Non-resonant Higgs pair production in the $b\bar{b}b\bar{b}$ final state at the LHC

David Wardrope, Eric Jansen, Nikos Konstantinidis, Ben Cooper, Rebecca Falla, and Nurfikri Norjoharuddeen
Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, United Kingdom

We present a particle-level study of the Standard Model non-resonant Higgs-pair production process in the $b\bar{b}b\bar{b}$ final state, at the Large Hadron Collider at $\sqrt{s} = 14$ TeV. Each Higgs boson is reconstructed from a pair of close-by jets formed with the anti-$k_t$ jet clustering algorithm, with radius parameter $R = 0.4$. Given the kinematic properties of the produced Higgs bosons, we show that this reconstruction approach is more suitable than the use of two large-radius jets that capture all the decay products of a Higgs boson, as was previously proposed in the literature. We also demonstrate that the sensitivity for observing this final state can be improved substantially when the full set of uncorrelated angular and kinematic variables of the 4$b$ system is combined, leading to a statistical significance of $\sim 2$ per experiment with an integrated luminosity of $3 \text{ab}^{-1}$.

I. INTRODUCTION

The thorough investigation of the properties of the Higgs boson discovered by ATLAS and CMS [1, 2] is one of the highest priorities in particle physics for the next two decades. A crucial property is the trilinear Higgs self-coupling which can be measured by the observation of Higgs-pair production. At the Large Hadron Collider (LHC), this is considered to be one of the most challenging processes to observe, even with a data set corresponding to an integrated luminosity of $3 \text{ab}^{-1}$, the target for the proposed High Luminosity LHC (HL-LHC) programme. Several particle-level studies were published even before the Higgs discovery [3, 4] and more have been published since then, assessing the sensitivity of different decay channels such as $b\bar{b}\gamma\gamma$, $b\bar{b}\tau\tau$ and $b\bar{b}WW$ [5–9]. The $b\bar{b}b\bar{b}$ final state was examined in Ref. [10], where it was found to have very low sensitivity, and more recently in Ref. [11] where the use of a tighter kinematic selection and jet substructure techniques appeared to give some improved sensitivity, although that study considered only the $4b$ multijet process as background.

In this paper, we extend our previous work on resonant Higgs-pair production in the $b\bar{b}b\bar{b}$ final state [12]—which inspired the recent ATLAS analysis [13]—to the non-resonant case, considering all the relevant background processes, namely $b\bar{b}b\bar{b}$, $b\bar{b}c\bar{c}$, and $t\bar{t}$. The $b\bar{b}b\bar{b}$ final state benefits from the high branching fraction of Higgs decaying to $b\bar{b}$ (57.8% in the Standard Model (SM) for $m_H = 125.5 \text{GeV}$, leading to about one third of the Higgs pairs decaying to $b\bar{b}b\bar{b}$), but suffers from large backgrounds. However, like in the previously studied resonant case [12], the transverse momentum ($p_T$) of the Higgs bosons in the non-resonant process in the SM is relatively high, with the most probable value around 150 GeV [11]. By tailoring the event selection to focus on this high-$p_T$ regime, where the two Higgs bosons are essentially back-to-back, one has the benefits outlined in Ref. [12] for the resonant case. Requiring four $b$-tagged jets, paired into two high-$p_T$ dijet systems, is a very powerful way to reduce the backgrounds. This is particularly true for the dominant multijet production, which has a cross section that falls rapidly with increasing jet and dijet $p_T$. There is also negligible ambiguity in pairing the four $b$-jets to correctly reconstruct the Higgs decays. Finally, due to the high boost, the four jets will have high enough transverse momenta for such events to be selected with high efficiency at the first level trigger of ATLAS and CMS, with efficient high level triggering possible through online $b$-tagging [13]. We note that triggering will be a major challenge at the HL-LHC, but the substantial detector and trigger upgrade programmes proposed by the two experiments, should make it possible to maintain the high trigger efficiencies reported by ATLAS in the 8 TeV run [13] for channels that are essential for key measurements at the HL-LHC, such as the Higgs trilinear self-coupling.

II. SIMULATION OF SIGNAL AND BACKGROUND PROCESSES

Signal and background processes are modelled using simulated Monte Carlo (MC) event samples. The $HH \rightarrow b\bar{b}b\bar{b}$ signal events are generated with MadGraph [14] 1.5.12, interfaced to PyTHIA [15] 8.175 for parton showering (PS) and hadronisation, and using the CTEQ6L1 [16] leading-order (LO) parton-density functions (PDF). The signal is scaled to a cross-section of $11.6 \text{fb}$ [17]. The $t\bar{t}$ events are simulated using PowHEG [18, 19] interfaced to PyTHIA 8.185. Only hadronic $t\bar{t}$ events are considered in this study (including hadronic $\tau$ decays), as the semileptonic (dileptonic) decays are suppressed by a lower branching fraction and the need for one (two) additional $b$-tagged jet(s) to pass the event selection. The $b\bar{b}b\bar{b}$ and $b\bar{b}c\bar{c}$ backgrounds are generated by Sherpa [20] 2.1.1, using the CT10 [21] PDF set. These event samples are scaled to their next-to-leading order (NLO) cross-section by applying a $k$-factor of 1.5 [22]. For all the above background processes, there is an additional filtering requirement at the event generation level, that there be at least four jets with $p_T > 40 \text{GeV}$. In addition, we have considered the most relevant single-Higgs production channels to give an indication of their con-
TABLE I: Summary of the event generators used to model the signal and background processes. The quoted $\sigma \times BR$ in the last column includes the event filtering described in the text, for the $b\bar{b}b\bar{b}$, $b\bar{b}c\bar{c}$, and $t\bar{t}$ processes.

| Process           | Generator                  | PDF set     | $\sigma \times BR$ [pb] |
|-------------------|----------------------------|-------------|-------------------------|
| $HH \to b\bar{b}b\bar{b}$ | MadGraph + Pythia        | CTEQ6L1     | $1.16 \times 10^{-2}$   |
| $b\bar{b}b\bar{b}$   | Sherpa                    | CT10        | 219                     |
| $b\bar{b}c\bar{c}$   | Sherpa                    | CT10        | 477                     |
| $t\bar{t}$           | Powheg + Pythia           | CT10        | 212                     |
| $ZH \to b\bar{b}b\bar{b}$ | Pythia                  | CTEQ6L1     | $3.56 \times 10^{-2}$   |
| $t\bar{t}(\to b\bar{b})$ | Pythia                  | CTEQ6L1     | $1.36 \times 10^{-1}$   |
| $H(\to b\bar{b})b\bar{b}$ | MadGraph\_aMC@nlo + Pythia | CTEQ6L1     | $4.89 \times 10^{-1}$   |

distribution in comparison to the signal and the dominant backgrounds listed above. The $Hb\bar{b}$ background is generated using MadGraph\_aMC@nlo [23] 1.5.12 interfaced to Pythia. The $ZH$ and $t\bar{t}H$ background processes are both generated using Pythia 8.175. The Higgs mass has been fixed to 125 GeV. More details can be found in Table I.

III. DISCUSSION OF THE SIGNAL TOPOLOGY

In this Section, we discuss briefly the signal topology and motivate, to some extent, the choices for the event selection strategy described in Section IV.

Figure 1 shows the $p_T$ distributions of the leading (circles) and subleading (squares) Higgs bosons in signal events. As mentioned above, in a substantial fraction of signal events (36.6%), both Higgs bosons have $p_T > 150$ GeV, whereas only 16.6% (3.6%) of events have both Higgs bosons with $p_T > 200$ GeV (300 GeV). Figure 2 compares the efficiency for reconstructing the Higgs boson, as a function of its $p_T$, using (a) two anti-$k_t$ jets with $R = 0.4$, and (b) a single Cambridge-Aachen (CA) jet [24, 25] with $R = 1.2$. It can be seen that the efficiency of the first approach is above 95% for Higgs $p_T$ values up to about 450 GeV. In contrast, the CA approach starts with low efficiency because, for low Higgs $p_T$, a radius $R = 1.2$ CA jet often cannot capture the Higgs decay products, and only becomes more than 95% efficient for Higgs $p_T$ values above around 400 GeV, where very little SM non-resonant HH signal remains. The above features are demonstrated clearly also in Figure 3, which shows the distance $\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2}$ between the two $b$-quarks from the Higgs boson decay as a function of the Higgs boson $p_T$. On the grounds of these observations,
we base our reconstruction strategy on four anti-$k_t$ jets with $R = 0.4$.

![Diagram](image)

**FIG. 3:** The distance $\Delta R$ between the two $b$-quarks from the Higgs boson decay as a function of the Higgs boson $p_T$.

## IV. EVENT SELECTION

The event selection proceeds by requiring at least four $b$-tagged jets with $p_T > 40$ GeV and $|\eta| < 2.5$. Jets are formed using the anti-$k_t$ algorithm [26] with radius parameter $R = 0.4$, implemented in Fastjet [27]. In order to emulate the effect of $b$-tagging in this particle-level study, we adopt the following procedure: jets are labelled as $b$-jets, $c$-jets, $\tau$-jets or light jets depending on the flavour of particles within $\Delta R < 0.3$ of the jet axis. If a $b$-hadron is found, the jet is labelled a $b$-jet, otherwise if a $c$-hadron is found the jet is labelled a $c$-jet. If neither a $b$-hadron nor a $c$-hadron is found, but a $\tau$-lepton is found instead, the jet is labelled a $\tau$-jet. All other jets are classified as light jets. We then apply $b$-tagging efficiency weights inspired by the published ATLAS and CMS $b$-tagging performance [28, 29]: 70% for $b$-labelled jets, 20% for $c$-labelled and $\tau$-labelled jets (i.e. a rejection factor of 5) and 1% for light-labelled jets (rejection factor of 100). All jets in the event are ordered by $b$-tagging weight and subsequently by $p_T$. The leading four jets are then used to form dijets, requiring $p_{T,\text{dijet}} > 150$ GeV, $85 < m_{\text{dijet}} < 140$ GeV and $\Delta R < 1.5$ between the two jets of the dijet system. If more than two dijets satisfy the above criteria, the two which are most back-to-back in the plane transverse to the beam line are retained. The two dijets are ordered in $p_{T,\text{dijet}}$, and the leading dijet is required to have $100 < m_{\text{dijet}} < 140$ GeV, while the subleading one must satisfy $85 < m_{\text{dijet}} < 130$ GeV. This is because the subleading dijet is often the one where one of the $b$-hadrons has decayed semileptonically, hence the dijet invariant mass shifts to lower values than 125 GeV and has a larger low-mass tail. Finally, in order to reject $t\bar{t}$ events we use the TMVA framework [30] to train a Boosted Decision Tree (BDT) discriminant, $X_{t\bar{t}}$, using four input variables, two from each dijet system, calculated as follows. We search for a third jet with $\Delta R < 2$ from the jets of the dijet system, and then calculate: (a) the invariant mass of the three-jet system (which would be close to the top mass for a hadronic top quark decay); and (b) the invariant mass of the third jet with the subleading jet of the dijet system (giving often the $W$ mass in a hadronic top quark decay). Using $X_{t\bar{t}}$, the $t\bar{t}$ background is reduced by a factor of $\sim 2.5$ for a 10% reduction in the signal and the multijet background.

After the above selection, the remaining signal cross section is 0.19 fb, corresponding to about 570 events in 3 ab$^{-1}$. The multijet background cross section is 82 fb, dominated by $b\bar{b}b\bar{b}$, and the $t\bar{t}$ cross section is 29 fb, indicating that the $t\bar{t}$ is a sizeable fraction of the total background. The single-Higgs production $H(\to bb)b\bar{b}$, $t\bar{t}H$ and $ZH$ processes have a combined cross section of 0.33 fb, comparable to the signal, with the main contribution coming from $t\bar{t}H$. Therefore, the signal-to-background ($s/b$) ratio at this point is 0.17% and the expected statistical significance ($s/\sqrt{b}$) for 3 ab$^{-1}$ is 0.98. Clearly, with such a low $s/b$ ratio, it would be impossible to extract any signal sensitivity reliably.

Further to the above selection, any additional differences between the signal and background can be exploited using the following list of ten independent kinematic and angular variables:

- the decay angle of the Higgs bosons in the rest frame of the 4$b$ system, $\Theta^*$;
- the decay angles of the $b$-quarks in the rest frame of the Higgs bosons, $\theta_1$ and $\theta_2$;
- the angle between the decay planes of the two Higgs bosons, $\Phi$;
- the angle between one of the above decay planes and the decay plane of the two-Higgs system, $\Phi_1$;
- the two dijet invariant masses, $m_{12}$ and $m_{34}$;
- the invariant mass of the 4$b$ system, $m_X$;
- the $p_T$ of the 4$b$ system, $p_{T,X}$; and
- the rapidity of the 4$b$ system, $y_X$.

These variables have also been proposed [31] and used [32] in the context of the $H \to ZZ^* \to 4\ell$ analyses at the LHC. Figure [3] shows the distributions of these variables in the signal and background after the above event selection. It can be seen, that some of them have little discrimination following the event selection, but others
FIG. 4: Subfigures (a)-(j) show the kinematic and angular variables used to separate the signal and background processes, as described in the text. Subfigure (k) shows the shape of the $t\bar{t}$ discriminant, $X_{tt}$, after the top veto has been applied.
show significant differences between the signal and backgrounds.

We combine the above variables, together with $X_H$, in a single BDT discriminant, $D_{HH}$. The outputs of this discriminant for signal and background are shown in Figure 5, while the performance of this discriminant is shown in Figure 6 in terms of signal efficiency vs. background rejection.

![Figure 5: The BDT discriminant $D_{HH}$.](image)

![Figure 6: Performance of the BDT discriminant $D_{HH}$](image)

**FIG. 5:** The BDT discriminant $D_{HH}$.

**FIG. 6:** Performance of the BDT discriminant $D_{HH}$.

### V. RESULTS AND DISCUSSION

Figure 7 shows $s/b$ and $s/\sqrt{b}$ as a function of the relative signal efficiency when varying the cut on $D_{HH}$, for an integrated luminosity of $3 \text{ ab}^{-1}$.

![Figure 7: $s/b$ and $s/\sqrt{b}$ as a function of the relative signal efficiency when varying the cut on $D_{HH}$](image)

These results demonstrate that the $t\bar{t}$ and $b\bar{b}c\bar{c}$ processes together represent more than half of the total background. Most of the remaining $t\bar{t}$ background consists of events where the decay products from both $W$'s from the top decays include a charm jet or a jet from a hadronic tau decay. This gives additional motivation to improve the charm and tau jet rejection of $b$-tagging at the HL-LHC. While the increasing pile-up will make this task even more challenging, the significantly improved pixel tracking detectors proposed for both ATLAS [33] and CMS are likely to provide the necessary $b$-tagging performance improvements. In order to demonstrate the potential benefits to this analysis from an improved $c/\tau$-jet rejection, we repeated the above study assuming a $b$-tagging efficiency of 10% for $c/\tau$-labeled jets. When doing this, the maximum statistical significance obtained is 2.13 with $s/b \approx 2.4\%$. Finally, if it was possible to suppress completely all non-$4b$ backgrounds for the same signal efficiency as above, then the above selection would give a statistical significance of 2.81 with $s/b \approx 4.1\%$. These last results are directly comparable with those pre-
sent in Ref. [11], indicating that our approach has a much better sensitivity for this process than the jet substructure approach proposed in that reference.

It is worth pointing out that recent theoretical calculations of the SM Higgs-pair production cross section with various improvements [53][56] find it to be 20-30\% higher than the NLO value used here. Even if the cross sections of the background processes were up by a similar factor with more precise calculations, the $s/\sqrt{b}$ would eventually be 10-15\% better than what is quoted above.

As this is a particle-level study, it is expected that experimental resolution effects will reduce somewhat the discriminating power of the variables used in the above event selection. However, it is worth pointing out that our particle-level predictions in Ref. [12], appear to be in broad agreement with the ATLAS result [13] that includes all the experimental resolution effects and background estimation uncertainties. In addition, there is plenty of scope for further optimising the current analysis. Examples of possible avenues to explore for further optimisation include: fitting the distribution of $D_{HH}$ to extract more information from the data; the use of control regions and data-driven techniques for determining the various backgrounds, as in Ref. [13]; the use of kinematic fitting techniques to improve the angular resolution of the four jets and hence the discriminating power of the angular variables described above; or the use of the shape of the b-tagging discriminant for each jet, to suppress further the non-4b background events.

VI. CONCLUSIONS

In SM non-resonant Higgs-pair production at the LHC, the Higgs bosons are mostly produced back-to-back, with relatively large $p_T$. Selecting four b-tagged jets and forming two back-to-back pairs, with $p_T^{\text{dijet}} > 150\text{ GeV}$ and $\Delta R < 1.5$ between the two jets in each pair, leads to a drastic suppression of all background processes (particularly the dominant multijet production) while maintaining a good signal yield. We conclude that, given the $p_T$ spectrum of the Higgs bosons, the use of pairs of anti-

We further find that exploiting the full kinematic and angular information of the 4b system can provide very substantial additional improvement in the sensitivity for $HH \rightarrow b\bar{b}b\bar{b}$ and the measurement of the Higgs trilinear self-coupling. Our particle-level study yields a statistical significance of 1.79 (2.13) per experiment for an integrated luminosity of 3 ab$^{-1}$, assuming a $c/\tau$-jet b-tagging efficiency of 20\% (10\%). While experimental systematic uncertainties will tend to reduce the sensitivity of the measurement, there is still plenty of scope to optimise the analysis further, hence we expect that the sensitivity quoted here should be achievable eventually at the HL-LHC.

| Requirement | $HH$ [fb] | $b\bar{b}b\bar{b}$ [fb] | $b\bar{b}c\tau$ [fb] | $t\bar{t}$ [fb] | single-$H$ [fb] | $s/b$ | $s/\sqrt{b}$ (for $3\text{ ab}^{-1}$) |
|-------------|-----------|-----------------|-----------------|----------------|---------------|-------|-------------------------------|
| Two dijets  | 0.30      | 513             | 122             | 290            | 2.53          | $3.2 \times 10^{-4}$ | 0.54             |
| $m_H$ windows | 0.21      | 74              | 17              | 73             | 0.65          | $1.3 \times 10^{-3}$ | 0.89             |
| Top veto    | 0.19      | 67              | 15              | 29             | 0.33          | $1.7 \times 10^{-3}$ | 0.98             |
| $D_{HH}$    | 0.08      | 2.8             | 0.6             | 2.6            | 0.05          | $1.3 \times 10^{-2}$ | 1.79             |
| $\epsilon_{c/\tau\text{-jet}} = 10\%$ | 0.06 | 1.5             | 0.1             | 1.0            | 0.04          | $2.4 \times 10^{-2}$ | 2.13             |
| $\epsilon_{c/\tau\text{-jet}} = 0\%$ | 0.06 | 1.5             | 0.0             | 0.0            | 0.04          | $4.1 \times 10^{-2}$ | 2.81             |
