# Resonant di-Higgs production at gravitational wave benchmarks: a collider study using machine learning

**Alexandre Alves,**[a] **Tathagata Ghosh,**[b] **Huai-Ke Guo**[c] **and Kuver Sinha**[c]

[a]*Departamento de Física, Universidade Federal de São Paulo, UNIFESP,*
*Diadema, Brazil*

[b]*Department of Physics & Astronomy, University of Hawaii,*
*Honolulu, HI 96822, U.S.A.*

[c]*Department of Physics and Astronomy, University of Oklahoma,*
*Norman, OK 73019, U.S.A.*

*E-mail:* aalves@unifesp.br, tghosh@hawaii.edu, ghk@ou.edu,
kuver.sinha@ou.edu

Abstract: 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.

https://doi.org/10.1007/JHEP12(2018)070

## Contents

## 1 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: $pp \rightarrow h_2 \rightarrow h_1 h_1 \rightarrow b\bar{b}\gamma\gamma$, 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 [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` [8] to optimize cuts on kinematic variables in our study. The second tool from the machine learning community that we apply is `XGBoost` [9] (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 2, we introduce the xSM model and settle on the benchmarks that allow a first order phase transition. In section 3, we calculate the gravitational wave energy spectra and signal-to-noise ratio for several benchmark models. In section 4, we perform our collider analysis. We end with our Conclusions.

## 2 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

| | $\cos\theta$ | $m_{h_2}$ (GeV) | $v_s$ (GeV) | $\lambda$ | $a_1$ (GeV) | $a_2$ | $b_3$ (GeV) | $b_4$ | $\lambda_{111}$ (GeV) | $\lambda_{211}$ (GeV) | $\frac{\lambda_{111}}{\lambda_{111}^{SM}}$ | $\Gamma_{h_2}^{\text{tot}}$ (GeV) | BR($h_1h_1$) (%) | $T_c$ (GeV) | $T_n$ (GeV) | $v_h(T_n)$ (GeV) | $\alpha$ | $\beta/H_n$ | $v_w$ ($v_+ = 0.05$) | SNR(LISA) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| BM5 | 0.984 | 455. | 47.4 | 0.179 | -708. | 4.59 | -607. | 0.85 | 47.0 | 92.8 | 1.48 | 2.06 | 30.5 | 59.3 | 33.5 | 234. | 1.88 | 127. | 0.766 | 9133. |
| BM6 | 0.986 | 511. | 40.7 | 0.185 | -744. | 5.11 | -618. | 0.82 | 46.9 | 90.5 | 1.48 | 2.44 | 22.8 | 62.3 | 49.7 | 217. | 0.48 | 726. | 0.345 | 20. |
| BM7 | 0.988 | 563. | 40.5 | 0.188 | -845. | 5.82 | -151. | 0.08 | 47.3 | 103.0 | 1.49 | 2.90 | 23.2 | 57.3 | 28.4 | 237. | 3.45 | 67. | 0.861 | 6537. |
| BM8 | 0.992 | 604. | 36.4 | 0.175 | -900. | 7.48 | -424. | 0.28 | 45.3 | 120.4 | 1.43 | 2.72 | 31.9 | 56.3 | 33.9 | 232. | 1.92 | 444. | 0.770 | 7473. |
| BM9 | 0.994 | 662. | 32.9 | 0.171 | -978. | 9.19 | -542. | 0.53 | 44.4 | 133.9 | 1.40 | 2.84 | 35.2 | 54.6 | 34.0 | 230. | 1.97 | 141. | 0.774 | 10016. |
| BM10 | 0.993 | 714. | 29.2 | 0.186 | -941. | 8.05 | 497. | 0.38 | 45.1 | 108.3 | 1.42 | 3.31 | 18.5 | 61.2 | 52.8 | 205. | 0.41 | 1307. | 0.274 | 0.50 |
| BM11 | 0.996 | 767. | 24.5 | 0.167 | -922. | 10.35 | 575. | 0.41 | 41.6 | 118.0 | 1.31 | 2.59 | 26.4 | 63.3 | 58.3 | 186. | 0.29 | 2586. | 0.164 | 0.00048 |
| BM12 | 0.994 | 840. | 21.7 | 0.197 | -988. | 8.71 | 356. | 0.83 | 44.1 | 73.3 | 1.39 | 3.98 | 6.1 | 68.9 | 67.4 | 152. | 0.13 | 10730. | 0.078 | $6.48\times10^{-10}$ |

**Table 1**. 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. BM1-4 are neglected for reasons explained in the text. $\lambda_{111}$ and $\lambda_{211}$ are cubic couplings, given with the convention of refs. [1–3]: $\lambda_{111} = i\lambda_{h_1h_1h_1}/6$ and $\lambda_{112} = i\lambda_{h_1h_1h_2}/2$. 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.

defined with the convention following refs. [1–3]:

$$V(H, S) = -\mu^2 H^\dagger H + \lambda(H^\dagger H)^2 + \frac{a_1}{2}H^\dagger H S + \frac{a_2}{2}H^\dagger H S^2 + \frac{b_2}{2}S^2 + \frac{b_3}{3}S^3 + \frac{b_4}{4}S^4. \quad (2.1)$$

Here $H^{\mathrm{T}} = (G^+, (v + h + iG^0)/\sqrt{2})$ 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

$$\mu^2 = \lambda v^2 + \frac{1}{2}v_s(a_1 + a_2 v_s),$$
$$b_2 = -\frac{1}{4v_s}[v^2(a_1 + 2a_2 v_s) + 4v_s^2(b_3 + b_4 v_s)]. \quad (2.2)$$

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

$$m^2 = \begin{pmatrix} 2\lambda v^2 & \frac{1}{2}a_1 v + v_s v\, a_2 \\ \frac{1}{2}a_1 v + v_s v a_2 & v_s(b_3 + 2v_s b_4) - \frac{1}{4v_s}v^2 a_1 \end{pmatrix},$$

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 $(h, s)$:

$$h_1 = c_\theta h + s_\theta s, \qquad h_2 = -s_\theta h + c_\theta s. \quad (2.3)$$

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 $\theta$:

$$\lambda = \frac{m_{h_1}^2 c_\theta^2 + m_{h_2}^2 s_\theta^2}{2v^2},$$
$$a_1 = \frac{2v_s}{v^2}[2v_s^2(2b_4 + \tilde{b}_3) - m_{h_1}^2 - m_{h_2}^2 + c_{2\theta}(m_{h_1}^2 - m_{h_2}^2)],$$
$$a_2 = \frac{-1}{2v^2 v_s}[-2v_s(m_{h_1}^2 + m_{h_2}^2 - 4b_4 v_s^2) + (m_{h_1}^2 - m_{h_2}^2)(2c_{2\theta}v_s - vs_{2\theta}) + 4\tilde{b}_3 v_s^3], \quad (2.4)$$

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

$$v_s, \qquad m_{h_2}, \qquad \theta, \qquad b_3, \qquad b_4, \quad (2.5)$$

while keeping in mind that $v$ can be solved from the Fermi constant and $m_{h_1} = 125\text{GeV}$. 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

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

$$i\lambda_{h_1 h_1 h_2} = \frac{1}{2}\Big[-2c_\theta s_\theta^2 \left(2a_2 v_s + a_1 - 6b_4 v_s - 2b_3\right)$$
$$+4v\left(a_2 - 3\lambda\right)c_\theta^2 s_\theta + c_\theta^3 \left(2a_2 v_s + a_1\right) - 2a_2 v s_\theta^3\Big]. \tag{2.6}$$

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:

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

$$i\lambda_{h_1 h_1 h_2} = \theta\frac{-4(2b_4 + \tilde{b}_3)v_s^2 - 2m_{h_1}^2 + 3m_{h_2}^2}{v}. \tag{2.7}$$

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,

$$\lambda_{h1XX} = c_\theta\lambda_{h1XX}^{\text{SM}}, \qquad \lambda_{h2XX} = -s_\theta\lambda_{h2XX}^{\text{SM}}, \tag{2.8}$$

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

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. [3] 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 1.[1]

## 3 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

---

[1]These parameters and the couplings all agree with [3]. Note that due to the limited precision shown in their paper, some reproduced numbers here differ slightly from their values. It should also be mentioned that in [3], a different parametrization is used with the parameter $a_2$ replaced by $m_{h_1}$. Therefore the independent set of parameters is $v_s, \lambda, a_1, m_{h_1}, b_3, b_4$. However in this method, for benchmarks BM1-3 generated in [3], the roles of $h_1$ and $h_2$ are switched, and we do not consider them further.

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 milihertz 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.

### 3.1 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:

$$
V(h, s, T) = -\frac{1}{2}[\mu^2 - \Pi_h(T)]h^2 - \frac{1}{2}[-b_2 - \Pi_s(T)]s^2 \\
+ \frac{1}{4}\lambda h^4 + \frac{1}{4}a_1 h^2 s + \frac{1}{4}a_2 h^2 s^2 + \frac{b_3}{3}s^3 + \frac{b_4}{4}s^4,
\tag{3.1}
$$

with the thermal masses given by

$$
\Pi_h(T) = \left(\frac{2m_W^2 + m_Z^2 + 2m_t^2}{4v^2} + \frac{\lambda}{2} + \frac{a_2}{24}\right)T^2,
$$
$$
\Pi_s(T) = \left(\frac{a_2}{6} + \frac{b_4}{4}\right)T^2,
\tag{3.2}
$$

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,[2] it is the cubic terms that allow the realization of a first order EWPT by providing a tree level barrier. This fact also greatly mitigates the possible effect due to neglecting higher order terms in the approach of calculating effective potential here [2]. We further note that in the above effective potential, we have neglected a tadpole term proportional to $T^2 s$, coming from the terms proportional to $a_1$ and $b_3$ in the tree level potential. The effect of this term has been found to be numerically negligible [1] as it is suppressed by $v_s/v_{EW}$.

---

[2]Note that the above effective potential can also be written in cylindrical coordinates to be compared with the result in refs. [1–3].
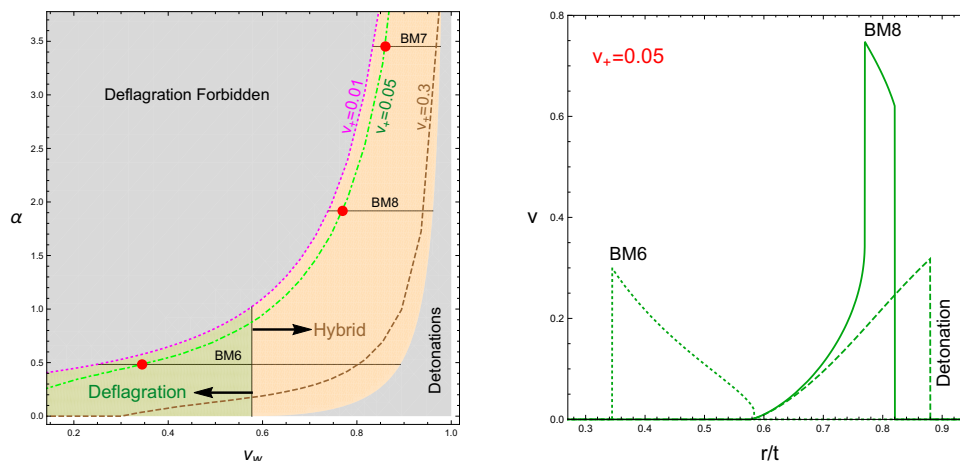
**Figure 1**. Left panel: constraint on the plane $(v_w, \alpha)$ from hydrodynamic considerations. Right panel: representative velocity profiles for plasma surrounding the bubble wall for each of the three modes with $r$ the distance from the bubble wall center and $t$ starting from the onset of the phase transition. See text for detailed explanations.

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

$$T_c, \quad T_n, \quad \alpha, \quad \beta, \quad v_w. \tag{3.3}$$

Here $T_c$ 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 $v_w$ 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 $\beta$.[3] These results are added to table 1 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 $v_h(T_n)/T_n \gtrsim 1$ [4, 30] needs to be met, which effectively quenches the sphaleron process inside the bubbles. All the benchmarks presented in table 1 satisfy this condition.

- Currently there is large uncertaintity with the determination of the bubble wall velocity $v_w$, 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.

---

[3]Aside from BM1, BM2 and BM3 in ref. [3] which we neglected for reasons explained earlier, we found that for BM4, the nucleation temperature $T_n$ cannot be obtained. This may be due to the limited precision presented there since it is known that tunneling calculations are very sensitive to input parameters.

- 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 surrounding the bubbles 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 $c_s < \xi_J(\alpha) < v_w$ where $\xi_J$ 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 $c_s < v_w < \xi_J(\alpha)$, 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 minimal 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 figure 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.[4] 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 figure 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 [34–39]). 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, 6] and it is $v_+$ rather than $v_w$ that should enter the calculations of EWBG. While a definitive justification of this argument would require analyzing the

---

[4]There may also be an additional excluded region on this plane from the consideration that for fixed $v_w$, $\alpha$ needs to be larger than a critical value to surmount a possible hydrodynamic obstruction [6, 33]. This mainly affects small values of $\alpha$ and is not considered here.

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. [40] 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 figure 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.

## 3.2 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 [41–46], bulk motion of the plasma in the form of sound waves [47, 48] and Magnetohydrodynamic (MHD) turbulence [49, 50](see refs. [12–14] for recent reviews). The total resulting energy spectrum can be written approximately as the sum of these contributions:

$$\Omega_{\mathrm{GW}}h^2 \simeq \Omega_{\mathrm{col}}h^2 + \Omega_{\mathrm{sw}}h^2 + \Omega_{\mathrm{turb}}h^2. \tag{3.4}$$

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 [47, 48] and the contribution from bubble collision can be neglected [51]. From such numerical simulations, an analytical formula has been obtained for this kind of gravitational wave energy spectrum [48]:

$$\Omega_{\mathrm{sw}}h^2 = 2.65 \times 10^{-6} \left(\frac{H_*}{\beta}\right) \left(\frac{\kappa_v \alpha}{1+\alpha}\right)^2 \left(\frac{100}{g_*}\right)^{1/3}$$
$$\times v_w \left(\frac{f}{f_{\mathrm{sw}}}\right)^3 \left(\frac{7}{4 + 3(f/f_{\mathrm{sw}})^2}\right)^{7/2}. \tag{3.5}$$

Here $g_*$ is the relativistic degrees of freedom in the plasma, $H_*$ is the Hubble parameter at $T_*$ when the phase transition has completed and has a value close to that evaluated at the nucleation temperature $H(T_n)$ for not very long EPWT. We take $T_* = T_n(1 + \kappa_T \alpha)^{1/4}$ where the fraction of vacuum energy goes to heating the plasma is given by $\kappa_T \approx 1 - \kappa_v$ [5]. Moreover, $f_{\mathrm{sw}}$ is the present peak frequency which is the redshifted value of the peak frequency at the time of EWPT($= 2\beta/(\sqrt{3}v_w)$):

$$f_{\mathrm{sw}} = 1.9 \times 10^{-5} \frac{1}{v_w} \left(\frac{\beta}{H_*}\right) \left(\frac{T_*}{100\mathrm{GeV}}\right) \left(\frac{g_*}{100}\right)^{1/6} \mathrm{Hz}. \tag{3.6}$$

The factor $\kappa_v$ 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
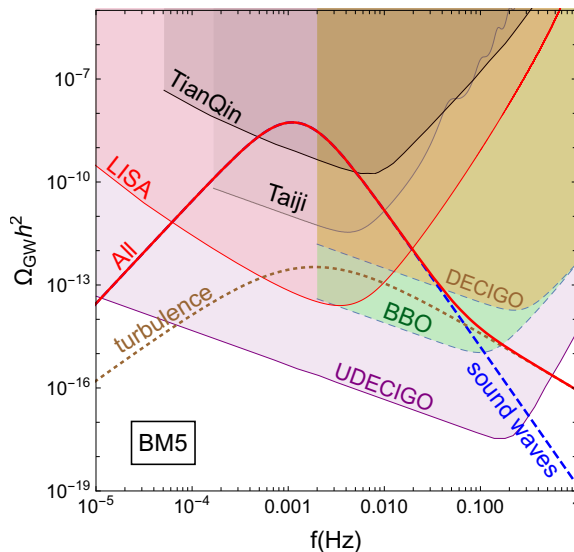
**Figure 2**. Gravitational wave energy spectrum for BM5 together with experimentally sensitive regions on the top. See text for more detailed explanations of this figure.

the EWPT [5]. We note that a more recent numerical simulation by the same group [52] obtained a slightly enhanced $\Omega_{\rm sw}h^2$ and a slightly reduced peak frequency $f_{\rm sw}$.

It should be noted that the above numerical simulations were performed under two important assumptions, which limit the possible applications here for some benchmarks. The first assumption is that the gravitational wave sourcing continues at the wavenumber corresponding to the thickness of the fluid shells, which is valid when the system is linear and requires the fluid velocity to be sufficiently smaller than unity. This is indeed what was adopted in the initial numerical simulations [47, 48] , in a more recent simulation [52] and in the recently proposed sound shell model [53]. The latter aims at understanding the origin of the shape of the gravitational wave spectra from previous simulations and linearly adds the fluid velocity profiles when calculating the velocity power spectra. This therefore puts doubts on the effectiveness in using the above formulae for our benchmarks with large velocities. Since there is currently no available result beyond current simulations, we assume the above results hold for these cases and remind the reader of this possible issue here. The second assumption is that the sourcing of gravitational waves continues until the Hubble time. This is important since the gravitational wave energy density is directly proportional to the lifetime of the sound waves. While there is no direct numerical simulation studies confirming this, it was found to be true in refs. [48, 53].

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 [54] is neglected, the resulting gravitational wave energy spectrum can be modeled in a similar way [49, 50],

$$\Omega_{\rm turb}h^2 = 3.35 \times 10^{-4} \left(\frac{H_*}{\beta}\right) \left(\frac{\kappa_{\rm turb}\alpha}{1+\alpha}\right)^{3/2} \left(\frac{100}{g_*}\right)^{1/3}$$
$$\times v_w \frac{(f/f_{\rm turb})^3}{[1+(f/f_{\rm turb})]^{11/3}(1+8\pi f/h_*)}, \tag{3.7}$$
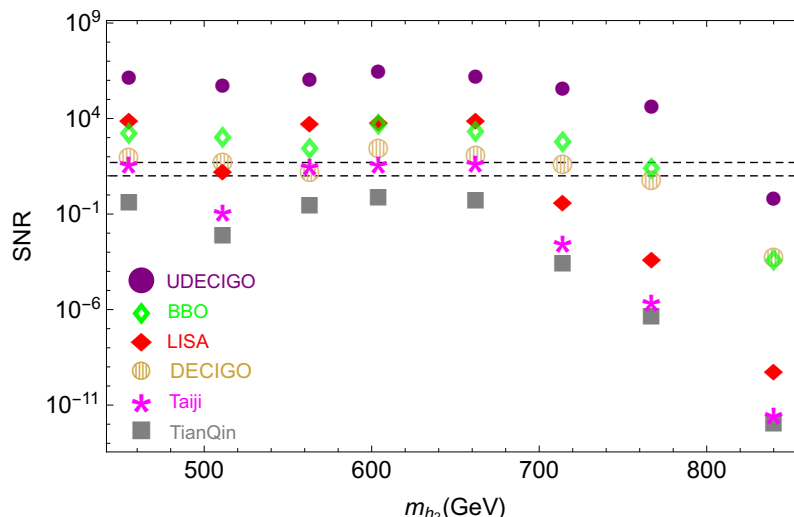
**Figure 3.** The SNR of the gravitational wave signals versus $m_{h_2}$ for the benchmarks shown in table 1 for proposed space-based gravitational wave detectors. The two horizontal dashed lines are suggested thresholds for detection: SNR = 10 and SNR = 50, depending on detector configurations.

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

$$f_{\rm turb} = 2.7 \times 10^{-5} \frac{1}{v_w} \left( \frac{\beta}{H_*} \right) \left( \frac{T_*}{100 {\rm GeV}} \right) \left( \frac{g_*}{100} \right)^{1/6} {\rm Hz}. \qquad (3.8)$$

Similar to $\kappa_v$, here the factor $\kappa_{\rm turb}$ is the fraction of latent heat that is transferred to MHD turbulence. A recent numerical simulation shows that when $\kappa_{\rm turb}$ is parametrized as $\kappa_{\rm turb} \approx \epsilon \kappa_v$, the numerical factor $\epsilon$ can vary roughly between $5 \sim 10\%$ [48]. 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. 3.5 and eq. 3.7, we can then obtain the total gravitational wave energy density spectrum. For example, the resulting gravitational wave energy spectrum for BM5 is shown in figure 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 [55] and TianQin [56] programs, Big Bang Observer (BBO), DECi-hertz Interferometer Gravitational wave Observatory (DE-CIGO) and Ultimate-DECIGO [57].[5] We note that astrophysical foregrounds, such as the unresolved stochastic gravitational waves from the population of white dwarf binaries in the Galaxy [58], might change the above sensitivity curves slightly. While a future precise modeling of these forgrounds is definitely important in discovering the stochastic gravitational wave of cosmological origin when the detector is online and taking data, we find it is sufficient to use above sensitivity curves in this study.

---

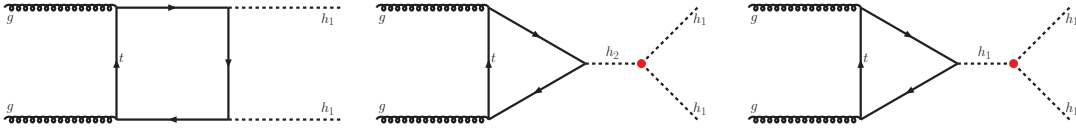[5]The BBO and DECIGO data are taken from the website http://rhcole.com/apps/GWplotter.

**Figure 4**. Representative Feynman diagrams for Di-Higgs production.

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

$$\text{SNR} = \sqrt{\delta \times \mathcal{T} \int_{f_{\min}}^{f_{\max}} df \left[ \frac{h^2 \Omega_{\text{GW}}(f)}{h^2 \Omega_{\text{exp}}(f)} \right]^2}, \tag{3.9}$$

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, assumed to be 5 here. The additional factor $\delta$ comes from the number of independent channels for cross-correlated detectors, which equals 2 for BBO as well as UDECIGO and 1 for the others [59]. 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 figure 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.

## 4 Di-Higgs analysis

Probing double Higgs production is a major goal of the HL-LHC [60–65]. Many theoretical studies of double Higgs production within the Standard Model have been conducted, for example in final states like $b\bar{b}\gamma\gamma$ [66–71], $b\bar{b}\tau^+\tau^-$ [72, 73], $b\bar{b}W^+W^-$ [74], and $b\bar{b}b\bar{b}$ [75, 76]. Moreover, resonant di-Higgs production has also been studied by various authors [3, 77–82] 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 $pp \to h_2 \to h_1 h_1 \to b\bar{b}\gamma\gamma$ which is the channel that we will explore. Double Higgs production occurs through the three contributions depicted in figure 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.

The non-resonant production cross section is strongly dependent on the size of $\lambda$, 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 $\lambda$ 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 figure 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 $b\bar{b}\gamma\gamma$ channel, which is currently the most promising channel to study the double Higgs production in the SM [7, 64, 66–69, 83, 84]. Recently, the fully leptonic $b\bar{b}W^+W^-$ channel was studied in the context of the xSM [3]. This channel presents better prospects than $b\bar{b}\tau^+\tau^-$ and $b\bar{b}\gamma\gamma$ 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 $b\bar{b}\gamma\gamma$ channel, on the other hand, is cleaner and permits the reconstruction of the Higgses, while its cross section is much smaller than the $b\bar{b}W^+W^-$ and $b\bar{b}\tau^+\tau^-$ 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` [85] at $\sqrt{s} = 14$ TeV and NN23LO1 PDFs [86]. We multiply the non-resonant LO rates by the NNLO QCD K-factor of 2.27 [87], the resonant one by the NNLL QCD K-factor of 2.5 [88] 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 2. The Higgs bosons are decayed into bottom quarks and photons with the `MadSpin` module of `MadGraph5`. We pass our simulated events to `Pythia8` [89] for hadronization and showering of jets. `FastJet` [90] is employed for clustering of jets and `Delphes` [91] for detector effects.

The backgrounds were also simulated within the same framework[6] and their total yield is shown in table 2. The backgrounds accounted for include $b\bar{b}\gamma\gamma$, $Zh$ ($Z \to b\bar{b}$ and $h \to \gamma\gamma$), $b\bar{b}h$ ($h \to \gamma\gamma$), $t\bar{t}h \to b\bar{b} + \gamma\gamma + X$, $jj\gamma\gamma$ (the light-jets $jj$ are mistaken for $b$-jets), $b\bar{b}jj$ (the light-jets $jj$ are mistaken for photons), $c\bar{c}\gamma\gamma$ (a $c$-jet is mistagged as a $b$-jet), $b\bar{b}\gamma j$ (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.

---

[6]The relevant backgrounds which contain a Higgs in the final state, the Higgs boson has been decayed within `Pythia8`.
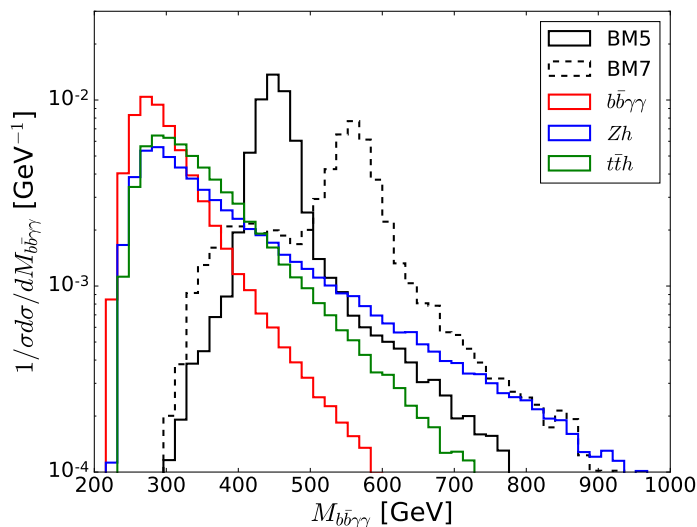
**Figure 5**. The $b\bar{b}\gamma\gamma$ invariant mass distribution for signals, BM5 (solid black) and BM7 (dashed black) benchmark points, and the main backgrounds.

The first four backgrounds are generated with one extra parton radiation to better simulate the kinematic distributions, and MLM scheme [92] of jet-parton matching is used to avoid double counting. Their cross section normalizations were taken from ref. [69]. All the other five backgrounds are normalized by their NLO QCD rates from [85] 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 $b\bar{b}\gamma j$, $c\bar{c}\gamma\gamma$ 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 $c(j)$-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 $b\bar{b}\gamma\gamma$ 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

| | BM5 | BM7 | Total Backgrounds |
|---|---|---|---|
| $\sigma$(fb) | 0.012 | $5.8 \times 10^{-3}$ | 0.83 |
| $\varepsilon_{\text{eff}}$ | 0.4 | 0.27 | $5.4(1.5) \times 10^{-3}$ |
| $\sigma \cdot \varepsilon_{\text{eff}} \cdot L$ | 14.4 | 4.7 | 13.5(3.7) |

**Table 2**. 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.

significance. Many other multivariate tools can be used to improve the classification of collision events as, for example, those employed in refs. [93–96].

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 $b\bar{b}\gamma\gamma$ invariant mass, totaling seven kinematic variables. The peak in the $b\bar{b}\gamma\gamma$ 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 $h_1 h_1 \to b\bar{b}\gamma\gamma$ 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 figure 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 benchmark points, assuming 3 ab$^{-1}$ of integrated luminosity and 10% systematic error in the total background rates, are the following

$$
\begin{aligned}
\underline{\text{BM5}:} \quad & p_T(b) > 47(30) \text{ GeV}, \; p_T(\gamma) > 86(49) \text{ GeV}, \\
& \Delta R_{\gamma\gamma} < 4.4, \; |M_{\gamma\gamma} - 125| < 5 \text{ GeV}, \\
& |M_{b\bar{b}\gamma\gamma} - 455| < 38 \text{ GeV} \\
\underline{\text{BM7}:} \quad & p_T(b) > 54(30) \text{ GeV}, \; p_T(\gamma) > 104(40), \text{ GeV}, \\
& \Delta R_{\gamma\gamma} < 3.3, \; |M_{\gamma\gamma} - 125| < 5 \text{ GeV}, \\
& |M_{b\bar{b}\gamma\gamma} - 563| < 46 \text{ GeV}.
\end{aligned}
$$

(4.1)

(4.2)

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 2, reaching a signal to background ratio slightly larger than 1 for both signal points.

The signal significance, assuming a 10% systematic error in the total background rate is 3.2$\sigma$ for the BM5, and 1.8$\sigma$ for BM7, respectively, as shown in the third column of table 3 where results for 5% and 15% systematics are also shown. Note that $S/B$, displayed in

the fourth column of table 3, increases to soften the degradation of significance with the systematics in BM5 and is kept constant for BM7. This is the job of the cut optimization program [97]. A public code of the algorithm used in this work to learn the cuts and run our multivariate analysis in an automatized way will be released in the future [97].

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 [98, 99] of all combinations of two particles, the $b\bar{b}\gamma\gamma$ 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 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 $b\bar{b}\gamma\gamma$ channel.

In the last column of table 3 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 figure 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 3(1) expected background events for the BM5(BM7) point assuming 3 ab$^{-1}$. We use the profile likelihood formula of ref. [100] which approximates well the true Poissonian statistics and embodies systematic uncertainties in the background rates to compute our signal significances.

## 5 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. 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

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

**Table 3**. The signal significance and the signal-to-background ratio of BM5 and BM7 benchmark points for three systematic uncertainties in the background total rates scenario — 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.
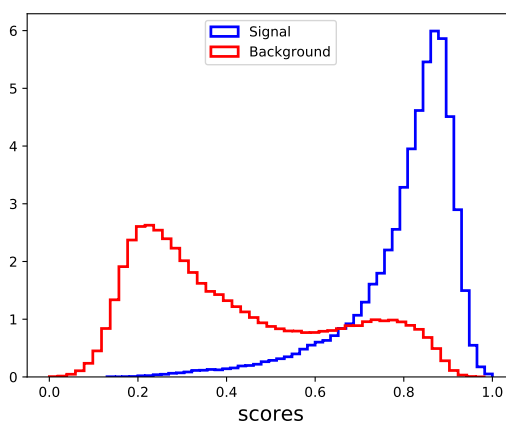


**Figure 6**. The normalized BDT output scores distribution for signal(BM5) and background events after joint cuts and hyperparameters optimization. A final selection is obtained by an additional cut on these distributions.

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 $b\bar{b}W^+W^-$, can be studied at these realistic benchmarks using the multivariate tools we have discussed.

## Acknowledgments

## References

[1] S. Profumo, M.J. Ramsey-Musolf and G. Shaughnessy, *Singlet Higgs phenomenology and the electroweak phase transition*, *JHEP* **08** (2007) 010 [arXiv:0705.2425] [INSPIRE].

[2] S. Profumo, M.J. Ramsey-Musolf, C.L. Wainwright and P. Winslow, *Singlet-catalyzed electroweak phase transitions and precision Higgs boson studies*, *Phys. Rev.* **D 91** (2015) 035018 [arXiv:1407.5342] [INSPIRE].

[3] T. Huang et al., *Resonant di-Higgs boson production in the $b\bar{b}WW$ channel: probing the electroweak phase transition at the LHC*, *Phys. Rev.* **D 96** (2017) 035007 [arXiv:1701.04442] [INSPIRE].

[4] D.E. Morrissey and M.J. Ramsey-Musolf, *Electroweak baryogenesis*, *New J. Phys.* **14** (2012) 125003 [arXiv:1206.2942] [INSPIRE].

[5] J.R. Espinosa, T. Konstandin, J.M. No and G. Servant, *Energy budget of cosmological first-order phase transitions*, *JCAP* **06** (2010) 028 [arXiv:1004.4187] [INSPIRE].

[6] J.M. No, *Large gravitational wave background signals in electroweak baryogenesis scenarios*, *Phys. Rev.* **D 84** (2011) 124025 [arXiv:1103.2159] [INSPIRE].

[7] A. Alves, T. Ghosh and K. Sinha, *Can we discover double Higgs production at the LHC?*, *Phys. Rev.* **D 96** (2017) 035022 [arXiv:1704.07395] [INSPIRE].

[8] J. Bergstra, *Hyperopt: distributed asynchronous hyper-parameter optimization*, https://github.com/jaberg/hyperopt.

[9] T. Chen and C. Guestrin, *XGBoost: a scalable tree boosting system*, https://github.com/dmlc/xgboost.

[10] VIRGO, LIGO SCIENTIFIC collaboration, B.P. Abbott et al., *Observation of gravitational waves from a binary black hole merger*, *Phys. Rev. Lett.* **116** (2016) 061102 [arXiv:1602.03837] [INSPIRE].

[11] LISA collaboration, H. Audley et al., *Laser Interferometer Space Antenna*, arXiv:1702.00786 [INSPIRE].

[12] C. Caprini et al., *Science with the space-based interferometer eLISA. II: gravitational waves from cosmological phase transitions*, *JCAP* **04** (2016) 001 [arXiv:1512.06239] [INSPIRE].

[13] R.G. Cai et al., *The gravitational-wave physics*, *Natl. Sci. Rev.* **4** (2017) 687 [arXiv:1703.00187] [INSPIRE].

[14] D.J. Weir, *Gravitational waves from a first order electroweak phase transition: a brief review*, *Phil. Trans. Roy. Soc. Lond.* **A 376** (2018) 20170126 [arXiv:1705.01783] [InSPIRE].

[15] P. Huang, A.J. Long and L.-T. Wang, *Probing the electroweak phase transition with Higgs factories and gravitational waves*, *Phys. Rev.* **D 94** (2016) 075008 [arXiv:1608.06619] [InSPIRE].

[16] K. Hashino, M. Kakizaki, S. Kanemura and T. Matsui, *Synergy between measurements of gravitational waves and the triple-Higgs coupling in probing the first-order electroweak phase transition*, *Phys. Rev.* **D 94** (2016) 015005 [arXiv:1604.02069] [InSPIRE].

[17] K. Hashino et al., *Gravitational waves and Higgs boson couplings for exploring first order phase transition in the model with a singlet scalar field*, *Phys. Lett.* **B 766** (2017) 49 [arXiv:1609.00297] [InSPIRE].

[18] A. Beniwal et al., *Gravitational wave, collider and dark matter signals from a scalar singlet electroweak baryogenesis*, *JHEP* **08** (2017) 108 [arXiv:1702.06124] [InSPIRE].

[19] D. Croon, V. Sanz and G. White, *Model discrimination in gravitational wave spectra from dark phase transitions*, *JHEP* **08** (2018) 203 [arXiv:1806.02332] [InSPIRE].

[20] S.R. Coleman and E.J. Weinberg, *Radiative corrections as the origin of spontaneous symmetry breaking*, *Phys. Rev.* **D 7** (1973) 1888 [InSPIRE].

[21] M. Quirós, *Finite temperature field theory and phase transitions*, in the proceedings of the *Summer School in High-energy physics and cosmology*, June 29–July 17, Trieste, Italy (1999), hep-ph/9901312 [InSPIRE].

[22] R.R. Parwani, *Resummation in a hot scalar field theory*, *Phys. Rev.* **D 45** (1992) 4695 [*Erratum ibid.* **D 48** (1993) 5965] [hep-ph/9204216] [InSPIRE].

[23] D.J. Gross, R.D. Pisarski and L.G. Yaffe, *QCD and instantons at finite temperature*, *Rev. Mod. Phys.* **53** (1981) 43 [InSPIRE].

[24] N.K. Nielsen, *On the gauge dependence of spontaneous symmetry breaking in gauge theories*, *Nucl. Phys.* **B 101** (1975) 173 [InSPIRE].

[25] H.H. Patel and M.J. Ramsey-Musolf, *Baryon washout, electroweak phase transition and perturbation theory*, *JHEP* **07** (2011) 029 [arXiv:1101.4665] [InSPIRE].

[26] W. Chao, H.-K. Guo and J. Shu, *Gravitational wave signals of electroweak phase transition triggered by dark matter*, *JCAP* **09** (2017) 009 [arXiv:1702.02698] [InSPIRE].

[27] L. Bian, H.-K. Guo and J. Shu, *Gravitational waves, baryon asymmetry of the universe and electric dipole moment in the CP-violating NMSSM*, *Chin. Phys.* **C 42** (2018) 093106 [arXiv:1704.02488] [InSPIRE].

[28] W. Chao, W.-F. Cui, H.-K. Guo and J. Shu, *Gravitational wave imprint of new symmetry breaking*, arXiv:1707.09759 [InSPIRE].

[29] C.L. Wainwright, *CosmoTransitions: computing cosmological phase transition temperatures and bubble profiles with multiple fields*, *Comput. Phys. Commun.* **183** (2012) 2006 [arXiv:1109.4189] [InSPIRE].

[30] J.M. Cline, *Baryogenesis*, talk given at the *Les Houches Summer School — Session 86: Particle Physics and Cosmology: The Fabric of Spacetime*, July 31–August 25, Les Houches, France (2006), hep-ph/0609145 [InSPIRE].

[31] H. Kurki-Suonio and M. Laine, *Supersonic deflagrations in cosmological phase transitions*, *Phys. Rev.* **D 51** (1995) 5431 [`hep-ph/9501216`] [ɪɴSPIRE].

[32] P.J. Steinhardt, *Relativistic detonation waves and bubble growth in false vacuum decay*, *Phys. Rev.* **D 25** (1982) 2074 [ɪɴSPIRE].

[33] T. Konstandin and J.M. No, *Hydrodynamic obstruction to bubble expansion*, *JCAP* **02** (2011) 008 [`arXiv:1011.3735`] [ɪɴSPIRE].

[34] P. John and M.G. Schmidt, *Do stops slow down electroweak bubble walls?*, *Nucl. Phys.* **B 598** (2001) 291 [*Erratum ibid.* **B 648** (2003) 449] [`hep-ph/0002050`] [ɪɴSPIRE].

[35] V. Cirigliano, S. Profumo and M.J. Ramsey-Musolf, *Baryogenesis, electric dipole moments and dark matter in the MSSM*, *JHEP* **07** (2006) 002 [`hep-ph/0603246`] [ɪɴSPIRE].

[36] D.J.H. Chung, B. Garbrecht, M. Ramsey-Musolf and S. Tulin, *Supergauge interactions and electroweak baryogenesis*, *JHEP* **12** (2009) 067 [`arXiv:0908.2187`] [ɪɴSPIRE].

[37] W. Chao and M.J. Ramsey-Musolf, *Electroweak baryogenesis, electric dipole moments and Higgs diphoton decays*, *JHEP* **10** (2014) 180 [`arXiv:1406.0517`] [ɪɴSPIRE].

[38] H.-K. Guo et al., *Lepton-Flavored Electroweak Baryogenesis*, *Phys. Rev.* **D 96** (2017) 115034 [`arXiv:1609.09849`] [ɪɴSPIRE].

[39] G.A. White, *A pedagogical introduction to electroweak baryogenesis*, IOP Concise Physics, Morgan & Claypool, U.K. (2016).

[40] J. Kozaczuk, *Bubble expansion and the viability of singlet-driven electroweak baryogenesis*, *JHEP* **10** (2015) 135 [`arXiv:1506.04741`] [ɪɴSPIRE].

[41] A. Kosowsky, M.S. Turner and R. Watkins, *Gravitational radiation from colliding vacuum bubbles*, *Phys. Rev.* **D 45** (1992) 4514 [ɪɴSPIRE].

[42] A. Kosowsky, M.S. Turner and R. Watkins, *Gravitational waves from first order cosmological phase transitions*, *Phys. Rev. Lett.* **69** (1992) 2026 [ɪɴSPIRE].

[43] A. Kosowsky and M.S. Turner, *Gravitational radiation from colliding vacuum bubbles: envelope approximation to many bubble collisions*, *Phys. Rev.* **D 47** (1993) 4372 [`astro-ph/9211004`] [ɪɴSPIRE].

[44] S.J. Huber and T. Konstandin, *Gravitational wave production by collisions: more bubbles*, *JCAP* **09** (2008) 022 [`arXiv:0806.1828`] [ɪɴSPIRE].

[45] R. Jinno and M. Takimoto, *Gravitational waves from bubble collisions: An analytic derivation*, *Phys. Rev.* **D 95** (2017) 024009 [`arXiv:1605.01403`] [ɪɴSPIRE].

[46] R. Jinno and M. Takimoto, *Gravitational waves from bubble dynamics: beyond the envelope*, `arXiv:1707.03111` [ɪɴSPIRE].

[47] M. Hindmarsh, S.J. Huber, K. Rummukainen and D.J. Weir, *Gravitational waves from the sound of a first order phase transition*, *Phys. Rev. Lett.* **112** (2014) 041301 [`arXiv:1304.2433`] [ɪɴSPIRE].

[48] M. Hindmarsh, S.J. Huber, K. Rummukainen and D.J. Weir, *Numerical simulations of acoustically generated gravitational waves at a first order phase transition*, *Phys. Rev.* **D 92** (2015) 123009 [`arXiv:1504.03291`] [ɪɴSPIRE].

[49] C. Caprini, R. Durrer and G. Servant, *The stochastic gravitational wave background from turbulence and magnetic fields generated by a first-order phase transition*, *JCAP* **12** (2009) 024 [`arXiv:0909.0622`] [ɪɴSPIRE].

[50] P. Binetruy, A. Bohe, C. Caprini and J.-F. Dufaux, *Cosmological backgrounds of gravitational waves and eLISA/NGO: phase transitions, cosmic strings and other sources*, *JCAP* **06** (2012) 027 [arXiv:1201.0983] [INSPIRE].

[51] D. Bödeker and G.D. Moore, *Electroweak bubble wall speed limit*, *JCAP* **05** (2017) 025 [arXiv:1703.08215] [INSPIRE].

[52] M. Hindmarsh, S.J. Huber, K. Rummukainen and D.J. Weir, *Shape of the acoustic gravitational wave power spectrum from a first order phase transition*, *Phys. Rev.* **D 96** (2017) 103520 [arXiv:1704.05871] [INSPIRE].

[53] M. Hindmarsh, *Sound shell model for acoustic gravitational wave production at a first-order phase transition in the early Universe*, *Phys. Rev. Lett.* **120** (2018) 071301 [arXiv:1608.04735] [INSPIRE].

[54] T. Kahniashvili et al., *Gravitational radiation from primordial helical inverse cascade MHD turbulence*, *Phys. Rev.* **D 78** (2008) 123006 [*Erratum ibid.* **D 79** (2009) 109901] [arXiv:0809.1899] [INSPIRE].

[55] X. Gong et al., *Descope of the ALIA mission*, *J. Phys. Conf. Ser.* **610** (2015) 012011 [arXiv:1410.7296] [INSPIRE].

[56] TianQin collaboration, J. Luo et al., *TianQin: a space-borne gravitational wave detector*, *Class. Quant. Grav.* **33** (2016) 035010 [arXiv:1512.02076] [INSPIRE].

[57] H. Kudoh, A. Taruya, T. Hiramatsu and Y. Himemoto, *Detecting a gravitational-wave background with next-generation space interferometers*, *Phys. Rev.* **D 73** (2006) 064006 [gr-qc/0511145] [INSPIRE].

[58] A. Klein et al., *Science with the space-based interferometer eLISA: Supermassive black hole binaries*, *Phys. Rev.* **D 93** (2016) 024003 [arXiv:1511.05581] [INSPIRE].

[59] E. Thrane and J.D. Romano, *Sensitivity curves for searches for gravitational-wave backgrounds*, *Phys. Rev.* **D 88** (2013) 124032 [arXiv:1310.5300] [INSPIRE].

[60] ATLAS collaboration, *ATLAS Higgs physics prospects at the high luminosity LHC*, PoS(ICHEP2016)426.

[61] ATLAS Collaboration, *Search for di-Higgs production with the ATLAS detector*, PoS(EPS-HEP 2017)272.

[62] CMS collaboration, D.M. Morse, *Latest results on di-Higgs boson production with CMS*, 2017, arXiv:1708.08249 [INSPIRE].

[63] D. Gonçalves et al., *Higgs boson pair production at future hadron colliders: from kinematics to dynamics*, *Phys. Rev.* **D 97** (2018) 113004 [arXiv:1802.04319] [INSPIRE].

[64] J.H. Kim, Y. Sakaki and M. Son, *Combined analysis of double Higgs production via gluon fusion at the HL-LHC in the effective field theory approach*, *Phys. Rev.* **D 98** (2018) 015016 [arXiv:1801.06093] [INSPIRE].

[65] J.H. Kim, K. Kong, K.T. Matchev and M. Park, *Measuring the triple Higgs self-interaction at the Large Hadron Collider*, arXiv:1807.11498 [INSPIRE].

[66] U. Baur, T. Plehn and D.L. Rainwater, *Probing the Higgs selfcoupling at hadron colliders using rare decays*, *Phys. Rev.* **D 69** (2004) 053004 [hep-ph/0310056] [INSPIRE].

[67] J. Baglio et al., *The measurement of the Higgs self-coupling at the LHC: theoretical status*, *JHEP* **04** (2013) 151 [arXiv:1212.5581] [INSPIRE].

[68] P. Huang, A. Joglekar, B. Li and C.E.M. Wagner, *Probing the electroweak phase transition at the LHC*, *Phys. Rev.* **D 93** (2016) 055049 [arXiv:1512.00068] [INSPIRE].

[69] A. Azatov, R. Contino, G. Panico and M. Son, *Effective field theory analysis of double Higgs boson production via gluon fusion*, *Phys. Rev.* **D 92** (2015) 035001 [arXiv:1502.00539] [INSPIRE].

[70] J. Chang et al., *Higgs-boson-pair production* $H(\to b\bar{b})H(\to \gamma\gamma)$ *from gluon fusion at the HL-LHC and HL-100 TeV hadron collider*, arXiv:1804.07130 [INSPIRE].

[71] J. Ellis, C.W. Murphy, V. Sanz and T. You, *Updated global SMEFT fit to Higgs, diboson and electroweak data*, *JHEP* **06** (2018) 146 [arXiv:1803.03252] [INSPIRE].

[72] U. Baur, T. Plehn and D.L. Rainwater, *Examining the Higgs boson potential at lepton and hadron colliders: A Comparative analysis*, *Phys. Rev.* **D 68** (2003) 033001 [hep-ph/0304015] [INSPIRE].

[73] M.J. Dolan, C. Englert and M. Spannowsky, *Higgs self-coupling measurements at the LHC*, *JHEP* **10** (2012) 112 [arXiv:1206.5001] [INSPIRE].

[74] A. Papaefstathiou, L.L. Yang and J. Zurita, *Higgs boson pair production at the LHC in the* $b\bar{b}W^+W^-$ *channel*, *Phys. Rev.* **D 87** (2013) 011301 [arXiv:1209.1489] [INSPIRE].

[75] D.E. Ferreira de Lima, A. Papaefstathiou and M. Spannowsky, *Standard model Higgs boson pair production in the* $(b\bar{b})(b\bar{b})$ *final state*, *JHEP* **08** (2014) 030 [arXiv:1404.7139] [INSPIRE].

[76] J.K. Behr et al., *Boosting Higgs pair production in the* $b\bar{b}b\bar{b}$ *final state with multivariate techniques*, *Eur. Phys. J.* **C 76** (2016) 386 [arXiv:1512.08928] [INSPIRE].

[77] ATLAS collaboration, L. Cerda Alberich, *Search for resonant and enhanced non-resonant di-Higgs production in the* $\gamma\gamma b\bar{b}$ *channel with data at 13 TeV with the ATLAS detector*, PoS(EPS-HEP 2017)687.

[78] M. Reichert et al., *Probing baryogenesis through the Higgs boson self-coupling*, *Phys. Rev.* **D 97** (2018) 075008 [arXiv:1711.00019] [INSPIRE].

[79] A. Adhikary et al., *Revisiting the non-resonant Higgs pair production at the HL-LHC*, *JHEP* **07** (2018) 116 [arXiv:1712.05346] [INSPIRE].

[80] C.-Y. Chen, S. Dawson and I.M. Lewis, *Exploring resonant di-Higgs boson production in the Higgs singlet model*, *Phys. Rev.* **D 91** (2015) 035015 [arXiv:1410.5488] [INSPIRE].

[81] I.M. Lewis and M. Sullivan, *Benchmarks for double Higgs production in the singlet extended standard model at the LHC*, *Phys. Rev.* **D 96** (2017) 035037 [arXiv:1701.08774] [INSPIRE].

[82] C.-Y. Chen, J. Kozaczuk and I.M. Lewis, *Non-resonant collider signatures of a singlet-driven electroweak phase transition*, *JHEP* **08** (2017) 096 [arXiv:1704.05844] [INSPIRE].

[83] V. Barger, L.L. Everett, C.B. Jackson and G. Shaughnessy, *Higgs-pair production and measurement of the triscalar coupling at LHC(8, 14)*, *Phys. Lett.* **B 728** (2014) 433 [arXiv:1311.2931] [INSPIRE].

[84] S. Dawson et al., *Working Group Report: Higgs Boson*, in the proceedings of the *2013 Community Summer Study on the Future of U.S. Particle Physics: Snowmass on the Mississippi (CSS2013)*, July 29–August 6, Minneapolis, U.S.A. (2013), arXiv:1310.8361 [INSPIRE].

[85] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections and their matching to parton shower simulations*, *JHEP* **07** (2014) 079 [arXiv:1405.0301] [INSPIRE].

[86] NNPDF collaboration, R.D. Ball et al., *Parton distributions with QED corrections*, *Nucl. Phys.* **B 877** (2013) 290 [arXiv:1308.0598] [INSPIRE].

[87] D. de Florian and J. Mazzitelli, *Two-loop virtual corrections to Higgs pair production*, *Phys. Lett.* **B 724** (2013) 306 [arXiv:1305.5206] [INSPIRE].

[88] S. Catani, D. de Florian, M. Grazzini and P. Nason, *Soft gluon resummation for Higgs boson production at hadron colliders*, *JHEP* **07** (2003) 028 [hep-ph/0306211] [INSPIRE].

[89] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159 [arXiv:1410.3012] [INSPIRE].

[90] M. Cacciari, G.P. Salam and G. Soyez, *FastJet user manual*, *Eur. Phys. J.* **C 72** (2012) 1896 [arXiv:1111.6097] [INSPIRE].

[91] DELPHES 3 collaboration, J. de Favereau et al., *DELPHES 3, A modular framework for fast simulation of a generic collider experiment*, *JHEP* **02** (2014) 057 [arXiv:1307.6346] [INSPIRE].

[92] M.L. Mangano, M. Moretti, F. Piccinini and M. Treccani, *Matching matrix elements and shower evolution for top-quark production in hadronic collisions*, *JHEP* **01** (2007) 013 [hep-ph/0611129] [INSPIRE].

[93] CDF collaboration, T. Aaltonen et al., *Observation of single top quark production and measurement of $|V_{tb}|$ with CDF*, *Phys. Rev.* **D 82** (2010) 112005 [arXiv:1004.1181] [INSPIRE].

[94] P. Baldi, P. Sadowski and D. Whiteson, *Searching for exotic particles in high-energy physics with deep learning*, *Nature Commun.* **5** (2014) 4308 [arXiv:1402.4735] [INSPIRE].

[95] A. Alves, *Stacking machine learning classifiers to identify Higgs bosons at the LHC*, 2017 *JINST* **12** T05005 [arXiv:1612.07725] [INSPIRE].

[96] A. Alves and K. Sinha, *Searches for dark matter at the LHC: a multivariate analysis in the mono-Z channel*, *Phys. Rev.* **D 92** (2015) 115013 [arXiv:1507.08294] [INSPIRE].

[97] A. Alves, T. Ghosh and K. Sinha, *CutOptimize: a Python package fot cut-and-count optimization*, to be relesead.

[98] A.J. Barr, *Measuring slepton spin at the LHC*, *JHEP* **02** (2006) 042 [hep-ph/0511115] [INSPIRE].

[99] A. Alves and O. Eboli, *Unravelling the sbottom spin at the CERN LHC*, *Phys. Rev.* **D 75** (2007) 115013 [arXiv:0704.0254] [INSPIRE].

[100] T.P. Li and Y.Q. Ma, *Analysis methods for results in gamma-ray astronomy*, *Astrophys. J.* **272** (1983) 317 [INSPIRE].