Further on up the road: $hhjj$ production at the LHC

Matthew J. Dolan, Christoph Englert, Nicolas Greiner, and Michael Spannowsky

1Theory Group, SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA
2SUPA, School of Physics and Astronomy, University of Glasgow, Glasgow, G12 8QQ, United Kingdom
3Max-Planck-Institut für Physik, Föhringer Ring 6, 80805 München, Germany
4Institute for Particle Physics Phenomenology, Department of Physics, Durham University, DH1 3LE, United Kingdom

A measurement of the $hh + 2j$ channel at the LHC would be particularly thrilling for electroweak physics. It is not only the leading process which is sensitive to the $W^+W^-hh$ and $ZZhh$ interactions, but also provides a potentially clear window to study the electroweak symmetry-breaking sector by probing Higgs-Goldstone interactions through the weak boson fusion component of the scattering process. Until now, a phenomenologically complete analysis of this channel has been missing. This is mostly due to the high complexity of the involved one-loop gluon fusion contribution and the fact that a reliable estimate thereof cannot be obtained through simplified calculations in the $m_t \to \infty$ limit.

In particular, the extraction of the Higgs trilinear coupling from this final state rests on a delicate $m_t$-dependent interference pattern which is not captured in an effective field theory approach. In this paper, we report on the LHC’s potential to access di-Higgs production in association with two jets in a fully-showered hadron-level analysis. Our study includes the finite top and bottom mass dependencies for the gluon fusion contribution. On the basis of these results, we also comment on the potential sensitivity to the trilinear Higgs and $VV^+hh$ ($V = W^\pm, Z$) couplings that can be expected from measurements of this final state.

Context and Introductory Remarks. After the recent discovery of the Higgs boson $H$ at the LHC and subsequent analyses of its interactions with known matter, a coarse-grained picture of consistency with the Standard Model (SM) expectation appears to be emerging. Taking this agreement at the level of Higgs couplings at face value, an immediate question that comes to mind is whether the realisation of electroweak symmetry breaking (EWSB) is in fact SM-like too.

Crucial to Higgs sector-induced EWSB is the presence of higher order monomials of the Higgs field in the potential which misalign the Higgs field from the electroweak symmetry-preserving direction, thus realising the gauge symmetry non-linearly via the Higgs mechanism. These terms are currently unknown and it is experimentally unclear whether they exist at all.

In the minimal approach of the SM the potential reads

$$V(H\dagger H) = \mu^2 H\dagger H + \eta (H\dagger H)^2 \overset{!}{=} \frac{1}{2} m_h^2 h^2 + \sqrt{\eta} m_h h^3 + \frac{\eta}{4} h^4 \eqref{1}$$

(after the Higgs doublet field is expanded around its vacuum expectation value (vev) in unitary gauge). Therefore, in the SM context, the quartic and trilinear Higgs couplings are directly related to the Higgs pole mass, the vev as set by, e.g., the $W$ mass, and the electroweak coupling.

When reconstructing the symmetry-breaking Higgs potential in a fully model-independent way by measuring the coefficients of the potential’s Taylor expansion about the symmetry breaking minimum $\langle H \rangle$, the impact of having finally observed the 125 GeV boson is rather limited. Discovering the Higgs with SM-compatible $W$ and $Z$ couplings does not provide any additional information other than the mere existence of a symmetry breaking vacuum (basically already known from observing massive gauge bosons) and the size of the curvature of the potential around the local minimum.* These are rather generic symmetry-breaking properties. Indeed, models which interpret the Higgs field as an iso-doublet of (pseudo-)Nambu-Goldstone fields generally involve Higgs self interactions significantly different from the SM. This typically has far-reaching phenomenological consequences for multi-Higgs production depending on the involved couplings and the presence of additional (often fermionic) matter. While details inevitably depend on the particular model, a measurement of the Higgs potential coefficient $\propto h^3$ undoubtedly sheds light on the source of EWSB by measuring the first quantity that reflects the dynamics of $V$. As a consequence, any knowledge about the trilinear Higgs coupling can be used to discriminate between various realisations of EWSB.

This has been the main motivation to study the LHC’s potential to reconstruct the Higgs trilinear coupling $\lambda_{SM} = \sqrt{\eta/2} m_h$ through measuring di-Higgs production.

*Interpreting the 125 GeV boson as a pseudo-dilation is formally the only exception to this argument with, however, typically little theoretical appeal and predictivity.
duction cross sections. Early analyses have revealed sensitivity to di-Higgs production in rare decays $hh \to b\bar{b}e\gamma$, which has been reviewed by ATLAS in Ref. \cite{12} only recently. The signal yield, however, is too small to tightly constrain the Higgs trilinear coupling in this channel. Hence, given the small production cross section of inclusive di-Higgs production at the LHC, it is imperative to apply and push state-of-the-art reconstruction and background rejection techniques for di-Higgs final states. For instance, the application of boosted $h \to b\bar{b}$ reconstruction and background rejection techniques as discussed in Ref. \cite{13} and used in $pp \to hZ + X$ analyses \cite{14} has also revealed a potentially large sensitivity to the trilinear coupling in the $b\bar{b}\tau^+\tau^-$ final states \cite{15,16,17}. In addition to new analysis strategies focussing on diverse phase space regions, it is also mandatory to extend the list of available hadron collider processes which can be included into a combined limit across various channels.

A process along this line which is at present of outstanding theoretical relevance is the production of a Higgs pair in association with two jets via weak boson fusion (WBF). This contribution to $pp \to hhjj + X$ production at the LHC is particularly interesting because the WBF component involves the quartic $VVhh$ ($V = W, Z$) vertices with couplings $g_{VVW hh} = e^2/(2s^2_w)$, $g_{ZZhh} = e^2/(2s^2_wc^2_w)$. The process obviously shares the QCD properties of single Higgs production via WBF, making (higher-order QCD) calculations straightforwardly adaptable from the latter process \cite{21,22}. However, a comprehensive signal vs. background investigation of the $hhjj$ final state and an analysis of the expected sensitivity to the $VVhh$ and trilinear Higgs couplings have not been performed so far. The purpose of the present work is to provide a first step in this direction.

One of the reasons for the lack of phenomenologically complete studies of this particular final state (apart from experimental issues that we are not qualified to comment on) is the highly involved modelling and up to now unknown size of the gluon fusion (GF) contribution to $pp \to hhjj + X$ at $\mathcal{O}(\alpha_s^2)$. On the one hand, the straightforward application of low-energy effective theorems to gluon-Higgs interactions \cite{24}:

$$
\mathcal{L}_{\text{eff}} = \frac{1}{4} \alpha_s \frac{1}{3\pi} G^a_{\mu\nu} G^{a\mu\nu} \log(1 + h/v),
$$

is fairly simple. On the other hand, however, since momentum transfers $p_{T,h} \sim m_t$ probe the kinematic region where interference with the Higgs trilinear diagrams becomes relevant for the integrated cross section \cite{17,22}, integrating out the top quark cannot be justified in phenomenological investigations. Destructive interference of the effective $gghh$ and $ggh$ vertices is encoded in the log's expansion of Eq. (2), yet all kinematic information is lost when employing the limit $m_t \to \infty$. Therefore, a targeted and reliable phenomenological analysis of the di-Higgs final state must not be based on Eq. (4).

Keeping the full quark mass dependencies in the gluon fusion component of $pp \to hhjj$ is a computationally intense task at the frontier of one-loop multi-leg calculations. Given the high complexity of this process we obtain a calculation time of up to $\sim 1$ minute per phase space point and per massive fermion in the loop for the pure gluon case $gg \to hhgg$, which exhibits the largest complexity with around one thousand diagrams (for details see below). Clearly, traditional Monte-Carlo event generation approaches do not promise a successful outcome unless the calculation time is significantly improved. In the following we choose to perform a phase space point-dependent re-weighting of the effective theory to overcome this predicament. This allows us to provide a first analysis of the $hhjj$ final state at the LHC. We also comment on the influence of modifications of the Higgs trilinear and $VVhh$ couplings on the resulting $pp \to hhjj + X$ phenomenology.

**Elements of the Analysis.** An apparent difference compared to single Higgs production studies in the two-jet category is the small cross section that is expected for $pp \to hhjj + X$ of inclusive $\mathcal{O}(10 \text{ fb})$. Typical GF and QCD background suppression tools for a 125 GeV Higgs boson such as e.g. a central jet veto (CJV) are not applicable because in order to observe a signal in the first place we have to rely on large Higgs branching ratios to bottom quarks (BR($h \to bb$) $\sim 58\%$), hadronically decaying Ws (BR($h \to W^+W^-$) $\sim 22\%$), and tau leptons (BR($h \to \tau^+\tau^-$) $\sim 6\%$). All these decay modes give rise to hadronic activity in the central detector region. (Semi-)leptonic Z boson decays are too limited by small branching ratios (BR($h \to ZZ$) $\sim 2.1\%$, BR($Z \to e^+e^- + \mu^+\mu^-$) $\sim 7\%$ and BR($Z \to h\text{adrons}$) $\sim 70\%$) to be of any phenomenological relevance in this case; e.g. the fully leptonic Higgs decay is clean but extremely rare BR($h \to 4\ell$) $\sim 7 \times 10^{-5}$. Hadronically-decaying Ws from a Higgs decay as considered in Ref. \cite{17} rely on extreme boosts of the recoiling Higgs decaying to bottoms and a resolution of the latter decay at a level comparable to the detector granularity. Such an approach does not seem promising in the present case due to the small cross section of $hhjj$ production and the larger systematic uncertainties of the multi-jet final state. Relaxing the CJV criterion in favour of a large invariant mass cut

\footnote{A measurement of the quartic Higgs interactions from triple Higgs final states appears impossible due to the tiny signal cross section \cite{13}.}

\footnote{Since then, analyses of (boosted) di-Higgs final states have gained at a lot of interest in the context of the SM and beyond \cite{14,15,16}. The Higgs boost is crucial to a successful analysis in the $b\bar{b}\tau^+\tau^-$ channel \cite{17}, inclusive analysis are dominated by the $t\bar{t}$ background \cite{21}.}

\footnote{The similarity of the WBF component with single Higgs production via WBF allows us to neglect interference effects of the two signal contributions \cite{24}.}
on the tagging jets \cite{26} is insufficient to tame the background contributions and is troubled by large combinatorial uncertainties and small statistics (see below). The most promising avenue is therefore a generalisation of the boosted final state analysis of Ref. \cite{13} to a lower \( p_T \) two-jet category: On the one hand, the signal cross section remains large by focussing on the \( hh \to b\bar{b}r^+\gamma^- \) final state and combinatorial issues can be avoided (i.e. through boosted kinematics and substructure techniques).

We generate signal events with \textsc{MadEvent} v4 \cite{28} and v5 \cite{29} for the WBF and GF contributions, respectively. The former event generation includes a straightforward add-on that allows to include the effect of modified Higgs trilinear coupling. The GF event generation employs the \textsc{FeynRules}/\textsc{Ufo} \cite{30} tool chain to implement the higher dimensional operators relevant for GF-induced \( hhjj \) production in the \( m_t \to \infty \) limit. We pass the events to \textsc{Herwig}++ \cite{31} for showering and hadronisation. For background samples we use \textsc{Sherpa} \cite{32} and \textsc{MadEvent} v5, considering \( tth, \, ttjj, \, ZWjj, \, ZHjj \) and \( ZZjj \). As in the \( hh \) and \( hhj \) cases the dominant background is due to \( t\bar{t} \). We normalise the background samples using NLO \( K \)-factors, namely 0.611 pb for \( tth \) \cite{33}, 300.5 pb for \( ttjj \) \cite{34}. We adopt the total flat \( K \) factor of 1.2 for \( Zh+2j \) motivated from Ref. \cite{35}. We have checked that all other backgrounds are completely negligible. The QCD corrections for the signal are known to be small for the WBF contribution \cite{21} \cite{22}. It is reasonable to expect that the corrections for the GF contributions will be similar to the \( pp \to hjj \) following the arguments of Ref. \cite{36}, however, we choose to remain conservative and do not include a NLO \( K \) factor guess for the GF contribution.

We correct the deficiencies of the GF event generation in the \( m_t \to \infty \) limit via an in-house re-weighting procedure (orory calculation), the effective theory overestimates the gluon fusion contribution by an order of magnitude.\footnote{However, it might be able to compensate this by folding in matrix elements to the analysis, generalizing the approach of Ref. \cite{27}.}

The maximum transverse momentum of the Higgs bosons is a good variable to compare effective with full theory. For inclusive \( hhjj \) production we find a re-weighted distribution as depicted in Fig. 1. Qualitatively, the re-weighting pattern follows the behaviour anticipated from \( pp \to hjj \) production \cite{13} and \( pp \to hjj \) \cite{26} \cite{47}. As expected, the shortcomings of the effective calculation for double Higgs production are more pronounced than for single Higgs production: Already for low momentum transfers the effective theory deviates from the full theory by factors of two, making the correction relevant even for low momenta, where one might expect the effective theory to be in reasonably good shape. It is precisely the competing and \( m_t \)-dependent contributions allowed to earlier which are not reflected in the effective theory causing this deviation. When the effective operators are probed at larger momentum transfers (and the massive quark loops are resolved in the full theory calculation), the effective theory overestimates the gluon fusion contribution by an order of magnitude.\footnote{A dedicated comparison of the full matrix element with the effec-}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{max\( p_{T,h} \) distribution and effective theory vs. full theory comparison as a function of the maximum Higgs transverse momentum of the fully showered and hadronised gluon fusion sample (satisfying the parton-level generator cuts \( p_{T,j} \geq 20 \text{ GeV} \) and \( |\eta_j| < 4.5 \)).}
\end{figure}
Due to the particular shape of the re-weighting in Fig. 1, we can always find a set of selection cuts for which effective theory and full calculation agree at the cross section level. Such an agreement, however, is purely accidental as it trades off a suppression against an excess in two distinct phase space regions. An effective field theoretic treatment of $hhjj$ production without performing the described re-weighting must never be trusted for neither inclusive nor more exclusive analyses.

In the hadron-level analysis we cluster jets from the final state using FASTJET [49] with $R = 0.4$ and $p_T \geq 25$ GeV and $|\eta_j| \leq 4.5$, and require at least two jets. We double $b$ tag the event (70% acceptance, 1% fake) and require the invariant mass of the $b$ jets to lie within 15 GeV of the Higgs mass of 125 GeV.

To keep matters transparent in the context of the highly involved $h \rightarrow \tau^+\tau^-$ reconstruction, we assume a perfect efficiency of 1 for demonstration purposes throughout.** We ask for two tau leptons that reproduce the Higgs mass of 125 GeV within ±25 GeV. The precise efficiencies for leptons in the busy hadronic environment of the considered process at a 14 TeV high luminosity are currently unknown, but we expect signal and background to be affect in similar fashion. We remind the reader that no additional requirements on missing energy or $m_{\tau\tau}$ are imposed, which are known to reconcile a smaller $\tau$ efficiency in the overall $S/B$.

The $b$ jets are removed from the event and jets that overlap with the above taus are not considered either. We require at least two additional jets which are termed “tagging jets” of the $hhjj$ event.

### Results

The cut flow of the outlined analysis can be found in Tab. I. There we also include analyses of signal samples with changed trilinear and $VVhh$ couplings. The latter modifications have to be interpreted with caution: The $VVhh$ couplings are purely electroweak and identical to the couplings of two Goldstone bosons to two gauge bosons. In the high energy limit the Goldstone equivalence theorem tells us that a modification of $VVhh$ away from its SM value is tantamount to unitarity violation, which explains the large growth of the WBF component for $\xi \neq 1$ (such an issue is not present for $\xi \neq 1$ even though the electroweak sector is ill-defined). The energy dependence of the matrix element is effectively cut-off by the parametric Bjorken-$x$ suppression of the parton distribution functions in the hadronic cross section. In models in which unitarising degrees are non-perturbative such a behavior is expected at least qualitatively. We leave an in depth theoretical discussion on approaches to parameterising such coupling deviations to future accords.

As can be seen from Tab. I the $hhjj$ analysis in the $b\bar{b}\tau^+\tau^-$ channel will be challenging. However, we remind the reader that no additional selection criteria have been employed that are known to improve $S/B$ in “ordinary” $hh \rightarrow bb\tau^+\tau^-$ analysis [15, 16]. The arguably straightforward strategy documented in Tab. I should rather be considered establishing a baseline for a more

### Table I: Cross sections in fb of the hadron-level analysis described in the text, including results with modified Higgs trilinear and $VVhh$ couplings. Signal cross sections already include the branching ratios to the $\tau\tau$ final states. The top four rows refer to the WBF sample and the last line includes the re-weighted GF contribution. For details see text.

| Event types | $\xi = 0$ | $\xi = 1$ | $\xi = 2$ | $t\bar{t}jj$ | Other BG | $S/B$ ratio to $\xi = 1$ |
|-------------|----------|----------|----------|-----------|---------|----------------|
| tau selection cuts | 0.212 | 0.091 | 0.100 | 3101.0 | 57.06 | $0.026 \times 10^{-3}$ |
| Higgs rec. from taus | 0.212 | 0.091 | 0.100 | 683.5 | 31.92 | $0.115 \times 10^{-3}$ |
| Higgs rec. from $b$ jets | 0.041 | 0.016 | 0.017 | 7.444 | 0.303 | $1.82 \times 10^{-3}$ |
| 2 tag jets | 0.024 | 0.010 | 0.012 | 5.284 | 0.236 | $1.65 \times 10^{-3}$ |
| incl. GF after cuts/re-weighting | 0.181 | 0.099 | 0.067 | 5.284 | 0.236 | 1/61.76 |
exhaustive investigation \cite{48} than the final verdict on $pp \to hhjj + X$ production.

The gluon fusion contribution dominates the signal component in the signal region, rendering the WBF contribution almost completely negligible for analysis with standard $VVhh$ coupling choices. The behaviour of the cross section as function of the Higgs trilinear interaction results from destructive interference as is anticipated for studies in $pp \to hh + X$ \cite{13, 22}.

With only about 30 expected WBF events in 3/ab, there is little leverage in the invariant dijet mass distribution to purify the selection towards WBF without jeopardising statistical power. On the other hand, depending on the mechanism of electroweak symmetry breaking, a large enhancement of the WBF contribution can outrun the dominant GF events. On a more positive note, if a trilinear Higgs coupling measurement is obtained from other channels such as $pp \to hh + X$, this information can in principle be used in the above analysis to obtain a confidence level interval for the quartic Higgs-gauge couplings in a simple hypothesis test.

A dedicated analysis which employs techniques motivated recently for di-Higgs final states \cite{16}, as well as methods to separate WBF from GF based on energy momentum flow observables and kinematic information \cite{50}, jet substructure \cite{13}, and/or matrix elements \cite{27}, is likely to significantly enhance S/B, especially when additional limiting factors of $b$ and tau tagging, smearing and trigger issues are treated more realistically. We are optimistic that this will eventually allow us to not only add $pp \to hhjj + X$ to the list of available di-Higgs final states even for a more realistic treatment of trigger issues, and $b$ and tau tagging efficiencies, but also provide an additional handle to measure Higgs trilinear and quartic Higgs-gauge couplings at a high luminosity LHC.

**Summary, Conclusions and Outlook.** Part of the electroweak physics agenda after the late Higgs discovery will be to phenomenologically reconstruct the symmetry breaking potential, as well as to precisely unravel the new particle’s role in TeV scale physics. Measurements of the Higgs trilinear and the quartic Higgs-gauge couplings are highly sensitive parameters in this context as they provide a clear picture of the Higgs sector dynamics and an independent cross-check of mechanism that enforces unitarity.

A lot of theoretical work has been devoted to the WBF contribution to $pp \to hhjj + X$, which is highly interesting for specifically these reasons and technically rather straightforward. Nonetheless, hardly any knowledge about the phenomenological relevance of this contribution as part of a full hadron-level analysis (including other signal sources as well as background estimate) has been gathered so far.

This letter summarises the beginning of a program which seeks to change this. We have presented a first complete and coherent phenomenological analysis of di-Higgs production in association with two jets. Exploiting the full bandwidth of state-of-the-art Monte Carlo tools, we have focused on what is probably the phenomenologically most attractive final state in terms of reconstruction potential, combinatorial limitations, relatively high signal yield and comparably large background rejection as a first step towards a more dedicated analysis. Indeed we find that WBF plays a completely subdominant role compared to GF, with little statistical handle to change this with traditional techniques even at high luminosity.

Also, we have showed that, independent of the particular phase space region that dedicated analysis targets, a reliable modelling of the signal crucially depends on the realistic generation of the gluon fusion signal contribution. Gluon fusion must not be based on effective field theory methods without applying a proper fully-differential correction procedure. To this end we have developed a stand-alone library based on MADGRAPH and GoSam, that implements a phase space point-dependent re-weighting procedure of the effective theory calculation, keeping all top and bottom mass dependencies.

The results indicate that such an analysis at the LHC will be challenging but not hopeless. In particular, recent developments in the context of multi-Higgs production have not been exploited in the present article. We leave this to future work \cite{48}.

**Acknowledgements.** We thank the organisers of 2013 Les Houches workshop for convincing us of the relevance and necessity of the presented work. MJD, CE, and MS thank Margarete Mühlleitner and Michael Spira for interesting conversations during Les Houches 2013. MJD, CE and MS also thank Mike Johnson, Peter Richardson, and Ewan Steele for computing support. NG wants to thank the other members of the GoSam collaboration for various useful discussions. CE is supported by the Institute for Particle Physics Phenomenology Associateship programme.

\cite{1} F. Englert and R. Brout, Phys. Rev. Lett. 13 (1964) 321; P. W. Higgs, Phys. Lett. 12 (1964) 132 and Phys. Rev. Lett. 13 (1964) 508; G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, Phys. Rev. Lett. 13 (1964) 585.
\cite{2} G. Aad et al. [ATLAS Collaboration], Phys. Lett. B 716 (2012) 1. S. Chatrchyan et al. [CMS Collaboration], Phys. Lett. B 716 (2012) 30.
\cite{3} S. Chatrchyan et al. [CMS Collaboration], JHEP 06 (2013) 081; G. Aad et al. [ATLAS Collaboration], arXiv:1307.1427 [hep-ex].
\cite{4} see e.g. G. F. Giudice, C. Grojean, A. Pomarol and R. Rattazzi, JHEP 0706 (2007) 045.
\cite{5} V. Barger, T. Han, P. Langacker, B. McElrath and P. Zerwas, Phys. Rev. D 67 (2003) 115001; T. Binoth, S. Karg, N. Kauer and R. Ruckl, Phys. Rev. D 74, 113008 (2006);
[hep-ph]. G. Cullen, H. van Deurzen, N. Greiner, G. Luisoni, P. Mastrolia, E. Mirabella, G. Ossola and T. Peraro et al., arXiv:1307.4737 [hep-ph]. T. Gehrmann, N. Greiner and G. Heinrich, JHEP 1306, 058 (2013)
H. van Deurzen, N. Greiner, G. Luisoni, P. Mastrolia, E. Mirabella, G. Ossola, T. Peraro and J. F. von Soden-Fraunhofen et al., Phys. Lett. B 721, 74 (2013)
[47] a numerical implementation of both cases is available in K. Arnold, M. Bahr, G. Bozzi, F. Campanario, C. Englert, T. Figy, N. Greiner and C. Hackstein et al., Comput. Phys. Commun. 180 (2009) 1661.
[48] M. J. Dolan, C. Englert, N. Greiner and M. Spannowsky, in preparation.
[49] M. Cacciari, G. P. Salam and G. Soyez, Eur. Phys. J. C 72 (2012) 1896.
[50] C. Englert, D. Goncalves, G. Nail and M. Spannowsky, Phys. Rev. D 88 (2013) 013016.