Citation for published item:

Further information on publisher’s website:
https://doi.org/10.1103/PhysRevD.98.095012

Publisher’s copyright statement:
Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI. Funded by SCOAP3.

Additional information:

Use policy
The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a link is made to the metadata record in DRO
- the full-text is not changed in any way
The full-text must not be sold in any format or medium without the formal permission of the copyright holders.
Please consult the full DRO policy for further details.
Probing electroweak precision physics via boosted Higgs-strahlung at the LHC

Shankha Banerjee,1,‡ Christoph Englert,2,† Rick S. Gupta,1,§ and Michael Spannowsky1,§

1Institute of Particle Physics Phenomenology, Durham University, Durham DH1 3LE, United Kingdom
2SUPA, School of Physics and Astronomy, University of Glasgow, Glasgow G12 8QQ, United Kingdom

(Received 26 July 2018; published 14 November 2018)

We study the process $pp \to Z(\ell^+\ell^-)h(b\bar{b})$ in the standard model effective field theory (SMEFT) at high energies using subjet techniques to reconstruct the Higgs boson. We show that at high energies this process probes four directions in the dimension-6 EFT space, namely the operators that contribute to the four contact interactions, $hZ_{\mu\nu}f\gamma^\mu f$, where $f = u_L, u_R, d_L$, and $d_R$. These four directions are, however, already constrained by the $Z$-pole and diboson measurements at LEP. We show that by utilizing the energy growth of this process in the SMEFT and the accuracy that can be achieved by using subjet techniques at the High Luminosity LHC, one can obtain bounds on these operators that are an order of magnitude better than existing LEP bounds.

DOI: 10.1103/PhysRevD.98.095012

I. INTRODUCTION

Characterizing the properties of the Higgs boson is arguably the most concrete particle physics goal of our time. This is further motivated by the dearth of any signs of physics beyond the standard model (BSM) in LHC data so far. One well-motivated course of action in this situation is to probe heavy new physics outside the reach of direct searches via precise indirect measurements. A historic example of constraining high energy physics even beyond the energy coverage of a collider is the LEP experiment, which was able to probe scales up to the few TeV via indirect precision measurements although it ran at a much smaller collision energy.

As the Higgs boson could not be produced before the LHC experiment under controlled conditions, one might naively think that any measurement of interactions involving the Higgs boson is complementary to past measurements. However, an effective field theory (EFT) perspective allows us to correlate measurements at different energy scales only on the basis of SM symmetry and matter content. In fact, there are fewer $SU(2)_L \times U(1)_Y$ invariant operators at the lowest order in the LHC-relevant EFT expansion at dimension 6 [1] than the number of (pseudo-)observables they contribute to. As a result, correlations between LHC and LEP measurements can be exploited. For instance the LEP measurements of $Z$-boson pole observables and anomalous triple gauge couplings (TGCs) inform the measurement of Higgs observables at the LHC as they can be related to a common gauge-invariant set of standard model effective field theory (SMEFT) operators. The analysis in Ref. [2] reveals, in fact, that apart from eight Higgs observables, the so-called “Higgs primaries,” all other Higgs interactions present in the dimension-6 Lagrangian can be constrained already by $Z$-pole and diboson measurements at LEP.

For the set of already constrained Higgs coupling deformations, the LHC has to compete with LEP’s precision to add new information in order to gain a more complete picture. This might seem challenging given that the LHC is intrinsically less accurate compared to LEP. The key advantage of the LHC (and other future colliders), however, is a much larger energy reach compared to LEP, thus allowing us to constrain new physics from a plethora of available kinematical information (see also [3–13]). As we will see, the high energy sensitivity of the LHC will allow us to strongly constrain EFT-induced anomalous couplings involved in processes that grow with energy,

$$\frac{\delta \sigma(\hat{s})}{\sigma_{SM}(\hat{s})} \sim \delta g_i \frac{\hat{s}}{m_Z^2}. \tag{1}$$

From Eq. (1), we see that the anomalous coupling $g_i$ can be measured/constrained at the per-mille to percent level even if the underlying sensitivity to the fractional cross-section deviation, $\delta \sigma/\sigma_{SM}$, is only $\mathcal{O}(30\%)$ at high energies.
The specific process we are interested in here is Higgs-strahlung, \( pp \rightarrow Z(\ell^+\ell^-)h(\bar{b}b) \). Studying the \( h \to bb \) mode instead of the \( h \to \gamma\gamma \) leads to a big enhancement in the rate but the Higgs-strahlung process still remains challenging with an \( O(50) \) background-to-signal ratio. Relating such a systemsatics limited result to the extraction of Higgs couplings can be at odds with the implicit assumption of perturbativity of the EFT expansion. We technically rely on the latter to perform proof-of-principle analyses and eventually full searches at ATLAS and CMS. As \( \delta\sigma/\sigma_{SM} \gtrsim 1 \) signals the breakdown of EFT validity for weakly coupled UV completions [14], a sensitivity to smaller values of \( \delta\sigma/\sigma_{SM} \) is essential. To gain such precision, we need high luminosities (at least 300 fb\(^{-1}\)) and advanced boosted Higgs tagging techniques which can reduce the ratio of the number of \( Z\bar{b}b \) to the SM \( Zh(\bar{b}b) \) events to an \( O(1) \) number as shown earlier in Refs. [15–17]. This work is, therefore, an example of a study at the “high energy-luminosity” frontier in the spirit of Refs. [18–20].

While adding the channel \( pp \rightarrow Z(\bar{\nu}\nu)h(\bar{b}b) \) can further improve the limits on the effective operators we study [15], this channel is subjected to backgrounds and employs observables with larger systematic uncertainties. We therefore leave an inclusion of this channel for future work.

As we will see, the leading high energy contribution to the \( pp \rightarrow Zh \) process comes from the four contact interactions \( hZ_{\mu}\bar{u}_{L,R}\gamma_{\mu}u_{L,R} \) and \( hZ_{\mu}\bar{d}_{L,R}\gamma_{\mu}d_{L,R} \) that are present in the dimension-6 extended Lagrangian. Thus, although many more operators contribute to the \( pp \rightarrow Zh \) process, the high energy limit isolates the four linear combinations of operators that generate the above contact terms. An interesting observation, first made in Ref. [20], is that the same four EFT directions (that the authors call “high energy primaries”) also control \( Wh \) and \( WW/WZ \) production. The reason is that at high energies these four final processes correspond to the production of different components of the Higgs doublet due to the Goldstone Boson equivalence theorem [21]. They are therefore related by \( SU(2)_L \) symmetry for \( \hat{s} \gg m_Z^2 \). Hence, although these four diboson processes may be very different from a collider physics point of view, they are intimately related by gauge symmetry, which stands at the heart of an EFT interpretation. This enables an elegant understanding of the connection of pseudo-observables in \( WW \) production (such as TGCs) with those in \( Zh \) production. It also allows us to present our results in a combined way with the projections for \( WZ \) production in Ref. [20].

II. THE HIGH ENERGY \textit{Vh}-AMPLITUDE IN THE SMEFT

Let us first study \( Vh \) production at high energy in the SMEFT where \( V = W, Z \). Although we focus on \( pp \rightarrow Zh \) production in the subsequent sections, here we keep the discussion more general considering also the \( Wh \) final state. We see that \( Vh \) production at hadron colliders at high energies isolates four independent directions in the full 59-dimensional space of dimension-6 operators. To derive this fact, consider first the vertices in the dimension-6 Lagrangian that contribute to the \( ff \rightarrow Zh \) process in unitary gauge,

\[
\Delta \mathcal{L}_6 \supset \sum_f \delta g_{Zf}^2 Z_{\mu} \bar{T} \gamma^\mu f + \delta g_{Vf}^W (W_{\mu}^+ \bar{u}_L) \gamma^\mu d_L + \text{H.c.}
\]

\[
+ g_\gamma^W h \left[ W^{+\mu} W_{\mu} + \frac{1}{2c_{b\bar{b}}} Z^{\mu} Z_{\mu} \right] + \delta g_{ZZ}^W h \left[ Z^{\mu} Z_{\mu} \right]
\]

\[
+ \sum_f \delta g_{Zf}^W \frac{h}{v} Z_{\mu} \bar{T} \gamma^\mu f + \delta g_{Wf}^W (W_{\mu}^+ \bar{u}_L) \gamma^\mu d_L + \text{H.c.}
\]

\[
+ \kappa_{Zf} h \frac{h}{v} A^{\mu\nu} Z_{\mu\nu} + \kappa_{Wf} h \frac{h}{v} W^{+\mu\nu} W_{\mu\nu} + \kappa_{Zf}^W h \frac{h}{v} Z^{\mu\nu} Z_{\mu\nu}.
\]

(2)

We are using the formalism presented in Refs. [2,22] where \( a_{cm}, m_Z, m_W \) have been used as input parameters and any corrections to the SM vector propagators, i.e., the terms \( V_{\mu} V^{\mu}, V_{\mu\nu} V^{\mu\nu} \) and \( V_{\mu} F^{\mu} \) have been traded in favor of the vertex corrections. Note that the above parametrization is equivalent to the one in Refs. [23,24] (see Ref. [25]). Keeping only the leading terms in \( \hat{s}/m_Z^2 \) in the BSM correction, we obtain for the amplitude \( \mathcal{M}(ff \rightarrow V_{T,L} h), \)

\[
Z_{T} h: \ g_f^e \frac{e \cdot J_f}{v} \frac{2m_Z^2}{\hat{s}} \left[ 1 + \left( \frac{g_{Zf}^e}{g_f^e} - \kappa_{ZZ} \right) \frac{\hat{s}}{2m_Z^2} \right],
\]

\[
Z_{L} h: \ g_f^q \frac{q \cdot J_f}{v} \frac{2m_Z^2}{\hat{s}} \left[ 1 + \left( \frac{g_{Zf}^q}{g_f^q} - \kappa_{Zf}^W \right) \frac{\hat{s}}{2m_W^2} \right],
\]

\[
W_{T} h: \ g_f^e \frac{e \cdot J_f}{v} \frac{2m_W^2}{\hat{s}} \left[ 1 + \left( \frac{g_{Wf}^e}{g_f^e} - \kappa_{Wf} \right) \frac{\hat{s}}{2m_W^2} \right],
\]

\[
W_{L} h: \ g_f^q \frac{q \cdot J_f}{v} \frac{2m_W^2}{\hat{s}} \left[ 1 + \left( \frac{g_{Wf}^q}{g_f^q} - \kappa_{Wf} \right) \frac{\hat{s}}{2m_W^2} \right],
\]

(3)

where \( g_f^e = g(T_f^e - Q_f s_{\theta_W})/c_{\theta_W} \), and \( g_f^q = g/\sqrt{2} \). \( J_f^T \) is the fermion current \( \bar{T} \gamma^\mu f \), the subscript \( L \) (T) denotes the longitudinal (transverse) polarization of the gauge boson, \( q \) denotes its fermion number and \( c \) denotes the polarization vector.

We see that only the \( g_{Zf}^q \) and \( \kappa_{Vf} \) couplings lead to an amplitude growing with energy. In the case of the \( \kappa_{Vf} \) couplings, the energy growth arises because of the extra powers of momenta in the \( hfV \) vertex, whereas for the contact interaction, \( g_{Zf}^q \), the energy growth is due to the fact that there is no propagator in the diagram involving this vertex. In fact for the latter interaction, the only difference in the amplitude with respect to the SM is the absence of the propagator. Thus, angular distributions are expected to
TABLE I. Dimension-6 operators that give dominant contribution to $pp \to Vh$ at high energies in the Warsaw [1] and strongly interacting light Higgs (SILH) [26] bases.

<table>
<thead>
<tr>
<th>SILH basis</th>
<th>Warsaw basis</th>
</tr>
</thead>
<tbody>
<tr>
<td>$O_W = \frac{ig}{2}(H^1 e^aD^a_{\mu\nu})D^a_{\mu\nu}W_{\mu\nu}$</td>
<td>$O_L^{(3)} = (\bar{Q}<em>L\sigma^\mu\rho Q_L)(iH)\sigma^\mu\rho D</em>{\rho\mu}H$</td>
</tr>
<tr>
<td>$O_R = \frac{ig}{2}(H^1 d^aD^a_{\mu\nu})\bar{D}<em>{\mu\nu}W</em>{\mu\nu}$</td>
<td>$O_L = (\bar{Q}_L\gamma^\mu Q_L)(iH)D^\mu H$</td>
</tr>
<tr>
<td>$O_{HW} = ig(D^a_{\mu\nu})^{\sigma\tau}D^a_{\mu\nu}W_{\mu\nu}$</td>
<td>$O_R = (\bar{u}<em>R\gamma^\mu u_R)(iH)\sigma^\mu D</em>{\mu\rho}H$</td>
</tr>
<tr>
<td>$O_{HB} = ig(D^a_{\mu\nu})^{\sigma\tau}(D^b_{\nu\mu})B_{\mu\nu}$</td>
<td>$O_R = (\bar{d}<em>R\gamma^\mu d_R)(iH)\sigma^\mu D</em>{\rho\mu}H$</td>
</tr>
<tr>
<td>$O_{2W} = -\frac{1}{2}(D^a_{\mu\nu}W_{\mu\nu})^2$</td>
<td>$O_{2H} = -\frac{1}{2}(\partial^\mu B_{\mu})^2$</td>
</tr>
</tbody>
</table>

be identical for BSM and SM production. Therefore, the only way to probe this interaction is through the direct energy dependence of differential cross sections.

On the other hand, the $\kappa_{VY}$ interactions contribute only to the transverse $V$ amplitude as a consequence of their vertex structure. Hence, they cannot interfere with the dominant longitudinal piece in the SM amplitude. As a result, after summing over all $V$-polarizations, the leading piece in the high energy cross-section deviation, is controlled only by the couplings $g_{Vf}^h$ whereas the $\kappa_{VV}$ contribution is suppressed by an additional $O(m_V^2/\delta)$ factor.

Note that the couplings, $\delta g_{Vf}^2$ and $\delta g_{2Zf}$ also lead to deviations from the SM amplitude but these corrections do not grow with energy and are also suppressed by an additional $O(m_V^2/\delta)$ factor with respect to the $g_{Vf}^h$ contribution. We have checked explicitly that including these couplings has no noticeable impact on our analysis.

At hadron colliders, the $pp \to Vh$ process at high energies and at leading order is therefore controlled by the five contact interactions: $g_{2f}^h$, with $f = u_L, u_R, d_L$ and $d_R$ and $g_{Wud}^h$. These five couplings correspond to different linear combinations of Wilson coefficients in any given basis. In Table I we show all operators in the Warsaw [1] and SILH [26] bases that generate these contact terms. As there are only four independent operators contributing to these five interactions in the Warsaw basis, there exists a basis independent constraint at the dimension-6 level,

$$g_{Wud}^h = c_{\theta_w}g_{Zu}^h - g_{Zd}^h \frac{\sqrt{2}}{2}.$$  (4)

leaving only the four independent $g_{Zf}^h$ couplings.

In Table II, we show the linear combinations of Wilson coefficients contributing to the four $g_{Zf}^h$ couplings in different EFT bases. The first row gives these directions in the Warsaw basis. The second row provides the aforementioned directions in the BSM primary basis of Ref. [2], where the Wilson coefficients can be written in terms of already constrained pseudo-observables. It is clear in this basis that the directions to be probed by high energy $Vh$ production can be written in terms of the LEP (pseudo-) observables. The couplings $\delta g_{Zf}^2$ defined in Eq. (2) are strongly constrained by Z-pole measurements at LEP, whereas the anomalous TGCs, $\delta \kappa_f$ and $\delta g_{Zf}^2$ (in the notation of Ref. [27]), were constrained by $WW$ production during LEP2.

For the physically motivated case where the leading effects of new physics can be parametrized by universal (bosonic) operators, the SILH Lagrangian provides a convenient formulation and we show the above directions in this basis in the third row of Table II. For this case, as shown in the fourth row of Table II, one can again write the directions in terms of only the “oblique”/universal pseudo-observables, viz., the TGCs $\delta \kappa_f$ and $\delta g_{Zf}^2$ and the Peskin-Takeuchi $\bar{S}$-parameter [28] in the normalization of Ref. [29]. For a definition of these observables we refer to the Lagrangian presented in Ref. [30] (see also Ref. [31]). As we already mentioned, upon using the Goldstone equivalence principle, one finds that the same four-dimensional subspace of operators also controls the

TABLE II. The linear combinations of Wilson coefficients contributing to the contact interaction couplings $g_{Zf}^h$ that control the $ff \to Vh$ process at high energies. The four directions relevant for hadron colliders (corresponding to $f = u_L, d_L, u_R, d_R$) can be read off from this table by substituting the value of the $SU(2)_L$ and $U(1)_Y$ quantum numbers $T_f^i$ and $Y_f$ for the corresponding $f$. Here $\hat{c}_W = c_W + c_{HW} - c_{2W}$ and $\hat{c}_R = c_R + c_{HB} - c_{2B}$. For the nomenclature of the operators, their corresponding Wilson coefficients and observables see e.g., Ref. [20].

<table>
<thead>
<tr>
<th>EFT directions probed by high energy $ff \to Vh$ production</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warsaw basis [1]</td>
</tr>
<tr>
<td>$-\frac{2g_{Zd}}{c_{\theta_w}}\left(T_f^i</td>
</tr>
<tr>
<td>BSM primaries [2]</td>
</tr>
<tr>
<td>$\frac{2g_{Zd}}{c_{\theta_w}}Y_fT_u^3\delta \kappa_f + 2\delta g_{Vf}^2 - \frac{2g_{Zd}}{c_{\theta_w}}(T_f^i</td>
</tr>
<tr>
<td>SILH Lagrangian [26]</td>
</tr>
<tr>
<td>$\frac{2g_{Zd}}{c_{\theta_w}}\left(2T_f^i\hat{c}_W - 2\hat{c}_R Y_f\hat{c}_R\right)$</td>
</tr>
<tr>
<td>Universal observables</td>
</tr>
<tr>
<td>$\frac{2g_{Zd}}{c_{\theta_w}}Y_fT_u^3\left(\delta \kappa_f - \bar{S} + Y_f\right) - \frac{2g_{Zd}}{c_{\theta_w}}(T_f^i</td>
</tr>
<tr>
<td>High energy primaries [20]</td>
</tr>
<tr>
<td>$-\frac{2g_{Zd}}{c_{\theta_w}}(T_f^i</td>
</tr>
</tbody>
</table>
longitudinal $VV$ production at high energies. This space is
defined in Ref. [20] in terms of the four high energy
primaries which are linear combinations of the four $g_{Vf}$
 couplings, as shown in the last row of Table II.

As it is not possible to control the polarization of the
initial state partons in a hadron collider, the process can, in
reality, only probe two of the above four directions. Taking
only the interference term, these directions are

\[
g_\theta^d = g_{2uz}^d + \frac{g_{2dh}^d}{g_{2lz}} g_{2uz}^l,
\]

\[
g_\theta^u = g_{2dh}^u + \frac{g_{2dl}^u}{g_{2dz}} g_{2dh}^d, \tag{5}
\]

where $g_\theta^d$ is defined below Eq. (3). Also, at a given energy,
the interference term for the $pp \rightarrow Zh$ process is sensitive
only to a linear combination of the up-type and down-type
coupling deviations, i.e., to the direction,

\[
g_\theta^p = g_\theta^u + \frac{\mathcal{L}_d(\delta)}{\mathcal{L}_u(\delta)} g_\theta^d, \tag{6}
\]

where $\mathcal{L}_{u,d}$ is the $u\bar{u}, d\bar{d}$ luminosity at a given partonic
center of mass energy. We find that the luminosity ratio
changes very little with energy (between 0.65 and 0.59 if
$\sqrt{s}$ is varied between 1 and 2 TeV). Thus, to a good
approximation, $pp \rightarrow Zh$ probes the single direction in
EFT space given by

\[
g_\theta^p = g_{2uz}^u - 0.76 g_{2dl}^u - 0.45 g_{2uz}^d + 0.14 g_{2dh}^d, \tag{7}
\]

where we have substituted the values for $g_\theta^d$ and evaluated
the luminosities at the energy $\sqrt{s} = (1.5 \text{ TeV})^2$. This can
now be written in terms of the LEP-constrained pseudo-
observables in the second and fourth row of Table II,

\[
g_{2dz}^p = 2 \delta g_{2uz}^u - 1.52 \delta g_{2dl}^u - 0.90 \delta g_{2uz}^d + 0.28 \delta g_{2dh}^d
\]

\[
- 0.14 \delta \kappa_y - 0.89 \delta g_1^2,
\]

\[
g_{2dh}^p = -0.14 (\delta \kappa_y - \delta + Y) - 0.89 \delta g_1^2 - 1.3 W \tag{8}
\]

where the first line applies to the general case and the second
line to the universal case.

Note that in the discussion so far we have not considered the
$gg \rightarrow Zh$ production channel at hadron colliders
[32–40]. While formally a higher order correction, after
all the cuts are applied, this subprocess contributes an
appreciable 15% of the total SM $pp \rightarrow Zh$ cross section
in our analysis due to the top-threshold inducing boosted final
states [37]. We find, however, that introduction of the EFT
operators does not lead to an energy growing amplitude
with this initial state, and thus this channel has a subdominant
contribution to the EFT signal. While we fully include
this contribution in our collider analysis, the introduction
of this channel does not alter the discussion so far in an
important way.

We now turn to the crucial issue of estimating the scale of new
physics (and thus the cutoff for our EFT treatment) for a
given size of the couplings, $g_{Vf}$. This also gives us an idea
of the new physics scenarios that our analysis can probe. As
is clear from the operators in Table I, the $g_{Vf}$ couplings
arise from current-current operators that can be generated,
for instance, by integrating out at tree level a heavy $SU(2)_L$
triplet (singlet) vector $W^a$ ($Z$) that couples to SM fermion
currents, $\bar{f} \sigma^{\mu\nu} f$ ($\bar{f} Y f$) with a coupling $g_f$ and to the Higgs
current $iH^\dagger Y D_H (iH^\dagger D^*_H)$ with a coupling $g_H$,

\[
g_{Zf}^F \sim \frac{g_H g_f Y_f v^2}{\Lambda^2}, \tag{9}
\]

where $\Lambda$ is the mass of the vector and therefore the
matching scale or cutoff of the low energy EFT. The
coupling to the SM fermions can be universal if the heavy
couplers to them only via kinetic mixing with the SM
 gauge bosons. This results in a coupling of the heavy vector
to the $SU(2)_L$ and hypercharge currents given by $g_w = g/2$ and
$g_B = g Y_f, Y_f$ being the SM hypercharge. As we want
our results to be applicable to the universal case, we assume
the coupling $g_f$ to be given by a combination of $g_h$ and $g_w$
to obtain

\[
g_{2uz,d_l}^h \sim \frac{g_H g^2 Y_{uz,d_l} v^2}{2 \Lambda^2},
\]

\[
g_{2uz,d_k}^h \sim \frac{g_H g y Y_{uz,d_k} v^2}{\Lambda^2}, \tag{10}
\]

and then further assume a weakly coupled scenario with
$g_H = 1$ [note that this is a bit larger than the corresponding
value $g_H = g/(2\epsilon_{0w})$ for the SM $hZZ$ coupling]. In
the above equation, we have ignored the smaller contributions
from $g_f$ to the left-handed couplings. For any set of
couplings $\{g_{2uz}^h, g_{2dl}^h, g_{2uz}^v, g_{2dh}^v\}$, we evaluate the cutoff
using Eq. (10) with $g_H = 1$ and take the smallest of the four
values. It is clear that for strongly coupled scenarios with
larger values of $g_H$, the cutoff assumed in our analysis is
smaller than necessary and thus our projected bounds are
conservative.

III. ANALYSIS

In order to probe the reach of the high luminosity runs of the
LHC in constraining the EFT directions in Table II, we
optimize a hadron-level analysis to obtain maximum
sensitivity to the BSM signal, which is well pronounced
in the high energy bins. To achieve this, we consider the
$Z(e^+ e^-)h$ production from a pair of quarks as well as from
a pair of gluons. As far as the decay of the Higgs boson is
concerned, we find that at an integrated luminosity of 300 fb\(^{-1}\), the diphoton mode yields less than five events at high energies \((p_{T,Z} > 150\text{ GeV})\) and is thus not sensitive to the effects we want to probe. We thus scrutinize the \(h(b\bar{b})Z(\ell^+\ell^-)\) final state where the dominant backgrounds are composed of \(Zb\bar{b}\) and the irreducible SM production of \(Z\). For the \(Zb\bar{b}\) process, we consider the tree-level production as well as the \(gg \rightarrow ZZ\) production at one loop. Reducible contributions arise from \(Z +\) jets production (c-quarks included but not explicitly tagged), where the light jets can be misidentified as \(b\)-jets, and the fully leptonic decay for \(\ell\). Instead of performing a standard resolved analysis, where one would demand two separate \(b\)-tagged jets, we demand a fat jet with a cone radius of \(R = 1.2\). We employ the so-called BDRS approach [15] with minor modifications to maximize sensitivity. In a nutshell, this technique helps in discriminating boosted electroweak-scale resonances from large QCD backgrounds.

We see that using this approach allows us to reduce the ratio of \(Zb\bar{b}\) to SM \(Z\) events from about 40 to an \(\mathcal{O}(1)\) number with about 40 SM events still surviving at 300 fb\(^{-1}\). This shows that the kind of analysis performed here would not be possible at integrated luminosities smaller than 300 fb\(^{-1}\).

The BDRS approach recombines jets using the Cambridge-Aachen (CA) algorithm [41,42] with a significantly large cone radius to contain all the decay products of the resonance. One then works backwards through the jet clustering and stops when a significant *mass drop*, \(m_{j_1} < \mu m_j\) with \(\mu = 0.66\), \(m_j\) being the mass of the fatjet) occurs for a not too asymmetric splitting,

\[
\text{min}(p_{T,j_1}^2, p_{T,j_2}^2) \Delta R_{j_1,j_2}^2 > y_{\text{cut}},
\]

with \(y_{\text{cut}} = 0.09\). If this condition is not met, the softer subjet \(j_2\) is removed and the subjets of \(j_1\) are tested for the aforementioned criteria to be satisfied in an iterative process. The algorithm stops as soon as one obtains two subjets, \(j_1\) and \(j_2\) abiding by the mass drop condition.

To improve the resonance reconstruction, the technique considers a further step called *filtering*. In this step, the constituents of \(j_1\) and \(j_2\) are again recombined using the CA algorithm with a cone radius \(R_{\text{filt}} = \text{min}(0.3, R_{bb}/2)\). Only the hardest three filtered subjets are retained to reconstruct the resonance. In the original work of Ref. [15], this resonance is the SM-like Higgs boson and thus the two hardest filtered subjets are \(b\) tagged. In our work, we find that the filtered cone radius \(R_{\text{filt}} = \text{max}(0.2, R_{bb}/2)\) works better in removing the backgrounds.\(^1\) The filtering step greatly reduces the active area of the initial fat jet.

---

\(^1\)The criteria \(R_{bb} = \text{max}(0.2, R_{bb}/2)\) followed by \(R_{\text{filt}} = \text{min}(0.3, R_{bb}/2)\) hardly change the results.

---

We use the FeynRules [43] and UFO [44] toolkits to implement the signal contributions (we comment on the effect of including squared dimension-6 interactions as compared to interference-only terms below). Both signal and background processes are generated including the full decay chain with MG5\_aMC@NLO [45], at leading order. For the gluon initiated part of the SM and BSM \(Z\) production, we employ the FeynArts/FormCalc/LoopTools [46,47] framework and the decays are performed using MadSpin [48,49]. We shower and hadronize the samples using Pythia 8 [50,51] and perform a simplified detector analysis.

Because our ultimate goal is to look for new physics effects in high energy bins, we generate the \(Z\), \(Zb\bar{b}\) and \(t\bar{t}\) samples with the following cuts: \(p_{T,j,b} > 15\text{ GeV}\), \(p_{T,c} > 5\text{ GeV}\), \(|y_j| < 4\), \(|y_{b/c}| < 3\), \(\Delta R_{b/c}\) > 0.2, \(\Delta R_{c/c} > 0.15\), 70 GeV < \(m_{c/c} < 110\text{ GeV}\), 75 GeV < \(m_b < 155\text{ GeV}\) and \(p_{T,c} > 150\text{ GeV}\). The former two processes are generated upon merging with an additional matrix element (ME) parton upon using the MLM scheme [52]. For the \(Z +\) jets process, we generate the samples without the invariant mass cuts on the jets; we further merge the sample up to three ME partons.

To account for higher order QCD corrections for the \(q\bar{q}\)-initiated \(Z\) processes, we apply a bin-by-bin (in \(M_{Zb}\), the invariant mass of the filtered double \(b\)-tagged fat jet and the reconstructed \(Z\)-boson) \(K\)-factor reweighting to the next-to-leading order (NLO)-accurate distribution both for the SM background and the EFT signal using Ref. [25]. For the \(gg\)-initiated \(Z\) process, we consider a conservative NLO \(K\) factor of 2 [38]. For the tree-level \(Zb\bar{b}\) and \(Z +\) jets processes, flat \(K\) factors of 1.4 (computed within MG5\_aMC@NLO) and 0.91 [53] are applied. For the \(gg \rightarrow ZZ\) production, a flat \(K\) factor of \(\sim 1.8\) [54] has been used.

We first test the power of a cut-based analysis. In doing so, we construct the fat jets with a cone radius of \(R = 1.2\), having \(p_{T} > 80\text{ GeV}\) and rapidity, \(|y| < 2.5\) using FastJet [55]. We isolate the leptons (\(e, \mu\)) by demanding that the total hadronic activity around a cone radius of \(R = 0.3\) must be less than 10% of its \(p_{T}\) and the leptons are required to have \(p_{T} > 20\text{ GeV}\) and \(|y| < 2.5\). All non-isolated objects are considered while constructing the fat jets. In selecting our events, we consider only those with exactly two isolated leptons having opposite charge and same flavor (OSSF). Moreover, we demand the invariant mass of the pair of leptons to lie in the range (80–100 GeV) in order to reconstruct the \(Z\)-peak. The reconstructed \(Z\) is required to be boosted with \(p_{T} > 160\text{ GeV}\) and the separation between the two isolated leptons is required to be \(\Delta R > 0.2\). In reconstructing the Higgs boson, we demand that each event has at least one fat jet containing no less than two \(B\)-meson tracks with \(p_{T} > 15\text{ GeV}\). The minimum transverse momentum of the fat jet is required to be \(p_{T} > 110\text{ GeV}\). After satisfying the mass drop and filtering criteria, we require exactly two subjets.
after the former step and at least two subjets after filtering. We proceed with $b$-tagging the two hardest subjets. We choose a $b$-tagging efficiency of 70% and a misidentification rate for light jets of 2%. After the filtering and $b$-tagging steps, we require events with exactly two $b$-tagged subjets, which are well separated from the isolated leptons: $\Delta(R(b_i, b_j)) > 0.4$ for both leptons $\ell_1, 2$ and $b$-tagged subjets $b_i$. We reconstruct the Higgs boson by requiring its invariant mass to lie in the range $(115-135 \text{ GeV})$.

In order to further reduce the backgrounds, we demand both the reconstructed $Z$ and the Higgs bosons to have $p_T > 200 \text{ GeV}$. The $t\bar{t}$ background can be removed almost entirely by requiring $E_T < 30 \text{ GeV}$. The cut flow affecting the most dominant background $Zb\bar{b}$ and the SM $Zh$ channel is summarized in Table III.

Before focusing on the very high energy effects by imposing cuts on $M_{Zh}$, we find that the ratio of cross section between SM $Zh$ and $Zb\bar{b}$ is $<0.26$. A multivariate implementation at this level strengthens this ratio further. In order to be quantitative, we impose looser cuts on the aforementioned variables 70 GeV $< m_\ell\ell < 110$ GeV, $p_T, \ell > 160$ GeV, $\Delta R_\ell \ell > 0.2$, $p_T, fat\text{jet} > 60$ GeV, 95 GeV $< m_h < 155$ GeV, $\Delta R_{b, b} > 0.4$ and $E_T < 30$ GeV. Because $Z + \text{jets} + t\bar{t}$ are much less significant than $Zb\bar{b}$, we train the boosted decision trees only with the SM $q\bar{q}$-initiated $Zh$ and $Zb\bar{b}$ samples using the following variables: $p_T$ of the two isolated leptons, $\Delta R$ between pairs of $b$-subjets and isolated leptons, between the two isolated leptons and between the hardest two $b$-subjets in the Higgs fat jet, the reconstructed $Z$-boson mass and its $p_T$, $\Delta \Phi$ separation between the fat jet and the reconstructed $Z$-boson, $E_T$, mass of the reconstructed Higgs jet and its $p_T, p_T$ of the two $b$-tagged filtered subjets, the ratio of their $p_T$ and the rapidity of the Higgs jet. We ensure that we do not have variables which are $\sim 100\%$ correlated but we retain all other variables. Because our final distribution of interest is

### Table III

<table>
<thead>
<tr>
<th>Cuts</th>
<th>$Zb\bar{b}$</th>
<th>$Zh$ (SM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>At least one fat jet with two $B$-mesons with $p_T &gt; 15$ GeV</td>
<td>0.23</td>
<td>0.41</td>
</tr>
<tr>
<td>Two OSSF isolated leptons</td>
<td>0.41</td>
<td>0.50</td>
</tr>
<tr>
<td>80 GeV $&lt; M_{\ell\ell} &lt; 100$ GeV, $p_{T,\ell\ell} &gt; 160$ GeV, $\Delta R_{\ell\ell} &gt; 0.2$</td>
<td>0.83</td>
<td>0.89</td>
</tr>
<tr>
<td>At least one fat jet with two $B$-meson tracks with $p_T &gt; 110$ GeV</td>
<td>0.96</td>
<td>0.98</td>
</tr>
<tr>
<td>Two Mass drop subjets and $\geq$ two filtered subjets</td>
<td>0.88</td>
<td>0.92</td>
</tr>
<tr>
<td>Two $b$-tagged subjets</td>
<td>0.38</td>
<td>0.41</td>
</tr>
<tr>
<td>$115$ GeV $&lt; m_\ell &lt; 135$ GeV</td>
<td>0.15</td>
<td>0.51</td>
</tr>
<tr>
<td>$\Delta R(b_i, b_j) &gt; 0.4$, $E_T &lt; 30$ GeV, $</td>
<td>y_h</td>
<td>&lt; 2.5$, $p_{T,h/Z} &gt; 200$ GeV</td>
</tr>
</tbody>
</table>

FIG. 1. The differential distribution of events at an integrated luminosity of 300 fb$^{-1}$ with respect to $M_{Zh}$ for the EFT signal as well as the different backgrounds. For the EFT signal we have taken the point $\{q^{2}_{Zu_R}, q^{2}_{Zu_L}, q^{2}_{Zd_R}, q^{2}_{Zd_L}\} = \{-0.005, 0.0001, -0.010, 0.005\}$ which is allowed by the LEP bounds.

The invariant mass of the $Zh$-system, we do not consider it as an input variable. We use the TMVA [56] framework to train our samples and always ensure that the Kolmogorov-Smirnov statistic is at least of the order $<0.1$ in order to avoid overtraining of the samples [57]. We find that the aforementioned ratio increases to $\sim 0.50$ upon using the boosted decision tree algorithm showing that a further optimization of the cut-based analysis was necessary. Finally, we test all our samples with the training obtained from the SM $q\bar{q}$-initiated $Zh$ and the $Zb\bar{b}$ samples.

To distinguish between the EFT signal and the irreducible SM $Zh(b\bar{b})$ background we utilize the growth of the EFT cross section at high energies. The effects are readily seen in the $M_{Zh}$ distribution, our observable of interest. In Fig. 1 we show the differential distribution with respect to this variable for the EFT signal as well as the different backgrounds for an integrated luminosity of 300 fb$^{-1}$. For the EFT signal we take a point that can be excluded in our analysis but is well within the LEP allowed region. We see that the EFT cross section keeps growing with energy, but much of this growth is unphysical at energies above the cutoff, i.e., $M_{Zh} \sim \Lambda$, where $\Lambda$ is the cutoff evaluated as described below Eq. (10) and shown by a vertical line in Fig. 1. For our analysis we dropped all events above this cutoff. For $M_{Zh} < \Lambda$, the EFT deviations are never larger than an $O(1)$ factor with respect to the SM background as expected on general grounds. Note, however, that even for $M_{Zh} < \Lambda$, even though the underlying anomalous couplings, $q^{2}_{Zf}$, are per-mille to percent level, the fractional deviations are much larger because of the energy growth of the BSM rate. To make full use of the shape deviation of the EFT signal with respect to the background, we perform a binned log likelihood analysis assuming a 5% systematic error. The likelihood function is taken to be the product of Poisson distribution functions for each bin with the mean
given by the number of events expected for a given BSM point. To account for the 5% systematic error we smear the mean with a Gaussian distribution. To obtain the projection for the 95% C.L. exclusion curve we assume that the observed number of events agrees with the SM.

### IV. DISCUSSION

Considering only the SM-BSM interference term, we find the per-mille level bounds,

\[
\begin{align*}
\delta g^h_{Zp} &\in [-0.004, 0.004] \quad (300 \text{ fb}^{-1}) \\
\delta g^h_{Zp} &\in [-0.001, 0.001] \quad (3000 \text{ fb}^{-1}).
\end{align*}
\]

Using Eq. (10) the above bounds can be translated to a lower bound on the scale of new physics given by 2.4 TeV (4.4 TeV) at 300 fb\(^{-1}\) (3000 fb\(^{-1}\)). One can now compare the above projections with existing LEP bounds by turning on the LEP observables contributing to \(\delta g^h_{Zp}\) in Eq. (8) one by one. This is equivalent to assuming that there are no large cancellations in Eq. (8) so that each individual term is bounded by Eq. (11). The results are shown in Table IV. We see that our projections are much stronger than the LEP bounds for the TGCs \(\delta g^Z_f\) and \(\delta \kappa_f\) and comparable in the case of the Z-pole observables \(\delta g^Z_f\) that parametrize the deviations of the Z coupling to quarks.

To obtain Eq. (11), we have used the cutoff as defined in Eq. (10) with \(g_h = 1\). One might expect a stronger bound by taking a larger \(g_h\). We find, however, that while taking stronger couplings can increase the cutoff many times, this does not lead to an appreciably higher sensitivity because the high energy bins have very few or no SM/BSM events being suppressed by the small PDFs at these energies.

For the universal case, the EFT directions presented in Table II can be visualized in the \(\delta \kappa_f - \hat{S}\) vs \(\delta g^Z_f\) plane as shown in Fig. 2 for the interesting class of models where \(W = Y = 0\) [20]. The flat direction related to the \(pp \to Zh\) interference term, i.e., \(\delta g^h_{Zp} = 0\), Eq. (7), is shown by the dashed blue line, where the direction \(\delta g^h_{Zp}\) is now given by the second line of Eq. (8). The grey shaded area shows the allowed region after the LEP II bounds [59] from the \(e^+e^- \to W^+W^-\) process are imposed. The results of this work are shown in blue [light (dark) blue for results at 300 (3000) fb\(^{-1}\)]. To understand the shape of the blue bands, note that along the dashed line, the SM-BSM interference term vanishes. If the interference was the only dominant effect, the projected allowed region would be a band along this direction. The BSM squared term thus plays a role in determining the shape of the blue region. To the left of the dashed blue line, the squared and the interference terms have the same sign while there is a partial cancellation between these two terms on the right-hand side of the dashed line. This results in the curvature of the blue band with stronger bounds to the left of the dashed line and weaker bounds to its right.

![FIG. 2. We show in light blue (dark blue) the projection for the allowed region with 300 fb\(^{-1}\) (3 ab\(^{-1}\)) data from the \(pp \to Zh\) process for universal models in the \(\delta \kappa_f - \hat{S}\) vs \(\delta g^Z_f\) plane. The allowed region after LEP bounds (taking the TGC \(\lambda_i = 0\), a conservative choice) are imposed is shown in grey. The pink (dark pink) region corresponds to the projection from the WZ process with 300 fb\(^{-1}\) (3 ab\(^{-1}\)) data derived in Ref. [20] and the purple (green) region shows the region that survives after our projection from the Zh process is combined with the above WZ projections with 300 fb\(^{-1}\) (3 ab\(^{-1}\)) data.](attachment:image.png)
We see that, as we move further from the origin, the effect of the squared term becomes more pronounced. This is expected, as along the dashed line, the interference term is accidentally 0, even for energies below the cutoff, and thus, the parametrically subdominant squared term is larger. To achieve a partial cancellation between these two terms one needs to deviate more and more from the dashed line. While EFT validity has been carefully imposed to derive our bounds, the fact that the interference term vanishes along the flat direction and the squared term becomes important does imply that for weakly coupled UV completions our bounds are susceptible to $\mathcal{O}(1)$ dimension-8 deformations in this direction. In the orthogonal direction shown by the dotted line, on the other hand, our projections are more robust and not sensitive to such effects. Such an ambiguity also exists in the results in Table IV, for the pseudo-observables such as $\delta g_{\mu\mu}$ and $\delta \kappa$, that are somewhat aligned to the above flat direction. This ambiguity can be resolved by performing a global fit upon combining analyses of all the $Vh, VV$ channels that will avoid such flat directions.

As we have emphasized already, $VV$ production constrains the same set of operators as the $Vh$ production. In Fig. 2, we also show the projected bound from the WZ process at 300 fb$^{-1}$ obtained in Ref. [20]. When both these bounds are combined, only the purple region remains. At 3000 fb$^{-1}$, this region shrinks further to the green region shown in Fig. 2. Thus, we see a drastic reduction in the allowed LEP region is possible by investigating $pp \rightarrow Zh$ at high energies.

**V. CONCLUSIONS**

As hints for new physics beyond the SM remain elusive with the LHC entering a new energy territory, model-independent approaches based on the assumption of no additional light propagating degrees of freedom are gaining ground. The power of effective field theory is that theoretical correlations between independent measurements can be exploited to formulate tight constraints on the presence of new physics, solely based on the SM symmetries and matter content.

The high precision measurements performed during the LEP era are therefore the driving forces behind combined constraints early in the LHC program. To enter new territory, the LHC has to push beyond the LEP sensitivity for interactions that relate the phenomenology at both collider experiments. The Higgs boson, as arguably the most significant TeV scale degree of freedom, can be placed at the core of such a program, which will naturally involve LHC measurements at high luminosity.

In this paper, we focused on the impact of associated Higgs production that provides complementary information to the diboson production modes observed at LEP2, which determine the precision of the associated coupling constraints. Using a dedicated investigation of expected signal and backgrounds, we find that the LHC will ultimately be able to improve the sensitivity expected from LEP measurements. Our results are summarized in Eq. (11), Table IV and Fig. 2. Higgs-strahlung is also complementary to diboson production at the LHC investigated in Ref. [20]. Combining Higgs-strahlung measurements with diboson results in the high energy limit will allow us to drastically improve the sensitivity to the underlying new physics parameters in an unparalleled way. Both high energies and luminosities are crucial for a study like ours. Potentially even higher new physics scales can thus be probed at the High Energy LHC or other future colliders.

**ACKNOWLEDGMENTS**

We thank Matthew McCullough for helpful discussions and collaboration on this project in its early stage. We also thank Biplob Bhattacharjee, Mikael Chala, Shilpi Jain, Giuliano Panico, Michael Peskin and Francesco Riva for helpful discussions at various stages of this work. C. E. is supported by the IPPP Associateship scheme and by the United Kingdom Science and Technology Facilities Council (STFC) under Grant No. ST/P000746/1. S. B. is supported by a Durham Junior Research Fellowship COFUNDED by Durham University and the European Union, under Grant No. 609412.
