Production of $hhjj$ at the LHC

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

1Theory Group, SLAC National Accelerator Laboratory, Menlo Park, California 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, Durham DH1 3LE, United Kingdom

(Received 19 October 2013; published 10 March 2014)

Until now, a phenomenologically complete analysis of the $hh+2j$ channel at the LHC 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 $m_h \to \infty$ limit. In this Letter, 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.

These facts are the main motivation to study the LHC’s potential to reconstruct the Higgs trilinear coupling $\lambda_{3\text{SM}} = \sqrt{\eta/2m_h}$ through measuring di-Higgs production cross sections [4,5]. (A measurement of the quartic Higgs interactions from triple Higgs final states appears impossible due to the tiny signal cross section [6].) Given the small production cross section of inclusive di-Higgs production at the LHC, it is imperative to apply state-of-the-art reconstruction and background rejection techniques for di-Higgs final states. For instance, the use of boosted $h \to b\bar{b}$ reconstruction techniques as discussed in Ref. [7] has revealed a potentially large sensitivity to the trilinear coupling in the $b\bar{b}\tau^+\tau^-$ final states [8,9]. In addition to these new analysis strategies focusing 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 Higgs decay channels [10].

Nonetheless, analyses of $pp \to hh+X$ do not fully constrain the Higgs sector when thought of as a general effective theory: Even when the trilinear Higgs interactions are known, we do not have information about the $WWhh$ and $ZZhh$ couplings in Eq. (1), which are as important for naturalness considerations as the Higgs boson itself.

Therefore, Higgs pair production in association with two jets via weak boson fusion (WBF) $O(a^3)$ is of outstanding theoretical relevance. This contribution to $pp \to hhhj + X$ production at the LHC is particularly interesting, because the WBF component involves the quartic $VV^*hh$ ($V = W^\pm, Z$) vertices. (At a hypothetical linear collider, measurements are rather straightforward [11], but it is entirely unclear if there is potential at the LHC.) This has motivated high precision QCD calculations [23] for the WBF component, although it remained entirely unclear whether these contributions are, in fact, relevant. Although it is theoretically imperative to discuss to what extent we

0031-9007/14/112(10)/101802(5) 101802-1 © 2014 American Physical Society
can access the $VV'hh$ couplings via direct measurements, a comprehensive signal vs background investigation of the $hhjj$ final state and an analysis of the expected sensitivity to the $VV'hh$ and trilinear Higgs couplings have not been performed so far. The purpose of this Letter is in not only providing the first complete and concise analysis of the $hhjj$ final state at the LHC ever, but we also show that previous work is largely irrelevant unless a large contribution from new physics is present. In this sense this work is not just a continuation of the earlier $pp \rightarrow hh$ program, as we give a realistic sensitivity estimate to new contributions not reflected in $pp \rightarrow hh$ analyses, putting existing work in a proper context.

One reason for the lack of phenomenologically complete studies of this particular final state is the highly involved modeling and (until now) unknown size of the gluon fusion (GF) contribution to $pp \rightarrow hhjj + X$ at $O(a_s^2 \alpha^2)$. While applying low-energy effective theorems to gluon-Higgs interactions [12] $\mathcal{L}_{\text{eff}} = \alpha_s / (12\pi) G_{\mu\nu} G^{\mu\nu}$ log$(1 + h/v)$ is fairly simple, 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 [8,13], so that integrating out the top quark cannot be justified in phenomenological investigations. Realistic modeling of these interference effects is at the heart of any attempt to extract the relevant couplings. A reliable targeted phenomenological analysis of the di-Higgs final state must therefore not be based on effective theory methods.

Keeping the full quark mass dependencies in the gluon fusion component of $pp \rightarrow hhjj$ is a computationally intense task at the frontier of one-loop multileg calculations. Given the high complexity of this process, we obtain a calculation time of up to $\sim 1$ min per phase space point and per massive fermion in the loop for the pure gluon case $gg \rightarrow hhgg$, which exhibits the largest complexity with around 1000 diagrams (details 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 perform a phase space point-dependent reweighting 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 present results on modifications of the Higgs trilinear and $VV'hh$ couplings on the resulting $pp \rightarrow hhjj + X$ 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 $pp \rightarrow hhjj + X$ of inclusive $O(10\text{pb})$. Typical GF and QCD background suppression tools for a 125 GeV Higgs boson such as, e.g., a central jet veto 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, hadronically decaying $W$'s, and tau leptons. 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 to be of any phenomenological relevance in this case.

Relaxing the central jet veto criterion in favor of a large invariant mass cut on the tagging jets [14] 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 generalization of the boosted final state analysis of Ref. [8] to a lower $p_T$ two-jet category: On the one hand, the signal cross section remains large by focusing on the $hh \rightarrow bb\tau^+\tau^-$ final state and combinatorial issues can be avoided (i.e., through boosted kinematics and substructure techniques).

We generate signal events with MADEVENT v4 [15] and v5 [16] for the WBF and GF contributions, respectively. The former event generation includes a straightforward add-on that allows us to include the effect of modified Higgs trilinear coupling. The GF event generation employs the FeynRules/UFO [17] tool chain to implement the higher dimensional operators relevant for GF-induced $hhjj$ production in the $m_t \rightarrow \infty$ limit. We pass the events to HERWIG++ [18] for showering and hadronization. For background samples we use SHERPA [19] and MADEVENT v5, considering $ttH$, $t\bar{t}jj$, $ZWWjj$, $ZHjj$, and $ZZjj$. As in the $hh$ and $hhjj$ cases, the dominant background is due to $t\bar{t}$. We normalize the background samples by using next-to-leading-order (NLO) $K$ factors, namely, 0.611 pb for $ttH$ [20] and 300.5 pb for $t\bar{t}jj$ [21]. We adopt a flat $K$ factor of 1.2 for $Zj + 2j$ motivated from Ref. [22]. 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 [23]. We 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 $m_t \rightarrow \infty$ limit via an in-house reweighting library which is called at run time of the analysis for the weighted events. Based on the unweighting efficiency of $pp \rightarrow W^+W^-jj$ as implemented in Ref. [24], we estimate a speed improvement of a factor of at least $10^3$ of our approach. (Similar reweighting techniques are used by the LHC experiments for calibrating Monte Carlo data against subsidiary measurements in control regions, e.g., for $W$-jets in $t\bar{t}$ analyses [25].) We include the effects of finite top and bottom quark masses, which are treated as complex parameters. The value of the Higgs trilinear coupling can be steered externally. For the generation of the matrix elements we used GoSam [26], a publicly available package for the automated generation of one-loop amplitudes. It is based on a Feynman diagrammatic approach using QGRAF [27] and FORM [28] for the diagram generation and SPINNEY [29], HAGGIES [30], and FORM to write an optimized FORTRAN output. The reduction of the one-loop amplitudes was done by using SAMURAI [31], which uses a $d$-dimensional integrand level.
decomposition based on unitarity methods [32]. The remaining scalar integrals have been evaluated by using OneLoop [33]. Alternatively, GoSam offers a reduction based on tensorial decomposition as contained in the GoLEM95 library [34]. The GoSam framework has been used recently for the calculation of signal and background processes important for Higgs boson searches at the LHC [35].

The maximum transverse momentum of the Higgs bosons is a good variable to compare effective with full theory. For inclusive $hhjj$ production we show a reweighted distribution in Fig. 1. Qualitatively, the reweighting pattern follows the behavior anticipated from $pp \rightarrow hhj$ production [8] and $pp \rightarrow hjj$ [14]. 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 $Higgs production: Already for low-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. (A dedicated comparison of the full matrix element with the effective theory is an interesting question in itself, which we save for a separate study [36].) Figure 1 also demonstrates that the phase space is well covered by the effective theory Monte Carlo implementation and phase space coverage does not result in an issue for our procedure (we find similar coverage for angular distributions).

Because of the particular shape of the reweighting 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 reweighting must never be trusted for either inclusive or more exclusive analyses.

In the hadron-level analysis we cluster jets from the final state by using FastJet [37], with $R = 0.4$, $p_T \geq 25$ 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 boson 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 find the tau leptons to be rather hard, which can be used to trigger the event via the two-tau trigger with little signal loss.) We ask for two tau leptons that reproduce the Higgs boson mass of 125 GeV within $\pm 25$ 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 the signal and background to be affected in a similar fashion. We remind the reader that no additional requirements on missing energy or $m_{T\tau}$ are imposed, which are known to reconcile a smaller $	au$ efficiency in the overall $S/B$ [9].

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 Table I. There we also include analyses of signal samples with changed trilinear and $VV'hh$ couplings.

As can be seen from Table I, the $hhjj$ analysis in the $b\bar{b}\tau^+\tau^- jj$ 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” $hh \rightarrow b\bar{b}\tau^+\tau^-$ analysis [8,9]. The arguably straightforward strategy documented in Table I should rather be considered as establishing a baseline for a more exhaustive investigation [36] than the final verdict on $pp \rightarrow hhjj + X$ production.

The gluon fusion contribution dominates the signal component in the signal region, rendering the WBF contribution almost completely negligible for analysis with
our 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 Letter. We leave this to future work [36].
We thank James Ferrando for helpful discussions. M. J. D., C. E., and M. S. thank Margarete Mühleitner and Michael Spira for interesting conversations during Les Houches. N. G. thanks the other members of the GoSam Collaboration for various useful discussions. C. E. is supported by the Institute for Particle Physics Phenomenology Associateship program.


