Physics Letters B 756 (2016) 103–108

Contents lists available at ScienceDirect

Physics Letters B

www.elsevier.com/locate/physletb

Measuring the Higgs-bottom coupling in weak boson fusion

Christoph Englert [a,*], Olivier Mattelaer [b], Michael Spannowsky [b]

[a] SUPA, School of Physics and Astronomy, University of Glasgow, Glasgow, G12 8QQ, United Kingdom
[b] Institute for Particle Physics Phenomenology, Department of Physics, Durham University, DH1 3LE, United Kingdom

ARTICLE INFO

Article history:
Received 22 January 2016
Received in revised form 23 February 2016
Accepted 28 February 2016
Available online 4 March 2016
Editor: A. Ringwald

ABSTRACT

We study Higgs production through weak boson fusion with subsequent decay to bottom quarks. By combining jet substructure techniques and matrix element methods in different limits we motivate this channel as a probe of the bottom-Yukawa interactions in the boosted regime. In particular we ameliorate the "no-go" results of cut-and-count analyses in this channel. After applying a data-driven reconstruction approach we find that the Higgs-bottom coupling can be limited to $0.82 < y_b/y_b^{\mathrm{SM}} < 1.14$ with 600 fb^{-1}.

1. Introduction

After the Higgs discovery [1,2] and a growing consistency of Higgs measurements by ATLAS and CMS with the Standard Model (SM) hypothesis [3], diversifying and extending Higgs to all available production and decay channels is of utmost importance. On the one hand, this strategy will us allow to over-constrain fits to, e.g., the dimension six extension of the Higgs sector, or, on the other hand, could facilitate a new physics discovery in non-standard and less "traditional" Higgs search channels.

The coupling of the Higgs boson to bottom quarks is outstandingly important in this regard, because we expect a Higgs decay to bottom final states at around 60% [4]. Yet, none of the currently available analyses is directly sensitive to this coupling. Even the smallest deviation of the Higgs coupling to bottom quarks has far reaching consequences for the Higgs lifetime; a modification of which might, e.g., point to a possible relation of the TeV scale with a hidden sector. Since a modified Higgs phenomenology can arise from multiple sources, fingerprinting the bottom Yukawa interaction is mandatory to experimentally verify mass generation of the third generation down sector, especially because the standard ways of looking for Yukawa interactions such as Higgs production or bottom-quark associated Higgs production suffer either from dominant virtual top-quark contributions or a small total rate in light of a huge background.

In fact, there are only a few processes that contribute to a direct measurement of the bottom-Yukawa coupling: associated Higgs production [5–7] and top-associated Higgs production [8–10], with subsequent decay $H \to b\bar{b}$, both of which are challenging to probe at the LHC, even with large statistics.

It is the purpose of this work to add another sensitive channel to this list: weak boson fusion (WBF)-like Higgs production with decay to bottom quarks. This channel has been studied in Ref. [11], which quoted a very small signal vs. background ratio, effectively removing this process from the list of interesting Higgs processes. This is mainly due to large backgrounds and little handle (such as a missing central jet veto [12–14]) to control them. In this work we extend the analysis of [11] by employing novel reconstruction and all-information approaches through combining shower deconstruction [15,16], an all-order matrix element method to analyse fat jets, with the fix-order matrix element method techniques [17,18] for the hard process. We show that the large backgrounds can be significantly reduced, while a major part of the signal can be retained. This allows us to ameliorate the no-go expectation of "traditional" cut-and-count analyses for WBF Higgs production with b-quark final states.

This work is organised as follows. In Sec. 2, we comment on our event generation and the used analysis tools. Specifically, we review the matrix element method and shower deconstruction in Secs. 2.2 and 2.3 to make this work self-contained. Sec. 3 is devoted to our results. We perform a naive cut-and-count analysis and show that kinematic handles alone do not provide enough discriminating power to sufficiently isolate signal from background. We show that the latter can be achieved with a combination of matrix element method and shower deconstruction techniques, leading to an expected sensitivity to the SM WBF contribution with around 100 fb^{-1} luminosity. Sec. 4 provides a summary and gives our conclusions.

* Corresponding author.
E-mail addresses: christoph.englert@glasgow.ac.uk (C. Englert), o.p.c.mattelaer@durham.ac.uk (O. Mattelaer), michael.spannowsky@durham.ac.uk (M. Spannowsky).

http://dx.doi.org/10.1016/j.physletb.2016.02.074

2. Event generation and analysis tools

2.1. Event generation

We generated events at Leading Order in the four flavor scheme with MadGraph5_aMC@NLO [19–21]/Pythia8.2 [22] using NNPDF2.3 [23] for the parton distribution functions. The generation was split into five independent samples: two for the signal and three for the background. The two signal samples are the Higgs production in association with two light jets via either weak boson fusion (WBF) or via gluon fusion (GF) with the Higgs decaying into a $b\bar{b}$ pair. The gluon fusion process was generated via the new extension of MadGraph5_aMC@NLO supporting loop induced processes [24] and includes the full top and bottom mass effects. For the background we split the $b\bar{b}jj$ final state into pure QCD production (referred to as $b\bar{b}jj$) and electroweak production (referred to as Zjj). The last background sample is the four light-flavor jet sample ($jjjj$), for which we limit ourselves to the pure QCD contribution. To avoid the double counting between the $jjjj$ and the $b\bar{b}jj$ samples related to b emission in the parton-shower, we ran a four flavor parton-shower for the $jjjj$ sample.

At parton level a couple of loose cuts are applied in order to gain in efficiency. For all the samples with two b quarks and two light jets in the final state, we apply the following cuts:

$$\begin{aligned} p_{T,b} &\geq 20\text{ GeV}\,,\\ p_{T,j} &\geq 35\text{ GeV}\,,\\ y_{j_1}\cdot y_{j_2} &< 0\,,\\ |y_{j_1}-y_{j_2}| &> 3.0\,,\\ m_{j_1,j_2} &\geq 500\text{ GeV}\\ o_{T,(b+\bar{b})} &\geq 150\text{ GeV}\,,\\ \Delta R_{all,all} &\geq 0.2\,. \end{aligned}$$

For the $jjjj$ sample, the same cuts are applied with the index "j_1, j_2" being identified as the two most forward jets and the index "b" refers to the two central jets.

2.2. The matrix element method

The matrix element method [17,18] is based on the Neyman–Person Lemma [25] stating that the best discriminant variable is the likelihood ratio where the likelihood is the product of probabilities evaluated on the sample. The probability of an event is computed in the matrix element method by calculating

$$\mathcal{P}_\alpha(p^{exp}) = \frac{1}{\sigma}\int d\Phi(p^{part})|M_\alpha(p^{part})|^2 P(p^{exp}|p^{part}),$$

where p^{exp} represents the measured momenta for a given event, p^{part} is the partonic phase–space point which we integrate over with a phase–space measure $d\Phi(p^{part})$ that also includes the parton distribution functions. $|M_\alpha(p^{part})|^2$ is the matrix element square for a given hypothesis α and $P(p^{exp}|p^{part})$, named the transfer functions, is the conditional probability to observe the experimental event under consideration for a given partonic phase–space point.

Using the best discriminant variable allows us to perform measurements for processes with extremely small cross-section or small event rate. However, even if the above integral can be computed via dedicated tools [26], this is very CPU intensive when performed for the full sample of events. In order to analyse large background samples efficiently, we therefore simplify the method by approximating the transfer function by a delta function, allowing us to drop the computation of the integral entirely [15,27]. This is conservative, since including such effect can only improve the sensitivity of the method.[1]

Therefore, for each event, the matrix element method is equivalent to computing the following likelihood ratio:

$$\chi_{MEM} = \frac{|M_{wbf}|^2 + |M_{gf}|^2}{|M_{jjjj}|^2 + |M_{bbjj}|^2 + |M_{Zjj}|^2}. \tag{1}$$

For additional speed efficiency, the gluon fusion matrix element is not computed using the one loop matrix element – like we did for the event generation – but at tree level with an effective vertex coming for the integrating out the top quark loop [30]. The signal matrix elements (*GF* and *WBF*) are computed for a three body final state (with the Higgs momentum being identified with the reconstructed $\bar{b}b$-pair momentum) while the backgrounds are computed for the four particle final state using the tagged b subjet from the fat jet (see Sec. 3). To avoid potential bias, we use different sets of PDFs for the analysis (CT10 [31]) compared to the one used for the event generation.

2.3. Shower deconstruction

Shower deconstruction [15,16] is an all-order matrix element method designed to discriminate hadronically decaying electroweak-scale resonances, i.e. tops, W/Z or Higgs boson, from QCD jets. First the constituents of a fatjet are reclustered into small inclusive jets, e.g. using the k_T jet algorithm [32] with $R = 0.2$ and $p_{T,j} > 5$ GeV. One obtains a configuration of N subjets with four-momenta $\{p\}_N = \{p_1, p_2, \cdots,\}$. Using these subjets as input to the method, a likelihood ratio is calculated from first-principle QCD, quantifying whether the observed distribution of subjets was initiated by the decay of a signal process, e.g. a Higgs boson, or background, e.g. a gluon. To calculate the likelihood ratio

$$\chi_{SD}(\{p\}_N) = \frac{P(\{p\}_N|S)}{P(\{p\}_N|B)}, \tag{2}$$

where $P(\{p\}_N|S)$ represents the probability of obtaining the subjet distribution $\{p\}_N$ given the signal hypothesis, and $P(\{p\}_N|B)$ is the probability for obtaining the same $\{p\}_N$ from background processes. To calculate $P(\{p\}_N|B)$ and $P(\{p\}_N|S)$ the method sums over all possible shower histories. In [33] it was shown that χ_{SD} is insensitive to pileup and shows good agreement between data and Monte-Carlo prediction. We follow loosely the approach described in [34] to combine shower deconstruction with the fix-order matrix element method of Sec. 2.2.

3. Results

Based on the event generation detailed above, we first establish a baseline cut scenario inspired by [11,35].

In the first step, we veto events with isolated leptons with $|y_l| \leq 2.5$ and $p_{T,l} > 10$ GeV. We then request a $R = 1.2$ Cambridge–Aachen [36] fat jet[2] with

$$\begin{aligned} p_{T,j_{\text{fat}}} &> 200\text{ GeV}\,,\\ |y_{j,\text{fat}}| &< 2.5\,,\\ \text{and}\quad m_{j,\text{fat}} &> 90\text{ GeV}\,. \end{aligned} \tag{3}$$

After having identified a fat jet, we remove its constituents from the final state, and the remaining constituents in the event are

[1] Further improvements could be achieved by evaluating the matrix elements at NLO accuracy [28,29].

[2] Jet finding and clustering is performed with FastJet [37].