

Probing vector-like quark models with Higgs-boson pair production

Giacomo Cacciapaglia,^a Haiying Cai,^a Alexandra Carvalho,^b Aldo Deandrea,^{a,c}
 Thomas Flacke,^d Benjamin Fuks,^{c,e,f} Devdatta Majumder^g and Hua-Sheng Shao^{e,f}

^aUniv Lyon, Université Lyon 1, CNRS/IN2P3, IPNL,
 F-69622, Villeurbanne, France

^bDipartimento di Fisica e Astronomia and INFN, Sezione di Padova,
 Via Marzolo 8, I-35131 Padova, Italy

^cInstitut Universitaire de France,
 103 boulevard Saint-Michel, 75005 Paris, France

^dCenter for Theoretical Physics of the Universe, Institute for Basic Science (IBS),
 Daejeon 34051 Korea

^eSorbonne Universités, UPMC Univ. Paris 06, UMR 7589, LPTHE,
 F-75005 Paris, France

^fCNRS, UMR 7589, LPTHE,
 F-75005 Paris, France

^gDepartment of Physics and Astronomy, University of Kansas,
 1082 Malott, 1251 Wescoe Hall Dr., Lawrence, KS 66045, U.S.A.

E-mail: g.cacciapaglia@ipnl.in2p3.fr, hcai@ipnl.in2p3.fr,
alexandra.oliveira@cern.ch, deandrea@ipnl.in2p3.fr, flacke@ibs.re.kr,
fuks@lpthe.jussieu.fr, Devdatta.Majumder@cern.ch,
huasheng.shao@lpthe.jussieu.fr

ABSTRACT: We investigate Higgs-boson pair production at the LHC when the final state system arises from decays of vector-like quarks coupling to the Higgs boson and the Standard Model quarks. Our phenomenological study includes next-to-leading-order QCD corrections, which are important to guarantee accurate predictions, and focuses on a detailed analysis of a di-Higgs signal in the four b -jet channel. Whereas existing Run II CMS and ATLAS analyses are not specifically designed for probing non-resonant, vector-like-quark induced, di-Higgs production, we show that they nevertheless offer some potential for these modes. We then investigate the possibility of distinguishing between the various di-Higgs production mechanisms by exploiting the kinematic properties of the signal.

KEYWORDS: Heavy Quark Physics, Higgs Physics, Beyond Standard Model, Technicolor and Composite Models

ARXIV EPRINT: [1703.10614](https://arxiv.org/abs/1703.10614)

Contents

1	Introduction	1
2	Theoretical framework	2
2.1	Model-independent implementation	2
2.2	Concrete examples	4
2.3	Di-Higgs production at the LHC	7
3	LHC searches at 13 TeV	10
3.1	Simulation setup	11
3.2	Resolved analyses	12
3.3	Boosted analyses	13
3.4	Efficiencies and expectations	14
4	Characterising the signal	16
5	Conclusions	19

1 Introduction

Vector-like quarks (VLQ), or quarks whose left-handed and right-handed components lie in the same representation of the Standard Model (SM) gauge symmetry group, are under strong scrutiny at the LHC due to their relevance in many extensions of the SM. They appear, for instance, in models featuring extra space-time dimensions, an extended gauge symmetry, or a composite Higgs sector [1–5]. VLQs in general mix with all three generations of SM quarks, and the mixing with the first generation can in particular be quite relevant for various new physics production mechanisms, even though such a mixing is constrained by low-energy experimental data. On the other hand, the structure of the theories containing vector-like quarks enables a wealth of options for describing the decays of these quarks into the SM particles, and the channel in which the VLQ decays into a Higgs boson often plays an important role. For instance, if the model particle content exhibits more than one VLQ multiplet that largely mixes with the SM quark sector, an almost exclusive decay into a Higgs boson (and not any other gauge boson) and an accompanying SM quark is possible [6, 7]. Such a possibility is especially realised in specific and well motivated new physics models containing two doublets of VLQs.

We focus in this work on the production of two Higgs bosons, or di-Higgs production, in the context of models featuring VLQs that always decay into a Higgs boson and a SM quark. Run II LHC searches for new physics in di-Higgs events have been performed by both the ATLAS and CMS collaborations, using final-state signatures made of four jets

issued from the fragmentation of bottom quarks, or b -jets, in particular when the pair of Higgs bosons could possibly originate from the decay of a heavier resonance [8–12]. Other signatures have also been considered, when, for instance, a resonantly produced di-Higgs system decays into a pair of b -jets and either a pair of taus [13, 14], a pair of photons [15, 16], or a pair of weak bosons [17], as well as into a pair of weak bosons and a diphoton system [18].

As the four- b -jet signature is associated with the largest branching ratio, we investigate the potential impact of the presence of VLQs on the searches by a study of di-Higgs production and decay into a final state containing four b -jets. The present study is meant to give an example of the potential of the di-Higgs mode for investigating physics beyond the SM, rather than being a complete analysis, which can only be done at the level of the LHC collaborations. Extra quarks can enhance the cross section [19, 20] related to the production of two Higgs bosons, the latter being produced either directly or through the VLQ decays. As a consequence, the LHC may be able to observe an excess and discover vector-like quarks, or to instead strongly constrain VLQ scenarios, by means of di-Higgs probes. We critically analyse existing bounds extracted from di-Higgs data, pointing out additional information which can be obtained from the present searches. We moreover discuss specific issues related to di-Higgs production in VLQ models, including the relevance of boosted and non-boosted topologies, as well as the relative strength of the QCD and electroweak (EW) production modes that could give rise to a di-Higgs system from the decay of VLQs. We in particular investigate the connection between the single and the pair production of VLQ particles. We extend our preliminary work performed in the context of the *2015 Les Houches Workshop on TeV Collider Physics* [19], and additionally include next-to-leading order (NLO) QCD effects in a study of scenarios in which the sole non-vanishing EW VLQ coupling involves a Higgs boson. Such effects are expected to be particularly important when the VLQs couple to the first generation quarks [19, 20].

This article is organised as follows. Section 2 contains a description of the effective framework that we employ for our VLQ study, and includes possibilities for linking it to ultraviolet-complete realisations. This section also investigates di-Higgs production in VLQ models, and presents a comprehensive analysis of the diverse production modes of two Higgs bosons in such models. In section 3, we detail our simulation setup and study the phenomenological consequences of our signal on existing LHC searches. In section 4, we compare and contrast the VLQ-induced QCD and EW production modes of a VLQ pair, and propose some variables that could be useful to distinguish them. We give our conclusions in section 5.

2 Theoretical framework

2.1 Model-independent implementation

We consider a simplified model where the SM is supplemented by new VLQs that only couple to the Higgs boson and the SM quarks, on top of the usual QCD gauge interactions. A model-independent parameterisation able to describe the main features of a generic VLQ was proposed in ref. [21]. Its main advantage consisted in the free coupling parameters that

were describing both the VLQ branching ratios and single production cross sections. This model has been recently extended so that QCD NLO effects can be included at the VLQ production level [20].

In the following, we use the above-mentioned NLO model after having turned off all unnecessary VLQ couplings to the EW gauge bosons. We restrict ourselves to the cases of a top-like heavy quark T (with an electric charge of $2/3$) of mass M_T and a bottom-like heavy quark B (with an electric charge of $-1/3$) of mass M_B , as any other extra quark would have an electric charge preventing it from coupling to the Higgs boson. The simplified Lagrangian that we consider is therefore given by

$$\begin{aligned} \mathcal{L}_{\text{VLQ}} = & i\bar{B}\not{D}B - M_B\bar{B}B + i\bar{T}\not{D}T - M_T\bar{T}T \\ & - h \left[\bar{B} \left(\hat{\kappa}_L^B P_L + \hat{\kappa}_R^B P_R \right) q_d + \bar{T} \left(\hat{\kappa}_L^T P_L + \hat{\kappa}_R^T P_R \right) q_u + \text{h.c.} \right], \end{aligned} \quad (2.1)$$

where the covariant derivatives only include QCD interactions, and P_L and P_R stand for the left-handed and right-handed chirality projectors. Although weak gauge couplings could have been also introduced, their contribution is always subleading, on top of being model-dependent, so that they have been ignored [21]. We have however retained the couplings of a single VLQ to the Higgs boson and a SM quark (SM flavour indices being understood) on the second line of the above equation. While the NLO implementation directly uses the $\hat{\kappa}$ parameters appearing in eq. (2.1) as free parameters, our phenomenological analysis relies on the conventions of ref. [21] where the couplings are fixed as

$$\hat{\kappa}_{L/R}^{T/B} = \sqrt{2} \frac{M_{T/B}}{v_{\text{SM}}} \kappa_{T/B} \sqrt{\frac{\zeta_{L/R}^i}{2\Gamma_h^0}}, \quad 2\Gamma_h^0 = \left(1 - \frac{M_h^2}{M_{T/B}^2} \right)^2 \sim 1, \quad (2.2)$$

v_{SM} being the SM Higgs vacuum expectation value and $\kappa_{T/B}$ a generic coupling strength. The rationale behind this normalisation choice is such that the parameter ζ^i is equal to the branching ratio associated with a T/B decay into a Higgs boson and the i^{th} generation SM quark. The kinematic factor Γ_h^0 only plays a role when a coupling to the Z and/or W boson is present, and it can be approximated to $\Gamma_h^0 \sim 1/2$ for the VLQ mass range of interest. The reason for the presence of a mass scaling in eq. (2.2) arises from the fact that $\hat{\kappa} v_{\text{SM}}/M_Q$ is proportional to the mixing angle between the VLQ and the SM quarks, which also affects the couplings of the SM quarks.

It can be proven, in general, that the mixing angles are chiral. In other words, the VLQ coupling to the Higgs boson and a specific chirality of SM fermion is suppressed with respect to the coupling to the other chirality by the mass of the SM fermion [21]. In any phenomenological analysis, one has thus the freedom to choose either left or right-handed VLQ couplings. The mixing between the extra quark and the SM quarks is however a constrained quantity. It is bounded by experimental data that includes precision measurements in the EW sector [22–24], flavour observables [22, 25] and the currently allowed room for deviations in the Higgs sector [26, 27]. These constraints on the mixing angles are approximately independent on the mass of the VLQ because they come from lower energy measurements. As an example, it was found that for a VLQ mixing to the first generation

only, as in the VLQ realisations considered in the rest of this paper, the bounds mainly stem from deviations of the Z -boson coupling to up quarks as measured in atomic parity violation experiments, and leads to $\kappa_T \lesssim 0.07$ [21]. Even though the mixing angle is in principle inversely proportional to the VLQ mass, we decided to fix its value independently on the VLQ mass. With this choice, it turns out to be easier to compare our findings with low energy bounds. The precise value of the coupling $\kappa_{T/B}$, however, crucially depends on the specific model.

As illustrative examples, we briefly describe in the rest of this section two scenarios where our model configuration featuring exclusive VLQ couplings to the Higgs boson and a light SM quark arises naturally. One of these examples concerns a weakly coupled theory and the other one a strongly coupled theory.

2.2 Concrete examples

The simplest example of a VLQ model featuring exclusive couplings to the Higgs boson and the first generation of SM quarks consists of an extension of the SM where two weak doublets of extra quarks are added to the SM. The first one, denoted by $Q_1 = (U_1, D_1)$, has a SM hypercharge of $Y_{SM} = 1/6$ and the other one, denoted by $Q_2 = (X_2, U_2)$, has an hypercharge $Y = Y_{SM} + 1 = 7/6$. Both doublets can couple to the SM fermions via Yukawa interactions with the Brout-Englert-Higgs field, so that the mass Lagrangian and the (physical) Higgs boson interaction terms read

$$-\mathcal{L} \supset M_1 \bar{U}_{1L} U_{1R} + M_2 \bar{U}_{2L} U_{2R} + \left(1 + \frac{h}{v_{SM}}\right) (y_1 \bar{U}_{1L} + y_2 \bar{U}_{2L}) \tilde{u}_R, \quad (2.3)$$

where we have neglected the Yukawa interactions of the first generation SM quark due to their smallness. In our notations, M_1 and M_2 denote the VLQ mass parameters, y_1 and y_2 the strengths of their interaction with the Higgs boson multiplied by the SM vacuum expectation value v_{SM} and \tilde{u}_R is the SM right-handed up-quark field.

The interest behind this model is that for degenerate VLQ masses, i.e. for the symmetric setup in which $M_1 = M_2 = M$ and $y_1 = y_2 = y$, the new physics impact on the Z -boson couplings to the SM quarks nicely cancels. Such cancellations can also be achieved in the more general case in which the mass parameters are different, but where the strength of the $Z\bar{u}u$ coupling and the electroweak precision observables agree with data at the price of introducing VLQ couplings both to the Z and Higgs bosons. The considered symmetric scenario has been first proposed in the context of Higgs-boson production [6], and then generalised [28, 29]. As a detailed study of the constraints on the free parameters can be found in ref. [24], we restrict ourself to a summary of the main features of the model.

The quark mass eigenbasis (T, T', u_R) is obtained by rotating the gauge eigenbasis (U_1, U_2, \tilde{u}_R) as follows,

$$\begin{aligned} \tilde{u}_R &= \sin \varphi_R T_R + \cos \varphi_R u_R, & U_{1R} + U_{2R} &= \sqrt{2}(\cos \varphi_R T_R - \sin \varphi_R u_R), \\ U_{1R} - U_{2R} &= \sqrt{2}T'_R, & U_{1L} + U_{2L} &= \sqrt{2}T_L, & U_{1L} - U_{2L} &= \sqrt{2}T'_L, \end{aligned} \quad (2.4)$$

where u_R is the SM (massless) quark field, T/T' are the VLQ mass eigenstates and φ_R is the mixing angle of the heavy quark with the SM quark. The physical masses are given by

$$M_T = \frac{M}{\cos \varphi_R} \quad \text{and} \quad M_{T'} = M = M_T \cos \varphi_R, \quad \text{with} \quad \tan \varphi_R = \frac{\sqrt{2}y}{M}. \quad (2.5)$$

After these redefinitions, it is found that the heavier state T couples to the SM up-type quark exclusively via the Higgs boson, with an interaction strength that reads

$$\hat{\kappa}_R^T = \frac{M_T}{v} \cos \varphi_R \sin \varphi_R. \quad (2.6)$$

We consequently derive from eq. (2.2),

$$\kappa_T \sim \frac{\cos \varphi_R \sin \varphi_R}{\sqrt{2}}, \quad (2.7)$$

in the notations of the simplified model that we have introduced in section 2.1. Moreover, the model contains three lighter and degenerate states, namely the T' up-type quark, a bottom-like heavy quark $B \equiv D_1$ and an $X \equiv X_2$ state with an electric charge of $5/3$. These three quarks couple to the SM quarks and the massive weak gauge bosons with a strength proportional to the sine of the mixing angle φ_R ,

$$\begin{aligned} T'-Z-u &\Rightarrow \tilde{\kappa}_R^{T'} = -\sin \varphi_R, \\ B-W^- -u &\Rightarrow \kappa_R^B = -\frac{\sin \varphi_R}{\sqrt{2}}, \\ X-W^+ -u &\Rightarrow \kappa_R^X = -\frac{\sin \varphi_R}{\sqrt{2}}, \end{aligned} \quad (2.8)$$

in the notations of ref. [21].

In order to assess the phenomenological viability of such a realisation, constraints on the masses and couplings of the light new resonances can be estimated from a CMS analysis that performed a search for single- and pair-produced light-flavour quark partners [30]. Considering the production of a pair of X and B quarks via strong interactions followed by an exclusive VLQ decay into a Wq system, the resulting WW plus jets signature can be used to extract a bound on the masses of these two B and X states independently of the mixing angle φ_R .

We derive bounds by comparing theoretical predictions for the total production rate $\sigma_{pp \rightarrow BB,XX}$ to the CMS exclusion shown in ref. [30], we get an $M_B = M_X > 945$ GeV constraint. For masses above this limit, the most stringent and sole constraint arises from searches for singly-produced B quarks subsequently decaying into a Wq system that is produced by $u\bar{u}$ annihilation. The reason is twofold. First, the corresponding CMS search is charge-sensitive and specifically targets the presence of W^- bosons in the final state. Next, the other relevant single VLQ production option giving rise to W^- bosons involves an \bar{X} resonance. The latter can only be produced via $\bar{u}u$ annihilation that is suppressed by the parton densities, which renders the search ineffective. This motivates the development of a future search for W^+q VLQ decays, which could benefit from a parton-density-enhanced uu initial state.

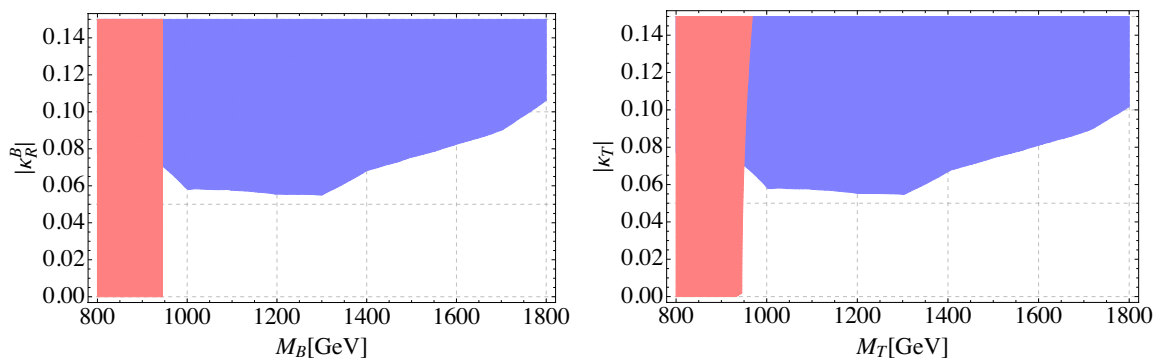


Figure 1. Existing constraints on the parameters of the simplest VLQ realisation featuring extra quarks coupling to the first generation of SM quark and the Higgs boson, as extracted from the CMS search of ref. [30]. The bounds are equivalently shown in the (κ_R^B, M_B) plane (left) and (κ_T, M_T) plane (right). The red area is excluded by the search for a strongly produced pair of VLQs, while the blue area shows the additional exclusion arising from the search for a singly-produced B -quark.

Making use of the results of ref. [30] and eq. (2.8), we derive a bound on the κ_R^B coupling that depends on the B quark mass (left panel of figure 1). In the right panel of figure 1, we re-express the results, via the mixing angle φ_R , in terms of the T coupling to the Higgs boson and the T -quark mass. The red area indicates the bounds derived from VLQ pair production, and the blue region reflects the single-VLQ production constraints. Further constraints could however arise from EW precision tests [24].

The second considered example of models originates from Composite Higgs setups with partial compositeness and an extended custodial symmetry [31]. Independently on the symmetry breaking pattern, the minimal top-partner content consists of a bi-doublet and a singlet field of extra quarks charged under the custodial $SO(4)$ symmetry. Both the bi-doublet and the singlet fields mix linearly with the SM elementary fields via two pre-Yukawa couplings that we denote y_L (for the bi-doublet) and y_R (for the singlet), while no SM Yukawa interaction is present. We choose a mixing with the first generation of SM quarks. As the small mass of the up quark is proportional to the product of the two y_L and y_R couplings, we assume that $y_L \ll y_R$ and hence decouple the bi-doublet from the rest of the spectrum [32].

In this limit, the mass and Yukawa interaction Lagrangian for the composite singlet field \tilde{U} reads

$$-\mathcal{L} \supset \bar{\tilde{U}}_L \left(M_1 \tilde{U}_R + y_R f \cos \epsilon \tilde{u}_R \right), \tag{2.9}$$

where \tilde{u}_R stands for the right-handed SM up-quark field and $\sin \epsilon = v_{SM}/f$ includes the effect of the EW symmetry breaking, f being the composite scale. A redefinition of the right-handed fields allows for the rewriting of the Lagrangian in the mass eigenbasis, the heavy quark mass M_T and mixing angle φ_1 being related to the Lagrangian parameters as

$$M_T = \sqrt{M_1^2 + y_R^2 f^2 \cos^2 \epsilon} = \frac{M_1}{\cos \varphi_1} \quad \text{with} \quad \tan \varphi_1 = \frac{y_R f \cos \epsilon}{M_1}. \tag{2.10}$$

As this redefinition only involves right-handed fields that are not sensitive to the weak interactions, no off-diagonal gauge interactions are generated. However, an off-diagonal coupling to the Higgs boson arises from the fact that the elementary right-handed up-quark couples to the composite object non-linearly, i.e. via a cosine function. The corresponding interaction term is in this case given by

$$\mathcal{L} \supset y_R \sin \epsilon h \bar{U}_L \tilde{u}_R = \frac{M_T}{v_{\text{SM}}} \frac{\sin^2 \epsilon \sin 2\varphi_1}{\cos \epsilon} h \bar{T}_L u_R + \dots + \text{h.c.} \quad (2.11)$$

Typically, the value of $\sin \epsilon$ is bounded to be small by EW precision tests, so that we consider a generic bound of $\sin^2 \epsilon \lesssim 0.1$ [24]. In the notations of eq. (2.2) and eq. (2.7), the coupling strength of the VLQ to the Higgs boson has therefore a natural upper bound of

$$\hat{\kappa}_R^T \lesssim \frac{M_T}{v_{\text{SM}}} \cdot 0.05 \quad \Rightarrow \quad \kappa_T \lesssim 0.035. \quad (2.12)$$

In the limit where the singlet field decouples from the spectrum, one can design an alternative composite realisation with the mixing of a bi-doublet to the SM elementary quark field. This can be mapped to the example introduced at the beginning of this section, with the equalities $M_1 = M_2$ and $y_1 = y_2$ being naturally enforced as the two doublets of the first example are now the components of a single multiplet of the custodial SO(4) symmetry [32].

2.3 Di-Higgs production at the LHC

Non-resonant di-Higgs production can result both from processes involving internal VLQ propagators, or from the production of a single VLQ (in association with a Higgs boson) or a pair of VLQs that then decay into a Higgs boson and a jet. We describe in this subsection all these relevant production modes in the context of VLQ setups featuring a VLQ coupling to the Higgs boson with an up-type quark exclusively. Di-Higgs states can be produced either alone or in association with jets, as illustrated on figure 2 where we show representative leading-order (LO) Feynman diagrams for the five different production mechanisms. In those diagrams, all external VLQs are understood to decay into a Higgs boson and a jet.

Di-Higgs production can be induced by the production of a pair of VLQs through strong interactions, as illustrated by the first diagram of figure 2. This mechanism, being only sensitive to gauge interactions, is independent of the value of the $\kappa_{T/B}$ parameters. In this case, the di-Higgs final state arises from two $Q \rightarrow qh$ decays and is thus produced in association with two additional jets, already at the LO accuracy. As mentioned in section 2.1, additional contributions could stem from EW gauge interactions, via, e.g., s -channel W/Z and photon exchanges. These model-dependent contributions are however typically small, and thus omitted.

Alternatively, the pair of VLQs could be produced electroweakly via diagrams that depend on $\kappa_{T/B}^2$, as illustrated by the second graph on figure 2. Although such contributions are naively thought to be much smaller than the QCD corresponding processes, VLQ carrying the same electric charge could be produced via such a mechanism, which benefits

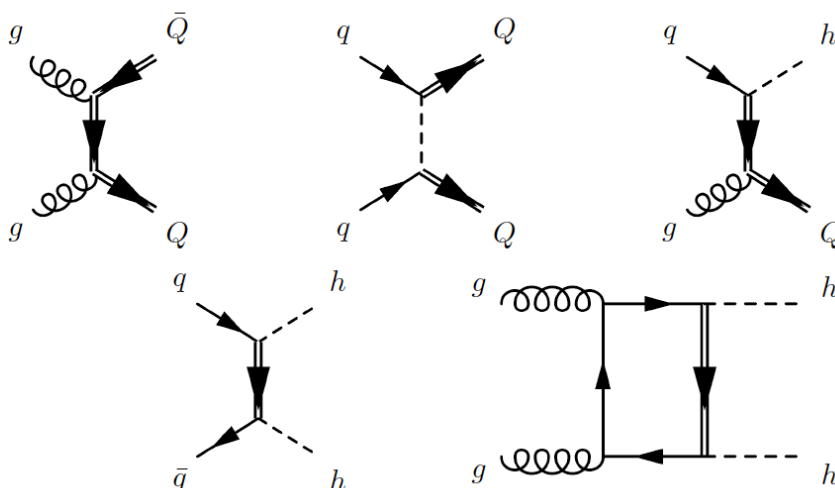


Figure 2. Sample LO Feynman diagrams for each of the five production channels described in the text and giving rise a VLQ-induced di-Higgs final states, with the decay $Q \rightarrow hq$ not explicitly indicated.

from an enhancement originating from the parton densities (PDF). We have furthermore verified that for the parameter space regions under consideration, interferences of EW and QCD contributions are always negligible.

A pair of Higgs boson can also be produced from the associated production of a single VLQ and a Higgs boson, the second Higgs boson arising from the VLQ decay. The di-Higgs system is thus produced together with an extra hard-scattering jet, and the corresponding amplitude, corresponding to the third Feynman diagram in figure 2, scales linearly with $\kappa_{T/B}$.

The di-Higgs system can also be produced directly, the VLQ only appearing as internal propagators. This is illustrated by the fourth and fifth diagrams of figure 2. The first of these diagrams is representative of di-Higgs production via a t -channel VLQ exchange and is proportional, at the level of the LO amplitude, to $\kappa_{T/B}^2$. In contrast, the second contribution is similar to the loop-induced SM production mode (that is included in the calculation), VLQs being additionally allowed to appear in the loop as studied in refs. [33–36].

In figure 3, we compare the total cross sections associated with these different di-Higgs production modes for $\kappa_{T/B} = 0.07$. For each production mechanism, we include LO (lighter colours) and NLO (dark colours) predictions in QCD, with the associated uncertainty band. For all subprocesses, NLO effects are found to largely enhance the rates and reduce the errors, as emphasized in ref. [20]. QCD-induced VLQ production contributions (in grey) are independent on the quark flavour and drop quickly with the VLQ mass. However, EW diagram contributions, that depend on the $\kappa_{T/B}$ coupling and are thus enhanced by the VLQ mass as shown by eq. (2.2), start to contribute for VLQ masses larger than about 1 TeV, as shown by the red bands that include all (i.e. both the QCD and EW components) VLQ-induced di-Higgs production processes $pp \rightarrow QQ, Q\bar{Q}, \bar{Q}\bar{Q}$. The effect is found more

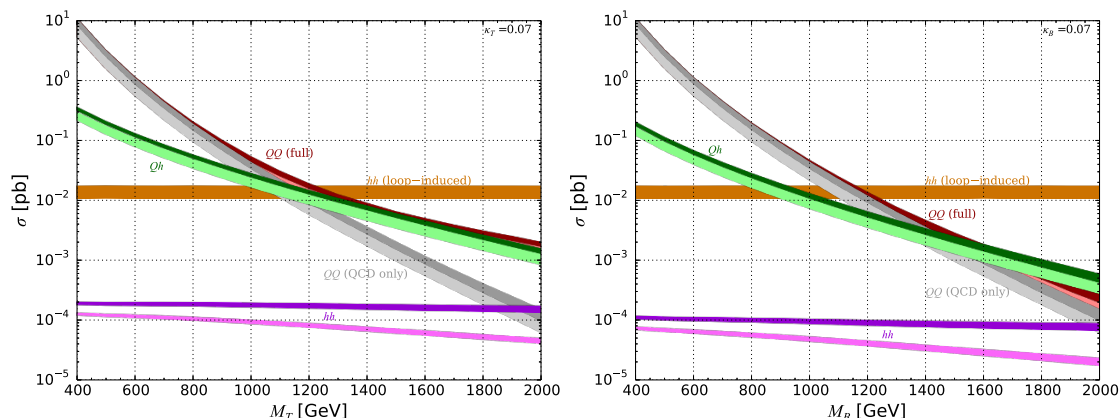


Figure 3. Di-Higgs production cross sections at the LHC, for proton-proton collisions at a centre-of-mass energy of 13 TeV and for a benchmark scenario in which $\kappa_{T/B} = 0.07$. The results are shown as a function of the VLQ mass and we consider up-type (left) and down-type (right) quark partners coupling to the first generation. A detailed explanation of the various curves is given in the text. LO results (light colours) and associated error bands can be compared to their NLO counterparts (darker colours).

dramatic for up-type quarks, which mainly results from the $pp \rightarrow QQ$ process that is PDF-enhanced as it could proceed via two up valence quarks.

The cross sections associated with Qh production are depicted by green bands, and benefits from a smaller phase space suppression for large VLQ masses than for VLQ pair production. For the considered benchmark scenario in which $\kappa_{T/B} = 0.07$, pair and single VLQ production yield cross sections of similar values when QCD contributions to VLQ pair-production become subleading, i.e. for $M_T \gtrsim 1$ TeV. The phase-space suppression associated with the EW contributions to VLQ pair-production is compensated by the coupling enhancement.

On the basis of the different cross section coupling scalings detailed above, Qh -induced di-Higgs production only dominates over the QQ mode for intermediate masses, EW pair production dominating instead for very large masses. This is further illustrated in figure 4 where total cross section contours, summed over all pair and single production modes, are shown in the (M_Q, κ_Q) plane. We distinguish the different regions of the parameter space in terms of the dominant di-Higgs production mode. Cases dominated by the QCD contributions to VLQ-pair production are shown in grey, by the EW contributions to VLQ pair production in red, and by Qh single production in green. The latter is, as expected, found to only dominate in the intermediate region where the VLQ mass is not too large (so that EW contributions are still subleading) and not too small (to prevent the strong contributions from dominating).

Loop-induced contributions to di-Higgs production by gluon fusion (last diagram of figure 2), as well as direct di-Higgs production via a t -channel VLQ exchange (fourth diagram of figure 2) are always negligible in the parameter space regions under consideration. The t -channel VLQ exchange is shown as violet bands in figure 3, and the gluon fusion loop-induced contributions are shown as a brown band. The former is subleading and

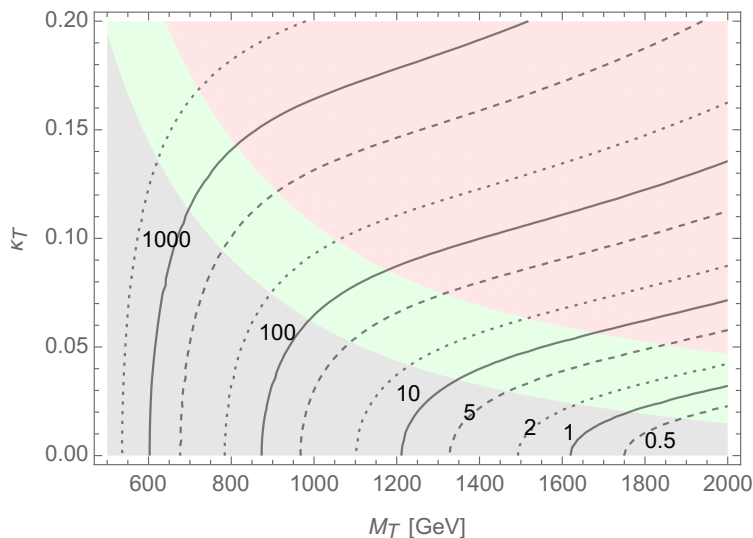


Figure 4. Di-Higgs (total) production cross section at the LHC shown as contours in the (M_Q, κ_Q) plane. The results, to be read in fb, contain both the single VLQ and VLQ-pair components and we represent as coloured regions the parameter space areas where a given production channel dominates (QCD-induced VLQ pair production in grey, EW-induced VLQ pair production in red and Qh -induced production in green). Direct (loop-induced and tree-level t -channel VLQ exchange) di-Higgs production has been neglected as the corresponding new physics contributions are always subleading for the parameter space regions under consideration.

hardly contributes, even for large masses for which all other modes start to be phase-space suppressed. This is still true even after accounting for the gigantic NLO K -factors due to the quark-gluon-initiated contributions that arise at NLO and that turn to dominate. The loop-induced results (brown) are dominated by the SM top-loop diagram and are thus not depending on the VLQ setup at all. These two direct di-Higgs production channels however yield Higgs bosons possessing a low transverse momentum p_T , so that the corresponding events are unlikely to populate the signal regions of the corresponding analyses that target high p_T Higgs boson pair-production (see section 3). We leave for future work the design of a specific study allowing to probe these production modes and their specific topology.

3 LHC searches at 13 TeV

Studying the pair-production of Higgs bosons at the LHC is an important topic of research. There exist a large number of dedicated ATLAS and CMS analyses at a centre-of-mass energy of $\sqrt{s} = 8$ and 13 TeV, and they differ by the final state they focus on. Various signatures have been considered, which include final states with two $h \rightarrow b\bar{b}$ decays [8–12], as well as those with one $h \rightarrow b\bar{b}$ and one $h \rightarrow \tau^+\tau^-$ decay [13, 14], one $h \rightarrow b\bar{b}$ and one $h \rightarrow \gamma\gamma$ decay [15, 16], one $h \rightarrow b\bar{b}$ and one $h \rightarrow W^+W^-$ decay [17], as well as one $h \rightarrow W^+W^-$ and one $h \rightarrow \gamma\gamma$ decay [18]. Here we focus on the ATLAS and CMS analyses using 13 TeV data in the four- b -jet final state and estimate the LHC sensitivity to the model

proposed in section 2.¹ This final state has the advantage of arising from Higgs boson decays with the largest branching fraction. In addition, b -jet tagging and, at high VLQ masses, jet substructure techniques make this channel promising and worthy to explore.

Both LHC collaborations have searched for new physics connected to di-Higgs production in cases where the Higgs-boson pair is resonantly produced from the decay of a heavier new particle, as well as when it is produced non-resonantly. Searches are typically organised into two classes depending on the p_T of the $h \rightarrow b\bar{b}$ system. The ‘resolved’ final state analyses are optimised for reconstructing low p_T Higgs boson pairs using four separate b -jets in the detector, whereas ‘boosted’ final state analyses target the high p_T $h \rightarrow b\bar{b}$ decays where the showered and hadronised b quarks are merged into one fat jet that is clustered using a larger distance parameter than a more conventional jet arising from a single hard parton.

In this section we use event selections inspired by the corresponding ATLAS [9] and CMS [11, 12] analyses at $\sqrt{s} = 13$ TeV to identify resolved and boosted $hh \rightarrow 4b$ configurations. As the experimental searches of refs. [9, 11, 12] are mainly aiming at the resonant production of a pair of Higgs bosons from a heavy new physics state decay, both Higgs bosons are studied in the same way, i.e. either both resolved or both boosted. The intermediate regime in which only one of the Higgs bosons would be boosted is not considered, although it may be interesting when the Higgs-boson system arises asymmetrically, as discussed in section 4.

We estimate the possible impact of the exotic production of a Higgs-boson pair through the decay of intermediate vector-like quarks. The simulations employed in this section are for an up-type vector-like quark generically denoted by Q . However, the VLQ flavour does not affect the signal topologies for the final states under consideration [19].

3.1 Simulation setup

For the simulation of the VLQ signal, we allow the VLQ mass M_Q to vary and be equal to 500 GeV, 650 GeV, 800 GeV, 1 TeV and 2 TeV. Signal simulation is performed within the MG5_aMC@NLO framework [38] where the entire event generation process is automated [39]. We make use of UFO model files [40] extracted from the Lagrangian presented in ref. [20] with the help of the FEYNRULES [41], NLOCT [42] and FEYNARTS [43] packages, as this Lagrangian embeds the model introduced in section 2. Hard scattering events are generated at the NLO accuracy in QCD, the virtual one-loop contributions being evaluated with the MADLOOP module [44] and then combined with the real contributions by means of the FKS subtraction method [45] as implemented in the MADFKS package [46]. VLQ decays are then handled automatically by means of the MADSPIN [47] and MADWIDTH [48] programs, which allow for retaining both off-shell and spin correlation effects. The fixed-order calculations, for which the NLO set of NNPDF 3.0 parton densities [49] is used, are matched with the parton shower and hadronization infrastructure of the PYTHIA 8 package [50], and we simulate the response of an LHC-like detector by using the DELPHES 3

¹For a recent study of the LHC high luminosity reach for SM di-Higgs searches in the four- b -jet final state we refer to ref. [37].

program [51] (using a CMS or an ATLAS description where relevant) that internally relies on the FASTJET package [52] for jet reconstruction.

3.2 Resolved analyses

We consider both an ATLAS-like and a CMS-like resolved di-Higgs boson analysis where all Higgs decay products can be entirely reconstructed [9, 11]. In the following, a *resolved jet* denotes a jet candidate reconstructed by means of the anti- k_t jet algorithm [53] with a distance parameter R set to

$$R = 0.4 \quad (\text{also known as an AK4-jet}). \quad (3.1)$$

We additionally impose that the jet transverse momentum p_T^j and pseudorapidity η^j satisfy

$$p_T^j > 20 \text{ GeV} \quad \text{and} \quad |\eta^j| < 2.5. \quad (3.2)$$

Jets are potentially tagged as b -jets with an efficiency extracted from the maps provided in ref. [54] and we additionally constrain the transverse momentum of all b -tagged jets to fulfil

$$p_T^b > 40 \text{ GeV} \quad (\text{ATLAS}) \quad \text{or} \quad p_T^b > 30 \text{ GeV} \quad (\text{CMS}) \quad (3.3)$$

in our ATLAS-like and CMS-like analysis, respectively. In both our resolved analyses, we select events that contain at least four resolved b -tagged jets,

$$N(b) \geq 4. \quad (3.4)$$

The four leading b -jets are then combined into two pairs of jets for which the angular distance in the transverse plane obeys

$$\Delta R(b, b) < 1.5, \quad (3.5)$$

each pair of b -jets being assumed to originate from the decay of a Higgs boson.

In the CMS-like analysis, we follow the *medium-mass region* selection [11] that has been designed to optimally probe resonantly produced di-Higgs systems where the resonance mass lies between 400 GeV and 1200 GeV. Denoting by M_{h_1} and M_{h_2} the invariant masses of the two reconstructed Higgs boson candidates, we select events for which these masses satisfy

$$\chi_{\text{CMS}}^2 = \left(\frac{M_{h_1} - \bar{M}_h}{\sigma_h} \right)^2 + \left(\frac{M_{h_2} - \bar{M}_h}{\sigma_h} \right)^2 < 1, \quad (3.6)$$

where \bar{M}_h is the average mean of the M_{h_1} and M_{h_2} distributions and is equal to 115 GeV. We moreover choose a width $\sigma_h = 23 \text{ GeV}$, as stemming from the CMS procedure for increasing the analysis sensitivity in the medium-mass region.

In contrast, our ATLAS-like analysis includes first a selection on the transverse momentum, pseudorapidity and invariant mass of the reconstructed Higgs bosons. The transverse momentum of the leading reconstructed Higgs-boson candidate $p_T(h_1)$ is constrained to fulfil

$$p_T(h_1) > \begin{cases} 200 \text{ GeV} & \text{for } m_{4j} < 600 \text{ GeV}, \\ 0.65 m_{4j} - 190 \text{ GeV} & \text{for } m_{4j} \in [600, 910] \text{ GeV}, \\ 400 \text{ GeV} & \text{for } m_{4j} > 910 \text{ GeV}, \end{cases} \quad (3.7)$$

where m_{4j} is the invariant mass of the system made of the two reconstructed Higgs bosons (or equivalently of the four leading b -jets), while the transverse momentum of the subleading Higgs-boson candidate $p_T(h_2)$ must obey

$$p_T(h_2) > \begin{cases} 150 \text{ GeV} & \text{for } m_{4j} < 520 \text{ GeV}, \\ 0.23 m_{4j} + 30 \text{ GeV} & \text{for } m_{4j} \in [520, 990] \text{ GeV}, \\ 260 \text{ GeV} & \text{for } m_{4j} > 990 \text{ GeV}. \end{cases} \quad (3.8)$$

Moreover, the two reconstructed Higgs bosons are required to be not too separated in pseudorapidity,

$$|\Delta\eta(h_1, h_2)| < \begin{cases} 1 & \text{for } m_{4j} < 820 \text{ GeV}, \\ 0.0016 m_{4j} - 0.28 & \text{for } m_{4j} > 820 \text{ GeV}. \end{cases} \quad (3.9)$$

A final selection is imposed on the two masses of the reconstructed Higgs bosons M_{h_1} and M_{h_2} ,

$$\chi_{\text{ATLAS}}^2 = \left(\frac{M_{h_1} - \bar{M}'_h}{0.1 M_{h_1}} \right)^2 + \left(\frac{M_{h_2} - \bar{M}_h}{0.1 M_{h_2}} \right)^2 < 2.56, \quad (3.10)$$

with $\bar{M}'_h = 124 \text{ GeV}$.

3.3 Boosted analyses

We describe in this section our CMS-like and ATLAS-like boosted analysis of a potential new physics di-Higgs signal. We denote as a CMS-like ‘fat jet’ (J_h^C with C standing for CMS) any jet candidate reconstructed by means of the anti- k_t algorithm with a distance parameter fixed to

$$R = 0.8 \quad (\text{also known as an AK8-jet}). \quad (3.11)$$

The J_h^C candidate is tagged as a Higgs boson jet if it further satisfies

$$|\eta^{J_h^C}| < 2.4, \quad \tau_{21}^{J_h^C} < 0.6 \quad \text{and} \quad m_{\text{pruned}}^{J_h^C} \in [105, 135] \text{ GeV}, \quad (3.12)$$

where $\tau_{21}^{J_h^C}$ is the ratio of the $N = 2$ and $N = 1$ N -subjettiness variables [55] and m_{pruned}^j is the pruned jet mass [56]. The efficiency associated with the b -tagging of a Higgs fat jet is implemented by assuming that the efficiency of tagging a subjet as a b -jet is the same as for a resolved jet with a transverse-momentum equal to half the fat jet transverse momentum. We have moreover optimistically assumed that any b -tagged fat jet contains two b -tagged subjets and we select events that feature at least two b -tagged Higgs fat jets B_h ,

$$N(B_h) \geq 2. \quad (3.13)$$

The two leading p_T CMS-like fat-jet are recognised as the two reconstructed Higgs bosons h_1 and h_2 , and we further enforce

$$|\eta(h_{1,2})| < 2.4, \quad p_T(h_{1,2}) > 200 \text{ GeV} \quad \text{and} \quad |\Delta\eta(h_1, h_2)| < 1.3. \quad (3.14)$$

The final discriminant variable in the CMS-like boosted analysis is the reduced mass

$$m_{\text{red}} = m_{4j} - \left(M_{h_1} - M_h \right) - \left(M_{h_2} - M_h \right) \quad (3.15)$$

with $M_h = 125$ GeV and m_{4j} abusively denoting the invariant mass of the di-Higgs system (for having consistent notations with the previous section). The reduced mass m_{red} is further imposed to be greater than 1 TeV.

For the ATLAS-like analysis, the ‘fat jet’ J_h^A (with A standing for ATLAS) is defined by once again using the anti- k_t jet algorithm but with this time a distance parameter set to

$$R = 1.0. \quad (3.16)$$

While CMS uses boosted jet algorithms based on particle-flow tracks [57, 58], the ATLAS collaboration reconstructs its fat jets (also called large- R jets [59]) from the information extracted from the topological clusters of the hadronic calorimeter. In our ATLAS-like boosted analysis, large- R jets are clustered within the FASTJET version embedded into DELPHES 3 and then trimmed [60], before we apply constraints on the invariant mass ($m^{J_h^A}$), pseudorapidity and transverse momentum of the two leading fat jets,

$$m^{J_h^A} > 50 \text{ GeV}, \quad |\eta(J_h^A)| < 2 \quad \text{and} \quad p_T(J_h^A) > 250 \text{ GeV}. \quad (3.17)$$

The signal region is defined by imposing extra constraints on the two leading fat jets that are identified as the two reconstructed Higgs bosons h_1 and h_2 ,

$$p_T(h_1) > 350 \text{ GeV}, \quad p_T(h_2) > 250 \text{ GeV} \quad \text{and} \quad |\Delta\eta(h_1, h_2)| < 1.7, \quad (3.18)$$

and by enforcing the χ_{ATLAS}^2 variable defined in eq. (3.10) to fulfill

$$\chi_{\text{ATLAS}}^2 < 2.56. \quad (3.19)$$

After these selection, we derive the invariant mass of the reconstructed Higgs bosons pair m_{2J} that is used in the ATLAS analysis, to characterise the signal.

3.4 Efficiencies and expectations

The analyses introduced in section 3.2 and section 3.3 are used to investigate the di-Higgs signal that arises from the presence of VLQs in the context of the model described in section 2. Figure 5 shows the different signal selection efficiencies that we present as a function of the VLQ mass. We consider both the CMS-like (solid lines) and the ATLAS-like (dashed lines) analyses, and both the boosted (right panel) and the resolved (left panel) selections. The different sets of curves correspond to the QCD (red) and EW (green) production of a VLQ pair (followed by two $Q \rightarrow hj$ decays), and to the associated production of a vector-like quark (that then decays into a Higgs boson plus jet system) with a Higgs boson (blue). In order to assess how well our analyses match what could be expected from the corresponding experimental analyses, we simulate a $pp \rightarrow G \rightarrow hh$ signal where G is a Kaluza-Klein graviton [61, 62] and compare our findings to the signal efficiencies presented in the original experimental publications. We obtain an agreement at the sub-percent level.

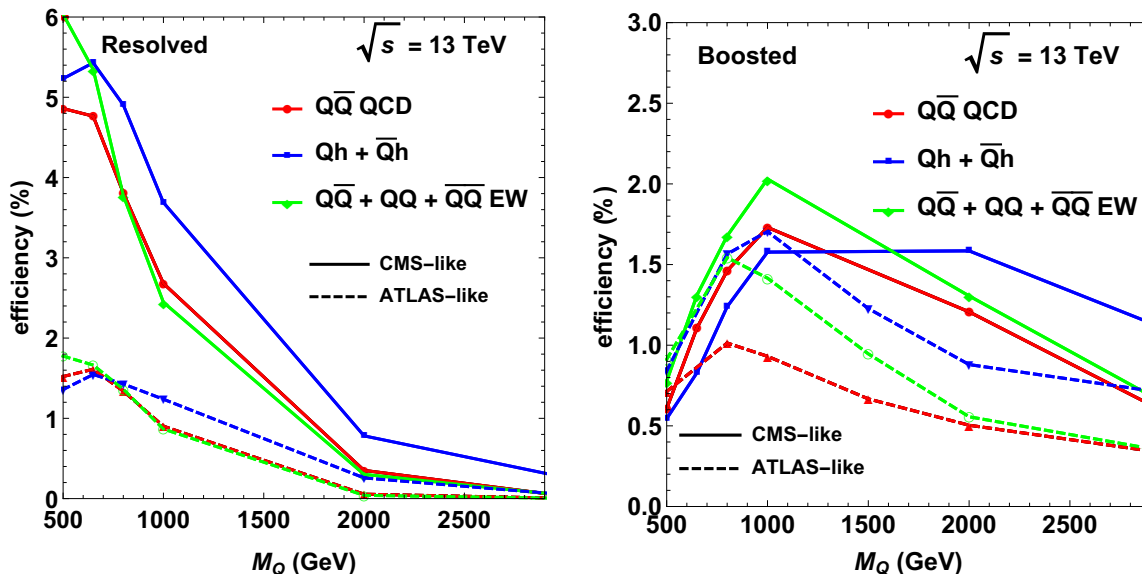


Figure 5. Selection efficiencies associated with the analyses introduced in section 3.2 and section 3.3 for the resolved (left panel) and boosted (right panel) cases. The CMS-like and ATLAS-like strategies are respectively depicted by solid and dotted lines, and the results are based on the (NLO-QCD accurate) simulation of VLQ-induced di-Higgs production in proton-proton collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV. We separately consider the QCD (red) and EW (green) production of a VLQ pair that then decays into a di-Higgs (plus two jets) system and the associated production of a single VLQ (that then decays into a Higgs boson plus jet system) with a Higgs boson (blue).

In the context of the resolved analysis, we observe that the efficiencies related to the QCD and electroweak VLQ pair production modes are almost identical for a given quark partner mass as long as the VLQ is not too light, so that the inclusion of the electroweak channels only modifies the total event rate. As expected, the efficiency is maximal for light VLQ masses, reaching up to 5% for the CMS-like analysis, and then decreases rapidly with increasing VLQ masses. The efficiencies associated with the ATLAS-like analysis are also found much lower, which is explained by the more severe selection strategy. For $M_Q > 800$ GeV, the single VLQ production channel leads to efficiencies that are higher than for pair production. The inclusion of this channel is therefore useful to assess the LHC sensitivity to VLQ models by means of di-Higgs probes more accurately. A stronger reach can hence be expected, as already suggested by the results shown in figure 4.

The efficiencies originating from the boosted analyses are in general smaller than for the resolved case, with a maximal values reached for $M_Q \sim 1$ TeV and a range spanning 1–2%. Genuine differences also appear when the di-Higgs system originates from QCD or EW VLQ pair production, as these two processes lead to a very different jet activity to which the boosted jet tagger is very sensitive to. We moreover observe smaller efficiencies for the ATLAS-like analysis due to the more severe selection strategy. These lower efficiencies are however not problematic as in the boosted regime, the background is also expected to be much more reduced than in a resolved context.

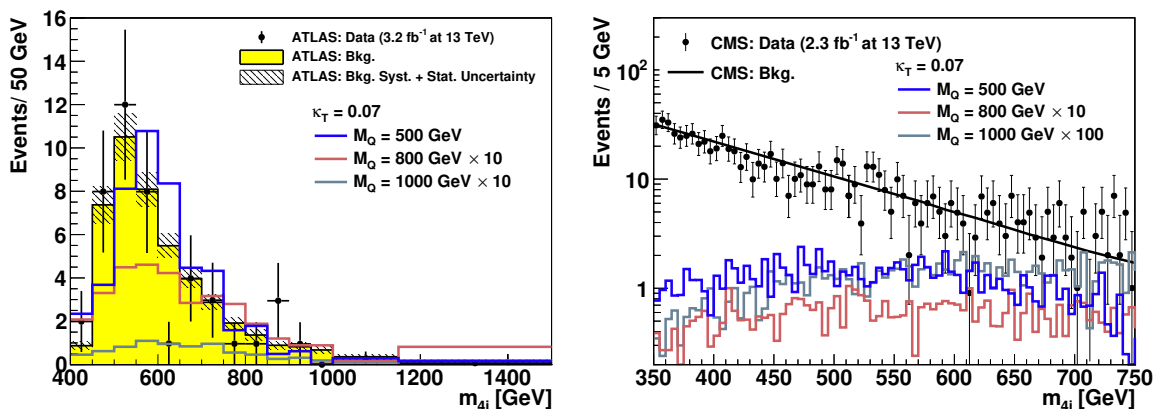


Figure 6. Distribution in the invariant mass of the reconstructed di-Higgs system both for the resolved ATLAS-like analysis (left) and CMS-like analysis (right). We select three benchmark setups with VLQ masses of 500 GeV (blue), 800 GeV (red) and 1 TeV (grey) and with $\kappa_T = 0.07$, and we overlay our predictions with the ATLAS [9] and CMS [11] data and background.

The four analyses that we have reimplemented have not been designed to target VLQ-induced di-Higgs production in the first place, as reflected in the comparison of the differential distributions in the analysis key observables with data illustrated in figure 6 in the resolved case. In this figure, we overlay our predictions for the spectrum in the reconstructed di-Higgs invariant-mass m_{4j} (built from the four leading jets) with the results obtained by the ATLAS (left panel) and CMS (right panel) collaborations in their resolved di-Higgs analysis of 3.2 fb^{-1} and 2.3 fb^{-1} of 13 TeV LHC data, respectively [9, 11]. In figure 7, we focus on the boosted analysis case and we respectively show the distribution in the invariant mass of the system made of the two boosted Higgs bosons, m_{2J} , in the case of the ATLAS-like boosted analysis (left panel) and the reduced mass m_{red} obtained in the context of the boosted CMS-like analysis (right panel) as defined in eq. (3.15). In both cases, the theoretical predictions and the data, respectively extracted from the ATLAS [8] and CMS [12] publications, are superimposed.

Although it may be challenging for a resolved analysis to be sensitive to VLQ-induced di-Higgs production for large VLQ masses, boosted analyses in principle offer extra handles for extending the sensitivity up to the TeV scale. The official numbers of data events and the predicted numbers of signal events for $\kappa_T = 0.07$ are given in table 1 for various VLQ masses. Comparing the magnitude of the yields, it turns out that ATLAS and CMS are already sensitive to a large fraction of the model parameter space with present data. This prevents us from requiring a specific design of a VLQ-dedicated di-Higgs search.

4 Characterising the signal

In this section, we study properties of the VLQ-induced di-Higgs signal that could be used to characterise it. Whereas the QCD production of a VLQ-pair (and thus of a di-Higgs system) is independent of κ_Q , the corresponding EW channel rate scales like κ_Q^4 , whilst the rate for the associated production of a VLQ with a Higgs boson scales like κ_Q^2 .

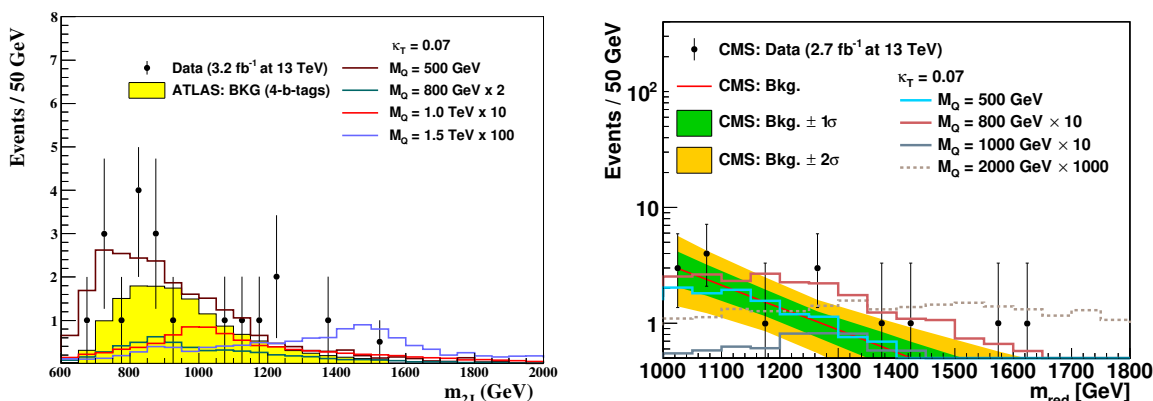


Figure 7. Distribution in the invariant mass of the system made of the two reconstructed boosted Higgs bosons after applying the boosted ATLAS-like analysis selection (left), and in the reduced mass resulting from the boosted CMS-like analysis (right). We select varied benchmark setups and overlay our predictions with the ATLAS [9] and CMS [12] data and backgrounds, respectively.

Analysis	Data	SM	Signals (for given VLQ masses)			
			0.5 TeV	0.8 TeV	1 TeV	2 TeV
ATLAS-resolved (3.2 fb^{-1})	44	47.6 ± 3.8	47.0	3.34	0.78	–
ATLAS-boosted (3.2 fb^{-1})	20	14.6 ± 2.4	23.5	2.82	0.933	0.024
CMS-resolved (2.3 fb^{-1})	797	n.a.	120	7.27	1.68	–
CMS-boosted (2.7 fb^{-1})	15	n.a.	17.6	2.99	1.10	0.04

Table 1. Number of data and predicted signal and background events for four VLQ masses of $M_Q = 500, 800, 1000,$ and 2000 GeV , as obtained in the four reinterpreted LHC analyses and for $\kappa_T = 0.07$.

The knowledge of the resolved or boosted regime is important, and we additionally consider a semi-boosted context where only one of the two Higgs bosons is boosted. This new analysis strategy is expected to be important in asymmetric cases where the transverse momenta of the two Higgs bosons are largely different, as it could happen from non-resonant Higgs-boson pair production. We therefore define, in this section, three categories that we denote by *boosted*, *semi-boosted* and *resolved*.

The fully boosted regime includes events where at least two Higgs fat jets are found after following the boosted CMS-like analysis of section 3.3 with the exception of the $\Delta\eta$ selection of eq. (3.14) that we omit. We tag as semi-boosted events in which only one Higgs fat jet is found, but that contains two additional b -jets that are consistent with a $h \rightarrow b\bar{b}$ decay as defined in the resolved CMS-like analysis of section 3.2. We further impose the two reconstructed Higgs bosons to satisfy eq. (3.6). Finally, the fully resolved category admits events with four well-identified b -jets as in the resolved CMS-like selection of section 3.2.

We analyse the properties of the two reconstructed Higgs bosons for VLQ masses of 500 GeV (left panel of the following figures), 1000 GeV (central panel of the following

figures) and 2000 GeV (right panel of the following figures). The results are summed over all the three above-mentioned categories. In figure 8, we present the distribution in the invariant mass of the reconstructed di-Higgs system. The spectra are indistinguishable for VLQ of mass equal to 500 GeV, as in this case the two Higgs bosons are mostly produced at rest. When one increases M_Q , the Higgs bosons start to become more boosted, which results in differences in the spectrum that depend on the production mode, the peak appearing at 2 TeV being related to a configuration where the two Higgs bosons are back-to-back.

More pronounced differences between the three different production channels can be observed by studying other kinematic variables like the transverse momentum of the di-Higgs system, while other observables exhibit very similar spectra. In all cases, it will not be easy to use the information to disambiguate the production modes. This is illustrated by figure 9 and figure 10 where we respectively show the distribution in the transverse momentum of the di-Higgs system and the separation in pseudorapidity between the two Higgs bosons. The p_T of the Higgs boson pair has a harder shape for the QCD pair production channel, while the difference in pseudorapidity between the two Higgs bosons features a softer shape for EW VLQ pair production. The jet properties are investigated in figure 11 where we investigate the AK4 jet multiplicity this time for the signal selected events. We present the fraction of selected events featuring n jets with n ranging from 0 to at least 5 jets as a function of the VLQ mass. Selected signal events arising from VLQ (QCD and EW) pair production feature most of the time between one and three extra jets, whereas di-Higgs production via a Qh pair in general leads to one or two extra jets only. The difference is however rather mild, so that jet vetoes would be poor handles to separate the different components of the signal.

Finally, we present in figure 12 the signal efficiencies as a function of the VLQ mass for the different analysis categories and the different components of the signal, and assess the impact of using NLO-accurate simulations (solid lines) instead LO-accurate simulations (dashed lines). While NLO effects are in general mild at the level of the efficiency for the EW processes (so that only the total rate turns to be modified), they are noticeable for the QCD production channel and can reach about 10%.

The results indicate that designing a semi-boosted category is useful as many events feature a single boosted Higgs boson and not a pair of them, in particular for VLQ masses of about 1 TeV. This effect is as expected more pronounced for Qh production where the two Higgs bosons are produced in an asymmetric fashion, one of them being directly produced and the second of them being originating from a VLQ decay. On different lines, the efficiency of the resolved analysis for what concerns EW VLQ-pair-induced di-Higgs production is very large in the low mass region, so that the inclusion of this channel can be really helpful to further constrain the low-mass region of the parameter space when the κ_Q coupling strength is not too small.

Our phenomenological analysis has moreover shown that the three different production channels have different kinematic properties so that further selection cuts could be implemented in the aim of distinguishing them once a signal is observed at the LHC. While no distribution seems sufficient by itself, the usage of a combination of all the available pieces of information simultaneously, via, e.g., a multivariate analysis technique, could

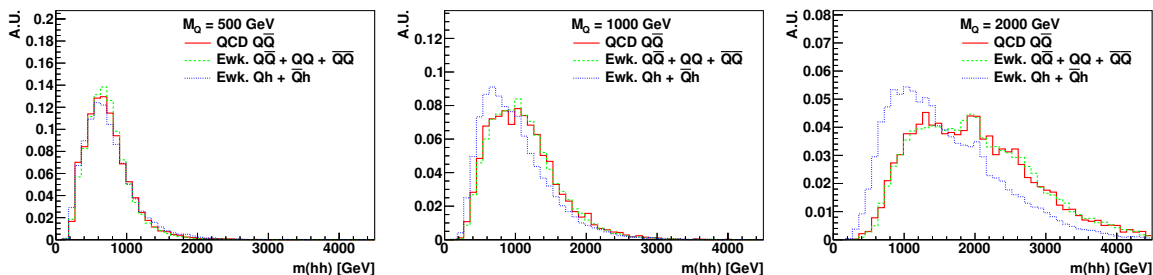


Figure 8. Normalised distribution, at the NLO accuracy in QCD, in the di-Higgs invariant mass in the context of VLQ-induced di-Higgs production. We distinguish QCD-induced (red) and EW-induced (green) VLQ-pair production (followed by to $Q \rightarrow hq$ decays) and the associated production of a single VLQ together with a Higgs boson (blue). We set the VLQ mass to 500 GeV (left panel), 1000 GeV (middle panel) and 2000 GeV (right panel).

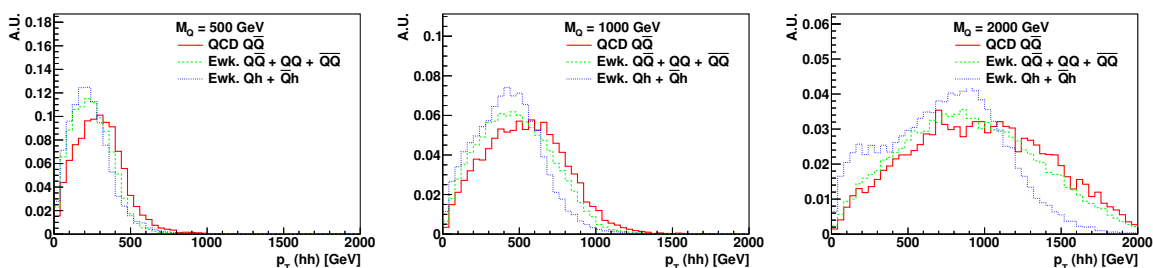


Figure 9. Same as figure 8 but for the transverse momentum of the di-Higgs system.

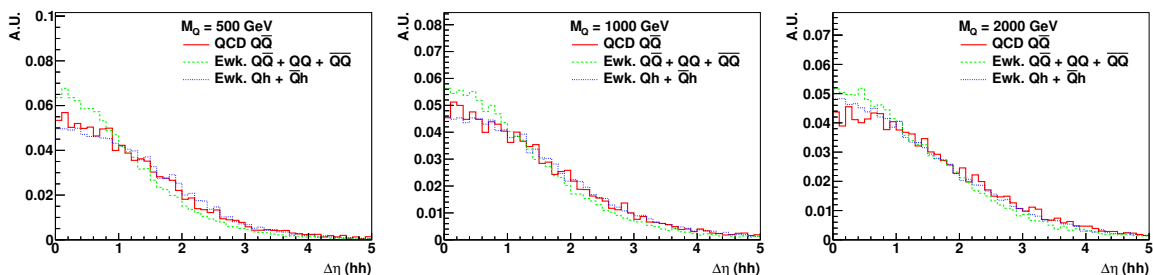


Figure 10. Same as figure 8 but for the separation in pseudorapidity of the di-Higgs system.

allow to achieve a sufficiently large discriminating power. Other channels and variables could nonetheless be useful for characterising any potential excess and disentangling the theoretical setups that could accommodate the excess [63].

5 Conclusions

Many extensions of the Standard Model feature one or more than one vector-like quark multiplets. The presence of several extra multiplets of quarks is motivated both by their impact on the predictions for various well-measured observables, which allows one to ease the corresponding experimental constraints, and by theoretical considerations. For instance, the combination of extra quark multiplets could tame down new physics effects on the couplings of the Z -boson to quarks, or could originate from the symmetry structure of a model

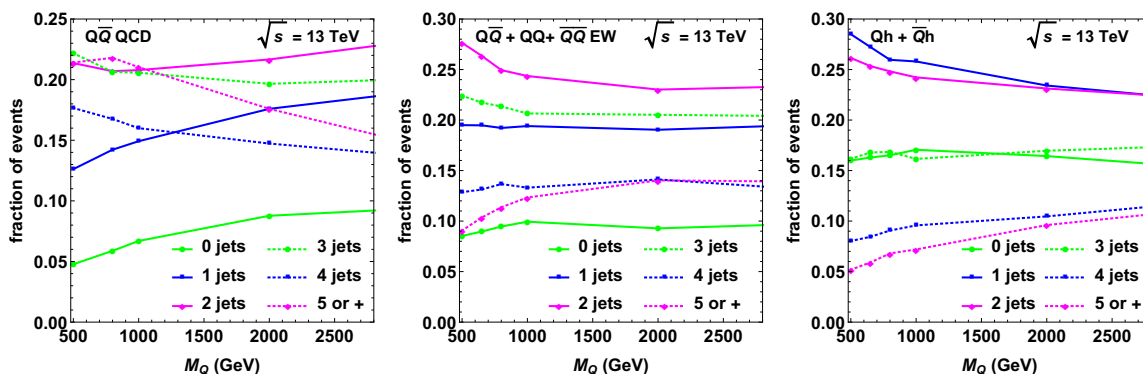


Figure 11. Fraction of selected signal events featuring a given number of AK4 jets that are not originating from a Higgs-boson decay. We separately consider QCD-induced (left panel) and EW-induced (middle panel) VLQ pair production, as well as associated Higgs and VLQ production (right panel).

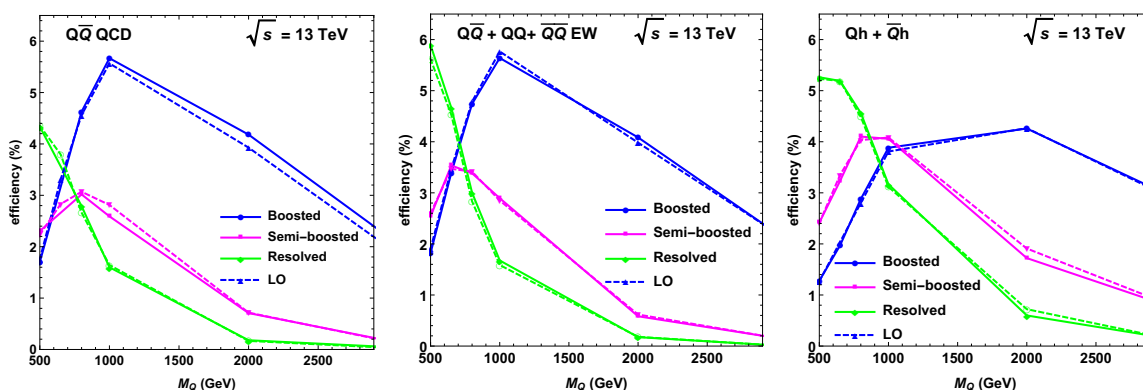


Figure 12. Signal selection efficiencies as a function of the VLQ mass for the boosted (blue), semi-boosted (magenta) and resolved (green) analysis categories and separated in terms of the different components of the signal, namely QCD (left panel) and EW (middle panel) VLQ-induced production and Qh (right panel) production. The solid lines refer to NLO simulations while the dashed lines refer to a LO simulation.

where the custodial symmetry is enforced. In such scenarios, the model may accommodate vector-like quarks in the low-energy part of the particle spectrum that dominantly couple to the Higgs boson and the SM quarks.

We have provided in the present study working examples of such a situation and developed a model-independent implementation in terms of an effective Lagrangian with parameters related to the physical observables, and we have connected this Lagrangian to two realistic ultraviolet-complete setups. We have further investigated the potential impact of di-Higgs probes for discovering or constraining the model, and we have detailed the variety of channels giving rise to a di-Higgs signature that could be mediated by (single and pair produced) vector-like quarks. Our predictions include next-to-leading order corrections in QCD and show their relevance not only in terms of cross-section values, but also for the increased precision with respect to the uncertainties.

We have furthermore considered the decay of the two Higgs bosons into four b quarks, and performed simulations corresponding in a close way to the existing Run II analyses performed by ATLAS and CMS both in the so-called resolved and boosted regimes. We have looked in details to the resulting signal efficiencies and to kinematical observables and their usage as handles on new physics. The current experimental analyses are mostly based on the study of resonant channels, which are not optimal in the case of a di-Higgs signal stemming from the interactions of vector-like quarks with the Higgs boson. Nonetheless, the Higgs bosons are typically sufficiently boosted to yield acceptable efficiencies for VLQ masses larger than 500 GeV.

Our study also allows to go to a further level of characterisation of a potential signal, using specific kinematical observables which can discriminate the VLQ-induced di-Higgs production mechanisms. The conclusion of such a study is that there is an opportunity of improvement with respect to the present bounds (or discovery) in the forthcoming LHC analyses, under the condition that a study tailored on the VLQ-induced di-Higgs channels is performed.

Acknowledgments

We thank the organisers of the 2015 ‘*Les Houches — Physics at TeV colliders*’ workshop where this work was initiated. GC, HC and AD acknowledge partial support from the Labex-LIO (Lyon Institute of Origins) under grant ANR-10-LABX-66 and FRAMA (FR3127, Fédération de Recherche ‘André Marie Ampère’). The work of HSS and BF is supported in part by French state funds managed by the Agence Nationale de la Recherche (ANR), in the context of the LABEX ILP (ANR-11-IDEX-0004-02, ANR-10-LABX-63), and the one of TF by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the ministry of Education, Science and Technology (No. 2013R1A1A1062597) and by IBS under the project code IBS-R018-D1. AC has been supported by the MIURFIRB RBF12H1MW grant. GC, HC, AD, BF and TF also acknowledge support from the France Korea Particle Physics LIA (FKPPL) and from the CNRS ‘Partenariat Hubert-Curien’ (PHC) STAR project number 34299VE. DM is supported by the National Science Foundation under award PHY-1607262.

Open Access. This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References

- [1] I. Antoniadis, *A possible new dimension at a few TeV*, *Phys. Lett. B* **246** (1990) 377 [[INSPIRE](#)].
- [2] D.B. Kaplan, *Flavor at SSC energies: A new mechanism for dynamically generated fermion masses*, *Nucl. Phys. B* **365** (1991) 259 [[INSPIRE](#)].
- [3] N. Arkani-Hamed, A.G. Cohen, E. Katz, A.E. Nelson, T. Gregoire and J.G. Wacker, *The minimal moose for a little Higgs*, *JHEP* **08** (2002) 021 [[hep-ph/0206020](#)] [[INSPIRE](#)].

- [4] R. Contino, L. Da Rold and A. Pomarol, *Light custodians in natural composite Higgs models*, *Phys. Rev. D* **75** (2007) 055014 [[hep-ph/0612048](#)] [[INSPIRE](#)].
- [5] O. Matsedonskyi, G. Panico and A. Wulzer, *Light Top Partners for a Light Composite Higgs*, *JHEP* **01** (2013) 164 [[arXiv:1204.6333](#)] [[INSPIRE](#)].
- [6] A. Atre, M. Chala and J. Santiago, *Searches for New Vector Like Quarks: Higgs Channels*, *JHEP* **05** (2013) 099 [[arXiv:1302.0270](#)] [[INSPIRE](#)].
- [7] A. Azatov et al., *Higgs boson production via vector-like top-partner decays: Diphoton or multilepton plus multijets channels at the LHC*, *Phys. Rev. D* **85** (2012) 115022 [[arXiv:1204.0455](#)] [[INSPIRE](#)].
- [8] ATLAS collaboration, *Search for pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, [ATLAS-CONF-2016-049](#).
- [9] ATLAS collaboration, *Search for pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Rev. D* **94** (2016) 052002 [[arXiv:1606.04782](#)] [[INSPIRE](#)].
- [10] CMS collaboration, *Search for non-resonant pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state with 13 TeV CMS data*, [CMS-PAS-HIG-16-026](#).
- [11] CMS collaboration, *Search for resonant pair production of Higgs bosons decaying to two bottom quark-antiquark pairs in proton-proton collisions at 13 TeV*, [CMS-PAS-HIG-16-002](#).
- [12] CMS collaboration, *Search for heavy resonances decaying to a pair of Higgs bosons in four b quark final state in proton-proton collisions at $\sqrt{s} = 13$ TeV*, [CMS-PAS-B2G-16-008](#).
- [13] CMS collaboration, *Search for resonant Higgs boson pair production in the $b\bar{b}\tau^+\tau^-$ final state using 2016 data*, [CMS-PAS-HIG-16-029](#).
- [14] CMS collaboration, *Search for non-resonant Higgs boson pair production in the $b\bar{b}\tau\tau$ final state using 2016 data*, [CMS-PAS-HIG-16-028](#).
- [15] ATLAS collaboration, *Search for Higgs boson pair production in the $b\bar{b}\gamma\gamma$ final state using pp collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector*, [ATLAS-CONF-2016-004](#).
- [16] CMS collaboration, *Search for $H(bb)H(\gamma\gamma)$ decays at 13 TeV*, [CMS-PAS-HIG-16-032](#).
- [17] CMS collaboration, *Search for Higgs boson pair production in the $b\bar{b}l\nu l\nu$ final state at $\sqrt{s} = 13$ TeV*, [CMS-PAS-HIG-16-024](#).
- [18] ATLAS collaboration, *Search for Higgs boson pair production in the final state of $\gamma\gamma WW^*(\rightarrow l\nu jj)$ using 13.3 fb^{-1} of pp collision data recorded at $\sqrt{s} = 13$ TeV with the ATLAS detector*, [ATLAS-CONF-2016-071](#).
- [19] G. Brooijmans et al., *Les Houches 2015: Physics at TeV colliders — new physics working group report*, [arXiv:1605.02684](#) [[INSPIRE](#)].
- [20] B. Fuks and H.-S. Shao, *QCD next-to-leading-order predictions matched to parton showers for vector-like quark models*, *Eur. Phys. J. C* **77** (2017) 135 [[arXiv:1610.04622](#)] [[INSPIRE](#)].
- [21] M. Buchkremer, G. Cacciapaglia, A. Deandrea and L. Panizzi, *Model Independent Framework for Searches of Top Partners*, *Nucl. Phys. B* **876** (2013) 376 [[arXiv:1305.4172](#)] [[INSPIRE](#)].
- [22] G. Cacciapaglia, A. Deandrea, L. Panizzi, N. Gaur, D. Harada and Y. Okada, *Heavy Vector-like Top Partners at the LHC and flavour constraints*, *JHEP* **03** (2012) 070 [[arXiv:1108.6329](#)] [[INSPIRE](#)].

- [23] S.A.R. Ellis, R.M. Godbole, S. Gopalakrishna and J.D. Wells, *Survey of vector-like fermion extensions of the Standard Model and their phenomenological implications*, *JHEP* **09** (2014) 130 [[arXiv:1404.4398](#)] [[INSPIRE](#)].
- [24] G. Cacciapaglia, A. Deandrea, N. Gaur, D. Harada, Y. Okada and L. Panizzi, *Interplay of vector-like top partner multiplets in a realistic mixing set-up*, *JHEP* **09** (2015) 012 [[arXiv:1502.00370](#)] [[INSPIRE](#)].
- [25] K. Ishiwata, Z. Ligeti and M.B. Wise, *New Vector-Like Fermions and Flavor Physics*, *JHEP* **10** (2015) 027 [[arXiv:1506.03484](#)] [[INSPIRE](#)].
- [26] S. Dawson and E. Furlan, *A Higgs Conundrum with Vector Fermions*, *Phys. Rev. D* **86** (2012) 015021 [[arXiv:1205.4733](#)] [[INSPIRE](#)].
- [27] N. Bizot and M. Frigerio, *Fermionic extensions of the Standard Model in light of the Higgs couplings*, *JHEP* **01** (2016) 036 [[arXiv:1508.01645](#)] [[INSPIRE](#)].
- [28] A. Atre, M. Carena, T. Han and J. Santiago, *Heavy Quarks Above the Top at the Tevatron*, *Phys. Rev. D* **79** (2009) 054018 [[arXiv:0806.3966](#)] [[INSPIRE](#)].
- [29] A. Atre et al., *Model-Independent Searches for New Quarks at the LHC*, *JHEP* **08** (2011) 080 [[arXiv:1102.1987](#)] [[INSPIRE](#)].
- [30] CMS collaboration, *Search for exotic light-flavor quark partners in pp collisions at $\sqrt{s} = 8$ TeV*, *CMS-PAS-B2G-12-016*.
- [31] K. Agashe, R. Contino, L. Da Rold and A. Pomarol, *A custodial symmetry for $Zb\bar{b}$* , *Phys. Lett. B* **641** (2006) 62 [[hep-ph/0605341](#)] [[INSPIRE](#)].
- [32] C. Delaunay, T. Flacke, J. Gonzalez-Fraile, S.J. Lee, G. Panico and G. Perez, *Light Non-degenerate Composite Partners at the LHC*, *JHEP* **02** (2014) 055 [[arXiv:1311.2072](#)] [[INSPIRE](#)].
- [33] C.O. Dib, R. Rosenfeld and A. Zerwekh, *Double Higgs production and quadratic divergence cancellation in little Higgs models with T parity*, *JHEP* **05** (2006) 074 [[hep-ph/0509179](#)] [[INSPIRE](#)].
- [34] S. Dawson, E. Furlan and I. Lewis, *Unravelling an extended quark sector through multiple Higgs production?*, *Phys. Rev. D* **87** (2013) 014007 [[arXiv:1210.6663](#)] [[INSPIRE](#)].
- [35] M.J. Dolan, C. Englert and M. Spannowsky, *New physics in LHC Higgs boson pair production*, *Phys. Rev. D* **87** (2013) 055002 [[arXiv:1210.8166](#)] [[INSPIRE](#)].
- [36] R. Grober, M. Muhlleitner and M. Spira, *Signs of Composite Higgs Pair Production at Next-to-Leading Order*, *JHEP* **06** (2016) 080 [[arXiv:1602.05851](#)] [[INSPIRE](#)].
- [37] 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](#)].
- [38] 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](#)].
- [39] N.D. Christensen et al., *A comprehensive approach to new physics simulations*, *Eur. Phys. J. C* **71** (2011) 1541 [[arXiv:0906.2474](#)] [[INSPIRE](#)].
- [40] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer and T. Reiter, *UFO — The Universal FeynRules Output*, *Comput. Phys. Commun.* **183** (2012) 1201 [[arXiv:1108.2040](#)] [[INSPIRE](#)].

- [41] A. Alloul, N.D. Christensen, C. Degrande, C. Duhr and B. Fuks, *FeynRules 2.0 — A complete toolbox for tree-level phenomenology*, *Comput. Phys. Commun.* **185** (2014) 2250 [[arXiv:1310.1921](#)] [[INSPIRE](#)].
- [42] C. Degrande, *Automatic evaluation of UV and R2 terms for beyond the Standard Model Lagrangians: a proof-of-principle*, *Comput. Phys. Commun.* **197** (2015) 239 [[arXiv:1406.3030](#)] [[INSPIRE](#)].
- [43] T. Hahn, *Generating Feynman diagrams and amplitudes with FeynArts 3*, *Comput. Phys. Commun.* **140** (2001) 418 [[hep-ph/0012260](#)] [[INSPIRE](#)].
- [44] V. Hirschi, R. Frederix, S. Frixione, M.V. Garzelli, F. Maltoni and R. Pittau, *Automation of one-loop QCD corrections*, *JHEP* **05** (2011) 044 [[arXiv:1103.0621](#)] [[INSPIRE](#)].
- [45] S. Frixione, Z. Kunszt and A. Signer, *Three jet cross-sections to next-to-leading order*, *Nucl. Phys.* **B 467** (1996) 399 [[hep-ph/9512328](#)] [[INSPIRE](#)].
- [46] R. Frederix, S. Frixione, F. Maltoni and T. Stelzer, *Automation of next-to-leading order computations in QCD: The FKS subtraction*, *JHEP* **10** (2009) 003 [[arXiv:0908.4272](#)] [[INSPIRE](#)].
- [47] P. Artoisenet, R. Frederix, O. Mattelaer and R. Rietkerk, *Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations*, *JHEP* **03** (2013) 015 [[arXiv:1212.3460](#)] [[INSPIRE](#)].
- [48] J. Alwall, C. Duhr, B. Fuks, O. Mattelaer, D.G. Öztürk and C.-H. Shen, *Computing decay rates for new physics theories with FeynRules and MadGraph5_aMC@NLO*, *Comput. Phys. Commun.* **197** (2015) 312 [[arXiv:1402.1178](#)] [[INSPIRE](#)].
- [49] NNPDF collaboration, R.D. Ball et al., *Parton distributions for the LHC Run II*, *JHEP* **04** (2015) 040 [[arXiv:1410.8849](#)] [[INSPIRE](#)].
- [50] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159 [[arXiv:1410.3012](#)] [[INSPIRE](#)].
- [51] 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](#)].
- [52] M. Cacciari, G.P. Salam and G. Soyez, *FastJet User Manual*, *Eur. Phys. J.* **C 72** (2012) 1896 [[arXiv:1111.6097](#)] [[INSPIRE](#)].
- [53] M. Cacciari, G.P. Salam and G. Soyez, *The anti- k_t jet clustering algorithm*, *JHEP* **04** (2008) 063 [[arXiv:0802.1189](#)] [[INSPIRE](#)].
- [54] CMS collaboration, *Identification of b-quark jets with the CMS experiment*, *2013 JINST* **8** P04013 [[arXiv:1211.4462](#)] [[INSPIRE](#)].
- [55] J. Thaler and K. Van Tilburg, *Identifying Boosted Objects with N-subjettiness*, *JHEP* **03** (2011) 015 [[arXiv:1011.2268](#)] [[INSPIRE](#)].
- [56] S.D. Ellis, C.K. Vermilion and J.R. Walsh, *Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches*, *Phys. Rev.* **D 81** (2010) 094023 [[arXiv:0912.0033](#)] [[INSPIRE](#)].
- [57] CMS collaboration, *Identification techniques for highly boosted W bosons that decay into hadrons*, *JHEP* **12** (2014) 017 [[arXiv:1410.4227](#)] [[INSPIRE](#)].

- [58] CMS collaboration, *Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus and MET*, [CMS-PAS-PFT-09-001](#).
- [59] ATLAS collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, [ATLAS-CONF-2014-018](#).
- [60] D. Krohn, J. Thaler and L.-T. Wang, *Jet Trimming*, *JHEP* **02** (2010) 084 [[arXiv:0912.1342](#)] [[INSPIRE](#)].
- [61] W.D. Goldberger and M.B. Wise, *Bulk fields in the Randall-Sundrum compactification scenario*, *Phys. Rev. D* **60** (1999) 107505 [[hep-ph/9907218](#)] [[INSPIRE](#)].
- [62] A. Oliveira, *Gravity particles from Warped Extra Dimensions, predictions for LHC*, [arXiv:1404.0102](#) [[INSPIRE](#)].
- [63] M. Backović, T. Flacke, J.H. Kim and S.J. Lee, *Boosted Event Topologies from TeV Scale Light Quark Composite Partners*, *JHEP* **04** (2015) 082 [[arXiv:1410.8131](#)] [[INSPIRE](#)].