.e. set $\Theta = 0$, the above simplifies to:

$$\frac{16}{x^6}\ell^2 m_H^8 y^2\Big[g_1(x^2X^2 + (x^2-4)P^2)(1-c_1^2c_2^2) + (x^2X^2 - (x^2-4)P^2)(g_{\sigma\sigma} - g_1c_1c_2)s_1s_2\cos(\varphi_1-\varphi_2)$$
$$+4XP\,x\,y\,s_1s_2(g_{\sigma\sigma} - g_1c_1c_2)\sin(\varphi_1-\varphi_2) + 2XP\,x\,y\,g_\sigma(c_2s_1^2 + c_1s_2^2)\Big] . \qquad \text{(D12)}$$

This agrees with the result of Buszello et al [29].

[1] T. C. Collaboration, t. D. Collaboration, t. T. N. Physics and H. W. Group, arXiv:0911.3930 [hep-ex].

[2] G. Aad *et al.* [The ATLAS Collaboration], arXiv:0901.0512 [hep-ex].

[3] G. L. Bayatian *et al.* [CMS Collaboration], J. Phys. G **34** (2007) 995.

[4] A. Djouadi, Phys. Rept. **457** (2008) 1 [arXiv:hep-ph/0503172].

[5] A. Datta, G. L. Kane and M. Toharia, arXiv:hep-ph/0510204.

[6] J. Hubisz, J. Lykken, M. Pierini and M. Spiropulu, Phys. Rev. D **78** (2008) 075008 [arXiv:0805.2398 [hep-ph]].

[7] M. Burns, K. Kong, K. T. Matchev and M. Park, JHEP **0810** (2008) 081 [arXiv:0808.2472 [hep-ph]].

[8] J. Alwall, P. Schuster and N. Toro, Phys. Rev. D **79** (2009) 075020 [arXiv:0810.3921 [hep-ph]].

[9] G. Hallenbeck, M. Perelstein, C. Spethmann, J. Thom and J. Vaughan, Phys. Rev. D **79** (2009) 075024 [arXiv:0812.3135 [hep-ph]].

[10] A. Djouadi, J. Kalinowski and M. Spira, Comput. Phys. Commun. **108** (1998) 56 [arXiv:hep-ph/9704448].

[11] J. F. Gunion and M. Soldate, "Overcoming a critical background to Higgs detection," Phys. Rev. D **34**, 826 (1986).

[12] B. A. Dobrescu and J. D. Lykken, arXiv:0912.3543 [hep-ph].

[13] M. Pivk and F. R. Le Diberder, Nucl. Instrum. Meth. A **555** (2005) 356 [arXiv:physics/0402083].

[14] J. R. Dell'Aquila and C. A. Nelson, Phys. Rev. D **33** (1986) 80.

[15] C. A. Nelson, Phys. Rev. D **37** (1988) 1220.

[16] T. Matsuura and J. J. van der Bij, Z. Phys. C **51** (1991) 259.

[17] A. Soni and R. M. Xu, Phys. Rev. D **48** (1993) 5259 [arXiv:hep-ph/9301225].

[18] D. Chang, W. Y. Keung and I. Phillips, Phys. Rev. D **48** (1993) 3225 [arXiv:hep-ph/9303226].

[19] V. D. Barger, K. m. Cheung, A. Djouadi, B. A. Kniehl and P. M. Zerwas, Phys. Rev. D **49** (1994) 79 [arXiv:hep-ph/9306270].

[20] A. Skjold and P. Osland, Phys. Lett. B **311** (1993) 261 [arXiv:hep-ph/9303294].

[21] B. Grzadkowski and J. F. Gunion, Phys. Lett. B **350** (1995) 218 [arXiv:hep-ph/9501339].

[22] T. Plehn, D. L. Rainwater and D. Zeppenfeld, Phys. Rev. Lett. **88** (2002) 051801 [arXiv:hep-ph/0105325].

[23] S. Y. Choi, D. J. . Miller, M. M. Muhlleitner and P. M. Zerwas, Phys. Lett. B **553** (2003) 61 [arXiv:hep-ph/0210077].

[24] S. S. Biswal, R. M. Godbole, R. K. Singh and D. Choudhury, Phys. Rev. D **73** (2006) 035001 [Erratum-ibid. D **74** (2006) 039904] [arXiv:hep-ph/0509070].

[25] R. M. Godbole, D. J. . Miller and M. M. Muhlleitner, JHEP **0712** (2007) 031 [arXiv:0708.0458 [hep-ph]].

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

[27] C. P. Buszello and P. Marquard, arXiv:hep-ph/0603209.

[28] C. P. Buszello, P. Marquard and J. J. van der Bij, arXiv:hep-ph/0406181.

[29] C. P. Buszello, I. Fleck, P. Marquard and J. J. van der Bij, Eur. Phys. J. C **32** (2004) 209 [arXiv:hep-ph/0212396].

[30] W. Bernreuther, P. Gonzalez and M. Wiebusch, arXiv:0909.3772 [hep-ph].

[31] K. Hagiwara, Q. Li and K. Mawatari, JHEP **0907** (2009) 101 [arXiv:0905.4314 [hep-ph]].

[32] R. Lafaye, T. Plehn, M. Rauch, D. Zerwas and M. Duhrssen, arXiv:0904.3866 [hep-ph].

[33] W. Y. Keung, I. Low and J. Shu, Phys. Rev. Lett. **101** (2008) 091802 [arXiv:0806.2864 [hep-ph]].

[34] H. Davoudiasl, S. Gopalakrishna, E. Ponton and J. Santiago, arXiv:0908.1968 [hep-ph].

[35] A. Bredenstein, A. Denner, S. Dittmaier and M. M. Weber, JHEP **0702** (2007) 080 [arXiv:hep-ph/0611234].

[36] N. Cabibbo and A. Maksymowicz, Phys. Rev. **137**, B438 (1965) [Erratum-ibid. **168**, 1926 (1968)].

[37] T. Han, J. D. Lykken and R. J. Zhang, Phys. Rev. D **59** (1999) 105006 [arXiv:hep-ph/9811350].

[38] A. De Rújula, J. M. K. Kaplan and E. de Rafael, Nuc. Phys. **B35**, 365 (1971).

[39] G. L. Bayatian *et al.* [CMS Collaboration], "CMS physics: Technical design report,"

[40] S. Frixione and B. R. Webber, JHEP **0206** (2002) 029 [arXiv:hep-ph/0204244].

[41] H. L. Lai *et al.* [CTEQ Collaboration], Eur. Phys. J. C **12** (2000) 375 [arXiv:hep-ph/9903282].

[42] T. Sjostrand, S. Mrenna and P. Skands, JHEP **0605**, 026 (2006) [arXiv:hep-ph/0603175].

[43] The CMS Collaboration, Analysis note EWK-08-006.

[44] G. J. Feldman and R. D. Cousins, Phys. Rev. D **57** (1998) 3873 [arXiv:physics/9711021].

[45] J. Neyman and E. S. Pearson Philosophical Transactions of the Royal Society of London. Series A Vol. 231, (1933), pp. 289-337

[46] A. Einstein, remark to Oscar Veblen at Princeton University, May 1921.

[47] Y. Gao, A. V. Gritsan, Z. Guo, K. Melnikov, M. Schulze and N. V. Tran, arXiv:1001.3396 [hep-ph].