Resonant Di-Higgs Production at Gravitational Wave Benchmarks: A Collider Study using Machine Learning

Alexandre Alves,1 Tathagata Ghosh,2 Huai-Ke Guo,3 and Kuver Sinha3

1Departamento de Física, Universidade de São Paulo, UNIFESP, Diadema, Brazil
2Department of Physics & Astronomy, University of Hawaii, Honolulu, HI 96822, USA
3Department of Physics and Astronomy, University of Oklahoma, Norman, OK 73019, USA

We perform a complementarity study of gravitational waves and colliders in the context of electroweak phase transitions choosing as our template the xSM model, which consists of the Standard Model augmented by a real scalar. We carefully analyze the gravitational wave signal at benchmark points compatible with a first order phase transition, taking into account subtle issues pertaining to the bubble wall velocity and the hydrodynamics of the plasma. In particular, we comment on the tension between requiring bubble wall velocities small enough to produce a net baryon number through the sphaleron process, and large enough to obtain appreciable gravitational wave production. For the most promising benchmark models, we study resonant di-Higgs production at the high-luminosity LHC using machine learning tools: a Gaussian process algorithm to jointly search for optimum cut thresholds and tuning hyperparameters, and a boosted decision trees algorithm to discriminate signal and background. The multivariate analysis on the collider side is able either to discover or provide strong statistical evidence of the benchmark points, opening the possibility for complementary searches for electroweak phase transitions in collider and gravitational wave experiments.

PACS numbers: 11.30.Er, 11.30.Fs, 11.30.Hv, 12.60.Fr, 31.30.jp

I. INTRODUCTION

Understanding the nature of the electroweak phase transition (EWPT) is a major goal in particle physics. A first order phase transition can be obtained by introducing new physics at the electroweak scale and this new physics can be explored at the high luminosity Large Hadron Collider (HL-LHC). On the other hand, a first order phase transition can generate gravitational waves that may be within the reach of future space-based detectors. It becomes important to understand how this complementarity plays out in concrete models - for example, can one obtain regions of parameter space where all conditions - first order phase transition, detectable gravitational waves, and a strong enough signal at the HL-LHC - are met?

The simplest template for studying these questions is the xSM model [1–3], which consists of the Standard Model (SM) extended by a real scalar. We make no comments about the completion of this model in the UV, the naturalness conflicts associated with introducing yet another scalar in addition to the Higgs, etc. Rather, our philosophy is to use the xSM as the simplest extension of the Higgs sector in which a complementary gravitational wave and collider study can be performed.

The purpose of the current paper is to first carefully explore gravitational wave signatures associated with the EWPT, and then study resonant di-Higgs production at the HL-LHC in the same context.

The new features of our study are the following:

(i) While the picture of complementarity presented above is appealing, making concrete connections from gravitational wave studies to particle physics at the electroweak scale faces many technical challenges in the calculations of electroweak baryogenesis (EWBG), EWPT and gravitational waves [4]. While we do not intend to target all these challenges in one strike, we initiate a process of making this connection more solid by presenting a careful treatment of the gravitational wave calculations.

We address several subtle issues pertaining to the bubble wall velocity and the hydrodynamics of the plasma, in particular the tension between requiring bubble wall velocities small enough to produce a net baryon number through the sphaleron process, and large enough to obtain appreciable gravitational wave production. The velocity that enters the calculations of EWBG might not be the bubble wall velocity for plasma in the modes of deflagrations and supersonic deflagrations ahead of the bubble wall, as demonstrated by hydrodynamic analysis and simulations [5]. This has the consequence that for a large wall velocity, a much smaller velocity for EWBG can be obtained and EWPT can be accompanied by a strong gravitational wave signal [6]. Therefore in our analysis, we make a clear distinction between these two velocities and determine their relation from a hydrodynamic analysis of the fluid profiles.

For our benchmark models, we compute the gravitational wave energy spectra and signal-to-noise ratio for future space-based gravitational wave experiments.

(ii) On the collider side, our objective is to apply the machine learning techniques initiated in [7] to resonant di-Higgs production, at benchmark points that are compatible with acceptable EWPT and that hold out the most optimistic prospects from gravitational wave observations. We conduct a di-Higgs study at the HL-LHC: $pp \rightarrow h_2 \rightarrow h_1 h_1 \rightarrow bb\gamma\gamma$, where $h_1$ denotes the SM Higgs. We carefully incorporate all relevant backgrounds in our study. In particular, we are careful to include contributions coming from jets being misidentified as photons, as well as light flavor jets or $c$-jets being
misidentified as $b$-jets.

We utilize two recent advances in the machine learning literature for our collider study. Firstly, recent results \cite{[8]} show that in terms of efficiency, Bayesian hyperparameter optimization of machine learning models tends to perform better than random, grid, or manual optimization. We use the Python library Hyperopt \cite{[9]} to optimize cuts on kinematic variables in our study. The second tool from the machine learning community that we apply is XGBoost \cite{[10]} (eXtreme Gradient Boosted Decision Trees), which has become increasingly popular among Kaggle competitors and data scientists in industry, especially since its winning performance in the HEP meets ML Kaggle challenge. Unlike a simple gradient boosting classifier, where classifiers (decision trees) are added sequentially, XGBoost is able to parallelize this task, leading to superior performance. Both cut thresholds and Boosted Decision Trees (BDT) hyperparameters are jointly optimized for maximum collider sensitivity.

Our paper is structured as follows. In Section II, we introduce the xSM model and settle on the benchmarks that allow a first order phase transition. In Section III, we calculate the gravitational wave energy spectra and signal-to-noise ratio for several benchmark models. In Section IV, we perform our collider analysis. We end with our Conclusions.

II. THE MODEL

The model “xSM” constitutes one of the simplest extensions of the SM where a real scalar gauge singlet $S$ is added to the particle content. The potential for the “xSM” model is defined with the convention following Ref. \cite{[1]}:

$$V(\phi, S) = -\mu^2 H^\dagger H + \lambda(H^\dagger H)^2 + \frac{a_1}{2} H^\dagger H S + \frac{a_2}{2} H^\dagger HS^2 + \frac{b_3}{3} S^3 + \frac{b_4}{4} S^4. \quad (1)$$

Here $H^\dagger = (G^+, (v + h + iG^0)/\sqrt{2})$ is the SM Higgs doublet and $S = v_s + s$ defines the real scalar singlet. All the parameters appearing here are real. The minimization conditions of this potential at the vacuum $(v, v_s)$ allows one to eliminate $\mu, b_2$ by

$$\mu^2 = \lambda v^2 + \frac{1}{2}v_s(a_1 + a_2v_s), \quad b_2 = -\frac{1}{4v_s}[v_s^2(a_1 + a_2v_s) + 4v_s^2(b_3 + b_4v_s)]. \quad (2)$$

With these substitutions, the mass matrix for $(h, s)$ is found to be:

$$m^2 = \begin{pmatrix} 2\lambda v^2 & \frac{1}{2}a_1v + v_s a_2 & v_s(b_3 + 2v_s b_4) - \frac{4v_s}{4v_s}a_1v^2 \\ \frac{1}{2}a_1v + v_s a_2 & v_s(b_3 + 2v_s b_4) - \frac{4v_s}{4v_s}a_1v^2 & v_s(b_3 + 2v_s b_4) - \frac{4v_s}{4v_s}a_1v^2 \\ v_s(b_3 + 2v_s b_4) - \frac{4v_s}{4v_s}a_1v^2 & v_s(b_3 + 2v_s b_4) - \frac{4v_s}{4v_s}a_1v^2 & v_s(b_3 + 2v_s b_4) - \frac{4v_s}{4v_s}a_1v^2 \end{pmatrix},$$

which can then be diagonalized by a rotation angle $\theta$. This results in the physical scalars $(h_1, h_2)$ in terms of the gauge eigenstates $(h, s)$:

$$h_1 = c_\theta h + s_\theta s, \quad h_2 = -s_\theta h + c_\theta s. \quad (3)$$

where $h_1$ is identified as the 125 GeV Higgs scalar and further $m_{h_2} > m_{h_1}$. Consequently, three of the potential parameters $(\lambda, a_1, a_2)$ can be replaced by three physical parameters $m_{h_1}, m_{h_2}$ and $\theta$:

$$\lambda = \frac{m_{h_1}^2 c_\theta^2 + m_{h_2}^2 s_\theta^2}{2v^2},$$

$$a_1 = \frac{2v_s}{v^2} [2v_s^2 (2b_4 + \tilde{b}_3) - m_{h_1}^2 - m_{h_2}^2 + c_\theta v (m_{h_1}^2 - m_{h_2}^2)],$$

$$a_2 = -\frac{1}{v^2}v_s [-2v_s (m_{h_1}^2 + m_{h_2}^2 - 4b_4 v_s^2) + (m_{h_1}^2 - m_{h_2}^2) (2v_s^2 v_s - v s_\theta) + 4\tilde{b}_3 v_s^3]. \quad (4)$$

where $\tilde{b}_3 \equiv b_3/v_s$. Then the full set of independent unknown parameters are

$$\nu_s, \quad m_{h_2}, \quad \theta, \quad b_3, \quad b_4. \quad (5)$$

while keeping in mind that $v$ can be solved from the Fermi constant and $m_{h_1} = 125$GeV. With the model parameters fully specified, the cubic scalar couplings that are relevant for di-Higgs production are $\lambda_{h_1 h_1 h_1}$ and $\lambda_{h_2 h_1 h_1}$, given by

$$i\lambda_{h_1 h_1 h_1} = 6 \left[ \lambda v c_\theta^3 + \frac{1}{4}v_s^2 (2a_2 v_s + a_1) + \frac{1}{2}a_2^2 v c_\theta s \theta + \frac{1}{3}v_s^3 (3b_4 v_s + b_3) \right],$$

$$i\lambda_{h_1 h_1 h_2} = \frac{1}{2} \left[ -2v_s c_\theta^2 (2a_2 v_s + a_1 - 6b_4 v_s - 2b_3) + 4v (a_2 - 3\lambda) c_\theta^2 s \theta + c_\theta^3 (2a_2 v_s + a_1) - 2a_2 s^2 \theta \right] \quad (6)$$

In the absence of mixing of the scalars when $\theta = 0$, the cubic Higgs coupling reduces to its SM value $i\lambda_{h_1 h_1 h_1} = 3m_{h_1}^2/v$ while $i\lambda_{h_1 h_1 h_2}$ vanishes. For small $\theta$ as suggested by experimental measurements, the following approximation is obtained for the cubic couplings through a Taylor expansion:

$$i\lambda_{h_1 h_1 h_1} = \frac{3m_{h_1}^2}{v} - \frac{3\theta^2}{2v} \left[ 4(b_4 + \tilde{b}_3) v_s^2 + 3m_{h_1}^2 - 4m_{h_2}^2 \right],$$

$$i\lambda_{h_1 h_1 h_2} = -\frac{4(b_4 + \tilde{b}_3) v_s^2 - 2m_{h_1}^2 + 3m_{h_2}^2}{v}. \quad (7)$$

The gauge and Yukawa couplings of $h_1$ are reduced by a factor $c_\theta$ and the couplings of $h_2$ are $-s_\theta$ times the SM values, that is:

$$\lambda_{h_1 XX} = c_\theta \lambda_{SM h_1 XX}, \quad \lambda_{h_2 XX} = -s_\theta \lambda_{SM h_2 XX}, \quad (8)$$

where $XX$ denotes $W^+ W^-, ZZ$ and $f \bar{f} f$.

Since it modifies the Higgs couplings, the mixing angle is constrained by experiments to be small. Moreover, direct searches for a heavier SM-like Higgs by ATLAS and CMS as well as electroweak precision measurements further constrain the parameter space of $(\theta, m_{h_2})$. Taking these phenomenological constraints into account, Ref. \cite{[9]} considered 12 benchmark points with $m_{h_2} \in \{ 250, 850 \}$ and studied the resonant di-Higgs production in the $b\bar{b}WW$ channel. Also imposed on these benchmarks is the strongly first order EWPT criterion, to be discussed in the next section. Several of these benchmarks are reproduced in the current work for gravitational wave and di-Higgs production studies. These are shown in Table.\cite{[9]}.
TABLE I: A subset of the benchmarks used in Ref. [3] (Table I) that can give a strongly first order EWPT as well as satisfying all phenomenological constraints. BM4-1 are neglected for reasons explained in the text. $\lambda_{111}$ and $\lambda_{112}$ are cubic couplings, given with the convention of Ref. [4]: $\lambda_{111} = i\lambda_{h_1h_1h_2}/6$ and $\lambda_{112} = \lambda_{h_1h_1h_2}/2$. Parameters that are relevant for EWPT and gravitational waves are also tabulated for each benchmark. The last column is the signal-to-noise ratio which quantifies the gravitational wave discovery prospect at LISA. See text for more detailed explanation.

III. ELECTROWEAK PHASE TRANSITION AND GRAVITATIONAL WAVES

Ever since the first detection of gravitational waves from binary black hole mergers by the LIGO and Virgo collaborations [10], gravitational waves have become an increasingly important new tool for studying astronomy and cosmology in addition to testing the general relativity of gravity in the strong field regime. More importantly, future space-based interferometer gravitational wave detectors, such as the Laser Interferometer Space Antenna (LISA) [11], can probe gravitational waves at the millihertz level, which is right the frequency range of the gravitational waves resulting from a first order EWPT [12-14]. Thus gravitational wave studies present a new window for looking into details of the mechanism of electroweak symmetry breaking, complementary to direct searches at colliders and precision measurements at the low energy intensity frontier [15-19]. This complementarity between traditional particle physics techniques and gravitational wave detections can then provide a more complete picture to understanding the physical mechanism for baryon number generation and solving the long standing baryon asymmetry problem of the universe.

A. Electroweak Phase Transition

The starting point for analyzing the EWPT is the calculation of the finite temperature effective potential, which typically involves the inclusion of the tree level effective potential, the conventional one loop Coleman-Weinberg term [20], the one loop finite temperature corrections [21] and the daisy resummation [22-23]. It is known that there is a gauge parameter dependence in the effective potential thus calculated [24]. However a gauge invariant effective potential can be obtained by doing a high temperature expansion with the result equivalent to including only the thermal mass corrections [25]. Here the gauge invariant effective potential is found to be:

$$V(h, s, T) = -\frac{1}{2}[\mu^2 - \Pi_h(T)]h^2 - \frac{1}{2}[-b_2 - \Pi_s(T)]s^2 + \frac{1}{4}\lambda h^4 + \frac{1}{4}a_1h^2s + \frac{1}{4}a_2h^2s^2 + \frac{b_3}{3}s^3 + \frac{b_4}{4}s^4, \quad (9)$$

with the thermal masses given by

$$\Pi_h(T) = \left(\frac{2m_{1W}^2 + m_{1Z}^2 + m_2^2}{4v^2} + \frac{\lambda}{2} + \frac{a_2}{24}\right)T^2, \quad (10)$$

$$\Pi_s(T) = \left(\frac{a_2}{6} + \frac{b_4}{4}\right)T^2,$$

where we have written the gauge and Yukawa couplings in terms of the physical masses of $W$, $Z$ and the $t$-quark. In the above effective potential, it is the cubic terms that allow the realization of a first order EWPT by providing a tree level barrier [95].

Among the physical parameters that characterize the dynamics of a first order EWPT, the following enter the calculation of the gravitational waves [12]:

$$T_c, \quad T_n, \quad \alpha, \quad \beta, \quad \nu_w,$$

where $T_c$ is the critical temperature at which the stable and metastable vacua become degenerate, $T_n$ is the nucleation temperature when a significant fraction of the space is filled with nucleated electroweak bubbles, $\alpha$ is the ratio between the released energy from the EWPT and the total radiation energy density at $T_n$, $\beta$ denotes approximately the inverse time duration of the EWPT and $\nu_w$ is the bubble wall velocity [26-28]. We use CosmoTransitions [29] to trace the evolution of the phases as temperature drops and solve the bounce solutions to determine $T_c, T_n, \alpha$ and $\beta$ [97]. These results are added to Table I for each benchmark. The following comments are important regarding these benchmarks:

- To avoid washout of the generated baryons inside the electroweak bubbles, the strongly first order EWPT criterion $\nu_w(T_n)/T_n > 1$ [4, 30] needs to be met, which effectively quenches the sphaleron process inside the bubbles. All the benchmarks presented in Table I satisfy this condition.

- Currently there is large uncertainty with the determination of the bubble wall velocity $\nu_w$, so it is usually taken as a free parameter in the calculations of EWBG, EWPT and gravitational waves. It is however not entirely free as there are constraints from admitting consistent hydrodynamic solutions of the plasma at the time of phase transition, to be discussed in the following.
FIG. 1: Left panel: constraint on the plane \((v_w, \alpha)\) from hydrodynamic considerations. Right panel: representative velocity profiles for plasma surrounding the bubble wall for each of the three modes with \(r\) the distance from the bubble wall center and \(t\) starting from the onset of the phase transition. See text for detailed explanations.

- Very strong phase transitions are observed for BM5, BM7, BM8 and BM9 as their values of \(\alpha\) are all larger than 1. A hydrodynamical analysis of the plasma shows that the profiles of the plasma can be classified into three categories [5]: deflagrations, detonations and supersonic deflagrations (aka hybrid) [31], depending on the value of the bubble wall velocity \(v_w\). For \(v_w\) smaller than the speed of sound in the plasma \((c_s = 1/\sqrt{3})\), the plasma takes the form of deflagrations with the following properties: (a) the plasma ahead of the phase front flows outward with non-zero velocity; (b) the plasma inside the bubbles are static. For \(c_s < \xi_f(\alpha) < v_w\) where \(\xi_f\) as a function of \(\alpha\) is the velocity corresponding to the Jouguet detonation [32], a detonation profile is obtained: (a) the plasma ahead of the wall is static; (b) the plasma inside the wall flows outward. For intermediate values of \(v_w\) with \(c_s < v_w < \xi_f(\alpha)\), a supersonic deflagration mode is obtained with the feature that both the plasma ahead of and behind the wall flow outward. An important implication relevant for the analysis here is that there is a minimum value of \(v_w\) when \(\alpha > 1/3\) for deflagration and hybrid modes [5], where \(v_w\) smaller than this value gives no consistent solution. For benchmarks BM11 and BM12 both with \(\alpha < 1/3\), \(v_w\) can take any value, while for BM5-10, there is a limited range for \(v_w\).

In the left panel of Fig. 1, we show on the plane of \((v_w, \alpha)\), the resulting ranges of \(v_w\) for BM6, BM7 and BM8, denoted by black horizontal lines that extend between the two gray region boundaries. We note that the value of \(\alpha\) for BM10 is close to that of BM6, while the values of \(\alpha\) for BM5 and BM9 are similar to BM8. We do not plot these cases to prevent the plot from being overcrowded. The left gray region is forbidden by the constraint mentioned above, while the right gray region gives a \(v_w\) too fast for EWBG to work [98]. The allowed regions in this plot are the light green region for deflagration and the brown region for supersonic deflagration. We also show three representative fluid profiles in each of the modes in the right panel of Fig. 1.

- The usual consensus for EWBG calculations is that the bubble wall velocity needs to be sufficiently small to allow diffusion of particles ahead of the wall and to produce net baryon number through the sphaleron process, with a typical value of \(v_w = 0.05\) (see for example [33–38]). However such small velocities would weaken gravitational wave production. The story changes when the hydrodynamic properties of the plasma surrounding the bubble wall are taken into account, and the dilemma between successful baryon number generation and a strong gravitational signal may be avoided. The reason is that the plasma ahead of the wall can be stirred by the expanding wall and gain a velocity in the deflagration and hybrid modes. This has the consequence that in the wall frame the plasma would hit the wall with a velocity \(v_+\) that is different from \(v_w\) [5] and it is \(v_w\) rather than \(v_w\) that should enter the calculations of EWBG. While a definitive justification of this argument would require analyzing the transport behavior of the particle species surrounding the wall in the above picture, we assume tentatively that this is true in this work (see Ref. [59] for a similar discussion on this point in the same model). The contours for a subsonic \(v_+\) with values of 0.3, 0.05 and 0.01 are shown in the left panel of Fig. 1. We can see that \(v_+\) decreases as \(\alpha\) increases for fixed \(v_w\), with the contour \(v_+ = 0\) coinciding with the boundary of the left gray region. Assuming \(v_+ = 0.05\) is used for EWBG calculations, we locate the value of \(v_w\), which corresponds to the intersection point of this contour with the horizon line of each benchmark, represented as a red point. The \(v_w\) found in this way is used to calculate the gravitational wave energy spectrum.

With above problems properly taken care of, we can now calculate the gravitational waves resulting from the EWPT.

B. Gravitational Waves

A stochastic background of gravitational waves can be generated during a first order EWPT from mainly three sources: collisions of the electroweak bubbles [40–45], bulk motion of the plasma in the form of sound waves [46, 47] and Magnetohydrodynamic (MHD) turbulence [48–49] (see Ref. [12–14] for recent reviews). The total resulting energy spectrum can be written approximately as the sum of these contributions:

\[
\Omega_{GW} h^2 \simeq \Omega_{rad} h^2 + \Omega_{ew} h^2 + \Omega_{tur} h^2.
\] (12)

While earlier studies of gravitational wave production from EWPT have focused on bubble collisions, recent advances in numerical simulations show that the long lasting sound waves during and after the EWPT give the dominant contribution to the gravitational wave production [46, 47], and the contribution from bubble collision can be neglected [50]. From such numerical simulations, an analytical formula has
been obtained for this kind of gravitational wave energy spectrum [47]:

\[ \Omega_{gw} h^2 = 2.65 \times 10^{-6} \left( \frac{H_n}{\beta} \right)^2 \left( \frac{\kappa_\nu \alpha}{1 + \alpha} \right)^2 \left( \frac{100}{g_*} \right)^{1/3} \]

\[ \times v_w \left( \frac{f}{f_{sw}} \right)^3 \left( \frac{7}{4 + 3(f/f_{sw})^2} \right)^{7/2}. \] (13)

Here \( g_* \) is the relativistic degrees of freedom in the plasma at the time of EWPT, \( H_n \) is the hubble parameter at \( T_n \) and \( f_{sw} \) is the present peak frequency which is the redshifted value of the peak frequency at the time of EWPT (\( = 2/\sqrt[3]{3v_w} \)):

\[ f_{sw} = 1.9 \times 10^{-5} \frac{1}{v_w} \left( \frac{\beta}{H_n} \right) \left( \frac{T_n}{100 \text{GeV}} \right) \left( \frac{g_*}{100} \right)^{1/6} \text{Hz}. \] (14)

The factor \( \kappa_\nu \) is the fraction of latent heat that is transformed into the bulk motion of the fluid and can be calculated as a function of \( (\alpha, v_w) \) by analyzing the energy budget during the EWPT [5]. We note that a more recent numerical simulation by the same group [51] obtained a slightly enhanced \( \Omega_{gw} h^2 \) and a slightly reduced peak frequency \( f_{sw} \).

Aside from the sound waves which give the dominant gravitational wave signals, the fully ionized plasma at the time of EWPT results in MHD turbulence, giving another source of gravitational waves. When a possible helical component [52] is neglected, the resulting gravitational wave energy spectrum can be modeled in a similar way [48,49]:

\[ \Omega_{turb} h^2 = 3.35 \times 10^{-4} \left( \frac{H_n}{\beta} \right)^2 \left( \frac{\kappa_{turb} \alpha}{1 + \alpha} \right)^{3/2} \left( \frac{100}{g_*} \right)^{1/3} \]

\[ \times \frac{v_w}{[1 + (f/f_{turb})^{11/3}(1 + 8\pi f/h_\nu)]}. \] (15)

where the peak frequency \( f_{turb} \) corresponding to MHD is given by:

\[ f_{turb} = 2.7 \times 10^{-5} \frac{1}{v_w} \left( \frac{\beta}{H_n} \right) \left( \frac{T_n}{100 \text{GeV}} \right) \left( \frac{g_*}{100} \right)^{1/6} \text{Hz}. \] (16)

Similar to \( \kappa_\nu \), here the factor \( \kappa_{turb} \) is the fraction of latent heat that is transferred to MHD turbulence. A recent numerical simulation shows that when \( \kappa_{turb} \) is parametrized as \( \kappa_{turb} \approx \epsilon \kappa_\nu \), the numerical factor \( \epsilon \) can vary roughly between 5 \( \sim \) 10% [47]. Here we take tentatively \( \epsilon = 0.1 \). As has been discussed in previous section, we take the value of \( v_w \) such that they all yield \( v_+ = 0.05 \), a good choice for EWBG calculations.

Adding the results given in Eq. (13) and Eq. (15) we can then obtain the total gravitational wave energy density spectrum. For example, the resulting gravitational wave energy spectrum for BM5 is shown in Fig. 2. The blue dashed line denotes the gravitational wave signal from sound waves and the brown dotted line from MHD turbulence, while the total contribution is shown with the solid red line. The color-shaded regions on the top are the experimentally sensitive regions for several proposed space-based gravitational wave detectors: LISA introduced earlier, the Taiji [53] and TianQin [54] programs, Big Bang Observer (BBO), DEci-hertz Interferometer Gravitational wave Observatory (DECIGO) and Ultimate-DECIGO [55,59].

To assess the discovery prospects of the generated gravitational waves, we calculate the signal-to-noise ratio with the definition adopted by Ref. [12]:

\[ \text{SNR} = \sqrt{\mathcal{T} \int_{f_{\text{min}}}^{f_{\text{max}}} df \frac{h^2 \Omega_{GW}(f)}{h^2 \Omega_{\text{exp}}(f)}}. \] (17)

where \( h^2 \Omega_{\text{exp}}(f) \) is the experimental sensitivity for the proposed experiments listed above and \( \mathcal{T} \) is the mission duration in years for each experiment. Here we assume \( \mathcal{T} = 5 \). For the LISA configurations with four links, the suggested threshold SNR for discovery is 50 [12]. For the six link configurations as drawn here, the uncorrelated noise reduction technique can be used and the suggested SNR threshold is 10 [12]. We show the SNR for the benchmarks versus \( m_{h_2} \) in Fig. 3. The SNR for LISA are also added in Table 1 for each of the benchmarks, where it shows that BM5, BM6, BM7, BM8, BM9 all have SNR larger than 10. In particular the SNR for BM5, BM7, BM8, BM9 are all much larger than 10 and for each of
these cases a very strong gravitational wave signal is expected. The last three benchmarks BM10-12 give gravitational wave signals too weak to be detected by LISA, Taiji and TianQin but some may be detected by other proposed detectors.

IV. DI-HIGGS ANALYSIS

Probing double Higgs production is a major goal of the HL-LHC \[56-61\]. Many theoretical studies of double Higgs production within the Standard Model have been conducted, for example in final states like $bb\gamma\gamma$ \[62-67\], $bb\tau^+\tau^-$ \[68-69\], $bbW^+W^-$ \[70\], and $bbbb$ \[71, 72\]. Moreover, resonant di-Higgs production has also been studied by various authors \[3, 73, 75\] in the context of EWBG \[4\].

In this Section, we study the collider prospects of probing the benchmark points for which a large SNR for proposed gravitational wave detectors has been calculated in the previous Section. The xSM model predicts a resonant di-Higgs production $pp \rightarrow h_2 \rightarrow h_1h_1 \rightarrow bb\gamma\gamma$ which is the channel that we will explore. Double Higgs production occurs through the three contributions depicted in Fig. 4. The non-resonant component involves the box diagram and the diagram with the trilinear Higgs coupling, while the resonant contribution corresponds to the diagram with $h_2$ in the $s$-channel.

FIG. 4: Representative Feynman diagrams for Di-Higgs production.

The non-resonant production cross section is strongly dependent on the size of $\lambda_{h_1h_1h_2}$, with a minimum at $\sim 0.31$ due to destructive interference between the box and the triangle diagrams. The benchmark points considered in this work all exhibit values of $\lambda_{h_1h_1h_2}$ between the SM value of 0.13 and 0.2, and for these points the non-resonant production cross section is suppressed compared to the SM. This suppression is partly compensated by the resonant contribution. We checked that the interference between the resonant and non-resonant contributions is negligible, so the contributions can be added incoherently.

While the resonant di-Higgs production cross section drops rapidly as the mass of $h_2$ is increased, the resonance peak of the $h_1h_1$ invariant mass becomes easier to identify in the tail of the background distribution as shown in Fig. 5. Taking this tradeoff into account, and noticing that BM5 and BM7 provide acceptable SNR in the gravitational waves calculation, we take these two benchmarks as the most promising ones to be probed at the HL-LHC.

We study the $bb\gamma\gamma$ channel, which is currently the most promising channel to study the double Higgs production in the SM \[7, 60, 62, 65, 76, 77\]. Recently, the fully leptonic $bbW^+W^-$ channel was studied in the context of the xSM \[3\]. This channel presents better prospects than $bb\tau^+\tau^-$ and $bb\gamma\gamma$ for scalar masses greater than around 450 GeV. However, the signal-to-background ratio for the BM5 and BM7 points is $\sim 0.1$ which may be an issue if the systematic uncertainties in $t\bar{t}$ backgrounds are not very well controlled. Moreover, the presence of two neutrinos precludes the reconstruction of the scalar resonance. The $bb\tau^+\tau^-$ channel, on the other hand, is cleaner and permits the reconstruction of the Higgses, while its cross section is much smaller than the $bbW^+W^-$ channels.

In Ref. \[7\], we found that the challenge of controlling the systematic uncertainties can be addressed by judiciously adjusting the selection criteria in order to raise the signal-to-background ratio. A full comparative study across different channels using our methods would be interesting, and is left for future study.

Inclusive di-Higgs production was simulated with MadGraph5_aMC \[78\] at $\sqrt{s} = 14$ TeV and NN23LO1 PDFs \[79\]. We multiply the non-resonant LO rates by the NNLO QCD K-factor of 2.27 \[80\], the resonant one by the NNLL QCD K-factor of 2.5 \[81\] and add them together to get the total cross section. This is justifiable once the contributions do not interfere. Besides the fact that the K-factors for the two contributions are similar, the kinematic cuts enhance the resonant contribution to eliminate backgrounds more efficiently. The total di-Higgs production cross section is thus approximated as described, and our signal events are weighted accordingly.

The signal cross sections are displayed in Table II. The Higgs bosons are decayed into bottom quarks and photons with the MadSpin module of MadGraph5. We pass our simulated events to Pythia8 \[82\] for hadronization and showering of jets. FastJet \[83\] is employed for clustering of jets and Delphes \[84\] for detector effects.

The backgrounds were also simulated within the same framework \[100\] and their total yield is shown in Table II. The backgrounds accounted for include $bb\gamma\gamma$, $Zb$ ($Z \rightarrow bb$ and $h \rightarrow \gamma\gamma$), $bbh$ ($h \rightarrow \gamma\gamma$), $t\bar{t}h \rightarrow bb + \gamma\gamma + X$, $jj\gamma\gamma$ (the light-jets $j$ are mistaken for b-jets), $bbjj$ (the light-jets $j$ are mistaken for photons), $c\bar{c}\gamma\gamma$ (a c-jet is mistagged as a b-jet), $bb\gamma j$ (the light-jet is mistaken for a photon), and $c\bar{c}\gamma j$ (the c-jets are mistagged as b-jets and the light-jet as a photon), nine in total.

The first four backgrounds are generated with one extra parton radiation to better simulate the kinematic distributions, and MLM scheme \[85\] of jet-parton matching is used to avoid double counting. Their cross section normalizations were taken from Ref. \[65\]. All the other five backgrounds are normalized by their NLO QCD rates from \[78\] but their simulation do not involve extra jets. The probability of a light-jet to be mistagged as a photon is taken to be $1.2 \times 10^{-4}$, although this may be an underestimate if pileup is taken into account.

We note that several previous studies underestimated the background, and/or did not take into account light flavor jets or c-jets being misidentified as b-jets, or jets being misidentified as photons. We correctly take into account $bb\gamma j$, $c\bar{c}\gamma\gamma$ and $c\bar{c}\gamma j$ backgrounds in our work. We assume a 70% b-tagging efficiency for jet $p_T > 100$ GeV, a photon efficiency of 90% and a 20(5)% mistagging factor for $c(\bar{c})$-jets. We refer to \[7\] for further details of the background simulation and normalization.
The basic event selection requirements are two b-tagged jets with $p_T(b) > 30$ GeV, and two photons with $p_T(\gamma) > 20$ GeV, all within $|\eta| < 2.5$. Bottom jets and photons pairs are further required to reconstruct a 125 GeV Higgs boson with $|M_{bb(\gamma\gamma)} - 125| < 25$ GeV, and all identified particles are isolated from any other reconstructed object within a cone of $\Delta R = 0.4$ around the particle’s 3-momentum.

In order to improve the statistical significance of the signal hypothesis against the background hypothesis, we used machine learning tools. First, we used an algorithm to learn the best kinematic cut thresholds in order to maximize the significance metric. This algorithm was shown to increase the significance of the non-resonant di-Higgs study in the $bb\gamma\gamma$ channel up to 50% without relying on any other multivariate analysis [7]. It is based on a Gaussian process algorithm built upon the backend program Hyperopt [8]. We refer to [7] for a detailed description of the algorithm and its usefulness in increasing the signal significance. Many other multivariate tools can be used to improve the classification of collision events as, for example, those employed in Refs. [86–89].

The kinematic variables chosen are: (1) the transverse momentum of the two leading bottom jets and the two leading photons, (2) the $\gamma\gamma$ invariant mass, (3) $\Delta R(\gamma, \gamma)$, the distance between the two leading photons, and (4) the $bb\gamma\gamma$ invariant mass, totaling seven kinematic variables. The peak in the $bb\gamma\gamma$ mass is helpful in isolating the signal events, and, in contrast to the standard analysis of non-resonant SM double Higgs production in this channel, makes the search efficient without many more variables. One interesting kinematic feature helps to explain the larger efficiency of the BM5 point. The heavy Higgs mass is right on the bulk of the non-resonant $h_1h_2 \to bb\gamma\gamma$ invariant mass after the basic selections, around 450 GeV, but for the BM7 point, it is displaced to 563 GeV as we can see in Fig. 5. Requiring a cut around the mass peak thus retains more non-resonant di-Higgs events for the BM5 point, raising its cut efficiency compared to BM7.

The cuts that maximize the signal significance for BM5 and BM7 benchmarks, assuming 3 ab$^{-1}$ of integrated luminosity and 10% systematic error in the total background rates, are the following

**BM5:**

- $p_T(b) > 47(30)$ GeV, $p_T(\gamma) > 86(49)$ GeV, $\Delta R_{\gamma\gamma} < 4.4$, $|M_{bb\gamma\gamma} - 125| < 5$ GeV, $|M_{bb\gamma\gamma} - 455| < 38$ GeV (18)

**BM7:**

- $p_T(b) > 54(30)$ GeV, $p_T(\gamma) > 104(40)$ GeV, $\Delta R_{\gamma\gamma} < 3.3$, $|M_{bb\gamma\gamma} - 125| < 5$ GeV, $|M_{bb\gamma\gamma} - 563| < 46$ GeV (19)

The cut selections for other systematics are similar. Because BM7 has a smaller production rate, the cuts learned by the algorithm were harder than the BM5 case in order to raise the significance. The cut efficiencies for the signals are almost three orders of magnitude larger than the backgrounds, as we can see in Table II, reaching a signal to background ratio slightly larger than 1 for both signal points.

| $\sigma(b\bar{b})$ | $S/B$ | $\varepsilon_{eff}$ | $L$ |
|------------------|--------|--------------------|-----|
| BM5              | 0.012  | 5.8 $\times$ 10$^{-4}$ | 0.83 |
| BM7              | 0.27   | 5.4(1.5) $\times$ 10$^{-3}$ | 4.7 |
| $\varepsilon_{eff} \cdot L$ | 14.4   | 13.5(3.7)          |     |

**TABLE II:** The signal benchmarks, BM5 and BM7 are displayed at the first two columns and the total background is displayed in the last column. In the first row, we show the cross sections, in fb, after the basic selection discussed in the text. The second and third rows show the cut efficiencies and the number of events after optimization assuming 3 ab$^{-1}$. The numbers in parenthesis in the last column represents the backgrounds for the cuts that maximize the BM7 point.

Further improvement of the study was achieved by training boosted decision trees with XGBoost [9] using the full representation of the events which comprise 28 kinematic variables: the transverse momentum of the two leading bottom jets $p_T(bb)$ and two leading photons $p_T(\gamma\gamma)$, the $\Delta R$ distance, the invariant masses and the Barr variable [91][92] of all combinations of two particles, the $bb\gamma\gamma$ mass, the azimuthal angle between the leptons and the bottoms pairs $\Delta\phi(bb, \ell\ell)$, and the missing transverse energy of the event. In order to tag $t\bar{t}h$ events we also used the number of leptons of the event. We averaged the results of a 10-fold cross validation to assess the robustness of our BDT training procedures. In order to obtain
We first calculated their gravitational wave energy spectra and signal-to-noise ratio for proposed experiments, being careful about subtle issues pertaining to the bubble wall velocity and the hydrodynamics of the plasma. Then, we took the most optimistic benchmarks and performed a collider study of double Higgs production using machine learning tools for two learning tasks: (1) to search for optimum cut thresholds and BDT hyperparameters, and (2) discriminate signal and background events with BDTs. Our results show that state-of-the-art machine learning tools can be quite powerful in probing these processes, even assuming substantial systematic uncertainties.

There are several future directions. The tension between requiring bubble wall velocities small enough to produce a net baryon number through the sphaleron process, and large enough to obtain appreciable gravitational wave production, merits further study and a more comprehensive understanding of the parameter space in concrete models. A deeper understanding of the mechanism of gravitational wave production will be needed to obtain more realistic benchmark models. On the collider side, other final states of di-Higgs, such as $bbW^+W^-$, can be studied at these realistic benchmarks using the multivariate tools we have discussed.

### V. CONCLUSIONS

Understanding the EWPT is an important goal of current and future experiments. We have explored the complementarity of the HL-LHC and proposed space-based gravitational wave detectors in achieving this goal.

We have taken the simplest template where this complementarity can be probed - the xSM model - and studied several benchmarks that are compatible with a first order EWPT.

| BM point | $\varepsilon_{\text{sys}}(\%)$ | optimized cuts($\sigma$) | S/B | BDT($\sigma$) |
|----------|-------------------------------|--------------------------|-----|---------------|
| BM5      |                               |                          |     |               |
| 5        | 3.4                           | 0.9                      | 6.3(6.4) |
| 10       | 3.2                           | 1.2                      | 6.1(6.4) |
| 15       | 2.9                           | 1.4                      | 5.9(6.4) |
| BM7      |                               |                          |     |               |
| 5        | 1.9                           | 0.8                      | 3.3(3.4) |
| 10       | 1.8                           | 0.8                      | 3.2(3.4) |
| 15       | 1.7                           | 0.8                      | 3.1(3.4) |

TABLE III: The signal significance and the signal-to-background ratio of BM5 and BM7 benchmark points for three systematic uncertainties in the background total rates scenarios – 5, 10 and 15%. The results of the optimized cut-and-count and the corresponding $S/B$ achieved are displayed in the third and fourth columns, and the BDT analysis in the last column. Also, in the last column, we show in parenthesis the results for the joint BDT+cuts optimization.

The best result possible, we tuned the BDT hyperparameters and the cut thresholds jointly. By doing this, we find the best compromise between cut-and-count and the multivariate analysis.

The BDT classification increases the signal significance for both benchmark scenarios, predicting discovery for the BM5 and evidence for BM7 for systematics ranging from 5 to 15%. In the case of BM5, a $5\sigma$ discovery might be possible for around 2 $ab^{-1}$ in the $b\bar{b}\gamma$ channel.

In the last column of Table III, we display two significances: one by cutting on the BDT scores distributions of signal and background after tuning only the kinematic cuts but keeping BDT hyperparameters fixed, reaching $6.3\sigma$ and $3.3\sigma$ for BM5 and BM7, respectively, in the 5% systematics scenario. The other number, in parenthesis, represents the significance achieved by jointly optimizing cuts and BDT hyperparameters. In this case, the significances increase slightly for all systematics but the joint optimization algorithm learns to soften $S/B$ even further, making the significance prospects insensitive to systematic uncertainties in the background rates. The final cut on the BDT scores shown in Fig. 6 is also optimized in order to get the maximum significance possible. The typical best BDT score cut is around 0.7 which corresponds approximately to a 80% efficiency for signals and 80% rejection for backgrounds resulting in around 12(4) signal events against approximately to a 80% efficiency for signals and 80% rejection for backgrounds.

The other number, in parenthesis, represents the results for the joint BDT+cuts optimization. A final selection is obtained by an additional cut on these distributions.

FIG. 6: The normalized BDT output scores distribution for signal (BM5) and background events after joint cuts and hyperparameters optimization. A final selection is obtained by an additional cut on these distributions.

### VI. ACKNOWLEDGMENTS

A. Alves thanks Conselho Nacional de Desenvolvimento Científico (CNPq) for its financial support, grant 307265/2017-0. K. Sinha is supported by the U. S. Department of Energy grant de-sc0009956. T. Ghosh is supported by U. S. Department of Energy grant de-sc0010504 and in part by U. S. National Science Foundation grant PHY-125057. HG would like to thank Hao-Lin Li for helpful discussions.

A. Alves thanks Conselho Nacional de Desenvolvimento Científico (CNPq) for its financial support, grant 307265/2017-0. K. Sinha is supported by the U. S. Department of Energy grant de-sc0009956. T. Ghosh is supported by U. S. Department of Energy grant de-sc0010504 and in part by U. S. National Science Foundation grant PHY-125057. HG would like to thank Hao-Lin Li for helpful discussions.
