

Higgs boson look-alikes at the LHC

A. De Rújula,^{1,2,3} Joseph Lykken,⁴ Maurizio Pierini,³ Christopher Rogan,⁵ and Maria Spiropulu^{3,5}

¹*Instituto de Física Teórica, Univ. Autónoma de Madrid, Madrid and CIEMAT, Madrid, Spain*

²*Physics Department, Boston University, Boston, Massachusetts 02215, USA*

³*Physics Department, CERN, CH 1211 Geneva 23, Switzerland*

⁴*Fermi National Accelerator Laboratory, P.O. Box 500, Batavia, Illinois 60510, USA*

⁵*Lauritsen Laboratory of Physics, California Institute of Technology, Pasadena, California 91125, USA*

(Received 2 March 2010; published 14 July 2010)

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 resonance mass of $200 \text{ GeV}/c^2$, we achieve a median discrimination significance of 3σ with as few as 19 events, and even better discrimination for the off-shell decays of a $145 \text{ GeV}/c^2$ resonance.

DOI: 10.1103/PhysRevD.82.013003

PACS numbers: 14.80.Bn

I. INTRODUCTION

The CDF and D0 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 undiscovery 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 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 quan-

tum numbers and couplings. We shall refer to these other possibilities as Higgs look-alikes (HLLs).

The purpose of this paper is to quantify the degree to which one can discriminate a standard model Higgs boson from HLLs at, or close to, the moment of discovery at the LHC. 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 physics, namely, the Higgs decay $H \rightarrow ZZ^* \rightarrow \ell_1^+ \ell_1^- \ell_2^+ \ell_2^-$, where $\ell_{1,2}^\pm$ denotes an electron or a muon, and Z^* denotes that one of the Z s may be strongly off-shell. 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

Higgs decay into two photons, where the photon polarizations are not measured.

The branching fraction for the golden mode is small; example values for an SM Higgs $\rightarrow ZZ^* \rightarrow 4\ell$ are 0.0011 for $m_H = 200 \text{ GeV}/c^2$, 0.0014 for $m_H = 350 \text{ GeV}/c^2$, and 0.00036 for $m_H = 145 \text{ GeV}/c^2$ [10]. Even for favorable Higgs masses, this branching fraction is 2 orders of magnitude smaller than that for semileptonic $H \rightarrow W^+W^- \rightarrow \ell\nu jj$, a channel that, 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 smaller than that for the fully leptonic SM Higgs decay $H \rightarrow W^+W^- \rightarrow \ell^+\nu\ell^-\bar{\nu}$. 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 and HLL production also introduces new theoretical and measurement uncertainties involving associated hadronic jets and the parton distribution functions 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 Sec. IV, where we parametrize the relevant efficiencies, resolutions and acceptances for an LHC detector. Because both the ATLAS and CMS detectors in general measure muons and electrons with exquisite precision, the resolution and efficiency for detecting the four leptons can be significantly degraded with 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 biases in the characterization of a Higgs signal.

Our analysis depends on five distinct angles that describe the $H \rightarrow ZZ^* \rightarrow 4\ell$ 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 „Plots technique [13]. This produces an effectively background-subtracted data sample where, instead of making stringent requirements that reduce the signal yield available for characterization, 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 [4,17–33]. 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 the angular distributions.

In our analysis we compare an SM Higgs signal to a variety of Higgs look-alikes. We consider the most general Lorentz invariant couplings 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 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 of a massive neutral spin 1 boson to ZZ or ZZ^* . The off-shell case has not been presented before, to our knowledge. 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 then also part of a Z' look-alike analysis, which is interesting in its own right [34].

We also discuss as one of our HLLs a massive spin 2 resonance coupling to the ZZ energy-momentum tensor, not necessarily with the universality of a gravitonlike coupling. Although universally-coupled massive gravitons are already experimentally excluded in the relevant mass range [35], general spin 2 HLLs are a natural example of our study of spin discriminations.

In Sec. II we define our notation for the observables of the four-lepton final state. Section III contains the general gauge and Lorentz invariant couplings of an HLL to ZZ or ZZ^* , with a discussion of other symmetry properties. We describe in Sec. IV event generation, detector simulation, and the construction of effectively background-subtracted

samples using „Plots; here also we show the sculpting of the angular distributions and correlations by the finite phase space acceptance of the detector. In Sec. V we describe our statistical approach to HLL discrimination using hypothesis testing with likelihood ratios. We demonstrate in Sec. VI the consistency of our methods by applying them to the discrimination of signal from SM ZZ background. In Sec. VII we detail many examples quantifying our ability to discriminate an 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/c². We summarize, in Sec. VIII, our results and outlook for further improvements. Here we explicitly quantify the extent to which 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 an 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: their 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, as illustrated in Fig. 1.

The H azimuthal angle plays no physical role, while the p_T and η distributions influence the way the detector selects events, sculpting the distributions of the final-state lepton's directions and energies. Once an event is boosted back to the 4 ℓ rest-system (the rest-system of the two initial-state fusing partons), the memory of p_T and η is lost, modulo these phase space acceptance effects.

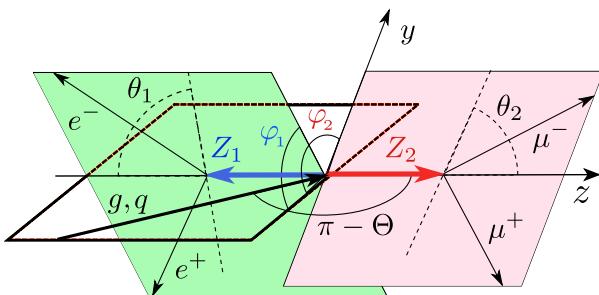


FIG. 1 (color online). The Cabibbo-Maksymowicz angles [53] in the $H \rightarrow ZZ$ decays.

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 reexpress 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 $\cos\Theta \rightarrow -\cos\Theta$.

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 $\varphi_1 \rightarrow \Phi + \phi$, $\varphi_2 \rightarrow \Phi$. Thus ϕ then represents the azimuthal rotation between the Z_2 and Z_1 decay planes.

In summary, the 4-momenta of the process $gg \rightarrow H \rightarrow Z_1 Z_2 \rightarrow \ell_1^- \ell_1^+ \ell_2^- \ell_2^+$ are explicitly parametrized in the H rest frame as

$$\begin{aligned}
p_{g_2} &= \frac{m_H}{2}(1, S \cos\Phi, S \sin\Phi, C), \\
p_{g_1} &= \frac{m_H}{2}(1, -S \cos\Phi, -S \sin\Phi, -C), \\
k &= m_H(1, 0, 0, 0), \\
p_2 &= m_2(\gamma_2, 0, 0, \beta_2 \gamma_2), \\
p_1 &= m_1(\gamma_1, 0, 0, -\beta_1 \gamma_1), \\
p_{\ell_2^-} &= \frac{m_2}{2}(\gamma_2(1 + \beta_2 c_2), 0, s_2, \gamma_2(\beta_2 + c_2)), \\
p_{\ell_2^+} &= \frac{m_2}{2}(\gamma_2(1 - \beta_2 c_2), 0, -s_2, \gamma_2(\beta_2 - c_2)), \\
p_{\ell_1^-} &= \frac{m_1}{2}(\gamma_1(1 + \beta_1 c_1), -ss_1, -cs_1, -\gamma_1(\beta_1 + c_1)), \\
p_{\ell_1^+} &= \frac{m_1}{2}(\gamma_1(1 - \beta_1 c_1), ss_1, cs_1, -\gamma_1(\beta_1 - c_1)).
\end{aligned} \tag{3}$$

Here k denotes the 4-momentum of H , while p_1, p_2 are the 4-momenta of Z_1, Z_2 . We 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

$$\vec{\Omega} = \{\Phi, \cos\Theta\}, \quad \vec{\omega} = \{\phi, \cos\theta_1, \cos\theta_2\}. \tag{4}$$

For a SM Higgs, the distributions in Θ and Φ are flat if we ignore the phase space acceptance effects inherent in any experimental analysis. In previous studies these two angles have typically been 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.* [36].

The decay amplitudes defined in the next section depend on two combinations of the boost parameters γ_1 and γ_2 , defined by

$$\gamma_a = \gamma_1 \gamma_2 (1 + \beta_1 \beta_2), \tag{5}$$

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

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

$$\gamma_a^2 - \gamma_b^2 = 1. \tag{7}$$

More explicitly, we have

$$\gamma_a = \frac{1}{2m_1 m_2} (m_H^2 - (m_1^2 + m_2^2)). \tag{8}$$

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 Z s of four-momenta p_1 and p_2 are given by the expression:

$$L_{\mu\alpha} = X g_{\mu\alpha} - (Y + iZ) \frac{k_\alpha k_\mu}{M_Z^2} + (P + iQ) \epsilon_{\mu\alpha} \frac{p_1 p_2}{M_Z^2}, \tag{9}$$

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 $i g M_Z / \cos\theta_W$, 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

$$L^{\rho\mu\alpha} = X(g^{\rho\mu} p_1^\alpha + g^{\rho\alpha} p_2^\mu) + (P + iQ) \epsilon^{\rho\mu\alpha} (p_1 - p_2), \tag{10}$$

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

$$\begin{aligned}
L^{\rho\sigma\mu\alpha} &= X_0 m_H^2 g^{\mu\rho} g^{\alpha\sigma} \\
&\quad + (X_1 + iY_1)(p_1^\alpha p_2^\rho g^{\sigma\mu} + p_1^\rho p_2^\mu g^{\sigma\alpha}) \\
&\quad + (X_2 + iY_2)p_1^\rho p_2^\sigma g^{\mu\alpha},
\end{aligned} \tag{11}$$

where we have dropped contributions that have more than two derivatives or are odd under parity, and again with all coefficients real. The special case of tree-level gravitonlike couplings corresponds to

$$X_0 = -\frac{1}{2}\kappa, \quad X_1 = \kappa, \quad X_2 = -\kappa, \tag{12}$$

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 HLs with nonvanishing 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 Appendices C and D, where

$$\eta \equiv \frac{2c_v v_a}{(c_v^2 + c_a^2)} \simeq 0.15, \quad (13)$$

denotes the quantity 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 polarization ϵ_1 and ϵ_2 is $\propto \epsilon_1 \cdot \epsilon_2$, see Eq. (9). The angular distribution of the leptons in $H \rightarrow ZZ \rightarrow 4l$ decay, for on- or off-shell Z 's of mass m_1 and m_2 , is

$$\begin{aligned} \frac{d\Gamma[0^+]}{dc_1 dc_2 d\phi} \propto & m_1^2 m_2^2 m_H^4 [1 + c_1^2 c_2^2 + (\gamma_b^2 + c^2) s_1^2 s_2^2 \\ & + 2\gamma_a c_1 s_1 s_2 c_1 c_2 + 2\eta^2 (c_1 c_2 + \gamma_a c_1 s_1 s_2)]. \end{aligned} \quad (14)$$

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

The coupling of a $J^{PC} = 0^{-+}$ pseudoscalar to two Z 's of polarization ϵ_1 and ϵ_2 and four-momenta p_1 and p_2 is proportional to $\epsilon[\epsilon_1, \epsilon_2, p_1, p_2]$, see Eq. (9). The angular distribution of the leptons in its $ZZ \rightarrow 4l$ decay is

$$\frac{d\Gamma[0^-]}{dc_1 dc_2 d\phi} \propto m_1^4 m_2^4 \gamma_b^2 (1 + c_1^2 c_2^2 - c^2 s_1^2 s_2^2 + 2\eta^2 c_1 c_2). \quad (15)$$

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 polarization 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. (10). 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. Performing this symmetrization, we also introduce the notation

$$m_d^2 \equiv m_1^2 - m_2^2, \quad (16)$$

and find the angular distribution of the leptons in $H \rightarrow ZZ^* \rightarrow 4l$ decay as follows:

$$\begin{aligned} \frac{d\Gamma[1^-]}{dCdc_1 dc_2 d\Phi d\phi} \propto & 4m_1^2 m_2^2 \gamma_b^2 [S^2 s_1^2 s_2^2 (2m_d^4 - m_H^2 [m_1^2 \cos(2(\Phi + \phi)) + m_2^2 \cos(2\Phi)]) \\ & + m_H^2 (1 + C^2) [2m_2^2 s_1^2 + 2m_1^2 s_2^2 - (m_1^2 + m_2^2) s_1^2 s_2^2] \\ & + 4m_H m_d^2 CS [m_1 c_1 s_1 s_2 \sin(\Phi + \phi) - m_2 c_2 s_2 s_1 \sin\Phi] \\ & - 2m_H^2 m_1 m_2 s_1 s_2 ((1 + C^2)(c_1 c_2 - \eta^2) c + S^2 (c_1 c_2 + \eta^2) \cos(2\Phi + \phi)]]. \end{aligned} \quad (17)$$

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 polarization ϵ_1 and ϵ_2 and four-momenta p_1 and p_2 is proportional to $\epsilon[\epsilon_H, \epsilon_1, \epsilon_2, p_1 - p_2]$, see Eq. (10). After the same symmetrization in $\cos\Theta$ described above, and introducing the notation

$$M_1^2 \equiv m_H^2 - 3m_1^2 - m_2^2, \quad M_2^2 \equiv m_H^2 - m_1^2 - 3m_2^2, \quad (18)$$

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

$$\begin{aligned} \frac{d\Gamma[1^+]}{dCdc_1dc_2d\Phi d\phi} \propto & m_H^2 S^2 s_1^2 s_2^2 [M_2^4 m_1^2 \cos(2(\Phi + \phi)) + M_1^4 m_2^2 \cos(2\Phi)] + 8m_1^2 m_2^2 m_d^4 S^2 [c_1^2 + c_2^2 + s_1^2 s_2^2 s^2 + 2\eta^2 c_1 c_2] \\ & + m_H^2 (1 + C^2) [2M_1^4 m_2^2 s_1^2 + 2M_2^4 m_1^2 s_2^2 - (M_2^4 m_1^2 + M_1^4 m_2^2) s_1^2 s_2^2] \\ & - 8m_H m_d^2 m_1 m_2 CS [M_2^2 m_1 s_2 (c_2 s_1^2 c \sin(\Phi + \phi) + c_1 (c_1 c_2 + \eta^2) \sin\Phi) \\ & - M_1^2 m_2 s_1 (c_1 s_2^2 c \sin\Phi + c_2 (c_1 c_2 + \eta^2) \sin(\Phi + \phi))] \\ & + 2m_H^2 M_1^2 M_2^2 m_1 m_2 s_1 s_2 [(1 + C^2) (c_1 c_2 - \eta^2) c - S^2 (c_1 c_2 + \eta^2) \cos(2\Phi + \phi)]. \end{aligned} \quad (19)$$

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 the example of a positive parity spin 2 with gravitonlike couplings produced by gluon fusion and decaying to two on-shell Z 's. Defining the on-shell ratio $x \equiv m_H/M_Z$ and using the massive graviton formalism of [37], we obtain the tree-level angular distribution:

$$\begin{aligned} \frac{d\Gamma[gg \rightarrow \text{graviton} \rightarrow ZZ]}{dCdc_1dc_2d\Phi d\phi} \propto & 16x^4 C^2 + 2(x^4 + 16)S^4 + s_1^2 s_2^2 [(x^4 + 16)S^4 - 4x^2(x^2 + 4)S^2 + 4x^4] \\ & + 8x^2 S^2 [[2 + S^2 + (2 - 3S^2)c_2^2]s_1^2 \cos(\Phi + \phi)^2 + [2 + S^2 + (2 - 3S^2)c_1^2]s_2^2 \cos^2\Phi] \\ & + S^4 s_1^2 s_2^2 [x^4 \cos(2\Phi + \phi)^2 + 16c^2] - (s_1^2 + s_2^2)[(x^2 + 4)^2 C^4 + 2(3x^4 - 16)C^2] \\ & + (x^2 - 4)^2] + 2S^2 c_1 c_2 s_1 s_2 [x^2 [2(x^2 + 4) - (x^2 + 12)S^2] \cos(2\Phi + \phi) \\ & + 4[4x^2 - (3x^2 + 4)S^2]c]. \end{aligned} \quad (20)$$

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

C. Tests of symmetries

Now we discuss the behavior 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. (9). In computing the ensuing $H \rightarrow ZZ^* \rightarrow 4\ell$ process one finds that the XP interference term is of the form

$$\frac{d\Gamma[0, \text{Todd}]}{dc_1dc_2d\phi} \propto 2m_1^3 m_2^3 m_H^2 \gamma_b s_1 s_2 s [s_1 s_2 c + \gamma_a (c_1 c_2 + \eta^2)], \quad (21)$$

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-momenta 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. (9). To this order, a \tilde{T} -odd observable must vanish if CP is a symmetry, as shown in [38]. Thus, a nonvanishing \tilde{T} -odd observable

such as that of Eq. (21) can only arise if CP -invariance is violated.

The XQ interference term resulting from Eq. (9) is

$$\begin{aligned} \frac{d\Gamma[0, \text{Codd}]}{dc_1dc_2d\phi} \propto & -2\eta m_1^3 m_2^3 m_H^2 \gamma_b [c_1 + c_2] \\ & \times (1 + c_1 c_2 + \gamma_a s_1 s_2 c). \end{aligned} \quad (22)$$

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.

Bose-Einstein statistics

The general coupling, up to two derivatives, of a $J = 1$ particle to two Z 's is that of Eq. (10). 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], whose authors were the first to emphasize the importance of M_Z^* as a discriminating variable, 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 independent of any on- or off-shell considerations. The two Z 's in an $H \rightarrow ZZ^*$ 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 an HLL conserve P and CP , but the object is not pointlike, 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 nonvanishing Y in Eq. (9), 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 nonzero Y/X which is governed by the shapes of angular distributions. Contributions to Y can arise from gauge-invariant operators of dimension 5 containing a non-SM-like spin 0 HLL (Appendix A) or from higher dimension operators containing the SM Higgs [39–41].

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

$$d\Gamma_C = \cos^2\xi d\Gamma_{XX} + \cos\xi \sin\xi d\Gamma_{XY} + \sin^2\xi d\Gamma_{YY}, \quad (23)$$

where $d\Gamma_{XX}$ is the standard result of Eq. (14). The interference term is

$$\frac{d\Gamma_{XY}}{dc_1 dc_2 d\phi} \propto -2m_1^3 m_2^3 m_H^2 \gamma_b^2 s_1 s_2 (c_1 c_2 c + \gamma_a s_1 s_2 + \eta^2 c), \quad (24)$$

and the last term is

$$\frac{d\Gamma_{YY}}{dc_1 dc_2 d\phi} \propto m_1^4 m_2^4 \gamma_b^4 s_1^2 s_2^2. \quad (25)$$

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 pointlike, 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 $H \rightarrow ZZ$ branching fraction, the integral of Eq. (23) over $\cos\theta_1$, $\cos\theta_2$, and ϕ :

$$\Gamma_C \propto m_1^2 m_2^2 [2\cos^2\xi + (\gamma_a \cos\xi - m_1 m_2 \gamma_b^2 \sin\xi)^2]. \quad (26)$$

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 parametrized reconstruction resolutions and efficiencies based on the published CMS performance results [42]. A similar study can be performed with parametrizations 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 multidimensional 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 leptons from the five- or six-dimensional probability density functions (*pdfs*) of

$$\vec{X} \equiv \{\vec{\omega}, \vec{\Omega}\} \quad \text{for } ZZ, \quad \vec{X} \equiv \{\vec{\omega}, \vec{\Omega}, M_{Z^*}\} \quad \text{for } ZZ^*, \quad (27)$$

where $\vec{\Omega}$, $\vec{\omega}$ are given in Eq. (4). 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 [43]. We consider proton-proton collisions at $\sqrt{s} = 10$ TeV, and we model the parton distribution functions (PDFs) using CTEQ5L [44].

In this analysis we do not assume a specific signal production mechanism and cross section, instead relying on the discrimination provided by the angular distributions of the leptons in the final state. Figure 2 has the p_T vs η *pdfs* for a spin-0 and a spin-1 HLL. As discussed in Sec. I, for all the signal generation we use the p_T vs η *pdfs* of the scalar. For the SM ZZ background the p_T , η and $m_{4\mu}$ are taken from a three-dimensional *pdf* generated using the PYTHIA [45] 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 parametrized as a function of the muon p_T and η according to [42], where the muon reconstruction efficiency is close to

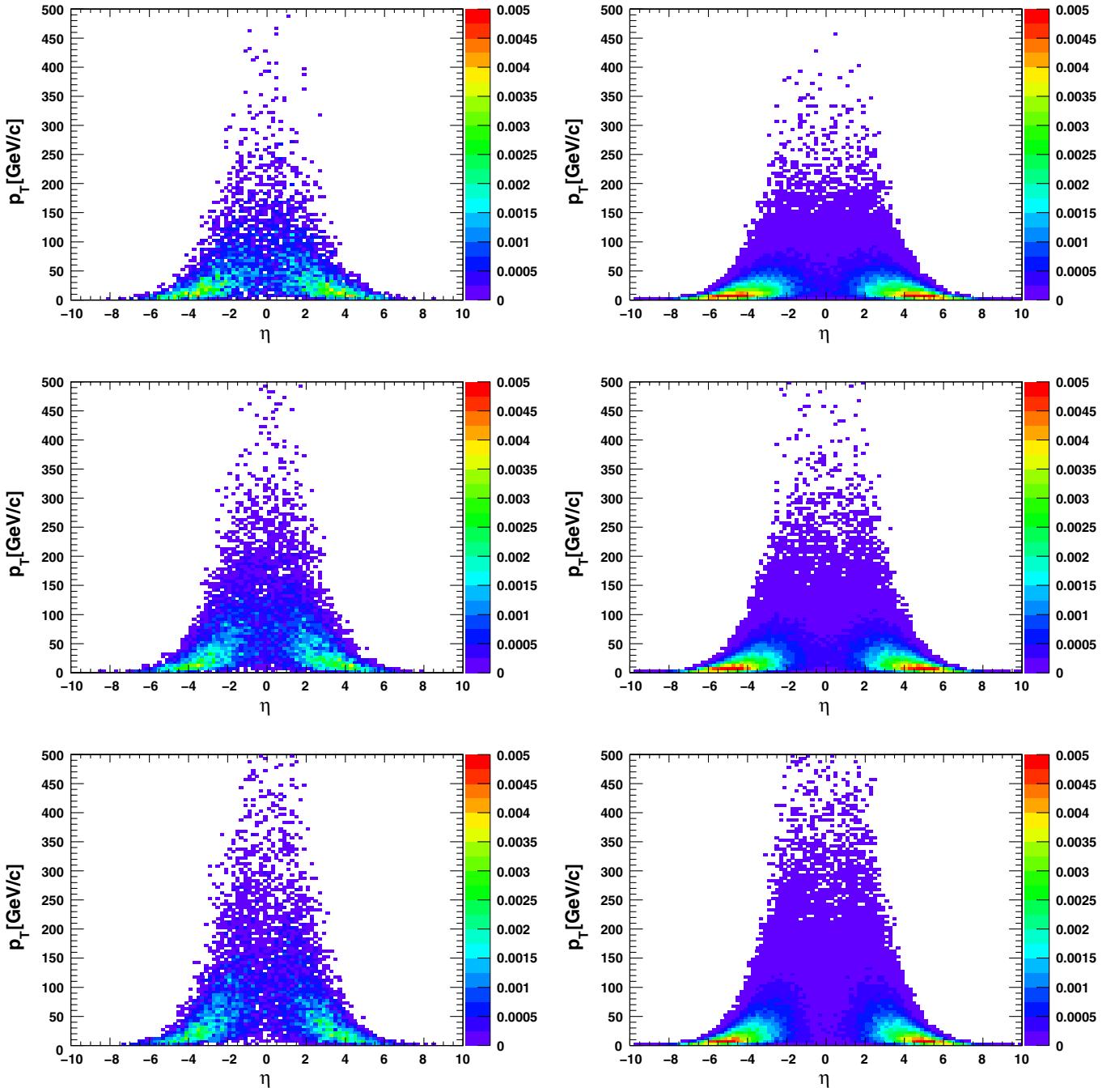


FIG. 2 (color online). 2D p_T - η pdf of a 0^+ HLL resonance (left) and a 1^- one (right) for $\sqrt{s} = 10$ TeV collisions, obtained using PYTHIA and the CTEQL5 parton density functions and for $m_H = 145, 200, 350$ GeV/ c^2 (top, middle and bottom).

100% for muons with $p_T \geq 10$ GeV/ c and $|\eta| \leq 2.3$, corresponding to the event selection in our 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, see Fig. 3, with the corresponding ones from a published full-simulation analysis [46].

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 negligible. This shows that the sculpting of the observ-

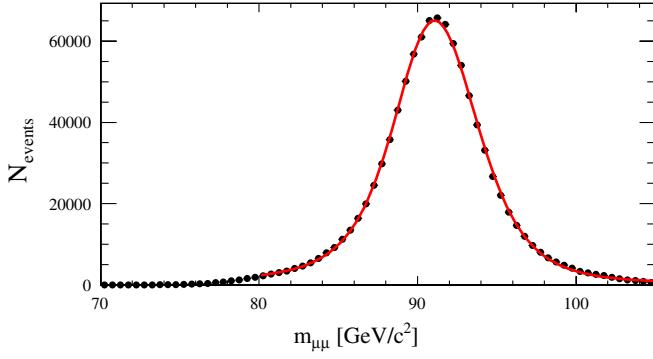


FIG. 3 (color online). Distribution of the dimuon invariant mass for a sample of signal $H \rightarrow ZZ$ events, generated using our very-fast muon simulation. The parameters of the superimposed fit are extracted from [46].

ables' *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 particular 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 $145 \text{ GeV}/c^2$ generated with no explicit angular correlations. Adding the angular correlations 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 $200 \text{ GeV}/c^2$ ($350 \text{ GeV}/c^2$) as shown in Fig. 7. The same figure demonstrates that the efficiency has a nontrivial dependence on the nature of the spin correlations. Specifically, for a 0^- resonance of $200 \text{ GeV}/c^2$ ($350 \text{ GeV}/c^2$) the efficiency is 60% (69%). With an absence of explicit spin correlations the efficiency for a $350 \text{ GeV}/c^2$ 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 correspond to the kinematic configurations where at least one of the muons is soft. When the correlations between the variables $\vec{\omega}$ and $\vec{\Omega}$ appear explicitly in the differential cross sections, as is

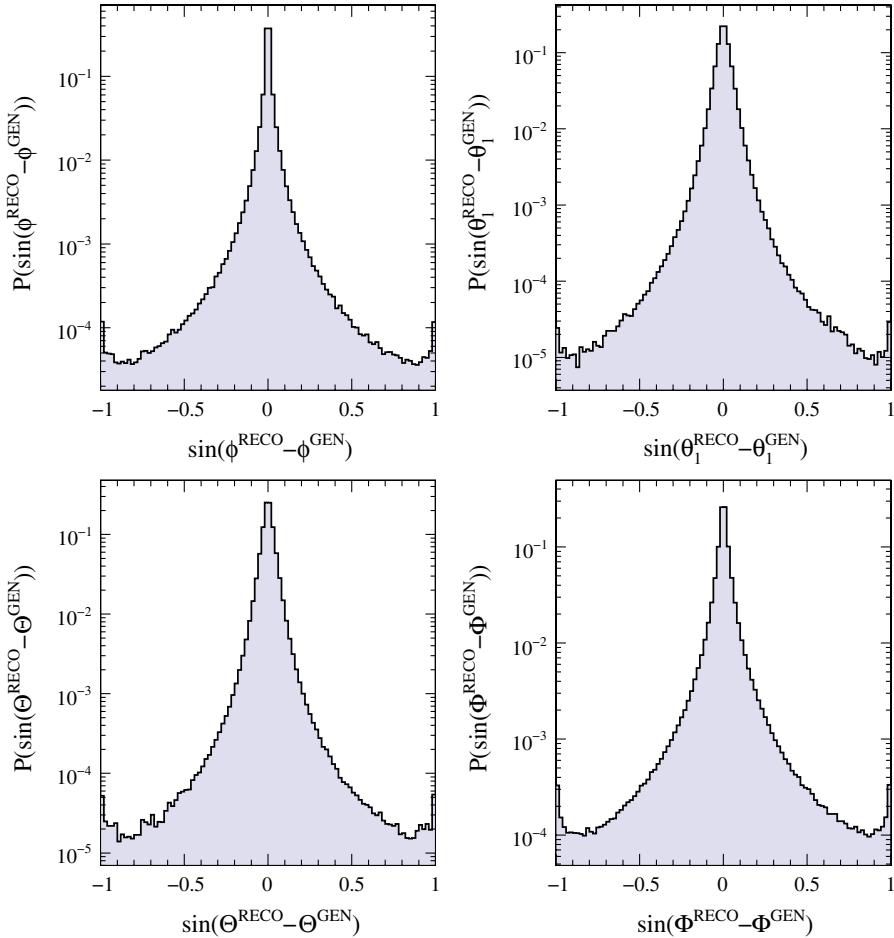


FIG. 4 (color online). Reconstruction resolution for the angular variables of \vec{X} shown here for a resonance with mass $145 \text{ GeV}/c^2$. The $\cos\theta_2$ and $\cos\theta_1$ distributions are very similar in this case. Only events surviving the signal selection are included. All distributions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

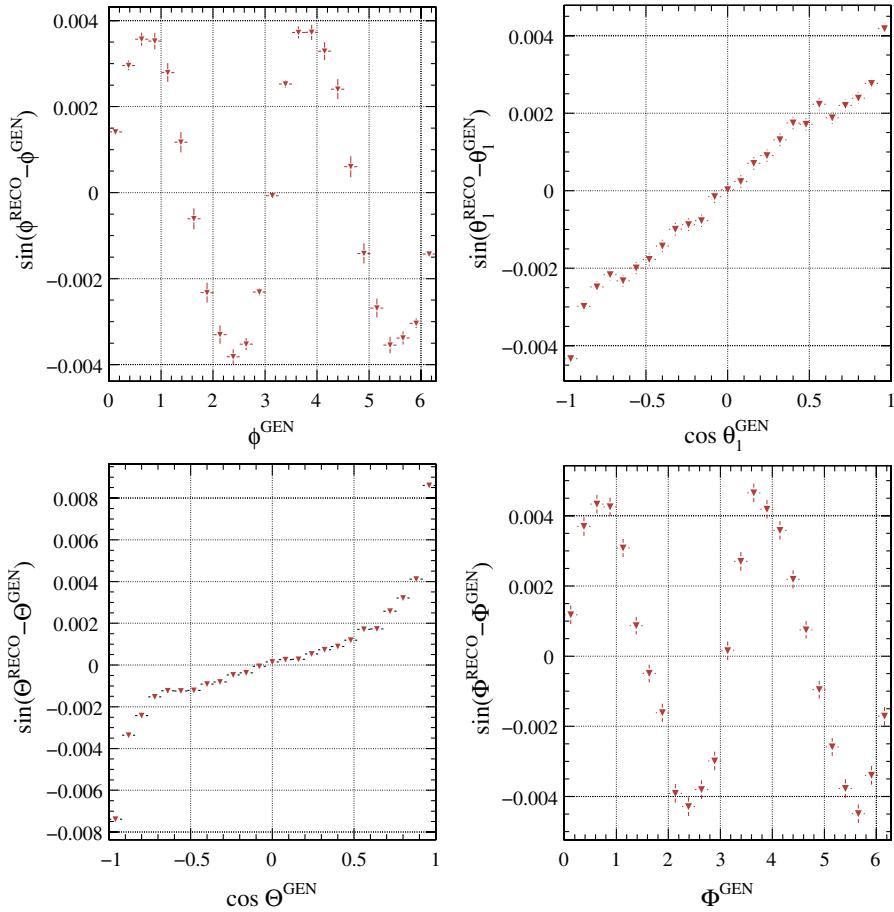


FIG. 5 (color online). Distributions showing systematic biases for a subset of the reconstructed variables \vec{X} for a resonance with mass $145 \text{ GeV}/c^2$. Only events that survive the signal selection are included. All biases are negligible.

the case for $J = 1^\pm$, the phase space acceptance effects are amplified. The consequences on model discrimination are discussed in Sec. VII B.

The shapes of the reconstructed $\vec{\omega}$ and $\vec{\Omega}$ distributions depend on the phase space acceptance both for electron and muon final states ($H \rightarrow ZZ \rightarrow 2e2\mu$ or $4e$). Figure 8 shows the relevant kinematic distributions. All the 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 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 $H \rightarrow ZZ$ 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

$$\mathcal{L} = \frac{1}{N!} \exp\left(-\sum_j N_j\right) \prod_{i=1}^N (N_S P_S[m_{4\mu}^i] + N_B P_B[m_{4\mu}^i]), \quad (28)$$

where N_j ($j = S, B$) represents the yield of each component, $m_{4\mu}^i$ 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 $m_H = 250 \text{ GeV}/c^2$. This fit configuration is appropriate for the HLL characterization.

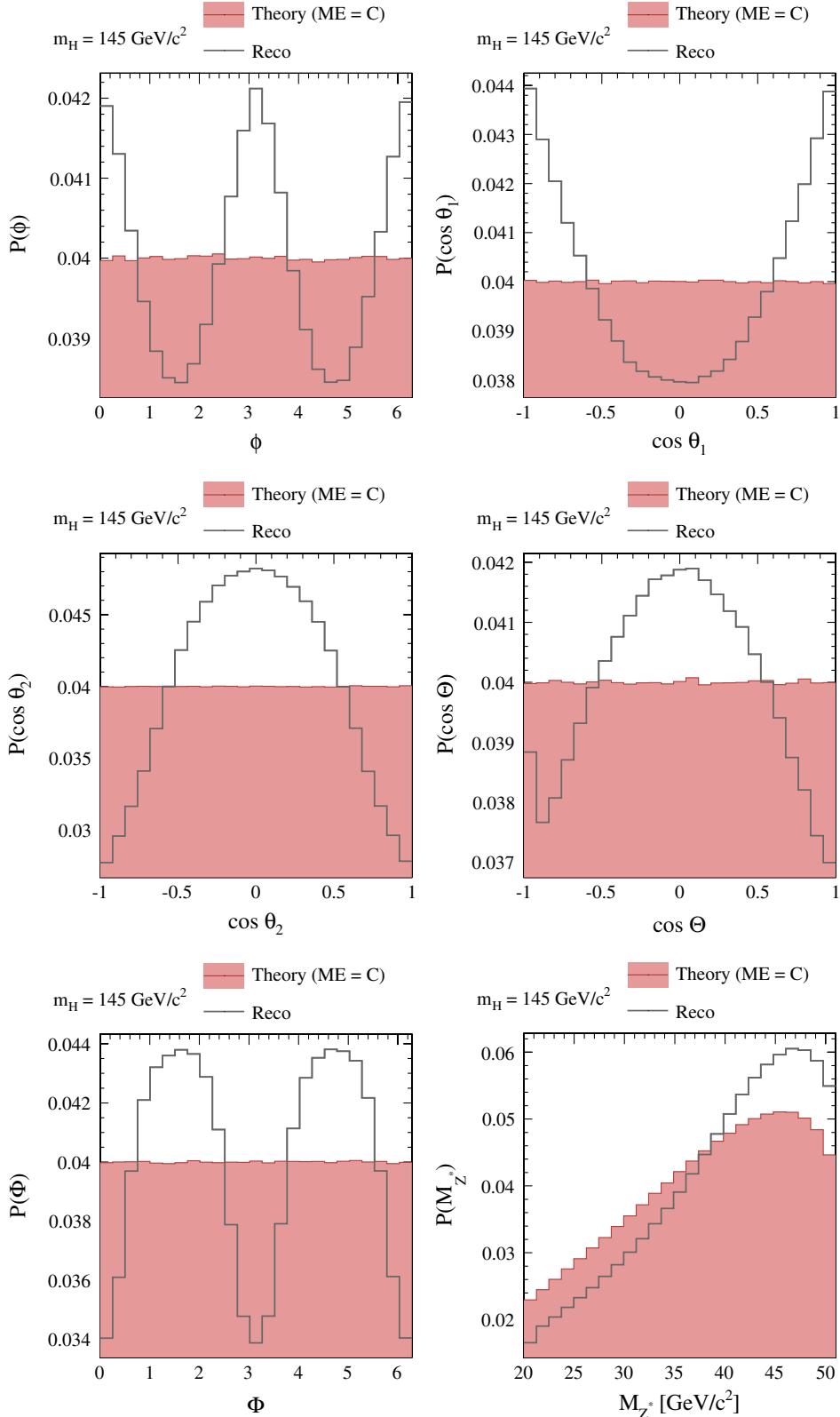


FIG. 6 (color online). The variables \vec{X} used in this analysis for a $145 \text{ GeV}/c^2$ resonance. The off-shell M_{Z^*} is required to lie a window between 20 and $50 \text{ GeV}/c^2$. The shaded histograms are the 1D distributions using a constant matrix element (i.e. no angular correlations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

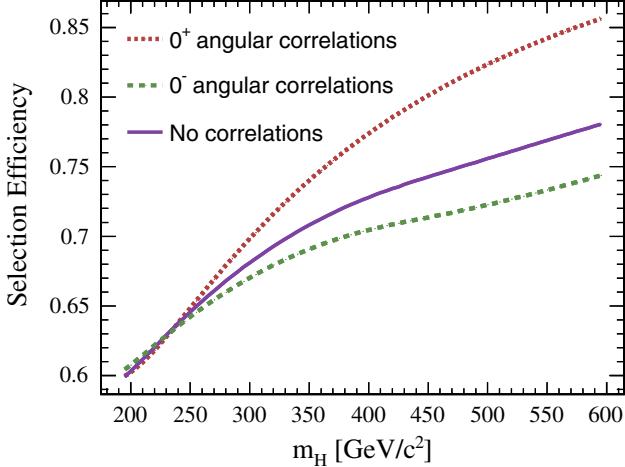


FIG. 7 (color online). The analysis efficiency for 0^+ , 0^- as a function of the resonance mass. The case with no correlations is also shown for comparison.

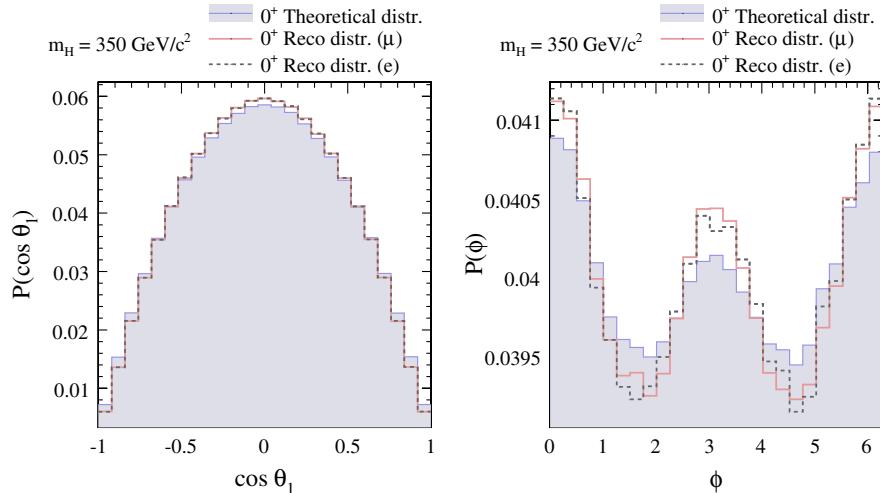


FIG. 8 (color online). Kinematic distributions for the variables $\cos\theta_1$ (left) and ϕ (right) for a 0^+ resonance with mass $350 \text{ GeV}/c^2$. The shaded histograms show the 1D projections of the variables as described by the analytic *pdfs*. The overlaid histograms (blue, red) show the same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

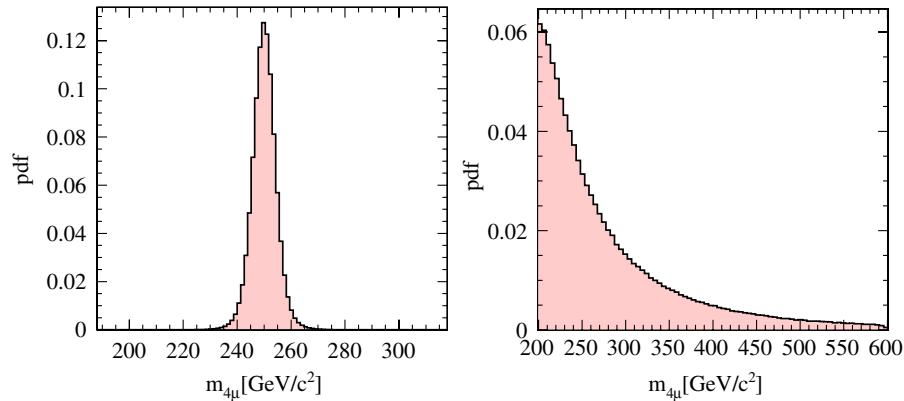
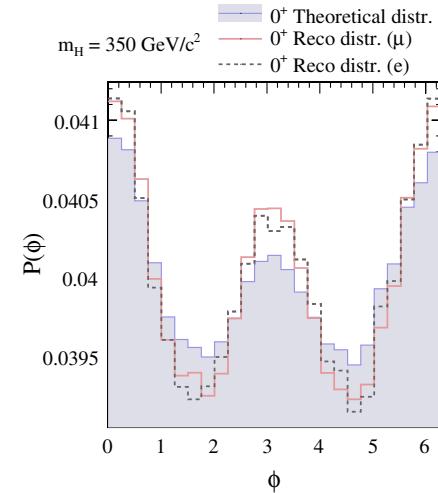


FIG. 9 (color online). Distribution of the 4μ invariant mass for a sample of signal with $m_H = 250 \text{ GeV}/c^2$ (left), and background (right) ZZ events.

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 data set is an important prerequisite. We use the s -Weight [13] technique and reweight the selected data set according to *how likely* each event is considered to be signal by the fit. The s -Weight 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). On the upper plot of Fig. 10, the 4μ invariant-mass distribution is shown for a sample of $N_S = 70H \rightarrow ZZ$ events (with $m_H = 250 \text{ GeV}/c^2$) on top of $N_B = 1000$ continuum ZZ background events, corresponding to a $\simeq 5\sigma$ deviation from the background-only hypothesis. The superimposed curves represent the likelihood function re-



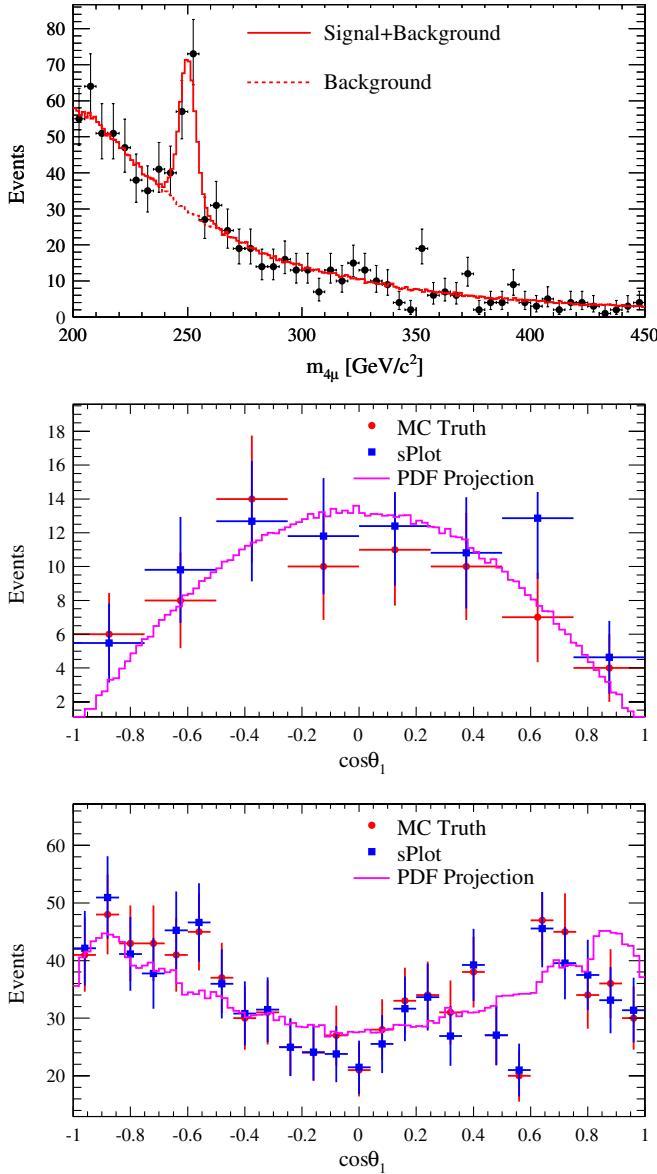


FIG. 10 (color online). The 4μ invariant mass distribution for a sample of $N_S = 70H \rightarrow ZZ$ events with $m_H = 250 \text{ GeV}/c^2$ and $N_B = 1000ZZ$ background events. The superimposed curves represent the likelihood function returned by an ML fit, with N_S , N_B , and $m_{4\mu}$ as free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

turned by an ML fit (with N_S , N_B , and $m_{4\mu}$ as free parameters). The middle plot shows the signal s -Weighted $\cos\theta_1$ distribution. Similarly, the bottom plot shows the background s -Weighted $\cos\theta_1$ 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 *pdfs* 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 for the 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) involve attempting to establish the nature of a newly discovered particle. Guided by our results on s -Plots, we contend that it is a very good approximation to confront two different “signal” hypotheses in the absence of background—the latter having been statistically subtracted. This assumes that a resonance mass peak 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 fixed. In the (2b) case the unknown mixing angles for the impure hypothesis, referred to as $\vec{\xi}$ (and including e.g. various mixing angles ξ and δ as discussed in Sec. VII F), 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 [47]. 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 ξ , Eq. (23) 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,

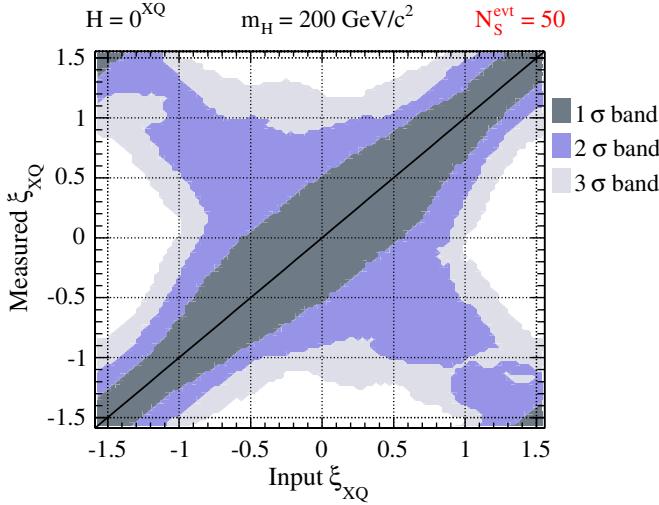


FIG. 11 (color online). Confidence intervals for measured values of ξ_{XQ} for a C -violating $J = 0$ resonance with a mass $200 \text{ GeV}/c^2$.

generated with the theoretical *pdf* of Eq. (23). This experimental *pdf*, referred to as P , is a function $P = P_{m_H}(\xi, \vec{X})$ of m_H (which is kept fixed through this exercise) ξ , 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 $\mathcal{L}(\xi) = \prod_{i=1}^{N_S} P_i$. The experimentally measured value of the ξ parameter, $\hat{\xi}$ corresponds to the value that maximizes $\mathcal{L}(\xi)$. The simulation is repeated many times, as a function of the true value of the mixing angle ξ . Running many experiments one can derive the confidence interval, i.e. the range covering the true value of ξ for some confidence level and some measured value $\hat{\xi}$ [47].

It is customary to estimate the error (or the number n of standard deviations σ) in the measured ξ from the expression $\mathcal{L}(\xi_{\max} \pm n\sigma) = \mathcal{L}(\xi_{\max}) - n^2/2$. 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 $\xi = \xi_{XQ}$ as given in Eq. (33) and in the similar figures of Sec. VII). Instead, the confidence level (CL) is evaluated measuring the frequency of a given result in the set of generated pseudoexperiments.

B. Confronting 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 s Plot method. Large numbers of events are generated assuming each hypothesis and used to construct two unbinned experimental *pdfs*: $P_{\mathbb{H}_{0,1}} \equiv P_{m_H}(\vec{X} | \mathbb{H}_{0,1})$. For our pure spin-

parity cases, the simple nature of the hypotheses considered guarantees through the Neyman-Pearson (NePe) lemma [48] that the hypothesis test is *universally most powerful*. Next, we explicitly identify 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-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 pseudoexperiments 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 pseudoexperiments 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 $P(\Lambda | \mathbb{H}_0)$ 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 pseudoexperiments 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

$$\alpha = \int_{\hat{\Lambda}(\alpha)}^{\infty} P(\Lambda | \mathbb{H}_0) d\Lambda. \quad (29)$$

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

$$\beta(\alpha) = \int_{-\infty}^{\hat{\Lambda}(\alpha)} P(\Lambda | \mathbb{H}_1) d\Lambda \quad (30)$$

is the probability of accepting \mathbb{H}_0 even though \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

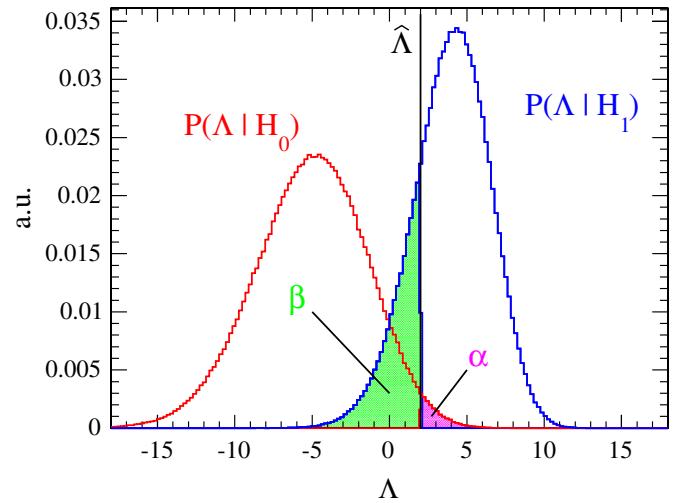


FIG. 12 (color online). Distribution of Λ for $m_H = 200 \text{ GeV}/c^2$ and $N_S = 23$, constructed with $\sim 10^9$ pseudoexperiments. The hypotheses being confronted are $\mathbb{H}_0 = 0^+$ and $\mathbb{H}_1 = 0^-$.

value Λ_{exp} , is obtained for Λ . The associated *p-value* = $\int_{\Lambda_{\text{exp}}}^{\infty} P(\Lambda \mid \mathbb{H}_0) d\Lambda$, is compared to α to determine if the measurement favors one hypothesis versus the other.

Instead of the α and β values, the significance σ 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 $\sigma = 1$. The number n of α -equivalent standard deviations is obtained by inverting

$$\alpha = \frac{1}{\sqrt{2\pi}} \int_n^{\infty} dx e^{-x^2/2}. \quad (31)$$

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 claim 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).

A pure vs impure HLL hypothesis test has an additional complication due to the dependence of the likelihood function on the mixing angles $\vec{\xi}$ in at least one of the two hypotheses. In this case, we are testing the simple (i.e. mixing angle 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 of $\vec{\xi}$ 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 mixing angle 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 mixing angles, 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 background, \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 the information in the theoretical *pdfs* 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 computing the significance (the number of σ 's) as a function of the signal yield and for different values of the ratio of signal-over-background yields. 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 $190 \text{ GeV}/c^2$ to $600 \text{ GeV}/c^2$. To address the uncertainties, we compare the two hypotheses for various preselected 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. (29), 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 be done using a different \vec{X} *pdf* for each bin of the template functions of Fig. 9. This step is straightforward when performing the real analysis, but CPU intensive when performing hundreds of billions of pseudoexperiments. The resonance mass is narrow enough for the peak to be determined, in which case the results are very close to the ones obtained with the full mass-dependence of the \vec{X} *pdf*. For simplicity we did not include the finite width of the resonance in the likelihood.

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 5σ 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 following a putative Higgs discovery. As noted already, the search analysis could be model-independent, relying only

on the reconstruction of a resonant excess over nonresonant backgrounds. In this case a discovery is completely factorized from its characterization.

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 data sets considered here, where we demonstrate HLL discrimination with data sets comparable to the original discovery sample.

A powerful check is to compare the signal significance of two nominal analyses:

- (i) 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).
- (ii) 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 HLLs. The *pdf* of \vec{X} corresponds to the value of m_H as extracted from the fit.

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 than determining the optimal statistical approach. The overall normalization is obtained by assuming $\sqrt{s} = 10$ TeV with a corresponding SM Higgs production cross section [42].

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 data sets 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 data sets.

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/ c^2 . The fit configuration is specified by the nominal expected signal-overbackground yield ratio $\langle N_S/N_B \rangle$ and by the nominal number of signal events N_S . We consider different scenarios by fixing different values of $\langle N_S/N_B \rangle$ and perform the study as a function of N_S .

For each fit configuration we run 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 background yields and the value of m_H . The sets of fits provide a distribution for the statistical significance obtained in a particular experiment.

This is summarized in Figs. 13–15. The two bands in the figures correspond to the 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”

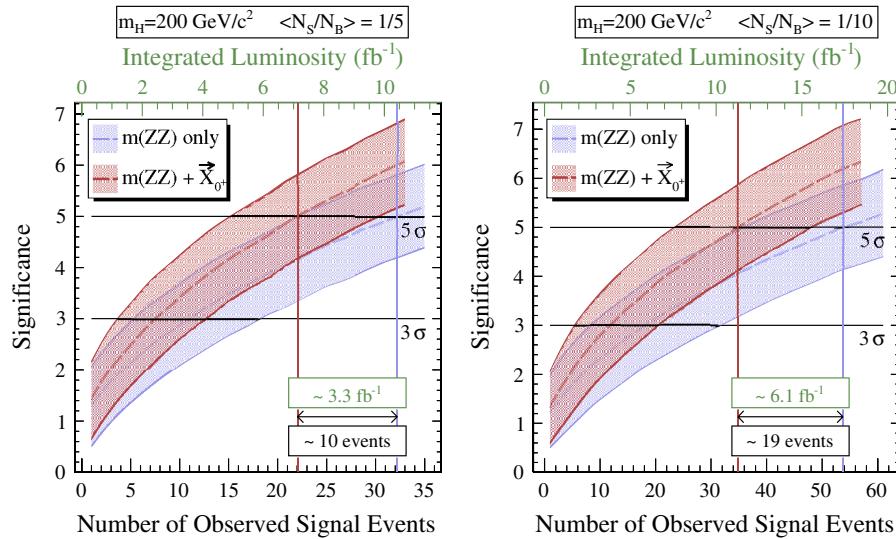


FIG. 13 (color online). Distribution of signal significance for a 200 GeV/ c^2 SM Higgs boson decaying in the $H \rightarrow ZZ \rightarrow 4\mu$ channel for pp collisions with $\sqrt{s} = 10$ TeV. The mean signal-to-background ratios used are $\langle N_S/N_B \rangle = 1/5$ (left) and $1/10$ (right).

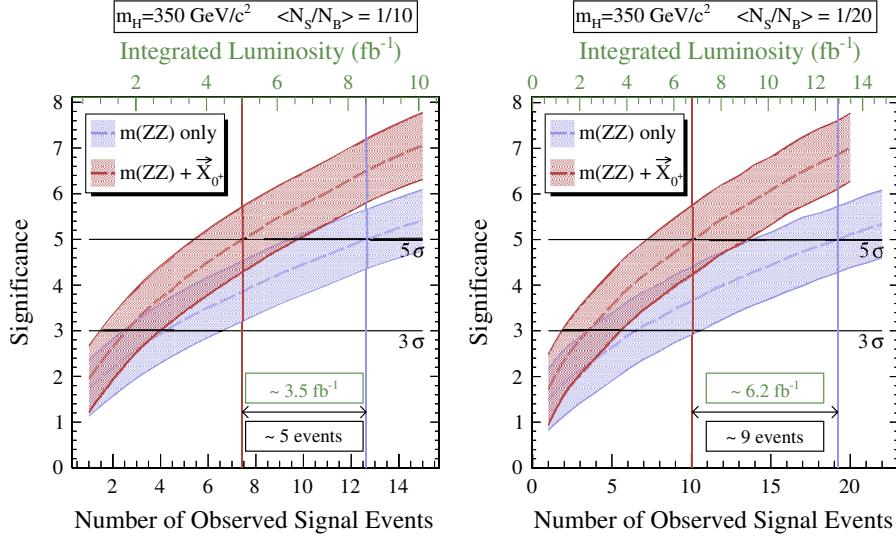


FIG. 14 (color online). Distribution of signal significance for a $350 \text{ GeV}/c^2$ SM Higgs boson decaying in the $H \rightarrow ZZ \rightarrow 4\mu$ channel for pp collisions with $\sqrt{s} = 10 \text{ TeV}$. The mean signal-to-background ratios used are $\langle N_S/N_B \rangle = 1/10$ (left) and $1/20$ (right).

fit (light band) and the “ $m(ZZ) + \vec{X}_0$ ” fit (dark band). The horizontal lines show the 3σ and 5σ thresholds (*evidence* and *discovery*, in the usual convention). 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. The vertical lines guide the eye to show the significant difference between the median number of expected events (or integrated luminosity) required to achieve 5σ significance.

Figure 13 has the case of an $m_H = 200 \text{ GeV}/c^2$ SM Higgs boson, while Fig. 14 illustrates similar results for an

$m_H = 350 \text{ GeV}/c^2$ SM Higgs boson. For each mass, different values for $\langle N_S/N_B \rangle$ are considered; we show here the results for $\langle N_S/N_B \rangle = 1/5, 1/10$ for $m_H = 200 \text{ GeV}/c^2$ and $\langle N_S/N_B \rangle = 1/10, 1/20$ for $m_H = 350 \text{ GeV}/c^2$. 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; despite the lower cross section this results in higher significance for the higher mass case for the same luminosity.

Similarly, Fig. 15 has the results for $m_H = 200 \text{ GeV}/c^2$ and $m_H = 350 \text{ GeV}/c^2$ pseudoscalar HLLs. 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

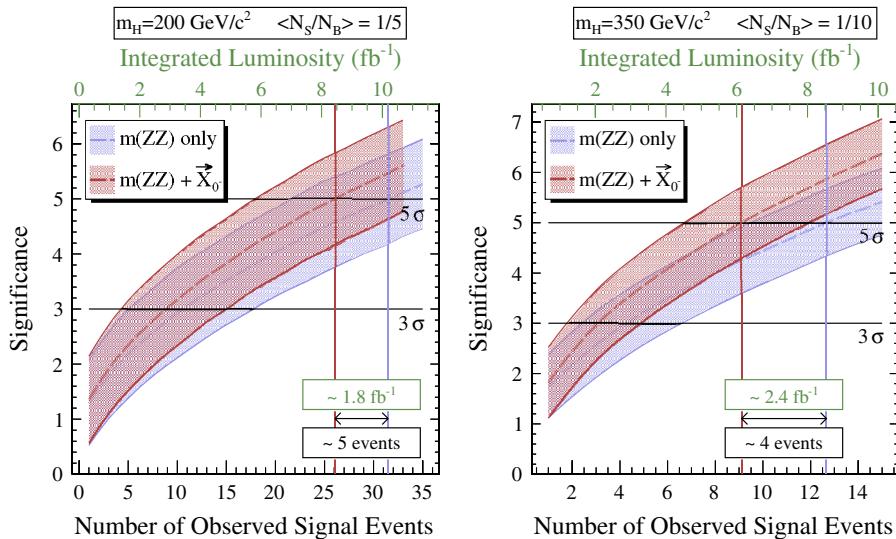


FIG. 15 (color online). Distribution of signal significance for a 200 (left) and 350 (right) GeV/c^2 pseudoscalar particle in the $0^- \rightarrow ZZ \rightarrow 4\mu$ channel for pp collisions with $\sqrt{s} = 10 \text{ TeV}$. The mean signal-to-background ratios used are $\langle N_S/N_B \rangle = 1/5$ (left) and $1/10$ (right), and we assume an NLO SM Higgs production cross section for the pseudoscalar.

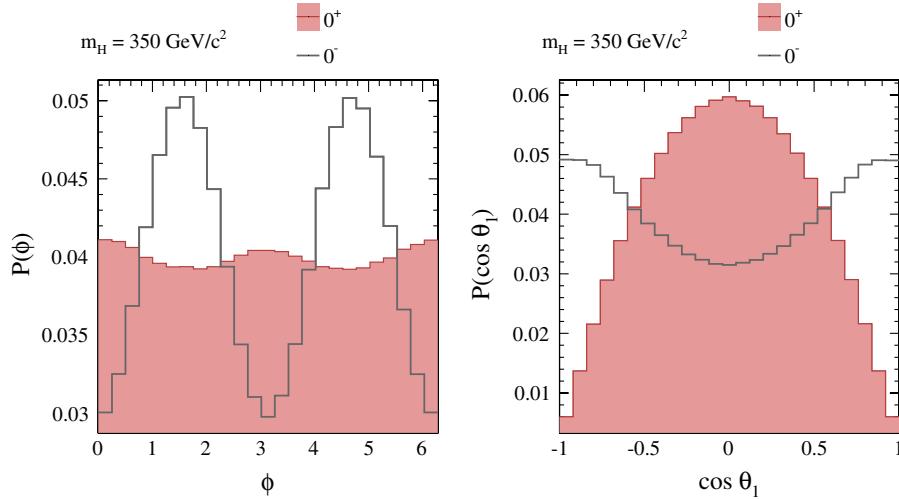


FIG. 16 (color online). Distributions of the variables ϕ (left) and $\cos\theta_1$ (right) for 0^+ and 0^- resonances with $m_H = 350 \text{ GeV}/c^2$. All distributions are normalized to a unit integral.

between the angles) are different from the SM case. The angular distributions and correlations for a pseudoscalar resonance are similar to those of 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.

VII. RESULTS

We present results for three HLL masses: $m_H = 145$, 200 , and $350 \text{ GeV}/c^2$, using pseudoexperiments built with the full \vec{X} pdf.

A. $0^+ \text{ vs } 0^-$

We consider here two different pure scalar hypotheses: 0^+ , corresponding to an SM Higgs, and 0^- , a pseudoscalar.

Neither of these possibilities has an 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, $m_2 = M_{Z^*}$, for $m_H < 2M_Z$) are used to discriminate between the two hypotheses.

In Fig. 16 we show the distributions in ϕ and $\cos\theta_1$ at $m_H = 350 \text{ GeV}/c^2$ 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 $2M_Z$ threshold, the kinematic factors in Eqs. (14) and (15) 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 illustrated in Fig. 17 (left) for $m_H = 145 \text{ GeV}/c^2$. For all the discriminating variables we consider, the ability to distinguish between

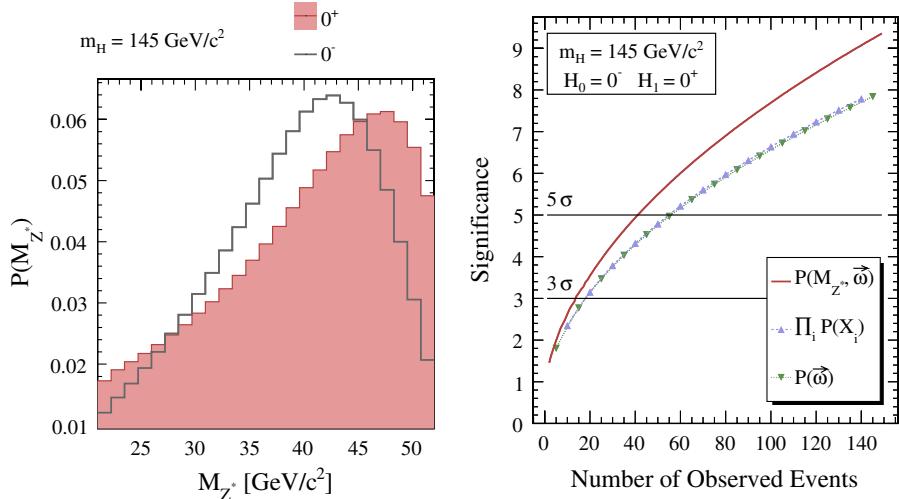


FIG. 17 (color online). Left: Distribution of M_{Z^*} for 0^+ and 0^- $H \rightarrow ZZ^*$ decays at $m_H = 145 \text{ GeV}/c^2$, normalized to a unit integral. Right: Median significance for rejecting 0^- in favor of 0^+ , assumed to be correct, as a function of N_S . The different likelihood constructions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

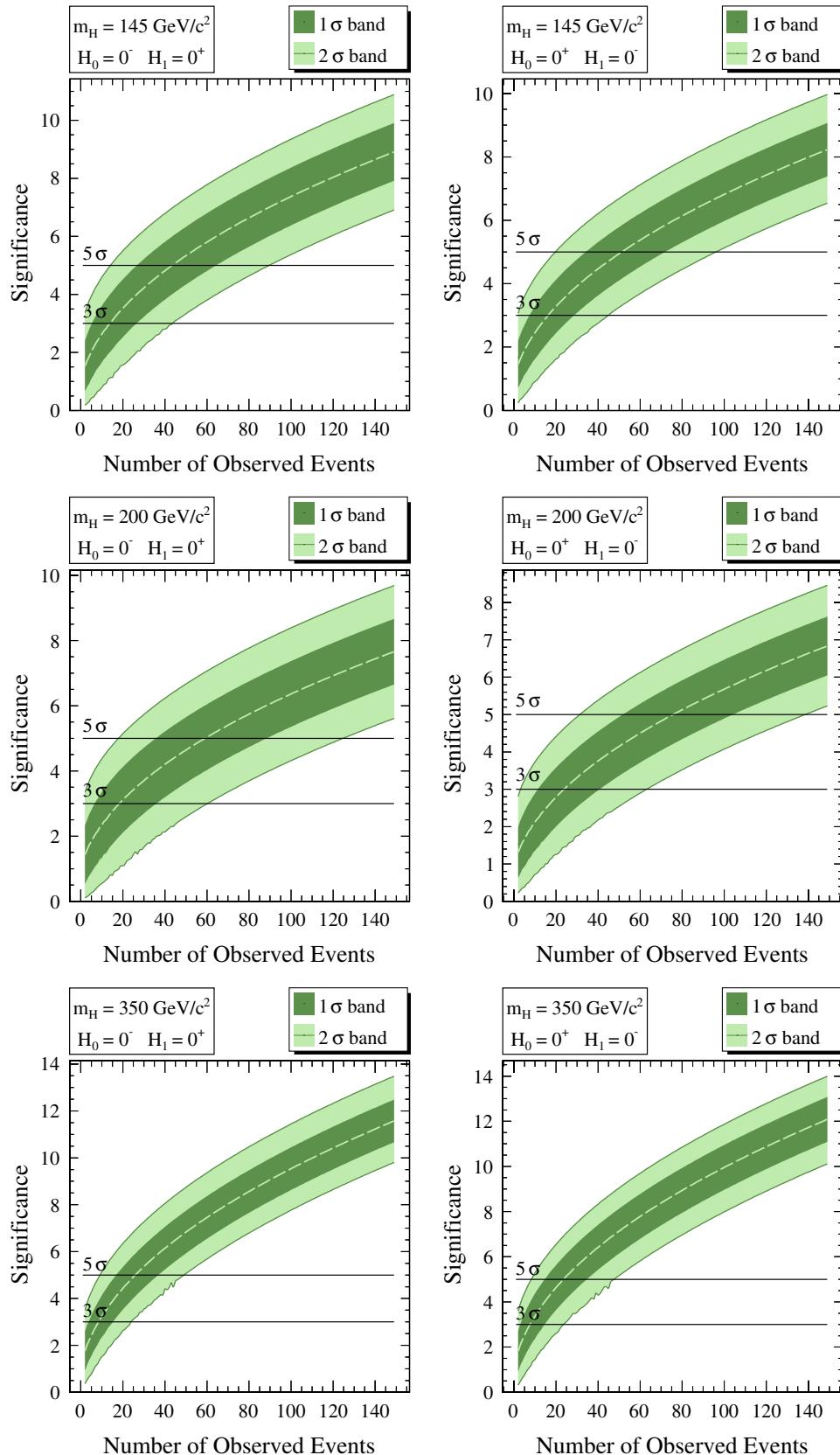


FIG. 18 (color online). Significance for rejecting 0^- in favor of 0^+ , assuming 0^+ is true (left), and vice-versa, $0^+ \leftrightarrow 0^-$ (right), for $m_H = 145, 200$ and $350 \text{ GeV}/c^2$ (top, middle and bottom). The dashed central line is the median significance. The 1 and 2σ bands correspond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

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 significance for discriminating between the 0^+ and 0^- hypotheses (assuming one or 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 $350 \text{ GeV}/c^2$. 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 a NePe test between these simple hypotheses with test statistic $\log(\mathcal{L}[0^+]/\mathcal{L}[0^-])$. The variables $\vec{\omega}$ (and M_{Z^*} , when applicable), along with their correlations, are used in the likelihood construction. The significance for rejecting one hypothesis in favor of the other at the time of 5σ excess (see Sec. VI) is better than 3σ for $m_H = 145, 200$, and $350 \text{ GeV}/c^2$ while a 5σ discrimination can be achieved with twice the observed signal events (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 nontrivial dependences on the Z -production angles $\vec{\Omega}$ that provide additional dis-

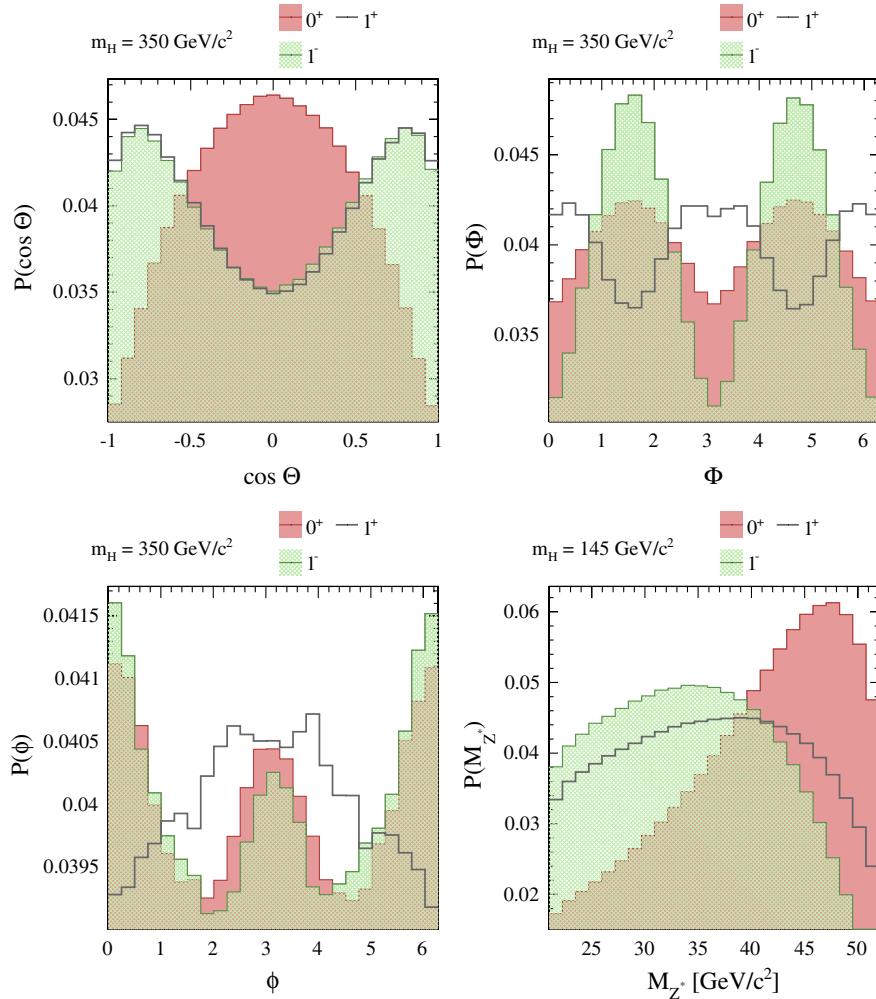


FIG. 19 (color online). Distributions of the variables $\cos\Theta$ (upper left), Φ (upper right), ϕ (lower left) and M_{Z^*} (lower right, $m_H = 145 \text{ GeV}/c^2$) for 0^+ , 1^- and 1^+ resonances. All distributions are normalized to a unit integral. The angular distributions are shown for $m_H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

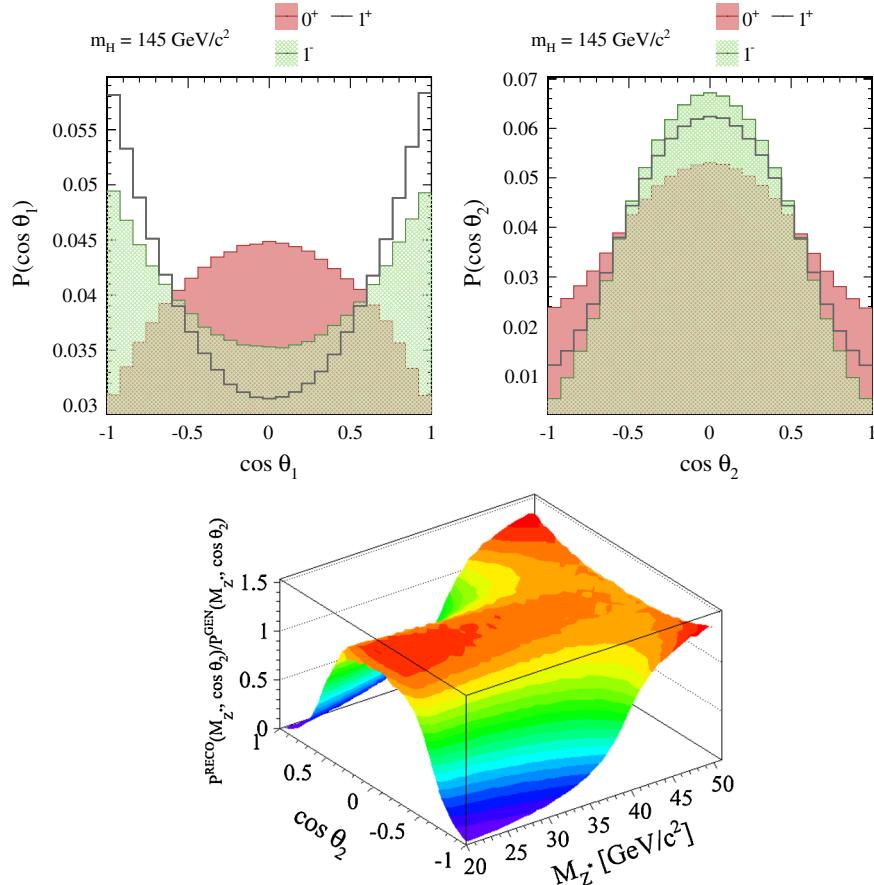


FIG. 20 (color online). Distributions in $\cos\theta_1$ (left) and $\cos\theta_2$ (right), for 0^+ , 1^- and 1^+ resonances with mass $145 \text{ GeV}/c^2$, normalized to a unit integral. The very strong correlation between $M_{Z''}$ and $\cos\theta_2$ is also shown (bottom) for the $J = 1$ case.

crimination 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 distributions in $c_1 \equiv \cos\theta_1$ 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). Figure 20 also shows the very strong correlation between $M_{Z''}$ and $\cos\theta_2$. 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 conclu-

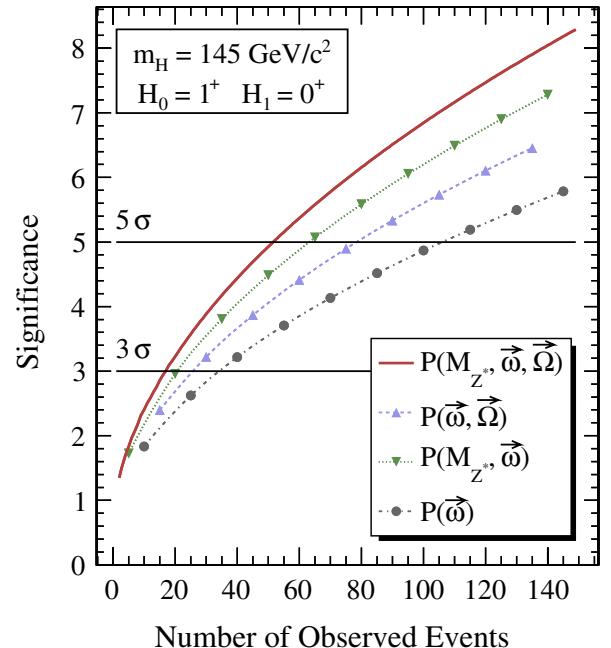


FIG. 21 (color online). Median significance for rejecting 1^+ in favor of 0^+ (assuming 0^+ is true), for different likelihood constructions used in the likelihood ratio test statistic.

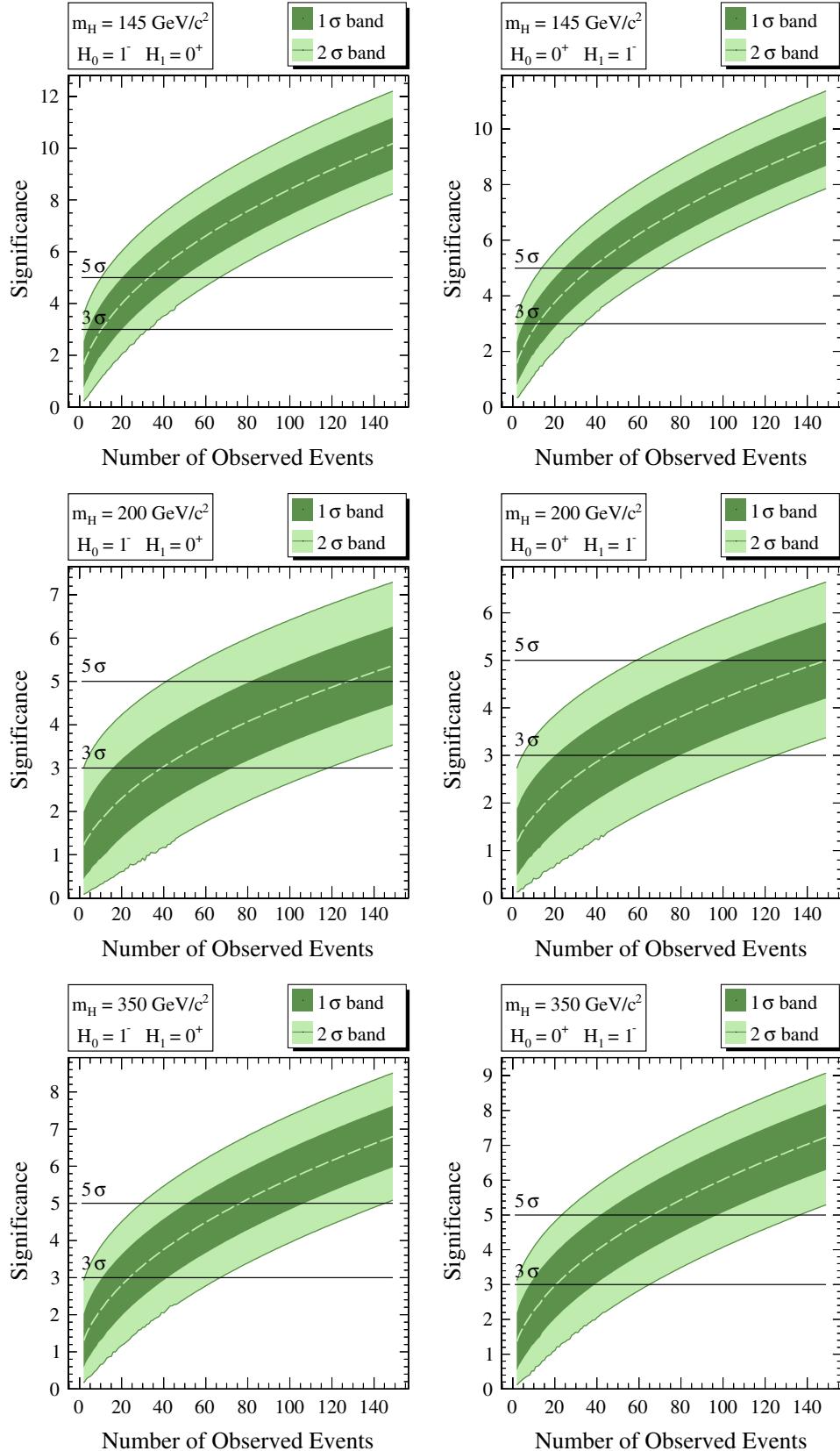


FIG. 22 (color online). Significance for rejecting 1^- in favor of 0^+ , assuming 0^+ is true (left), or vice-versa ($0^+ \leftrightarrow 1^-$, right), for $m_H = 145, 200$ and $350 \text{ GeV}/c^2$ (top, middle and bottom).

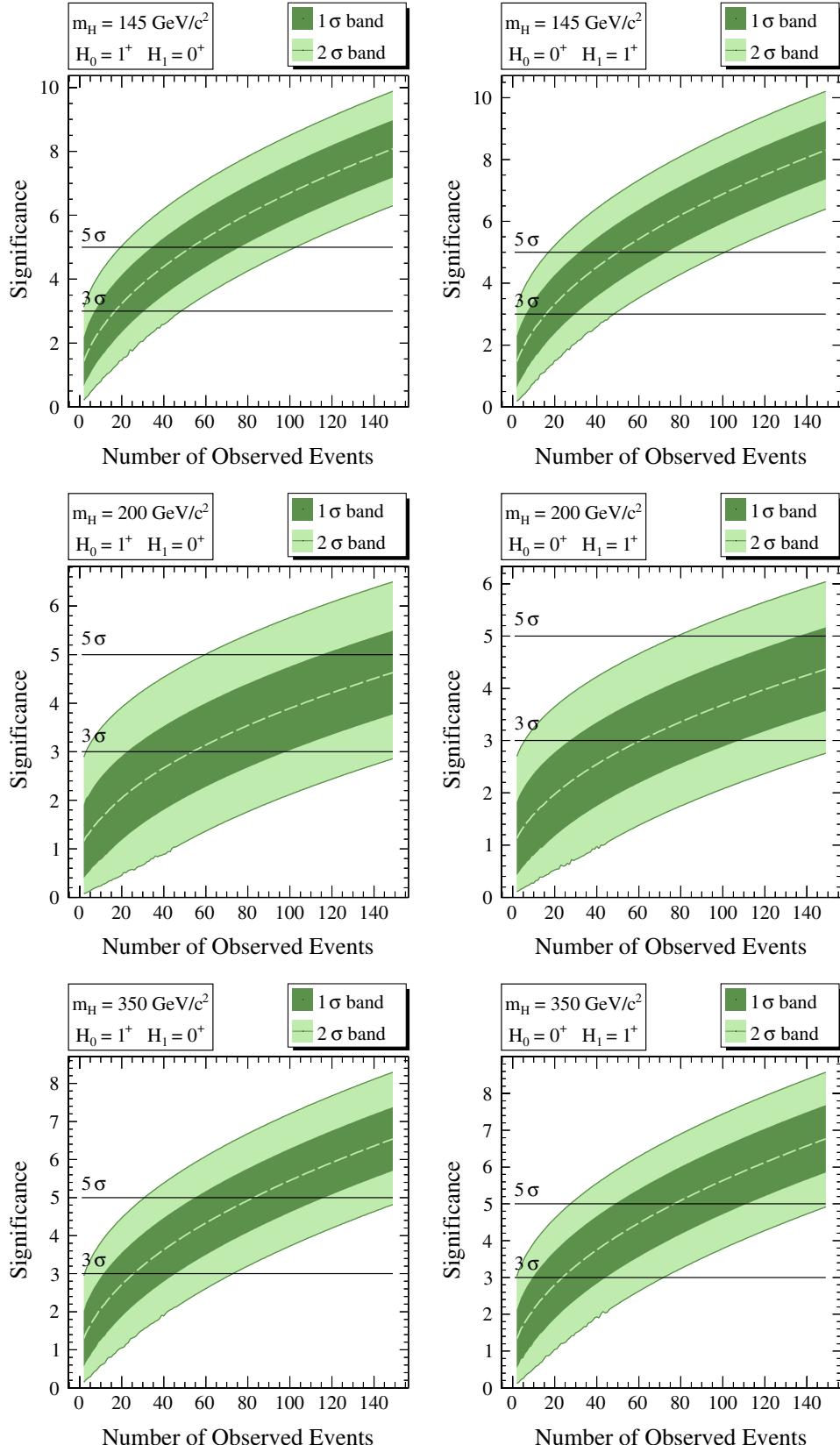


FIG. 23 (color online). Significance for rejecting 1^+ in favor of 0^+ , assuming 0^+ is true (left), or vice-versa ($0^+ \leftrightarrow 1^+$, right), for $m_H = 145, 200$ and $350 \text{ GeV}/c^2$ (top, middle and bottom).

sion 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: $P(a_1, \dots, a_N)$ denotes N -dimensional *pdfs* in the correlated variables $\{a_1, \dots, a_N\}$. In $P(M_{Z^*}, \vec{\omega}, \vec{\Omega})$ we employ all five angles and the Z^* invariant mass, with all their correlations; in $P(\vec{\omega}, \vec{\Omega})$ we use all five angles but not the invariant mass, giving a test statistic with significantly less power. Worse still is $P(M_{Z^*}, \vec{\omega})$, which ignores the Z pair production angles Θ and Φ , and worst of all is $P(\vec{\omega})$, which uses only the three “traditional” decay angles.

The significance for discriminating between the 0^+ and 1^- (1^+) hypotheses, as a function of N_S , is summarized in Figs. 22 and 23. The full correlated set 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 $\log(\mathcal{L}[0^+]/\mathcal{L}[1^-])$ ($\log(\mathcal{L}[0^+]/\mathcal{L}[1^+])$). The discrimination between 0^+ and 1^- or 1^+ is similar.

C. 0^+ vs 2^+

We consider one pure spin 2 model: a $J = 2^+$ heavy gravitonlike 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 for $m_H = 200$ and $350 \text{ GeV}/c^2$. 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 $2M_Z$, 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 these hypotheses as shown in Fig. 26.

The significance for discriminating between 0^+ and 2^+ as a function of N_S , is summarized in Fig. 27 for $m_H = 200$

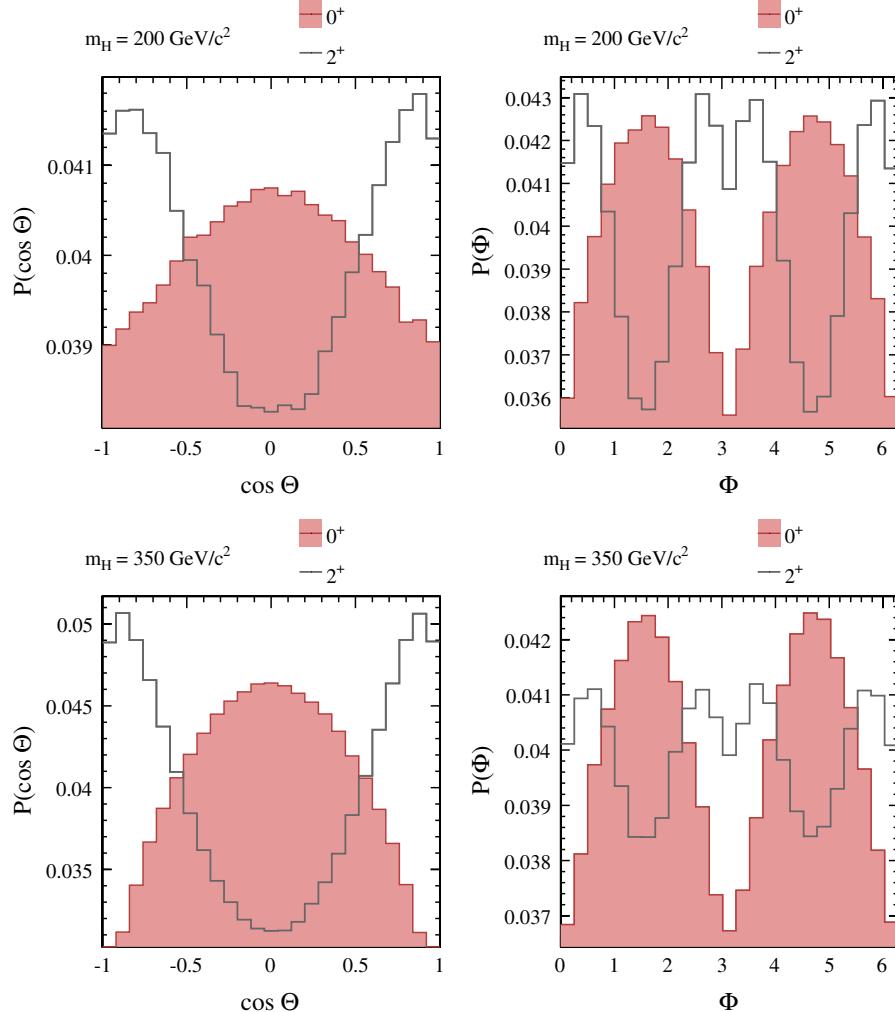


FIG. 24 (color online). Distributions of the variables $\cos\Theta$ (left) and Φ (right) for 0^+ , 2^+ resonances with masses of 200 and $350 \text{ GeV}/c^2$ (top, bottom). All distributions are normalized to a unit integral.

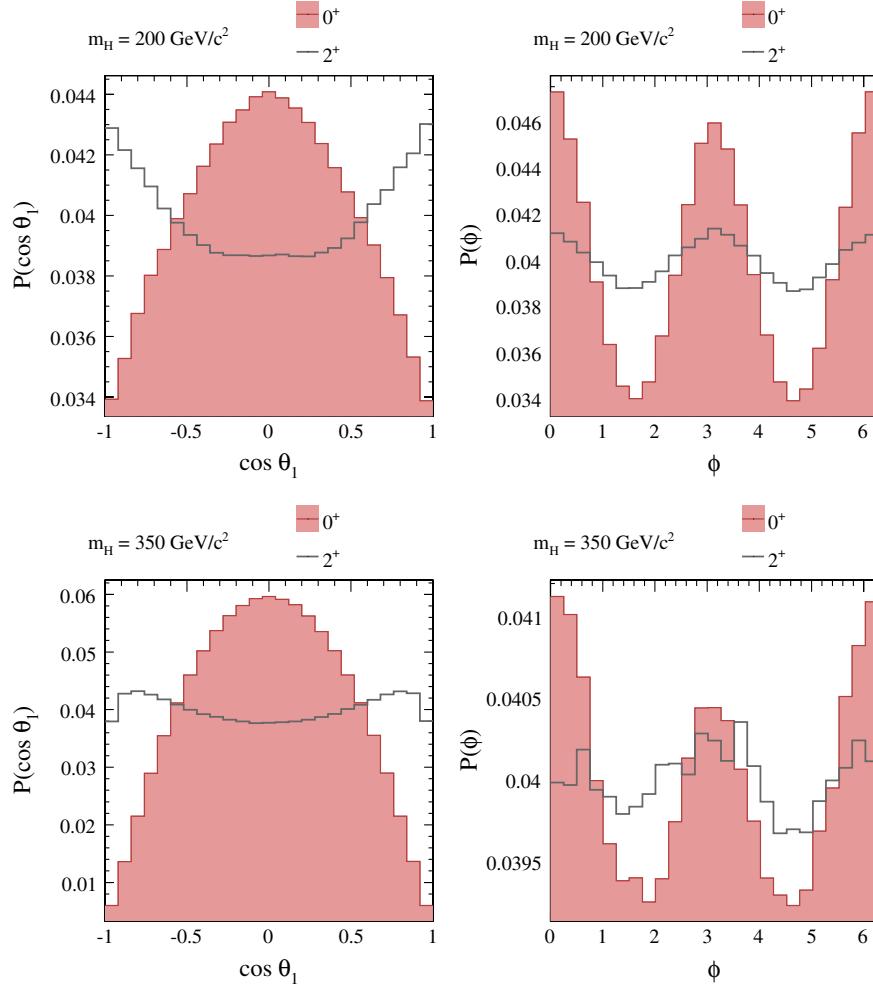


FIG. 25 (color online). Distributions of the variables $\cos\theta_1$ (left) and ϕ (right) for 0^+ , 2^+ resonances with masses of 200 and 350 GeV/c^2 (top, bottom). All distributions are normalized a unit integral.

and $350 \text{ GeV}/c^2$. 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 $\log(\mathcal{L}[0^+]/\mathcal{L}[2^+])$ and $\log(\mathcal{L}[0^+]/\mathcal{L}[2^-])$.

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 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 always use the full set of angular variables plus, when appropriate, M_{Z^*} , corresponding to the optimal statistic for model discrimination.

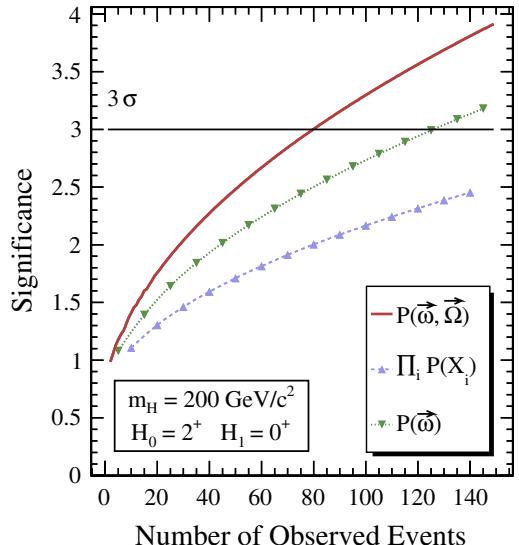


FIG. 26 (color online). Median significance for rejecting 2^+ in favor of 0^+ , assuming 0^+ is true, for the different likelihood constructions discussed in the text.

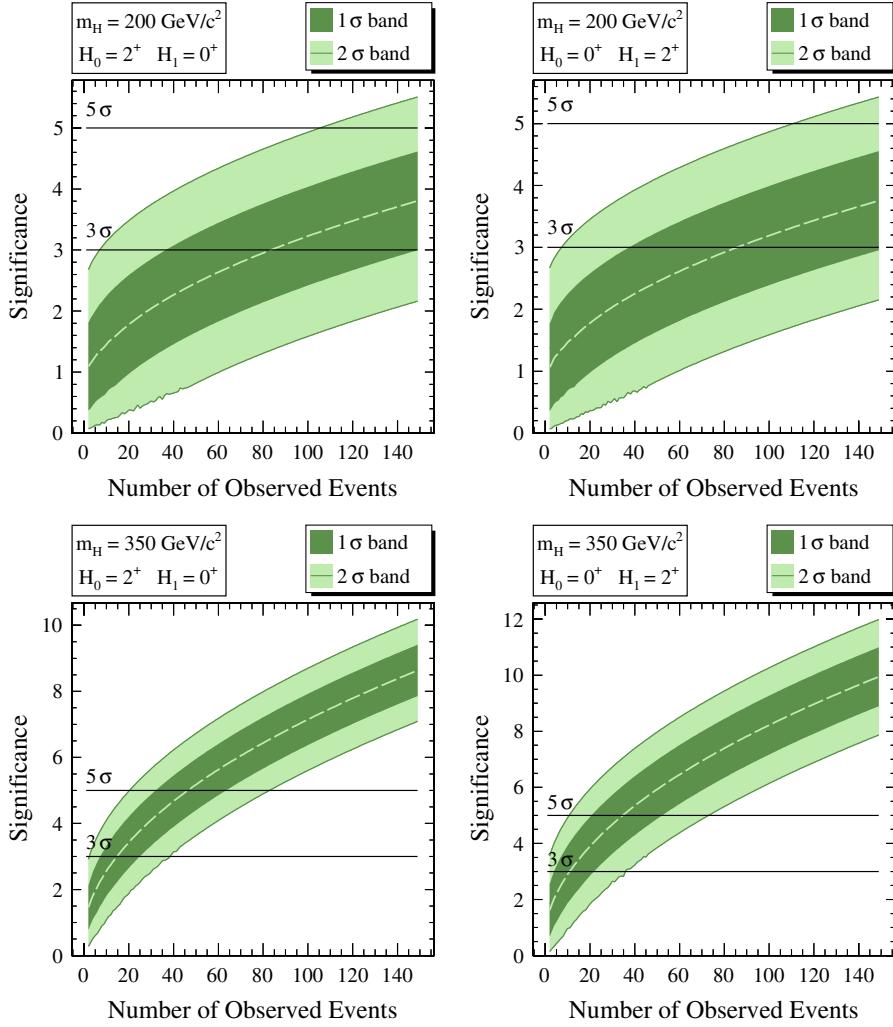


FIG. 27 (color online). Significance for rejecting 2^+ in favor of 0^+ , assuming 0^+ is true (left) or vice-versa ($0^+ \leftrightarrow 2^+$, right), for $m_H = 200$ and $350 \text{ GeV}/c^2$ (top, bottom).

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 $m_H < 2M_Z$, 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 Figs. 29 and 30. The $m_H = 145 \text{ GeV}/c^2$ 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 $m_H = 200 \text{ GeV}/c^2$ reveals that it is more difficult to discriminate between 0^+ and $J = 1$ at this mass. This is predominantly due to the *pdfs* for the angles $\cos\theta_{1,2}$ (which are similar for 0^+ and $J = 1$ for $m_H > 2M_Z$).

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 $m_H = 200 \text{ GeV}/c^2$. 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, exploiting the distributions shown in Figs. 31 and 33. If both of the $J = 0$ cases are excluded in favor of $J = 1$ or $J = 2$, one needs to discriminate between the latter. Relative to the $J = 0$ 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 $J = 2$ case reflects the same limitation, as shown in Figs. 34 and 35 for 1^- vs 2^+ (1^+ vs 2^+).

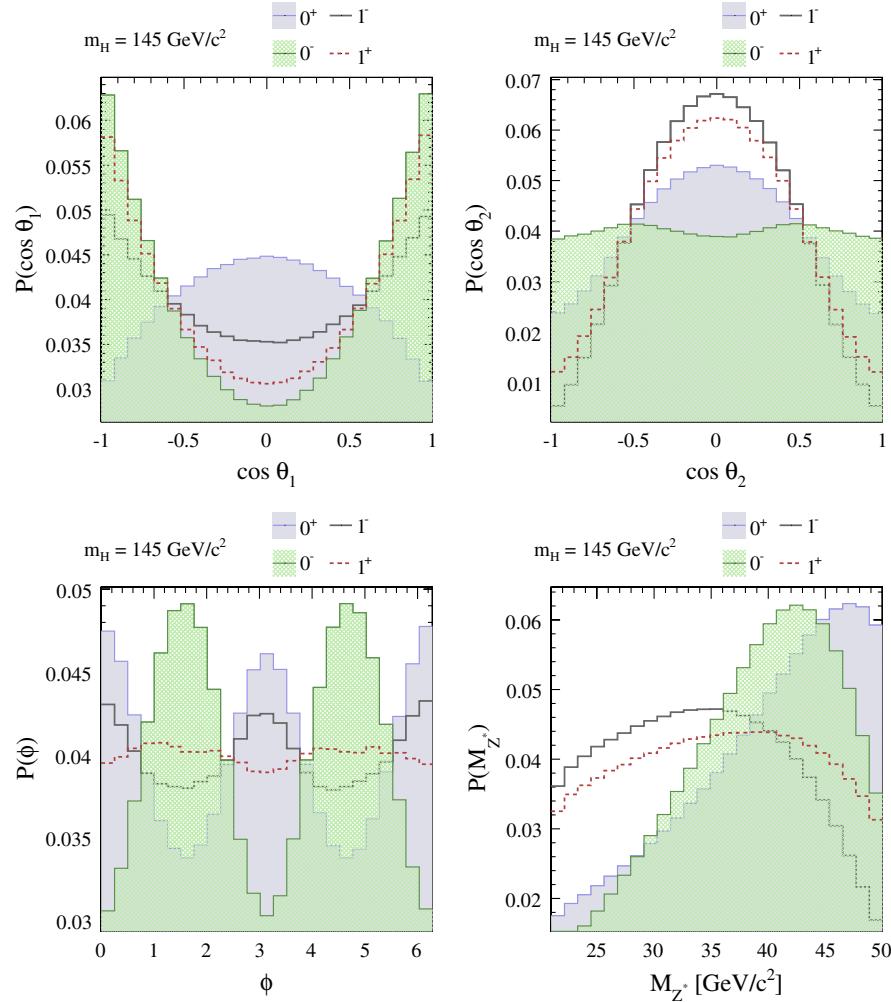


FIG. 28 (color online). Distributions, normalized to a unit integral, of the variables $\cos\theta_1$ (top left), $\cos\theta_2$ (top right), ϕ (bottom left) and M_{Z^*} (bottom right) for 0^+ , 0^- , 1^+ and 1^- resonances with $m_H = 145 \text{ GeV}/c^2$.

The hardest differentiation is between 1^- and 1^+ . Figs. 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 significance for distinguishing between the two $J = 1$ cases is shown in Fig. 36. We conclude that the discriminating potential 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. (9) 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 nonvanishing

types of coupling, resulting in one free mixing “angle” and an overall normalization (which we ignore):

- (i) $X \neq 0, P \neq 0$: A scalar whose ZZ coupling violates CP , described in terms of an angle ξ_{XP} as:

$$\mathcal{L}_{\mu\alpha} \propto \cos(\xi_{XP})g_{\mu\alpha} + \sin(\xi_{XP})\epsilon_{\mu\alpha}p_1p_2/M_Z^2 \quad (32)$$

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

$$\mathcal{L}_{\mu\alpha} \propto \cos(\xi_{XQ})g_{\mu\alpha} + i\sin(\xi_{XQ})\epsilon_{\mu\alpha}p_1p_2/M_Z^2 \quad (33)$$

- (iii) $X \neq 0, Y \neq 0$: A composite 0^+ , parametrized in terms of an angle as:

$$\mathcal{L}_{\mu\alpha} \propto \cos(\xi_{XY})g_{\mu\alpha} - \sin(\xi_{XY})k_\alpha k_\mu/M_Z^2 \quad (34)$$

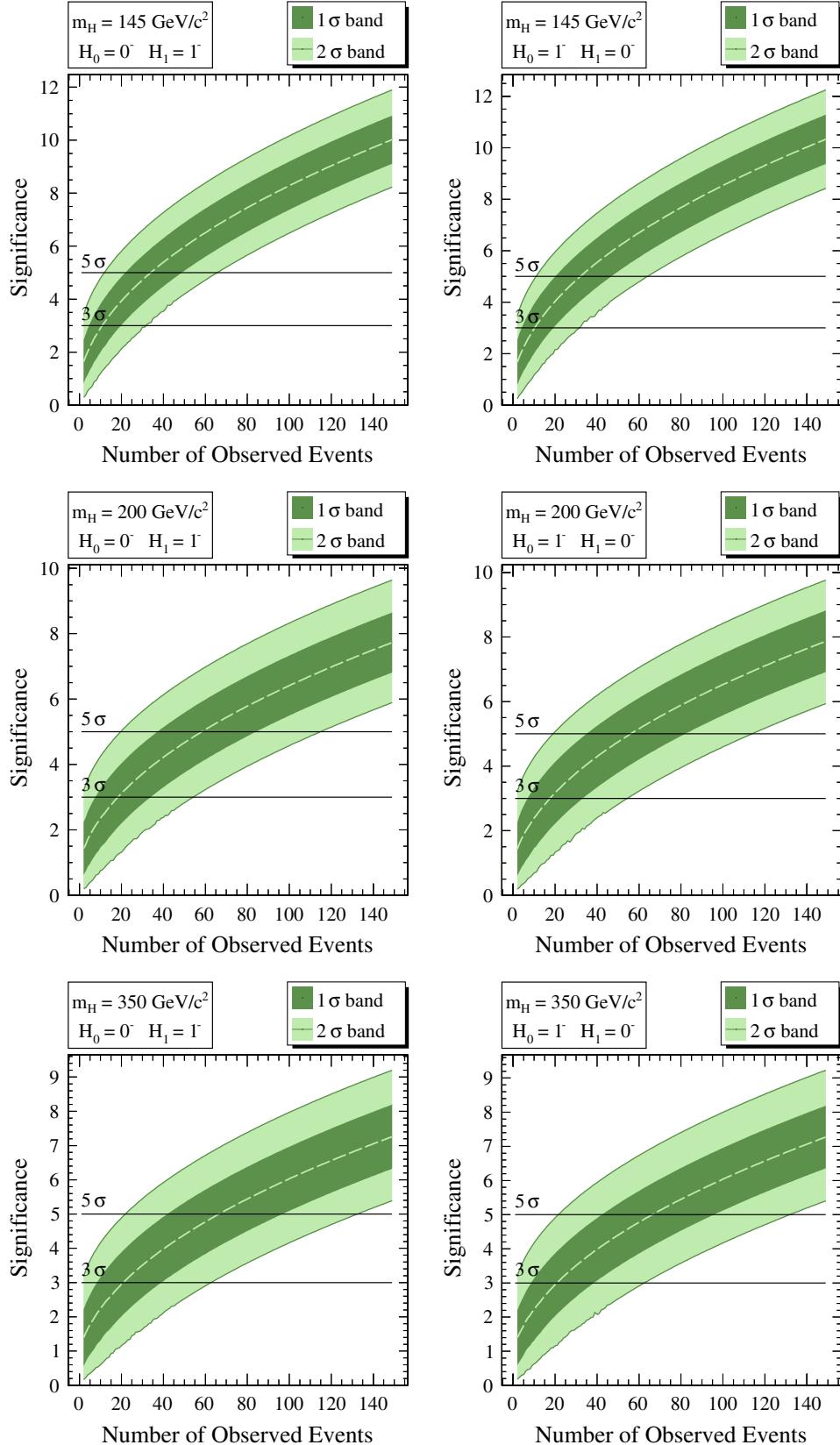


FIG. 29 (color online). Significance for rejecting 0^- in favor of 1^- , assuming 1^- is true (left) or vice-versa ($0^- \leftrightarrow 1^-$, right) for $m_H = 145, 200$ and 350 GeV/c^2 (top, middle and bottom).

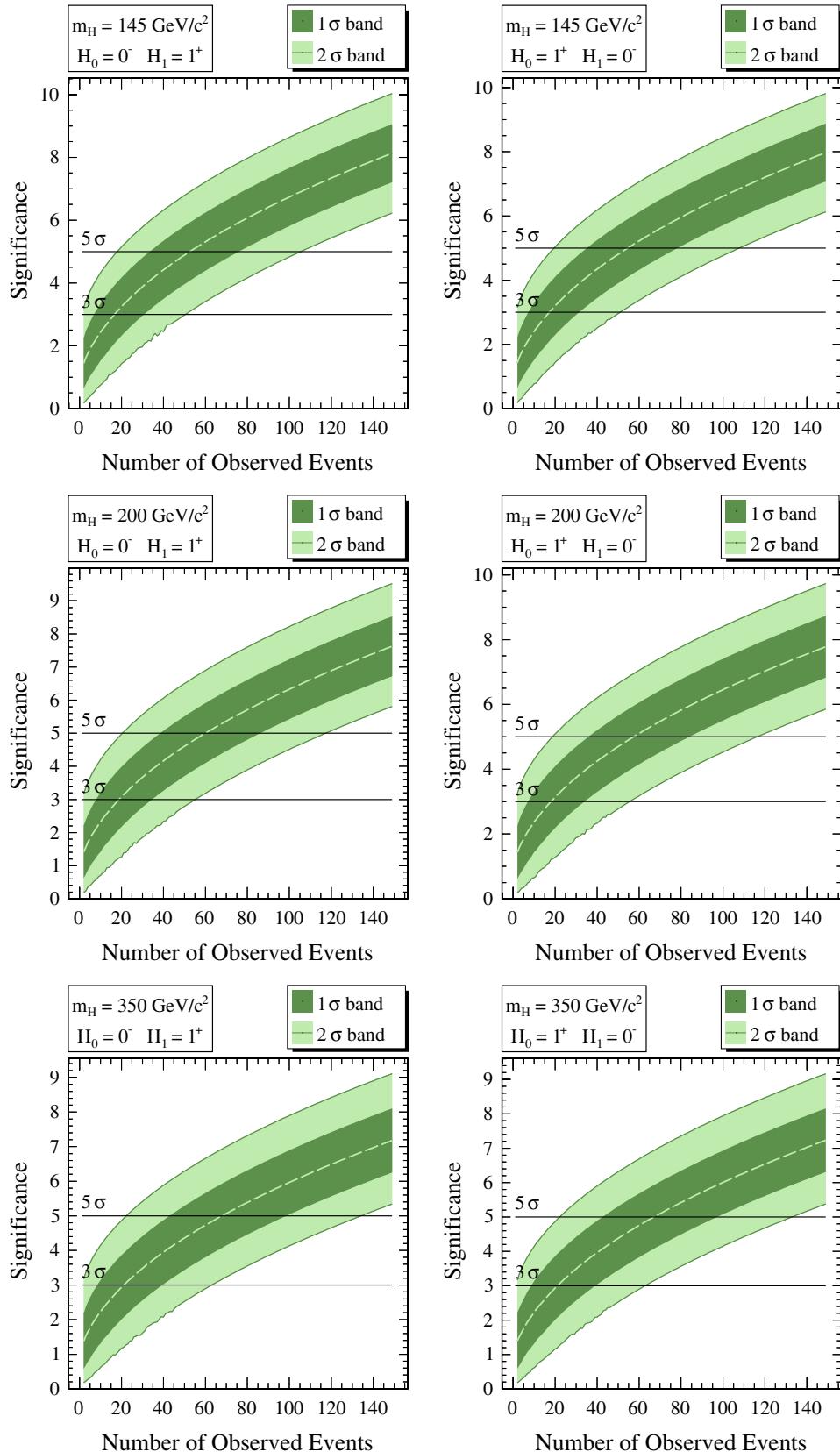


FIG. 30 (color online). Significance for rejecting 0^- in favor of 1^+ , assuming 1^+ is true (left) or vice-versa ($0^- \leftrightarrow 1^+$, right) for $m_H = 145, 200$ and $350 \text{ GeV}/c^2$ (top, middle and bottom).

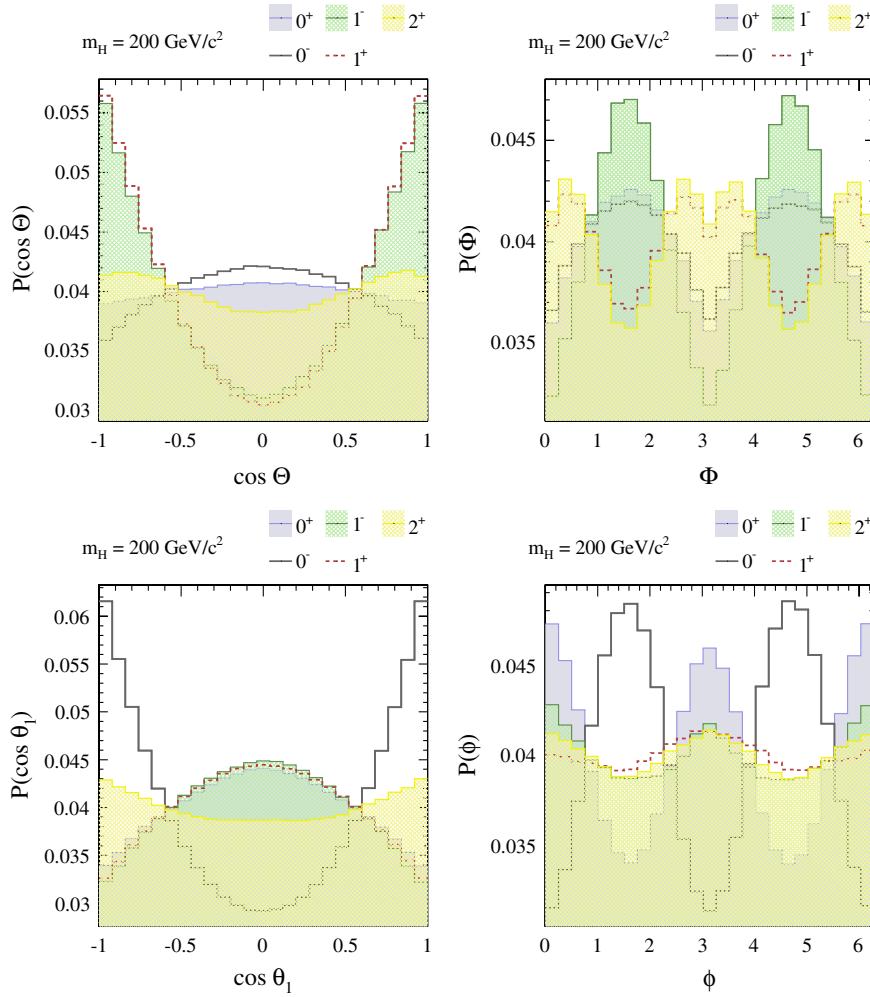


FIG. 31 (color online). Distributions of $\cos\Theta$ (top left), Φ (top right), $\cos\theta_1$ (bottom left) and ϕ (bottom right) for all the pure J^{PC} choices we study, for $m_H = 200 \text{ GeV}/c^2$. All distributions are normalized to a unit integral.

As a function of N_S we estimate the significance with which one can determine:

- (a) What range of values of the angles can be excluded in favor of a pure 0^+ for a SM-like resonance;
- (b) Whether a pure 0^+ can be excluded in favor of a nontrivial mixture when the resonance corresponds to one of the three mixed cases discussed above.

We consider first the example of a CP -violating HZZ coupling with $m_H = 350 \text{ GeV}/c^2$.

To address (a) we construct 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 $\log[\mathcal{L}^{XP}(\xi_{XP})/\mathcal{L}(0^+)]$, where $\mathcal{L}(0^+)$ and $\mathcal{L}^{XP}(\xi_{XP})$ denote the likelihoods for a set of events agreeing with the

hypotheses \mathbb{H}_1 and \mathbb{H}_0 , respectively. The test cannot be performed for $\xi_{XP} = 0$, since in this case the \mathbb{H}_0 CP -violating hypothesis we want to test reduces to the alternative \mathbb{H}_1 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 favor 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 significance for excluding a CP -violating coupling exceeds 3σ for $|\xi_{XP}| > 0.5$ and 5σ for $|\xi_{XP}| > 0.9$.

In addressing (b) we cannot construct a simple NePe test between 0^+ and a fixed- ξ_{XP} hypothesis. Instead, we treat

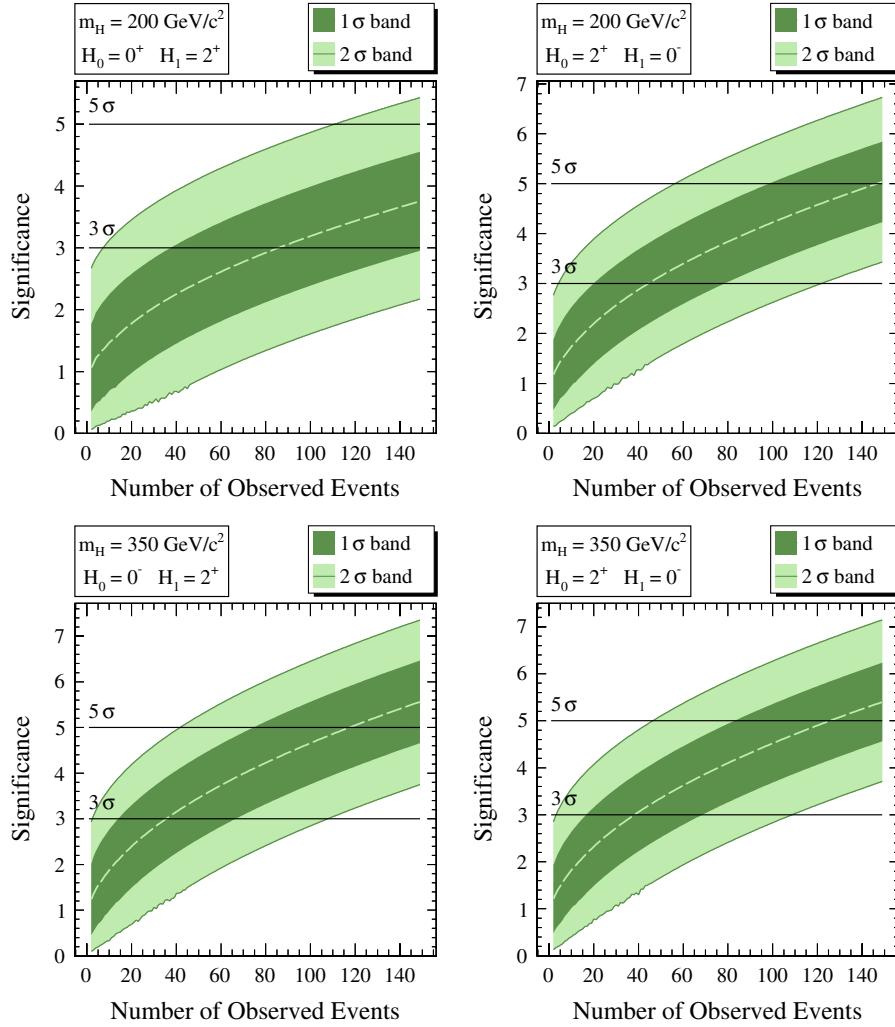


FIG. 32 (color online). Significance for rejecting 0^- in favor of 2^+ , assuming 0^- is true (left) or vice-versa ($0^- \leftrightarrow 2^+$, right) for $m_H = 200$ and 350 GeV/c^2 (top, bottom).

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 $\xi_{XP} = \hat{\xi}_{XP}$ (\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 question (a). What Fig. 38 shows is the expected significance with which one can exclude the SM hypothesis in favor of the CP -violating hypothesis with $\xi_{XP} = \hat{\xi}_{XP}$, 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 $\xi_{XP} \rightarrow -\xi_{XP}$, 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 $\Lambda^{\text{fix}} = \log[\mathcal{L}^{XP}(\xi_{XP})/\mathcal{L}(0^+)]$, with ξ_{XP} fixed at its true value, corresponding to a simple hypothesis test and $\Lambda^{\text{max}} = \log[\max \mathcal{L}^{XP}(\hat{\xi}_{XP})/\mathcal{L}(0^+)]$, profiled to the value $\hat{\xi}_{XP}$ at which it peaks. The distributions of Λ^{fix} and Λ^{max} are shown in Fig. 39.

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

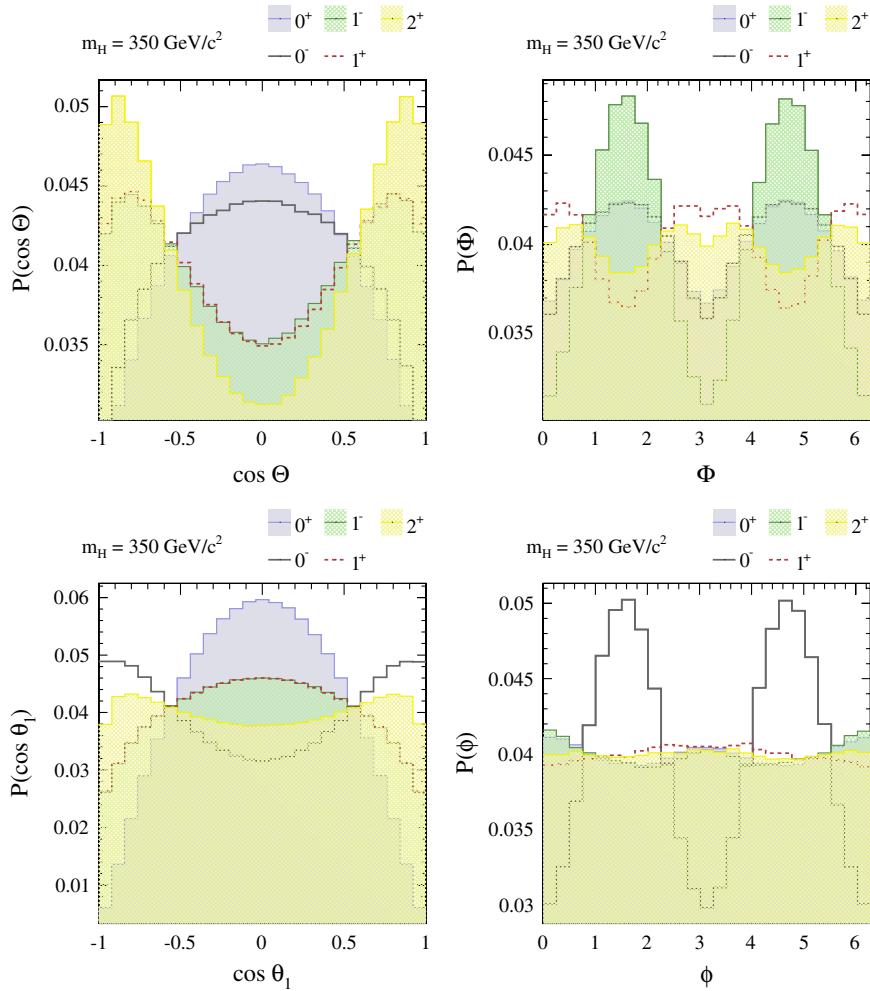


FIG. 33 (color online). Distributions of $\cos\Theta$ (top left), Φ (top right), $\cos\theta_1$ (bottom left) and ϕ (bottom right) for all the pure J^{PC} choices we study, for $m_H = 350 \text{ GeV}/c^2$. All distributions are normalized to a unit integral.

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 $P(\Lambda^{\text{max}}|0^{XP})$ 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 $P(\Lambda^{\text{max}}|0^+)$ coalesce for p -values exceeding 2σ 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 $m_H = 145$ and $200 \text{ GeV}/c^2$. For $m_H = 145 \text{ GeV}/c^2$, 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 $m_H = 145 \text{ GeV}/c^2$ and closer to unity for the higher m_H values. The corresponding results at $m_H = 350 \text{ GeV}/c^2$ 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 favor of both $\xi_{XP} \neq 0$ and $\xi_{XQ} \neq 0$, the next question would be whether it is possible to distinguish between these two cases. To

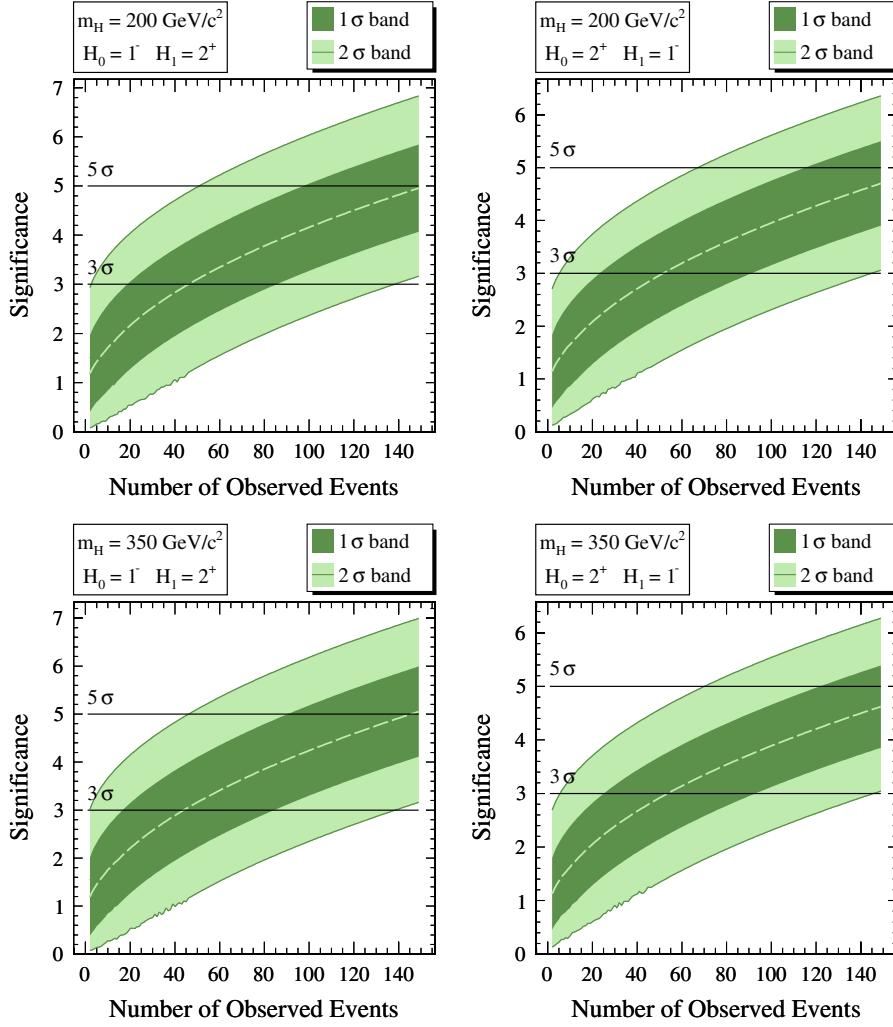


FIG. 34 (color online). Significance for rejecting 1^- in favor of 2^+ , assuming 2^+ is true (left) or vice-versa ($1^- \leftrightarrow 2^+$, right) for $m_H = 200$ and $350 \text{ GeV}/c^2$ (top, bottom).

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 favor 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 *profiled* “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 favor of $\xi_{XQ} \neq 0$. The results are shown in Fig. 42. The color-coded z -“axis” is the median of the significance for ruling out the hypothesis H_0 with the value of ξ_{H_0} given on the y -axis in favor of the H_1 hypothesis with $\xi_{H_1} \neq 0$, assumed to be correct for ξ_{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 5σ -level conclusions over most of the (ξ_{XP}, ξ_{XQ}) plane. For example, we show in Fig. 43 the 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 $m_H = 200 \text{ GeV}/c^2$. The ambiguity between ξ_{XP}^{meas} , $-\xi_{XP}^{\text{meas}}$, $\xi_{XQ} = \xi_{XP}^{\text{meas}}$ and $\xi_{XQ} = -\xi_{XP}^{\text{meas}}$ would be very hard to lift.

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 differ-

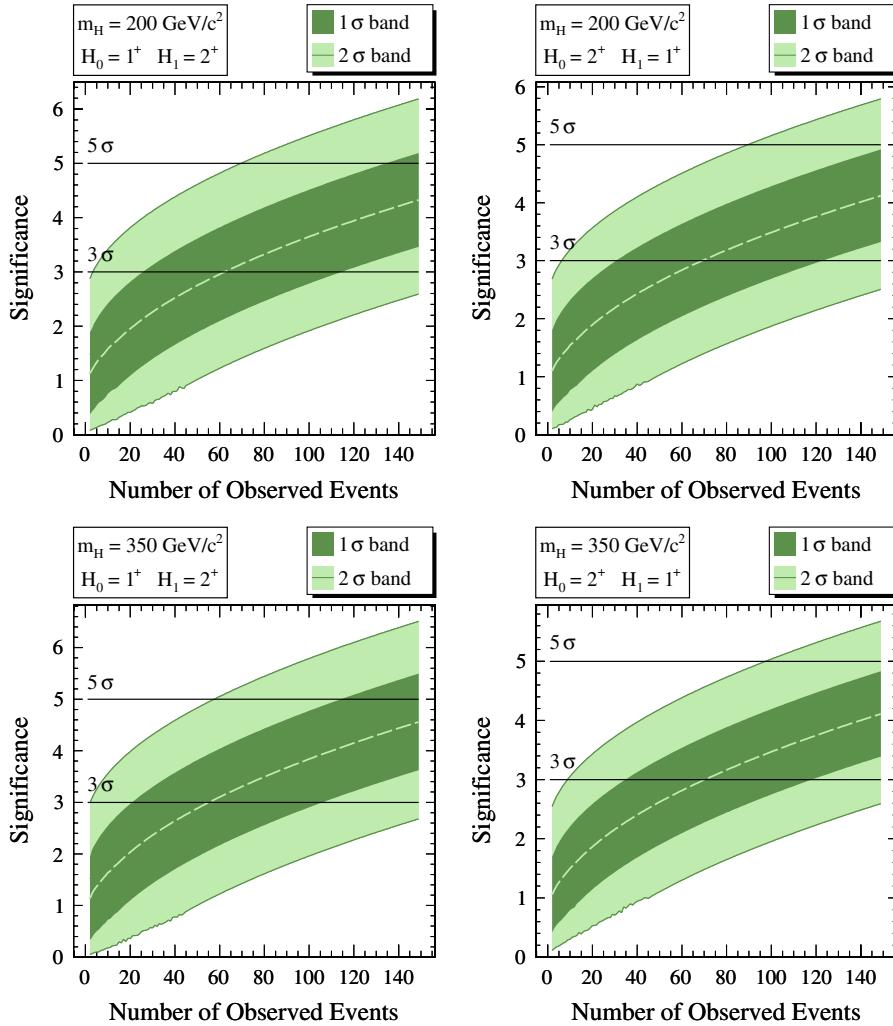


FIG. 35 (color online). Significance for rejecting 1^+ in favor of 2^+ , assumed to be correct (left) or vice-versa ($1^+ \leftrightarrow 2^+$, right) for $m_H = 200$ and $350 \text{ GeV}/c^2$ (top, bottom).

ent 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.

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. (9); if e.g. we instead chose the compositeness scale at $m_H = 350 \text{ GeV}/c^2$, this enhancement would be much smaller. Nevertheless the possible enhancement from a nonzero Y coupling, and the possible suppression from XY interference, signifies an interesting

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 an 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 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 nontrivial behavior of the significance values at and around the critical ξ_{XY} . Interestingly, the qualitative nature of these cancellations also changes with mass. For $m_H = 145 \text{ GeV}/c^2$ and $m_H = 200 \text{ GeV}/c^2$, the composite scalar with ξ_{XY} near the critical point is 0^+ -like, relative to nearby values of ξ_{XY} . For $m_H = 350 \text{ GeV}/c^2$, 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 cancel-

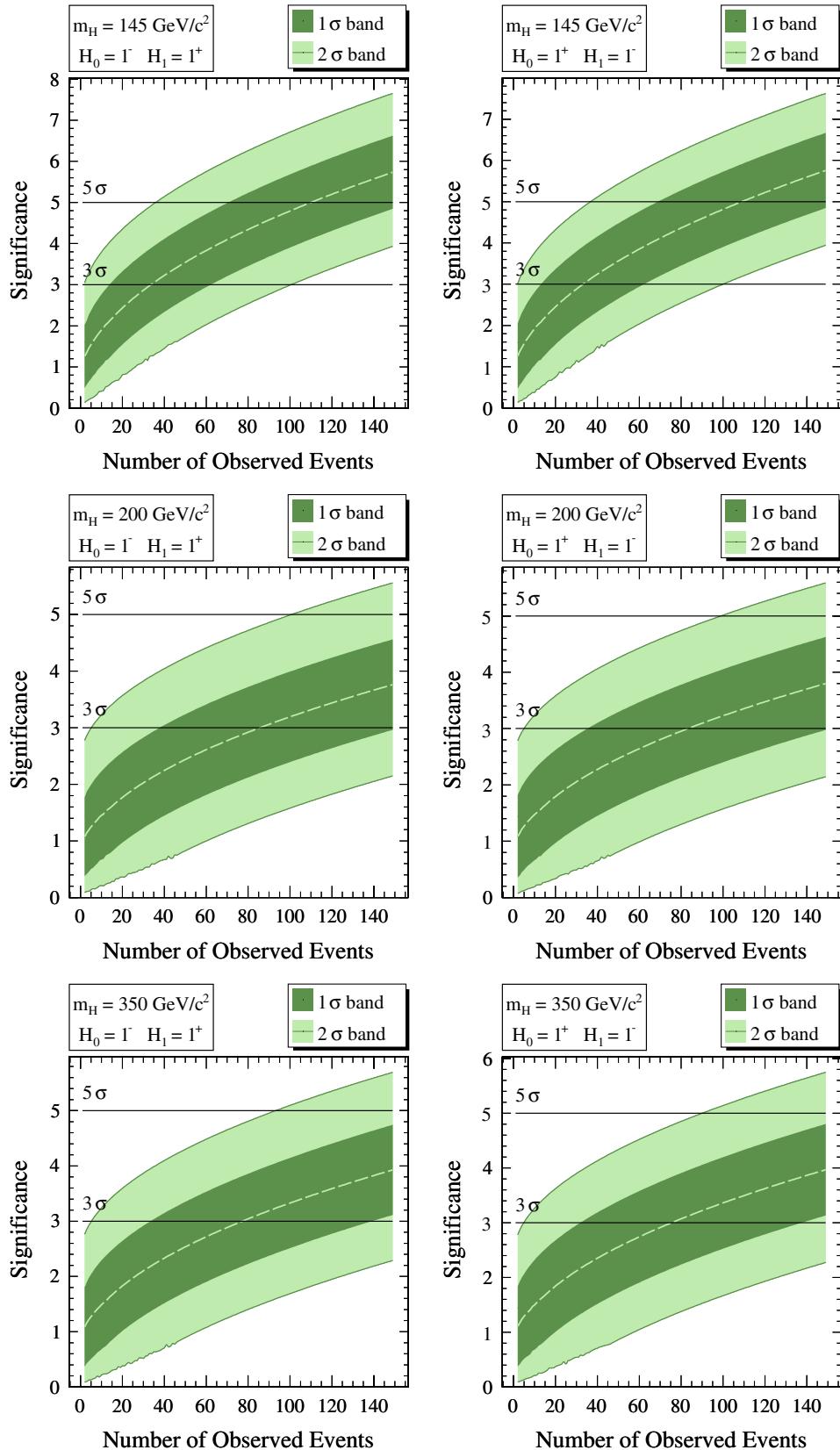


FIG. 36 (color online). Significance for rejecting 1^- in favor of 1^+ , assuming 1^+ is true (left) or vice-versa ($1^- \leftrightarrow 1^+$, right) for $m_H = 145, 200$ and $350 \text{ GeV}/c^2$ (top, middle and bottom).

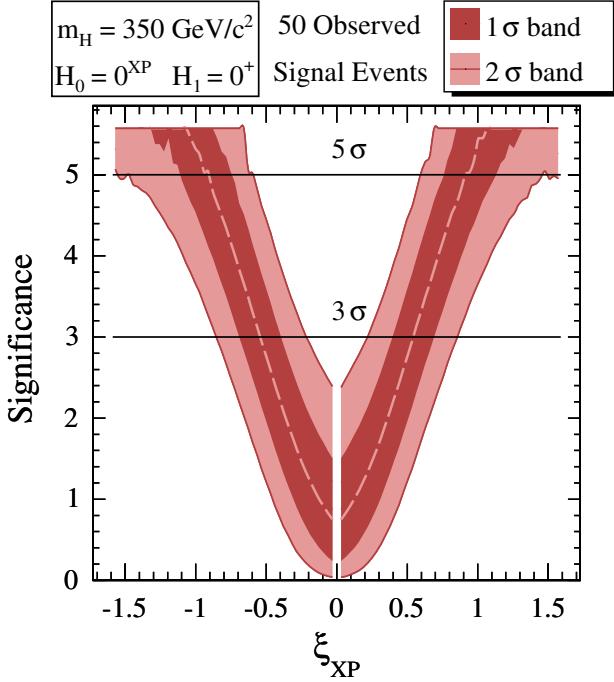


FIG. 37 (color online). Significance for excluding values of ξ_{XP} in the CP -violating $J = 0$ hypothesis in favor of the 0^+ one, assumed to be correct, for $m_H = 350 \text{ GeV}/c^2$ and $N_S = 50$. The dashed line corresponds to the median of the significance. The 1 and 2σ bands correspond to 68%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

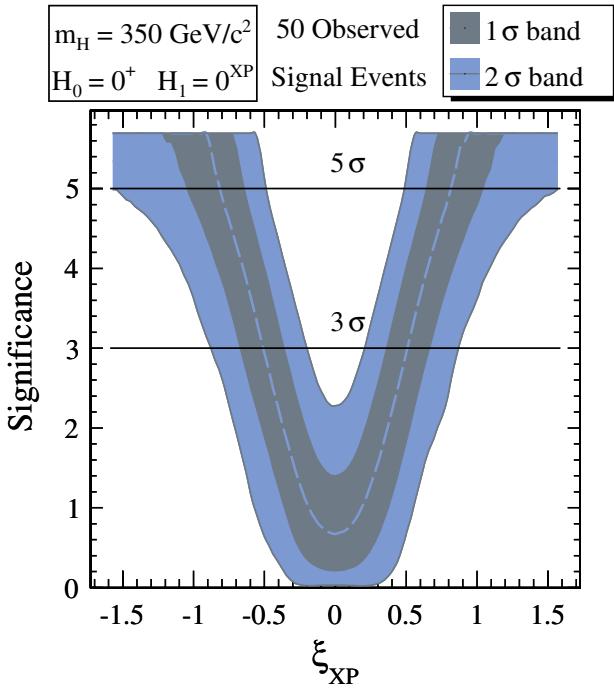


FIG. 38 (color online). The significance for excluding a pure 0^+ in favor of a CP -violating HZZ coupling ($\xi_{XP} \neq 0$), assuming the latter to be correct, with ξ_{XP} given by its x -axis values. Example for $N_S = 50$, $m_H = 350 \text{ GeV}/c^2$. Dashed line and bands as in Fig. 37.

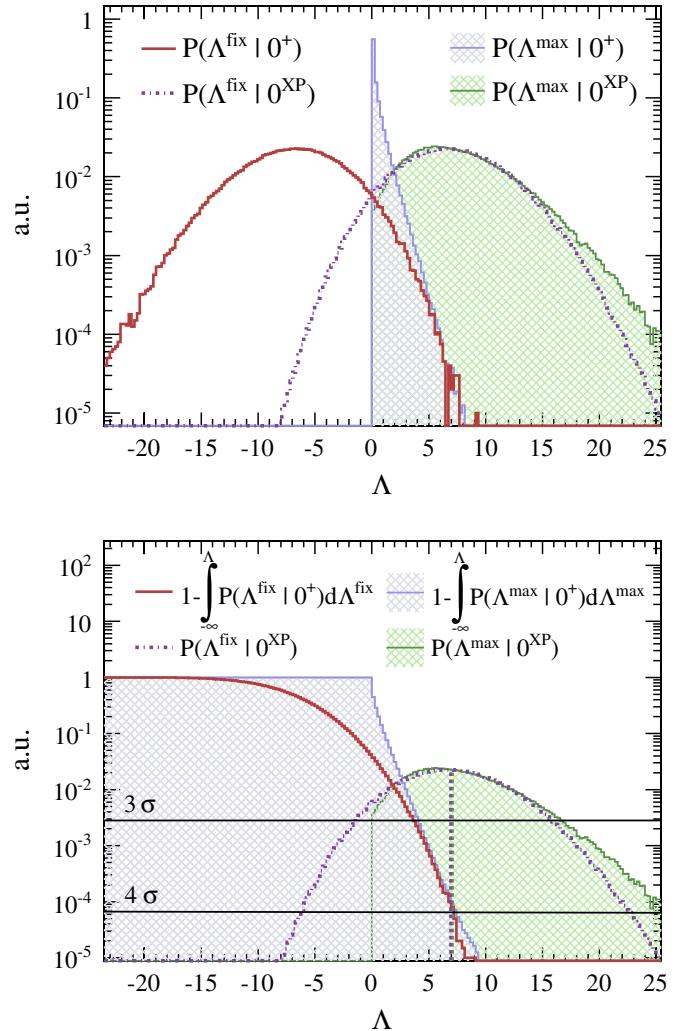


FIG. 39 (color online). Distributions of the two statistics Λ , defined in the text, for $m_H = 350 \text{ GeV}/c^2$ and $N_S = 50$. The hypotheses are $\mathbb{H}_0 = 0^+$, and $\mathbb{H}_1 = 0^{XP}$ with the CP -phase ξ_{XP} fixed at $\pi/5$. (Top) Probability distributions $P(\Lambda|\mathbb{H})$. (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.

lations the angular distributions of the pure 0^+ and the mixed case no longer resemble each other.

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 distinction 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 $m_H = 350 \text{ GeV}/c^2$, 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.

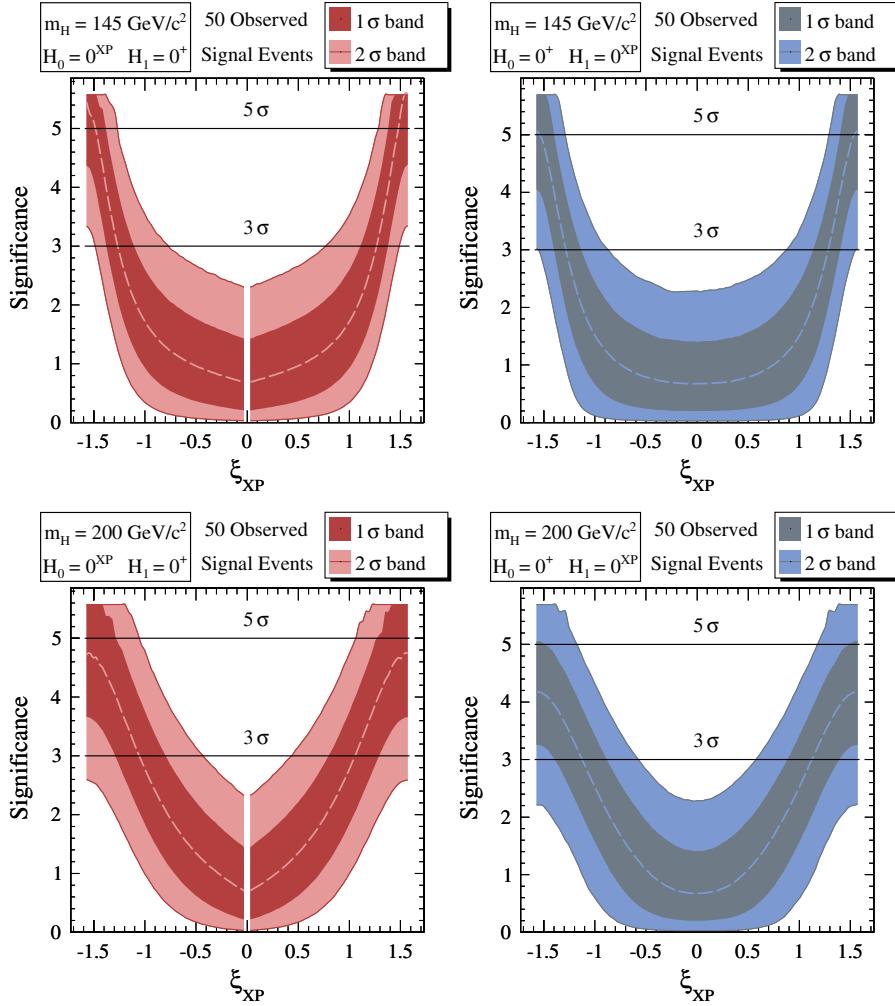


FIG. 40 (color online). Left: Significance for the exclusion of values of a CP -violating $\xi_{XP} \neq 0$ in favor of 0^+ ($\xi_{XP} = 0$), assumed to be correct. Right: Significance for excluding a pure 0^+ in favor of $\xi_{XP} \neq 0$, assumed correct with ξ_{XP} given by its x -axis values. Results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

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

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 to a Z pair can be parametrized, for non-vanishing X , P , and Q , as

$$\mathcal{L}^{\rho\mu\alpha} \propto \cos\xi(g^{\rho\mu}p_1^\alpha + g^{\rho\alpha}p_2^\mu) + e^{i\delta}\sin\xi\epsilon^{\rho\mu\alpha}(p_1 - p_2), \quad (35)$$

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.

Assuming a 0^+ resonance to be the correct choice, we determine the significance with which we can exclude values of ξ and δ for a $J = 1$ hypothesis. 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 $\Lambda = \log[\mathcal{L}(0^+)/\mathcal{L}(\xi, \delta)]$. The results, as a function of ξ for $\delta = \pi/2$ and $m_H = 350 \text{ GeV}/c^2$, are shown in Fig. 47. The points $\xi = 0$ and $|\xi| = \pi/2$ correspond to the pure vector and pure axial-vector limits,

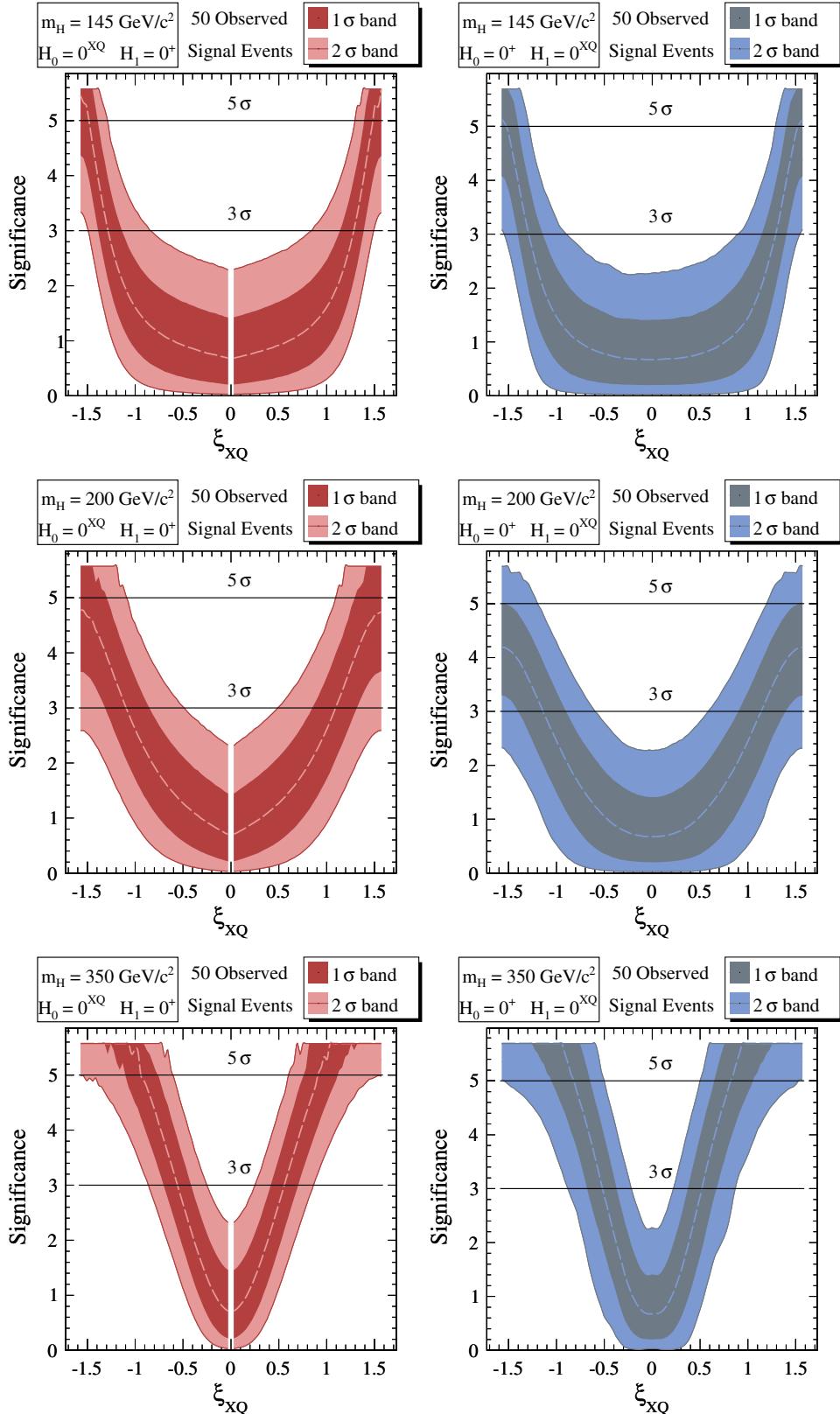


FIG. 41 (color online). Left: Significance for excluding values of a C -violating $\xi_{XQ} \neq 0$ in favor of 0^+ ($\xi_{XQ} = 0$), assumed to be correct. Right: Significance for excluding a pure 0^+ in favor of $\xi_{XQ} \neq 0$, assumed correct for the ξ_{XQ} -values on the x -axis. Hypothesis tests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

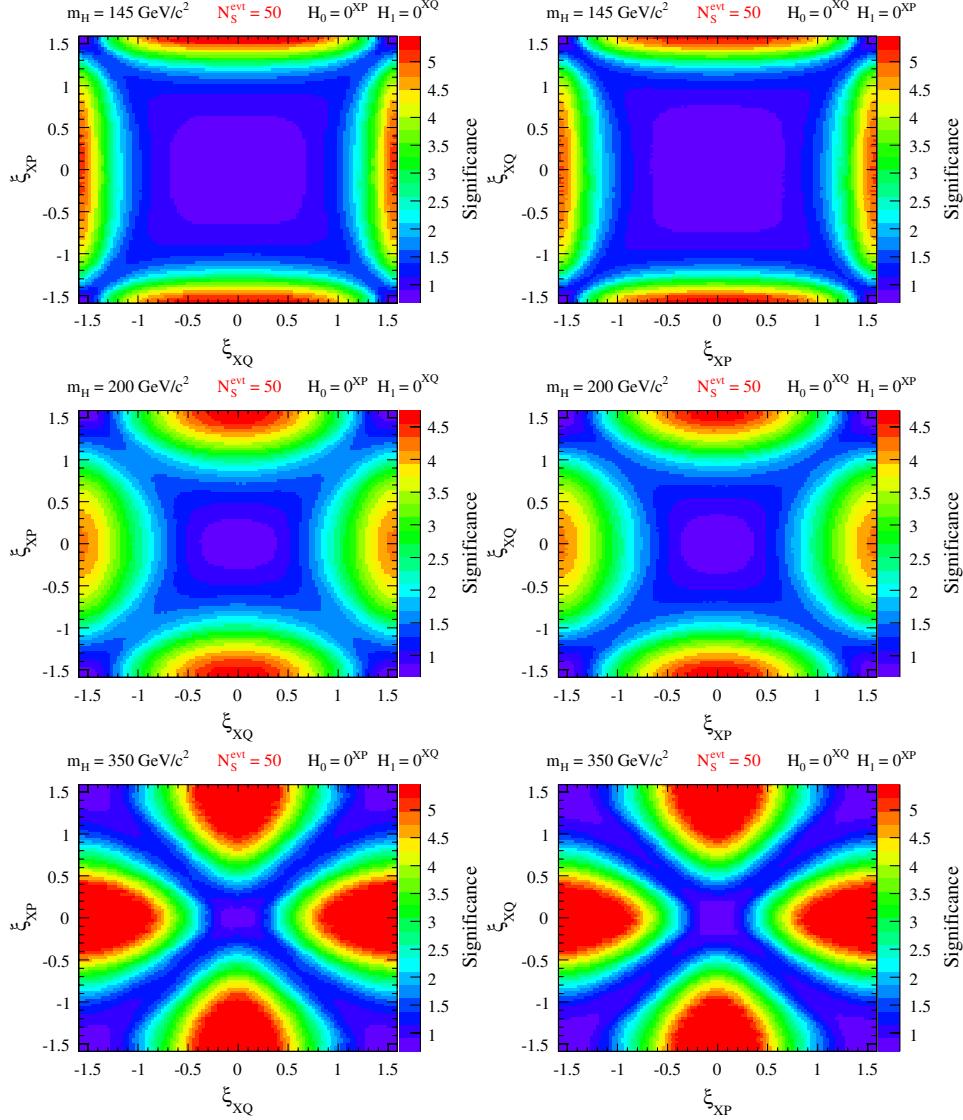


FIG. 42 (color online). The median of the significance (colored z -axis) for excluding values of $\xi_{\mathbb{H}_0}$ (y-axis) in favor of the $\xi_{\mathbb{H}_1} \neq 0$ hypothesis assuming as correct the values $\xi_{\mathbb{H}_1}$ of the x -axis. The tests are performed for $\mathbb{H}_1 = 0^{XP}$, $\mathbb{H}_0 = 0^{XQ}$ (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 $N_S = 50$.

respectively, and are consistent with Figs. 22 and 23 on these pure cases.

Assuming a $J = 1$ resonance with given ξ and δ to be the correct choice, we determine the significance with which we can exclude the 0^+ case in favor of $J = 1$. We have to treat ξ and δ as nuisance parameters, since we are considering the general $J = 1$ case. The statistic is $\log[\max \mathcal{L}(\hat{\xi}, \hat{\delta}) / \mathcal{L}(0^+)]$. The results, as functions of ξ for $\delta = \pi/4$ and $m_H = 200 \text{ GeV}/c^2$, 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.

In Fig. 49 we show the 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 $350 \text{ GeV}/c^2$. Notice that the significance levels color-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 favor 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

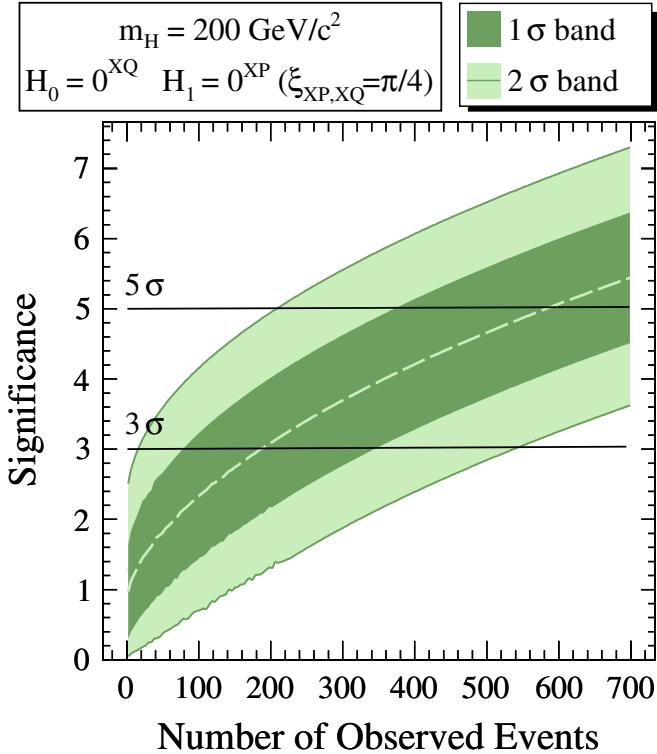


FIG. 43 (color online). The significance for excluding the C -violating $J = 0$ hypothesis in favor of a CP -violating case, assuming the latter to be correct, with $\xi_{XP,XQ} = \pi/4$. Example for $m_H = 200 \text{ GeV}/c^2$.

differentiating between 0^+ and $J = 1$ should exceed 5σ with $N_S \sim (60, 200, 85)$ events for $m_H = (145, 200, 350) \text{ GeV}/c^2$, respectively.

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

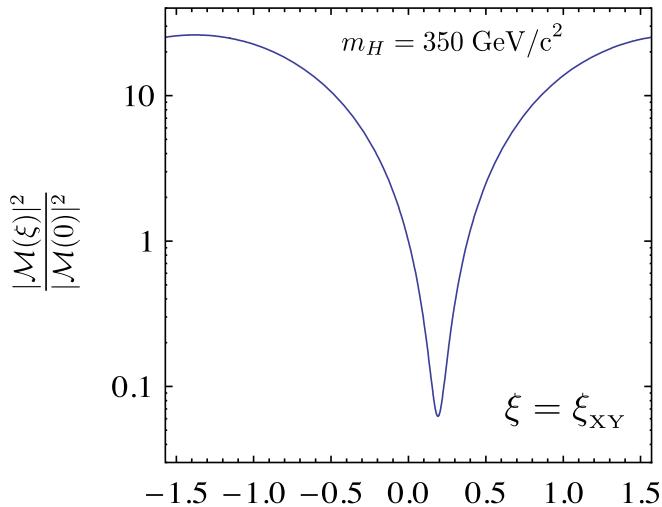


FIG. 44 (color online). The fully angularly-integrated matrix element squared for a composite 0^+ , showing a strong destructive interference at a given ξ_{XY} . The result, shown here for $m_H = 350 \text{ GeV}/c^2$, is normalized to $\xi_{XY} = 0$.

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 favor 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 disfavored 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 needs to be different.

More specifically, each mixed hypothesis family is characterized by mixing angles $\vec{\xi}$. 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 $\mathcal{L}(\vec{\xi}) = \prod_{e=1}^N P_e(\vec{x}_e, \vec{\xi})$. The measured values of the mixing angles, $\vec{\xi}_{\text{meas}}$, 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 $\vec{\xi} = \vec{\xi}_{\text{input}}$. For each experiment, the measured values of $\vec{\xi}$ 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 $P(\vec{\xi}_{\text{meas}} | \vec{\xi}_{\text{input}})$, estimated using this ensemble of experiments, the Feldman-Cousins unified approach [47] 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

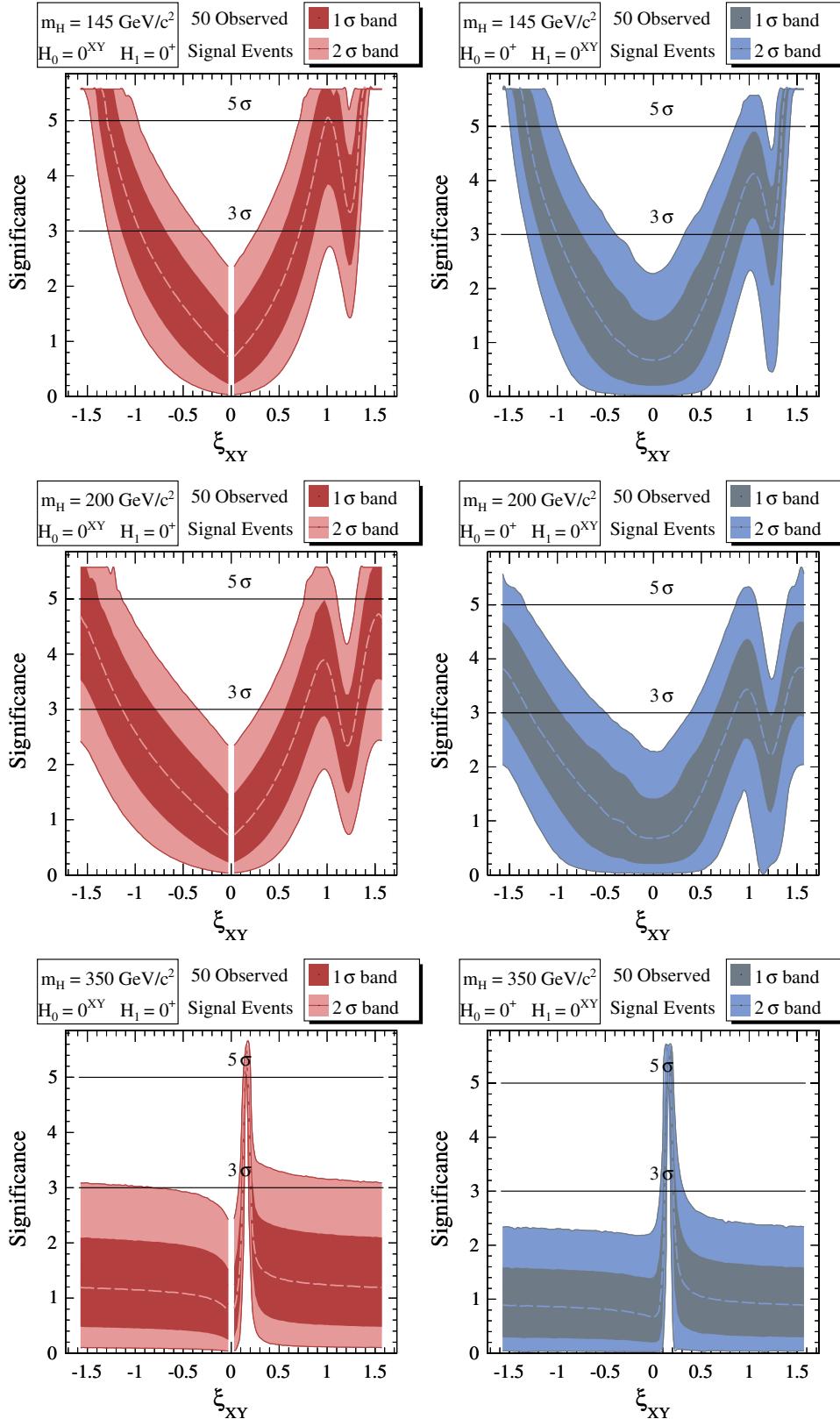


FIG. 45 (color online). Left: significance for excluding values of ξ_{XY} in favor of a pointlike 0^+ ($\xi_{XY} = 0$), assumed to be correct. Right: significance for excluding a pointlike 0^+ in favor of a composite one ($\xi_{XY} \neq 0$), assumed correct for the ξ_{XY} values on the x -axis, for $m_H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

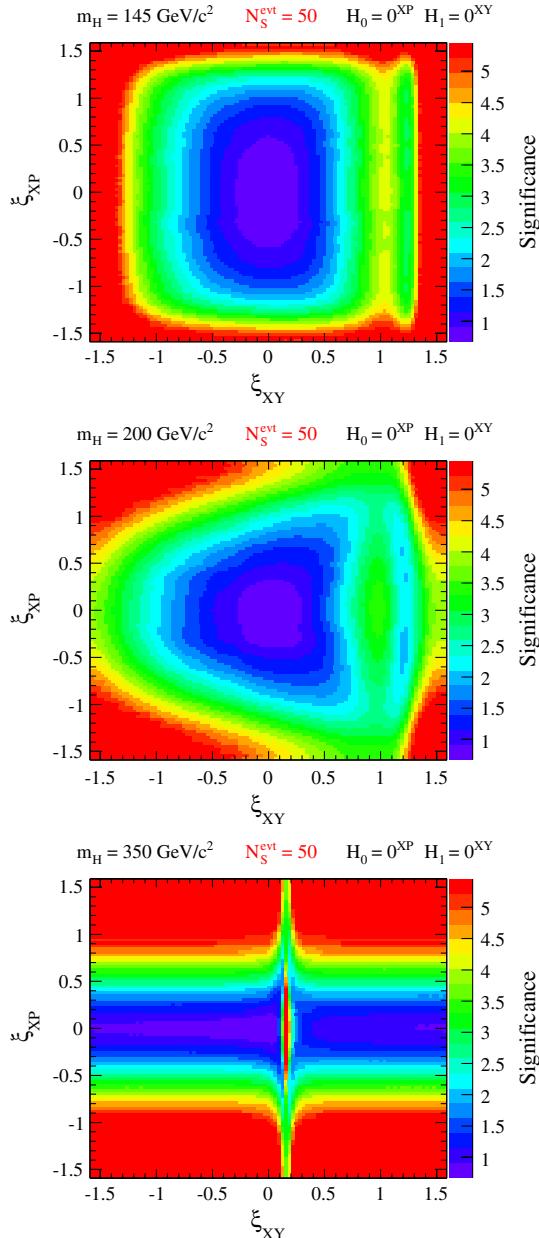


FIG. 46 (color online). The median of the significance (colored-labeled z -axis) for excluding values of ξ_{XP} (y-axis) in favor of the composite scalar assuming it to be correct with the ξ_{XY} values of the x -axis, for $m_H = 145, 200$ and $350 \text{ GeV}/c^2$ (top, middle and bottom) and $N_S =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

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 $m_H = 200 \text{ GeV}/c^2$ (middle of Fig. 50) we see that, if $\xi_{XP}^{\text{meas}} = 0$, the 3σ confidence interval is approximately $\xi_{XP} \in [-1, 1]$.

The 1σ bands in Fig. 50 are centered on the diagonal $\xi_{XP}^{\text{meas}} = \xi_{XP}^{\text{input}}$, implying that there is no significant bias in the measurement. In addition to this, the 2σ and 3σ bands also cover most of the diagonal $\xi_{XP}^{\text{meas}} = -\xi_{XP}^{\text{input}}$. This confirms our observation from Sec. VII F that our ability

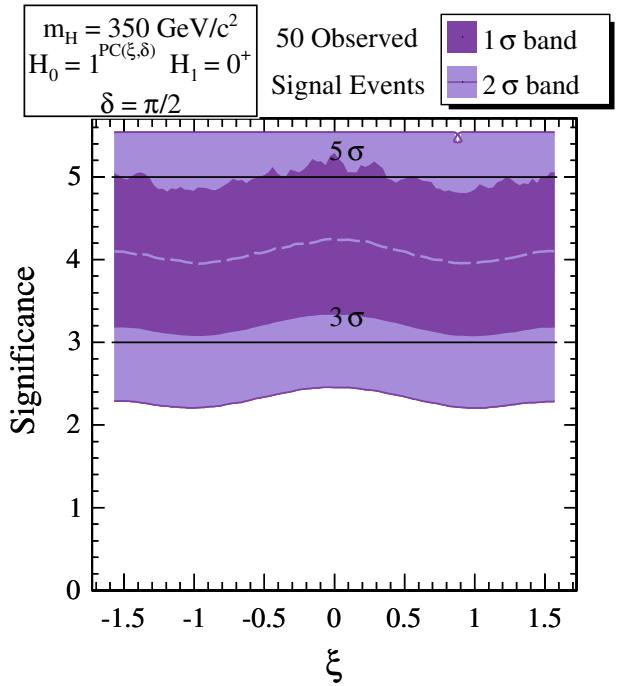


FIG. 47 (color online). Significance for excluding values of ξ , for $\delta = \pi/2$, in the general $J = 1$ hypothesis [dubbed $1^{PC}(\xi\delta)$] in favor of the 0^+ one, assumed to be correct. Results for $m_H = 350 \text{ GeV}/c^2$ and $N_S = 50$. The dashed line is the median of the significance.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

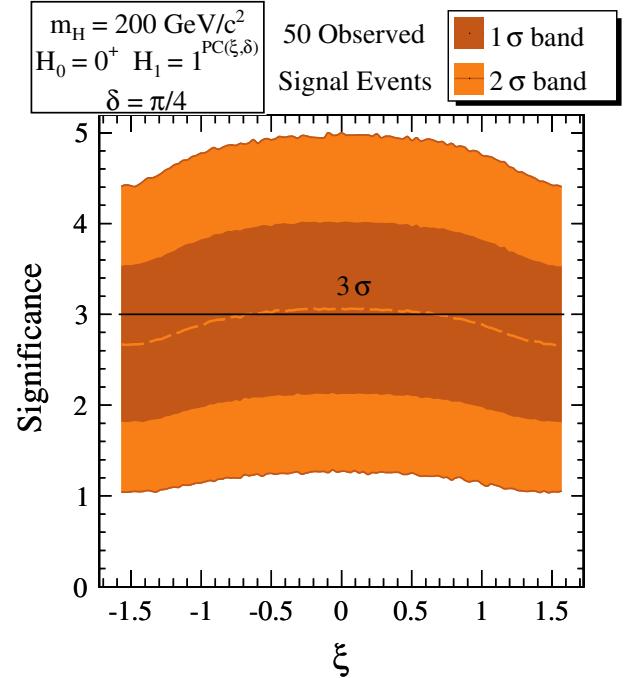


FIG. 48 (color online). Significance for excluding the 0^+ hypothesis in favor of the general $J = 1$ case [dubbed $1^{PC}(\xi\delta)$], assumed correct for ξ as in the x -axis and $\delta = \pi/4$. Results for $m_H = 200 \text{ GeV}/c^2$ and $N_S = 50$. The dashed line and bands are as in Fig. 47.

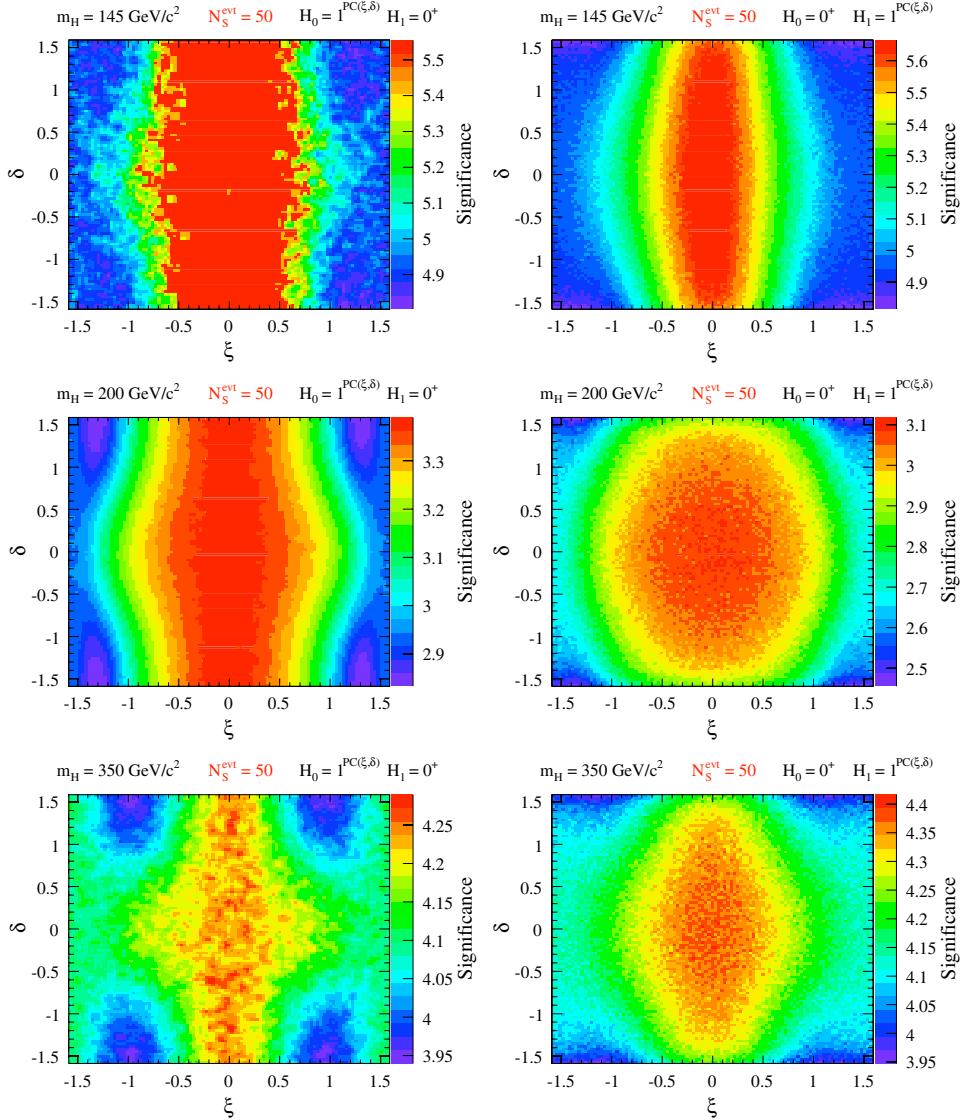


FIG. 49 (color online). Left: Median of the significance (colored z -axis) for excluding values of ξ and δ corresponding to a $J = 1$ hypothesis [dubbed $1^{PC(\xi,\delta)}$] in favor of 0^+ , if the latter is correct. Right: vice-versa, with values of ξ and δ indicated on the axes. Results for $m_H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

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 $m_H = 145 \text{ GeV}/c^2$ (but not for $m_H = 200 \text{ GeV}/c^2$) 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 $m_H = 145 \text{ GeV}/c^2$ 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 behaviors of similar magnitude.

Confidence intervals for measurements of the parameter ξ_{XQ} for a scalar with C -violating *HLL* couplings are

shown in Fig. 51; These are nearly identical to those in Fig. 50, reflecting the difficulty of discriminating the $\xi_{XP} \neq 0$ and $\xi_{XQ} \neq 0$ hypotheses, as discussed in Sec. VII E. For the C -odd case, the sign ambiguity of ξ_{XQ}^{meas} is slightly worse than for the \tilde{T} -odd one as demonstrated by the 1σ confidence bands appearing on the $\xi_{XQ}^{\text{meas}} = -\xi_{XQ}^{\text{input}}$ diagonal for $m_H = 350 \text{ GeV}/c^2$. This is also expected, since the C -odd interference term is proportional to the relatively small number $\eta \approx 0.15$, see Eq. (22). 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.

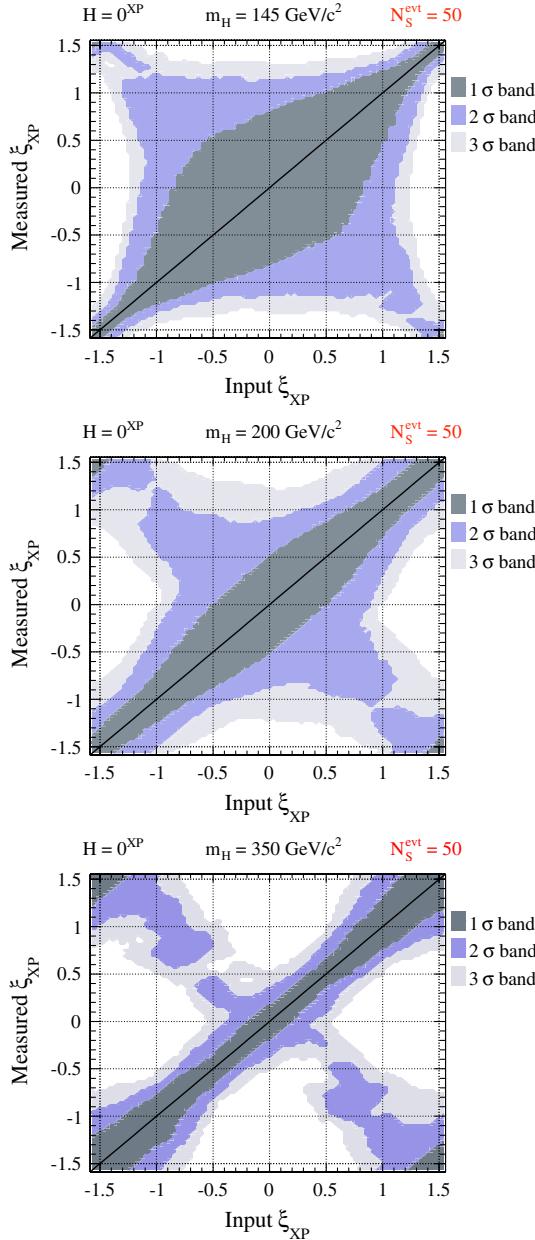


FIG. 50 (color online). Confidence intervals for measured values of ξ_{XP} for a CP -violating $J = 0$ resonance, for $m_H = 145$, 200 and 350 GeV/c^2 (top, middle and bottom), all for $N_S = 50$. For measured values of ξ_{XP} on the y -axis, confidence intervals should be read horizontally, see text.

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 $200 \text{ GeV}/c^2$, the 1σ intervals are centered on the diagonal $\xi_{XY}^{\text{meas}} = \xi_{XY}^{\text{input}}$. There are no bands along $\xi_{XY}^{\text{meas}} = -\xi_{XY}^{\text{input}}$, since the interference term is of a different nature than that of the discrete-symmetry violating cases. The extensions of the 2 and 3σ 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.

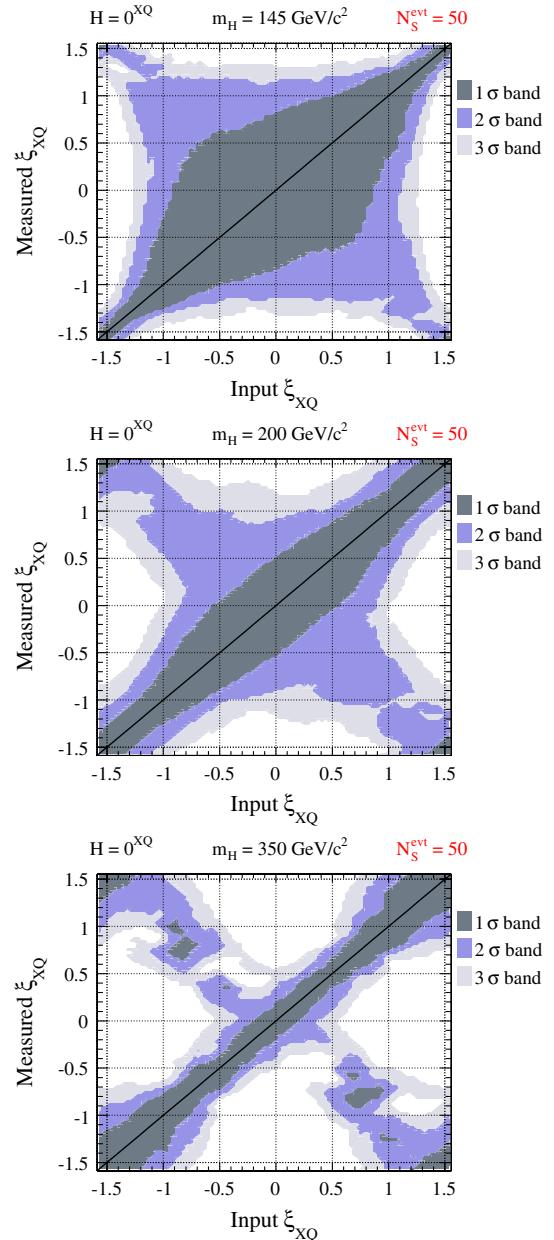


FIG. 51 (color online). Confidence intervals for measured values of ξ_{XQ} for a C -violating $J = 0$ resonance for $m_H = 145$, 200 and 350 GeV/c^2 (top, middle and bottom), all for $N_S = 50$. For measured values of ξ_{XQ} on the y -axis, confidence intervals should be read horizontally, see text.

The figure for $m_H = 350 \text{ GeV}/c^2$ 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 composite from a pointlike scalar *unless* it 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, parametrized by angles ξ and δ as described in

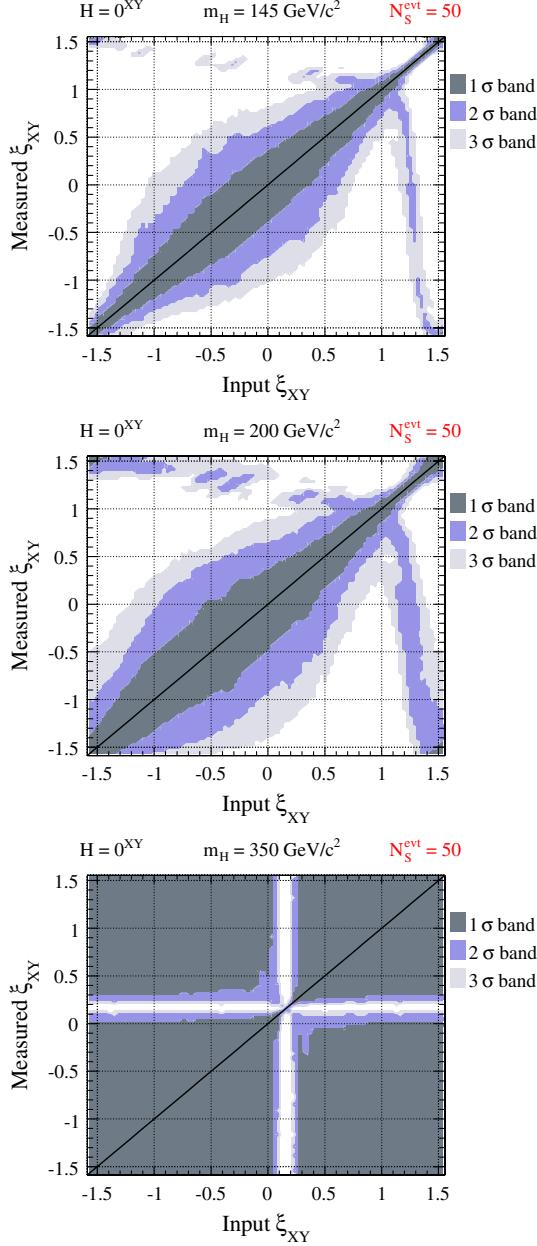


FIG. 52 (color online). Confidence intervals for measured values of ξ_{XY} for a composite $J = 0$ resonance, for $m_H = 145, 200$ and $350 \text{ GeV}/c^2$ (top, middle and bottom), all for $N_S = 50$. For measured values of ξ_{XY} on the y -axis, confidence intervals should be read horizontally.

Sec. VII F. We saw in Sec. VII D that the 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 $m_H = 145 \text{ GeV}/c^2$. 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

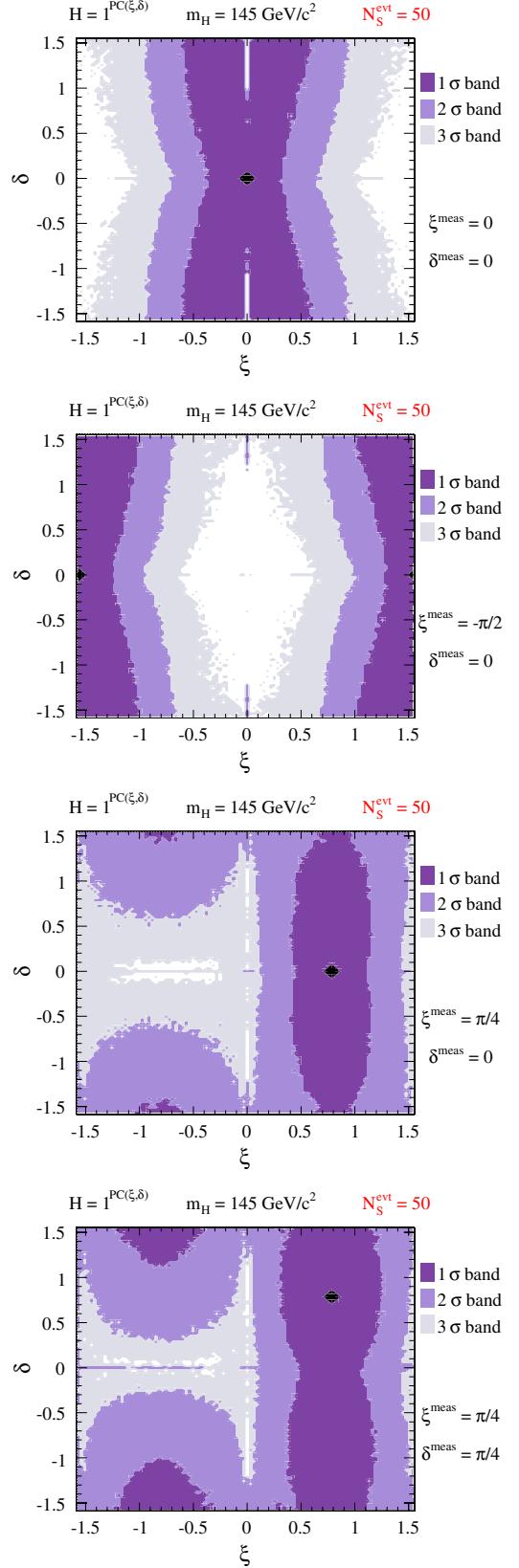


FIG. 53 (color online). Confidence intervals for measured values of ξ and δ for a $J = 1$ resonance with a mass $145 \text{ GeV}/c^2$ and $N_S = 50$ events. The input values, indicated by diamonds, are reported alongside the figures.

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.

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 suboptimal analysis strategies, as illustrated in Fig. 54 for the benchmark $m_H = 200 \text{ GeV}/c^2$.

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 $P(a_1, \dots, a_N)$ denotes N -dimensional *pdfs* in the correlated variables $\{a_1, \dots, a_N\}$. Here $\prod_i P(X_i)$ is constructed

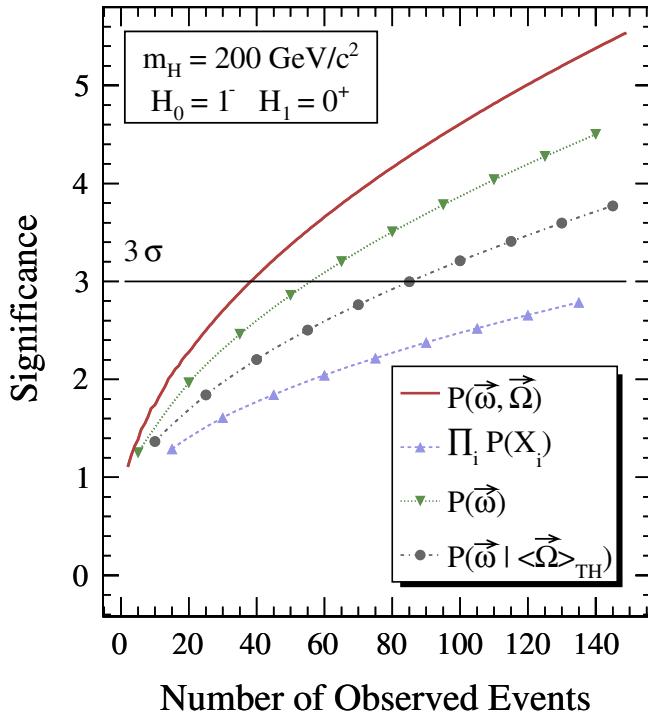


FIG. 54 (color online). Median significance for rejecting 1^- in favor of 0^+ , for different likelihood constructions used in the log-likelihood ratio test statistic. H_0 is always considered the true hypothesis.

from one-dimensional *pdfs* for all variables, ignoring (erroneously) their correlations. $P(\vec{\omega} | \langle \vec{\Omega} \rangle_{TH})$ 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 $P(\vec{\omega} | \langle \vec{\Omega} \rangle_{TH})$ 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 4π 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*, such that phase space acceptance sculpting of the Φ distribution 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 Fig. 54, not only degrades the discrimination significance but also produces a real chance of falsely labeling the quantum numbers of the new resonance. Assume, for example, the SM, with $m_H = 200 \text{ GeV}/c^2$. Let the data be fit to either a fully correlated or an uncorrelated one. The

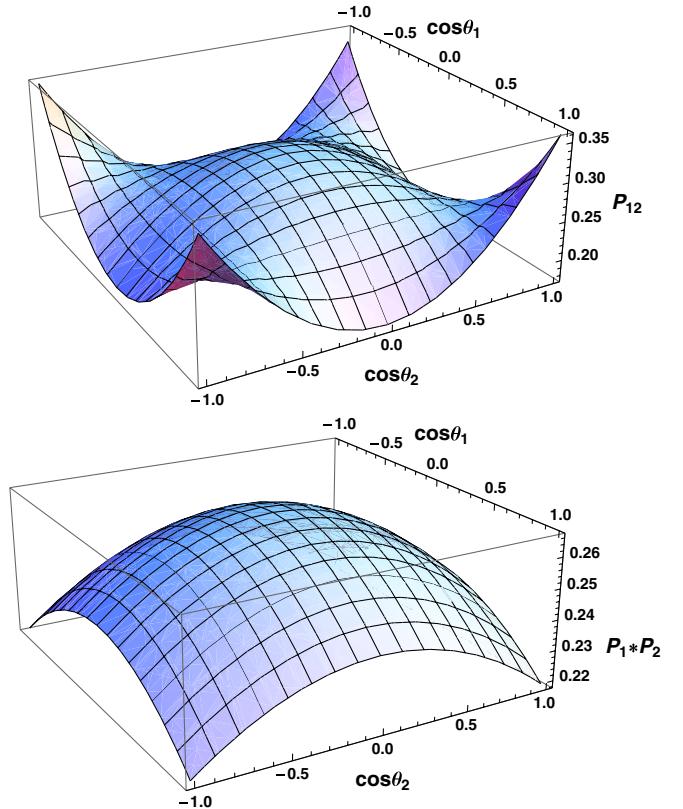


FIG. 55 (color online). The normalized theoretical *pdfs* in the variables $\cos\theta_1$ and $\cos\theta_2$ (integrated over ϕ) for $J = 0^+$ and $m_H = 200 \text{ GeV}/c^2$. Top: the 2D *pdf* $P_{12} = P[\cos\theta_1, \cos\theta_2]$. Bottom: the product $P_1 * P_2 = P[\cos\theta_1] \times P[\cos\theta_2]$ of the 1D *pdfs*.

projections of the corresponding theoretical *pdfs*, involving only the variables $\cos\theta_1$ and $\cos\theta_2$, are illustrated in Fig. 55. On the top (bottom) of the figure we show $P[\cos\theta_1, \cos\theta_2]$ ($P[\cos\theta_1] \times P[\cos\theta_2]$). With limited statistics—insufficient to distinguish between the correlated and uncorrelated distributions—the correct conclusion will be reached: the data are compatible with the SM. But, as the statistics are increased, the data will significantly deviate from the $P[\cos\theta_1] \times P[\cos\theta_2]$ distribution, and a false rejection of the SM hypothesis would become increasingly supported.

Overall, we have demonstrated that small signal samples in the $ZZ \rightarrow 4\ell$ or $ZZ^* \rightarrow 4\ell$ 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 and II for $m_H = 145 \text{ GeV}/c^2$ for all pure-case comparisons between $J = 0, 1$ parent particles, and in Tables III, IV, V, and

TABLE I. Minimum number of observed events such that the median significance for rejecting \mathbb{H}_0 in favor of the hypothesis \mathbb{H}_1 (assuming \mathbb{H}_1 is right) exceeds 3σ with $m_H = 145 \text{ GeV}/c^2$.

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

TABLE II. Same as Table I, but requiring that the median significance exceeds 5σ .

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

TABLE III. Minimum number of observed events such that the median significance for rejecting \mathbb{H}_0 in favor of the hypothesis \mathbb{H}_1 (assuming \mathbb{H}_1 is right) exceeds 3σ with $m_H = 200 \text{ GeV}/c^2$.

$\mathbb{H}_0 \Downarrow \mathbb{H}_1 \Rightarrow$	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 IV. Same as Table III, but requiring that the median significance exceeds 5σ .

$\mathbb{H}_0 \Downarrow \mathbb{H}_1 \Rightarrow$	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 V. Minimum number of observed events such that the median significance for rejecting \mathbb{H}_0 in favor of the hypothesis \mathbb{H}_1 (assuming \mathbb{H}_1 is right) exceeds 3σ with $m_H = 350 \text{ GeV}/c^2$.

$\mathbb{H}_0 \Downarrow \mathbb{H}_1 \Rightarrow$	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 VI. Same as Table V, but requiring that the median significance exceeds 5σ .

$\mathbb{H}_0 \Downarrow \mathbb{H}_1 \Rightarrow$	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	...

VI for $m_H = 200(350) \text{ GeV}/c^2$, for all pure-case comparisons between $J = 0, 1, 2$ parent particles.

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

The tables also show that the discriminating 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 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 entan-

lement between the decay variables. Some examples are: distinguishing the “natural-parity” $J = 0^+$ and 1^- hypotheses for $m_H = 145 \text{ GeV}/c^2$ requires only a dozen signal events for 3σ discrimination. For $200 \text{ GeV}/c^2$, discriminating 0^+ from 0^- is relatively easy, but distinguishing 0^+ from 2^+ is difficult. For $350 \text{ GeV}/c^2$, 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 $m_H = 350 \text{ GeV}/c^2$ 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 $\text{HLL} \rightarrow \text{ZZ}$ 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 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. Analysis limitations

In our analysis we focused on decay information, exploiting an approximate factorization between observables related to Higgs (or HLL) production and observables

related to decay. The factorization is only approximate because of phase space acceptance effects and, in the case of spin >0 HLLs, correlations between the initial and final-state particles. In a real data analysis one would want to include production information, which in turn would require a detailed knowledge of radiative corrections, PDFs, and full detector simulation for the HLLs. Such an analysis is beyond the scope of this paper. Within our narrower scope we have incorporated as much as we could all the issues that make a conceptual difference in the strategy. When we have made approximations or have neglected certain effects, it is because the detailed inclusion of these effects would not have a qualitative impact on our results.

The QCD corrections to the signal predictions for $d^2\sigma/dp_T d\eta$ are large, as is well-studied for the SM Higgs (see, for example, [4,49,50] and references therein.) The impact on the total cross sections is not relevant to our analysis, but the corrections to the (p_T, η) distributions will modify the phase space acceptance effects on the distributions of the final-state leptons. For the SM Higgs we have included these corrections at NLO, and a recent study shows that the effects of NNLO corrections on the final-state lepton distributions are not dramatic [51]. We have not included the differences between the phase space acceptance effects for the SM Higgs and those for the various HLLs, but we performed a comparison at LO to see that these differences are small compared to the acceptance effects themselves.

There are electroweak radiative corrections that directly involve the final-state leptons. For the SM Higgs these corrections have been computed and studied in detail [36]; the corrections are of the order of 5 to 10% and cause a mild distortion of the angular distributions. These effects should be included in a complete analysis, but they do not introduce anything conceptually new to the methodology proposed in this study, and their inclusion involves details of the experimental treatment of the vertex and subsequent radiations by electrons and muons.

We only considered the dominant ZZ background, and only at LO. It would be useful to include a more comprehensive treatment of the SM backgrounds to the golden channel and to use the full signal-to-background discrimination, e.g. by adding the Z mass distribution to the likelihood definition. A complete treatment of the backgrounds would require full detector simulation.

Our treatment of couplings and HLLs was not exhaustive, since we have ignored gauge-invariant operators with dimension greater than 6, have only examined one case of spin 2 HLL, and have not even mentioned the possibility of HLLs with spins higher than 2. At some point Occam’s razor obviates the need for such comparisons: “Raffiniert ist der Herr Gott, aber bosaft ist Er nicht,” to quote a known author [52].

The likelihood analyses pursued here are very computing intensive, since 5σ discrimination implies simulating

sufficiently many pseudoexperiments to fill out what amounts to the 5σ tails in multidimensional likelihood distributions, where they are typically highly non-Gaussian. The analysis presented here used more than 10^{14} pseudoexperiments 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.

ACKNOWLEDGMENTS

We especially acknowledge insights and inspiration from our late colleagues Andrew Lange and Juan Antonio Rubio. The authors are grateful to Andrew Cohen, Belén Gavela, Keith Ellis, Shelly Glashow, Ken Lane, Ken Lee, Michelangelo Mangano, Chiara Mariotti, Guido Martinelli, Sezen Sekmen, Riccardo Rattazzi, Raman Sundrum, Steven Weinberg, Jan Winter and Mark Wise for useful discussions. J. L. 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 No. DE-AC02-07CH11359 with the U.S. Department of Energy. C. R. and M. S. are supported in part by the U.S. Department of Energy under Contact No. DE-FG02-92-ER40701.

Note added.—While this manuscript was in preparation we received the preprint [54], reporting on an analysis similar to what we have presented here.

APPENDIX A: $SU(2)_L \times U(1)_Y$ GAUGE-INVARIANT COUPLINGS

To write Lagrangians generating the couplings of Sec. 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

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

with A_i arbitrary constants and Λ a mass parameter. This object generates, amongst others, the couplings of Eq. (9). The true dimensionality of the operators in Eq. (9) is that of the ones appearing in Eq. (A1), that is, dimension five.

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 $Y + iZ$ in Eq. (9). 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 nonstandard 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 $\Phi \rightarrow H + \nu$, 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. (10) is built of operators of dimension 6:

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

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

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

$$\begin{aligned} L = & \frac{1}{\Lambda} H_{\mu\nu} (A_1 \vec{W}_\alpha^\mu \vec{W}^{\nu\alpha} + A_2 B_\alpha^\mu B^{\nu\alpha}) \\ & + \frac{1}{\Lambda} H_\rho^\nu i \epsilon_{\mu\nu\alpha\beta} (A_3 \vec{W}^{\mu\alpha} \vec{W}^{\rho\beta} + A_4 B^{\mu\alpha} B^{\rho\beta}). \end{aligned} \quad (\text{A3})$$

The consideration of gauge-invariant constructions for HLLs with nontrivial 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 $p_{\text{cms}} \equiv |\vec{p}[Z]| = m_1 \gamma_1 \beta_1 = m_2 \gamma_2 \beta_2$ be the momentum of one or the other Z in the H center-of-mass system:

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

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

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

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

APPENDIX C: GENERAL RESULTS FOR SPIN 0 COUPLED TO ZZ*

In Sec. III we have already written the angular distributions $d\Gamma[0^+]$ and $d\Gamma[0^-]$ for the pure scalar and pseudo-scalar cases, see Eqs. (14) and (15). We also discussed the T -odd and C -odd interferences between the standard coupling—proportional to X in Eq. (9)—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. (21) and (22). 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. (24) and (25). This allows us to gather the results corresponding to the most general deviations from the SM Higgs couplings:

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

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

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

where

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

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

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

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 $Z_2 \rightarrow \mu^- \mu^+$ with muon helicity $\sigma_2/2$ and $Z_1 \rightarrow e^- e^+$ 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 $\sigma/2, \sigma = \pm 1$:

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

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

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

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

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

from which we derive the shorthand notation

$$\begin{aligned} g_1 & \equiv (c_v^2 + c_a^2)^2 (g_v^2 + g_a^2) \\ g_\sigma & \equiv -4c_v c_a (c_v^2 + c_a^2) (g_v^2 + g_a^2) \\ g_\tau & \equiv -2(c_v^2 + c_a^2)^2 g_v g_a \\ g_{\sigma\sigma} & \equiv 4c_v^2 c_a^2 (g_v^2 + g_a^2) \\ g_{\sigma\tau} & \equiv 8c_v c_a (c_v^2 + c_a^2) g_v g_a \\ g_{\sigma\sigma\tau} & \equiv -8c_v^2 c_a^2 g_v g_a. \end{aligned} \quad (\text{D4})$$

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

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

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 antiquark direction in a $q\bar{q}$ -initiated process. A solution is to symmet-

trize 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 same effect as setting g_τ , $g_{\sigma\tau}$, and $g_{\sigma\sigma\tau}$ to zero.

The standard convention in the literature for the three azimuthal angles is somewhat peculiar. The coordinate axes are 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 $\varphi_1 \rightarrow \Phi + \phi$, $\varphi_2 \rightarrow \Phi$ regain the previous convention. We will employ this notation in this appendix, which makes the formulas more symmetrical.

After the quark-antiquark symmetrization described above, the XX part of the full matrix element squared is given by

$$\begin{aligned} & 4m_1^2 m_2^2 X^2 \gamma_b^2 [g_1 S^2 s_1^2 s_2^2 (2\ell_0^2 m_d^4 - \ell^2 m_H^2 [m_1^2 \cos(2\varphi_1) + m_2^2 \cos(2\varphi_2)])] \\ & + g_1 \ell^2 m_H^2 (1 + C^2) [2m_2^2 s_1^2 + 2m_1^2 s_2^2 - (m_1^2 + m_2^2) s_1^2 s_2^2] + 4\ell \ell_0 g_1 m_H m_d^2 CS [m_1 c_1 s_1 s_2^2 \sin\varphi_1 - m_2 c_2 s_2 s_1^2 \sin\varphi_2] \\ & - 2\ell^2 m_H^2 m_1 m_2 s_1 s_2 [(1 + C^2)(g_1 c_1 c_2 - g_{\sigma\sigma}) \cos(\varphi_1 - \varphi_2) + S^2(g_1 c_1 c_2 + g_{\sigma\sigma}) \cos(\varphi_1 + \varphi_2)]]. \end{aligned} \quad (D6)$$

The PP part is given by

$$\begin{aligned} & P^2 [\ell^2 g_1 m_H^2 S^2 s_1^2 s_2^2 [M_2^4 m_1^2 \cos(2\varphi_1) + M_1^4 m_2^2 \cos(2\varphi_2)] + 8\ell_0^2 m_1^2 m_2^2 m_d^4 S^2 [g_1 (c_1^2 + c_2^2 + s_1^2 s_2^2 \sin(\varphi_1 - \varphi_2)^2) + 2g_{\sigma\sigma} c_1 c_2]] \\ & + (1 + C^2) \ell^2 g_1 m_H^2 [2M_2^4 m_2^2 s_1^2 + 2M_2^4 m_1^2 s_2^2 - (M_2^4 m_1^2 + M_1^4 m_2^2) s_1^2 s_2^2] \\ & - 8\ell \ell_0 m_H m_d^2 m_1 m_2 CS [M_2^2 m_1 s_2 (g_1 c_2 s_1^2 \sin\varphi_1 \cos(\varphi_1 - \varphi_2) + c_1 (g_1 c_1 c_2 + g_{\sigma\sigma}) \sin\varphi_2) \\ & - M_1^2 m_2 s_1 (g_1 c_1 s_2^2 \sin\varphi_2 \cos(\varphi_1 - \varphi_2) + c_2 (g_1 c_1 c_2 + g_{\sigma\sigma}) \sin\varphi_1)] \\ & + 2\ell^2 m_H^2 M_1^2 M_2^2 m_1 m_2 s_1 s_2 [(1 + C^2)(g_1 c_1 c_2 - g_{\sigma\sigma}) \cos(\varphi_1 - \varphi_2) - S^2(g_1 c_1 c_2 + g_{\sigma\sigma}) \cos(\varphi_1 + \varphi_2)]]. \end{aligned} \quad (D7)$$

The XP and XQ interference parts are given by

$$\begin{aligned} & 4m_1 m_2 X P \gamma_b [\ell^2 g_1 m_H^2 S^2 s_1^2 s_2^2 (M_1^2 m_2^2 \sin(2\varphi_2) - M_2^2 m_1^2 \sin(2\varphi_1)) \\ & + 2\ell \ell_0 g_1 m_H m_d^2 CS [m_2 s_1^2 c_2 s_2 (2m_1^2 \sin\varphi_1 \sin(\varphi_1 - \varphi_2) - M_1^2 \cos\varphi_2) - m_1 s_2^2 c_1 s_1 (2m_2^2 \sin\varphi_2 \sin(\varphi_1 - \varphi_2) + M_2^2 \cos\varphi_1)] \\ & - 2m_1 m_2 s_1 s_2 [(1 + C^2) \ell^2 m_H^2 M_3^2 (g_1 c_1 c_2 - g_{\sigma\sigma}) \sin(\varphi_1 - \varphi_2) + m_d^2 s^2 (g_1 c_1 c_2 + g_{\sigma\sigma}) (\ell^2 m_H^2 \sin(\varphi_1 + \varphi_2) \\ & + 2\ell_0^2 m_d^2 \sin(\varphi_1 - \varphi_2))] - 4\ell \ell_0 m_H m_1 m_2 m_d^2 CS [m_2 s_1 (g_1 c_1 + g_{\sigma\sigma} c_2) \cos\varphi_1 + m_1 s_2 (g_1 c_2 + g_{\sigma\sigma} c_1) \cos\varphi_2]], \end{aligned} \quad (D8)$$

$$\begin{aligned} & 4m_1 m_2 X Q \gamma_b [\ell \ell_0 g_\sigma m_H m_d^2 CS (m_2 s_1^2 s_2 (2m_1^2 \cos(\varphi_1 - \varphi_2) \sin\varphi_1 - M_1^2 \sin\varphi_2) \\ & - m_1 s_2^2 s_1 (2m_2^2 \cos(\varphi_1 - \varphi_2) \sin\varphi_2 - M_2^2 \sin\varphi_1)) + \ell^2 g_\sigma m_H^2 (1 + C^2) (M_1^2 m_2^2 s_1^2 c_2 + M_2^2 m_1^2 s_2^2 c_1) \\ & + m_1 m_2 s_1 s_2 [(1 + C^2) \ell^2 g_\sigma m_H^2 m_d^2 (c_1 - c_2) \cos(\varphi_1 - \varphi_2) - g_\sigma s^2 (c_1 + c_2) (\ell^2 m_H^2 M_3^2 \cos(\varphi_1 + \varphi_2) \\ & + 2\ell_0^2 m_d^4 \cos(\varphi_1 - \varphi_2))] + 2\ell \ell_0 g_\sigma m_H m_d^2 m_1 m_2 CS (1 + c_1 c_2) (m_2 s_1 \sin\varphi_1 - m_1 s_2 \sin\varphi_2)]. \end{aligned} \quad (D9)$$

Without the quark-antiquark symmetrization, one adds:

$$\begin{aligned}
& 8m_H m_1^2 m_2^2 X^2 \gamma_b^2 g_{\sigma\tau} [\ell^2 m_H C(m_2^2 c_2 s_1^2 - m_1^2 c_1 s_2^2 - m_1 m_2 (c_1 - c_2) s_1 s_2 \cos(\varphi_1 - \varphi_2)) \\
& - \ell \ell_0 m_d^2 S(m_2 s_1^2 s_2 \sin \varphi_2 + m_1 s_2^2 s_1 \sin \varphi_1)] + 2m_H P^2 g_{\sigma\tau} [\ell^2 m_H C(M_1^4 m_2^2 c_2 s_1^2 - M_2^4 m_1^2 c_1 s_2^2 \\
& + M_1^2 M_2^2 m_1 m_2 (c_1 - c_2) s_1 s_2 \cos(\varphi_1 - \varphi_2)) + 2\ell \ell_0 m_1 m_2 m_d^2 S((1 + c_1 c_2)(M_1^2 m_2 s_1 \sin \varphi_1 + M_2^2 m_1 s_2 \sin \varphi_2) \\
& - \cos(\varphi_1 - \varphi_2)(M_2^2 m_1 s_1^2 s_2 \sin \varphi_1 + M_1^2 m_2 s_2^2 s_1 \sin \varphi_2))] \\
& - 4m_H m_1 m_2 X P \gamma_b g_{\sigma\tau} [2\ell^2 m_H M_3^2 m_1 m_2 C(c_1 - c_2) s_1 s_2 \sin(\varphi_1 - \varphi_2) \\
& + \ell \ell_0 m_d^2 S[m_2 s_1^2 s_2 (M_4^2 \cos \varphi_2 - 2m_1^2 \sin \varphi_1 \sin(\varphi_1 - \varphi_2)) - m_1 s_2^2 s_1 (M_4^2 \cos \varphi_1 + 2m_2^2 \sin \varphi_2 \sin(\varphi_1 - \varphi_2))] \\
& + 2m_1 m_2 (1 + c_1 c_2)(m_2 s_1 \cos \varphi_1 - m_1 s_2 \cos \varphi_2)] \\
& + 4m_H m_1 m_2 X Q \gamma_b [2\ell^2 g_\tau m_H C(2M_1^2 m_2^2 s_1^2 - 2M_2^2 m_1^2 s_2^2 + M_4^2 m_d^2 s_1^2 s_2^2) \\
& - 4\ell^2 m_H m_d^2 m_1 m_2 c(g_{\sigma\sigma} - g_\tau c_1 c_2) s_1 s_2 \cos(\varphi_1 - \varphi_2) \\
& + 2\ell \ell_0 g_\tau m_d^2 S[m_2 c_2 s_1^2 s_2 (2m_1^2 \cos(\varphi_1 - \varphi_2) \sin \varphi_1 - M_4^2 \sin \varphi_2) + m_1 c_1 s_2^2 s_1 (2m_2^2 \cos(\varphi_1 - \varphi_2) \sin \varphi_2 - M_4^2 \sin \varphi_1)] \\
& + 4\ell \ell_0 m_1 m_2 m_d^2 S(m_2 s_1 (g_{\sigma\sigma} c_2 + g_\tau c_1) \sin \varphi_1 + m_1 s_2 (g_{\sigma\sigma} c_1 + g_\tau c_2) \sin \varphi_2)]. \tag{D10}
\end{aligned}$$

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 $m_d \rightarrow 0$, $M_1 = M_2 = M_3 \rightarrow 4m_H^2 y^2/x^2$, $M_4 \rightarrow m_H^2(x^2 - 2)/x^2$, and $\gamma_b \rightarrow xy$. For the full symmetrized matrix element squared the result is

$$\begin{aligned}
& \frac{4}{x^6} \ell^2 m_H^8 y^2 [2(x^2 X^2 + (x^2 - 4)P^2)[g_1(1 + C^2)(1 - c_1^2 c_2^2) - S^2(g_{\sigma\sigma} + g_1 c_1 c_2)s_1 s_2 \cos(\varphi_1 + \varphi_2)] \\
& - (x^2 X^2 - (x^2 - 4)P^2)[g_1 S^2 s_1^2 s_2^2 (\cos(2\varphi_1) + \cos(2\varphi_2)) - 2(1 + C^2)(g_{\sigma\sigma} - g_1 c_1 c_2)s_1 s_2 \cos(\varphi_1 - \varphi_2)] \\
& - 4XPxy s_1 s_2 [g_1 S^2 s_1 s_2 (\sin(2\varphi_1) - \sin(2\varphi_2)) - 2(1 + C^2)(g_{\sigma\sigma} - g_1 c_1 c_2)\sin(\varphi_1 - \varphi_2)] \\
& + 4XPxy g_\sigma [(1 + C^2)(c_2 s_1^2 + c_1 s_2^2) - S^2(c_1 + c_2)s_1 s_2 \cos(\varphi_1 + \varphi_2)]]. \tag{D11}
\end{aligned}$$

If we simply set $\Theta = 0$, the above simplifies to:

$$\begin{aligned}
& \frac{16}{x^6} \ell^2 m_H^8 y^2 [g_1(x^2 X^2 + (x^2 - 4)P^2)(1 - c_1^2 c_2^2) + (x^2 X^2 - (x^2 - 4)P^2)(g_{\sigma\sigma} - g_1 c_1 c_2)s_1 s_2 \cos(\varphi_1 - \varphi_2) \\
& + 4XPxy s_1 s_2 (g_{\sigma\sigma} - g_1 c_1 c_2)\sin(\varphi_1 - \varphi_2) + 2XPxy g_\sigma (c_2 s_1^2 + c_1 s_2^2)]. \tag{D12}
\end{aligned}$$

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

- [1] The CDF Collaboration, The D0 Collaboration, The Tevatron New Phenomena and Higgs Working Group, [arXiv:0911.3930](#).
- [2] G. Aad *et al.* (The ATLAS Collaboration), [arXiv:0901.0512](#).
- [3] G. L. Bayatian *et al.* (CMS Collaboration), *J. Phys. G* **34**, 995 (2007).
- [4] A. Djouadi, *Phys. Rep.* **457**, 1 (2008).
- [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**, 075008 (2008).
- [7] M. Burns, K. Kong, K. T. Matchev, and M. Park, *J. High Energy Phys.* **10** (2008) 081.
- [8] J. Alwall, P. Schuster, and N. Toro, *Phys. Rev. D* **79**, 075020 (2009).
- [9] G. Hallenbeck, M. Perelstein, C. Spethmann, J. Thom, and J. Vaughan, *Phys. Rev. D* **79**, 075024 (2009).
- [10] A. Djouadi, J. Kalinowski, and M. Spira, *Comput. Phys. Commun.* **108**, 56 (1998).
- [11] J. F. Gunion and M. Soldate, *Phys. Rev. D* **34**, 826 (1986).
- [12] B. A. Dobrescu and J. D. Lykken, *J. High Energy Phys.* **04** (2010) 83.
- [13] M. Pivk and F. R. Le Diberder, *Nucl. Instrum. Methods Phys. Res., Sect. A* **555**, 356 (2005).
- [14] J. R. Dell'Aquila and C. A. Nelson, *Phys. Rev. D* **33**, 80 (1986).
- [15] C. A. Nelson, *Phys. Rev. D* **37**, 1220 (1988).
- [16] T. Matsuura and J. J. van der Bij, *Z. Phys. C* **51**, 259 (1991).
- [17] A. Soni and R. M. Xu, *Phys. Rev. D* **48**, 5259 (1993).

- [18] D. Chang, W. Y. Keung, and I. Phillips, *Phys. Rev. D* **48**, 3225 (1993).
- [19] V.D. Barger, K. m. Cheung, A. Djouadi, B. A. Kniehl, and P. M. Zerwas, *Phys. Rev. D* **49**, 79 (1994).
- [20] A. Skjold and P. Osland, *Phys. Lett. B* **311**, 261 (1993).
- [21] B. Grzadkowski and J. F. Gunion, *Phys. Lett. B* **350**, 218 (1995).
- [22] T. Plehn, D. L. Rainwater, and D. Zeppenfeld, *Phys. Rev. Lett.* **88**, 051801 (2002).
- [23] S. Y. Choi, D. J. Miller, M. M. Muhlleitner, and P. M. Zerwas, *Phys. Lett. B* **553**, 61 (2003).
- [24] S.S. Biswal, R.M. Godbole, R.K. Singh, and D. Choudhury, *Phys. Rev. D* **73**, 035001 (2006); **74**, 039904(E) (2006).
- [25] R.M. Godbole, D.J. Miller, and M. M. Muhlleitner, *J. High Energy Phys.* **12** (2007) 031.
- [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**, 209 (2004).
- [30] W. Bernreuther, P. Gonzalez, and M. Wiebusch, arXiv:0909.3772.
- [31] K. Hagiwara, Q. Li, and K. Mawatari, *J. High Energy Phys.* **07** (2009) 101.
- [32] R. Lafaye, T. Plehn, M. Rauch, D. Zerwas, and M. Duhrssen, *J. High Energy Phys.* **08** (2009) 009.
- [33] Q. H. Cao, C. B. Jackson, W. Y. Keung, I. Low, and J. Shu, *Phys. Rev. D* **81**, 015010 (2010).
- [34] W. Y. Keung, I. Low, and J. Shu, *Phys. Rev. Lett.* **101**, 091802 (2008).
- [35] H. Davoudiasl, S. Gopalakrishna, E. Ponton, and J. Santiago, arXiv:0908.1968.
- [36] A. Bredenstein, A. Denner, S. Dittmaier, and M. M. Weber, *Phys. Rev. D* **74**, 013004 (2006).
- [37] T. Han, J. D. Lykken, and R. J. Zhang, *Phys. Rev. D* **59**, 105006 (1999).
- [38] K. M. Watson, *Phys. Rev.* **88**, 1163 (1952); A. De Rújula, J. M. K. Kaplan, and E. de Rafael, *Nucl. Phys.* **B35**, 365 (1971).
- [39] K. Hagiwara, T. Hatsukano, S. Ishihara, and R. Szalapski, *Nucl. Phys.* **B496**, 66 (1997).
- [40] L. J. Hall and C. F. Kolda, *Phys. Lett. B* **459**, 213 (1999).
- [41] R. Barbieri and A. Strumia, *Phys. Lett. B* **462**, 144 (1999).
- [42] G. L. Bayatian *et al.* (CMS Collaboration), “CMS physics: Technical design report”.
- [43] S. Frixione and B. R. Webber, *J. High Energy Phys.* **06** (2002) 029.
- [44] H. L. Lai *et al.* (CTEQ Collaboration), *Eur. Phys. J. C* **12**, 375 (2000).
- [45] T. Sjostrand, S. Mrenna, and P. Skands, *J. High Energy Phys.* **05** (2006) 026.
- [46] The CMS Collaboration, Analysis note Report No. EWK-08-006.
- [47] G. J. Feldman and R. D. Cousins, *Phys. Rev. D* **57**, 3873 (1998).
- [48] J. Neyman and E. S. Pearson, *Philosophical Transactions of the Royal Society of London, Series A* Vol. 231 (1933), pp. 289–337.
- [49] V. Ahrens, T. Becher, M. Neubert, and L. L. Yang, *Eur. Phys. J. C* **62**, 333 (2009).
- [50] D. de Florian and M. Grazzini, *Phys. Lett. B* **674**, 291 (2009).
- [51] M. Grazzini, *J. High Energy Phys.* **02** (2008) 043.
- [52] A. Einstein, remark to Oscar Veblen at Princeton University, 1921.
- [53] N. Cabibbo and A. Maksymowicz, *Phys. Rev.* **137**, B438 (1965); **168**, 1926(E) (1968).
- [54] Y. Gao, A. V. Gritsan, Z. Guo, K. Melnikov, M. Schulze, and N. V. Tran, *Phys. Rev. D* **81**, 075022 (2010).