Quark flavour-violating Higgs decays at the ILC

Daniele Barducci[a] **and Alexander J. Helmboldt**[b]

[a]*SISSA and INFN, Sezione di Trieste,*
via Bonomea 265, 34136 Trieste, Italy

[b]*Max-Planck-Institut für Kernphysik,*
Saupfercheckweg 1, 69117 Heidelberg, Germany

E-mail: daniele.barducci@sissa.it, alexander.helmboldt@mpi-hd.mpg.de

Abstract: Flavour-violating Higgs interactions are suppressed in the Standard Model such that their observation would be a clear sign of new physics. We investigate the prospects for detecting quark flavour-violating Higgs decays in the clean ILC environment. Concentrating on the decay to a bottom and a light quark j, we identify the dominant Standard Model background channels as coming from hadronic Standard Model Higgs decays with mis-identified jets. Therefore, good flavour tagging capabilities are essential to keep the background rate under control. Through a simple cut-based analysis, we find that the most promising search channel is the two-jet plus missing energy signature $e^+e^- \to bj + \not{E}_T$. At $500\,\mathrm{GeV}$, the expected $95\,\%$ CL upper limit on $\mathcal{B}(h \to bj)$ is of order 10^{-3}. Correspondingly, a $5\,\sigma$ discovery is expected to be possible for branching ratios as low as a few 10^{-3}.

Contents

1 Introduction

With the discovery of the Higgs boson at the Large Hadron Collider (LHC) in July 2012 [1, 2] the high energy particle physics program has entered a new exciting phase aiming at understanding whether the properties of the newly discovered state are consistent with the predictions of the Standard Model (SM). The 7 and 8 TeV runs of the LHC have constrained the Higgs bosons couplings to WW, ZZ, $\gamma\gamma$ and $\tau\tau$ pairs to be approximately within 10 % of their SM values, while weaker limits have been derived for the interactions with bottom and top quarks as well as gluons [3].

Obviously, more exotic Higgs couplings are of interest as well, in particular those non-diagonal in flavour space. Such interactions enable flavour-changing neutral currents (FCNC) and are thus forbidden at tree-level and, in the quark sector, GIM-suppressed at loop-level in the absence of new physics (NP). The corresponding processes in the SM are therefore very rare[1] and beyond the reach of today's experiments. Hence any observation would constitute a clear sign of physics beyond the SM. Indeed, there are various well-motivated models which predict sizable flavour-violating Higgs couplings, such as two Higgs doublet models [5–9], supersymmetric models [10], models with a composite Higgs [11], models of extra dimensions [12] and others.

Consequently, the ATLAS and CMS collaborations have put forward a program aiming at the search for such couplings. In particular, flavour-violating Higgs decays in the leptonic sector have received great attention due to possible indications of a non-zero $h \to \tau\mu$ branching ratio in 8 TeV data [13], which was, however, not confirmed by more recent

---

[1]In the SM one has for example $\mathcal{B}(h \to b\bar{s}) \sim 10^{-7}$ [4].

analyses performed with the first tenths of inverse femtobarn of integrated luminosity accumulated with 13 TeV proton-proton collisions [14]. In contrast, the situation is somewhat more complicated in the quark sector. On the one hand, flavour-violating Higgs couplings involving a top quark can be tested either via exotic top quark decays [15, 16] or in principle, given the mass of the Higgs boson, via Higgs decays to an off-shell top, a process which however seems to be outside the LHC reach. On the other hand, processes involving bottom quarks suffer from a huge QCD background which makes a measurement of decays such as $h \to bj$, with j representing a light quark, extremely challenging at the LHC, even at the end of the high luminosity phase when $\mathcal{O}(3000\,\text{fb}^{-1})$ of data will have been collected.

It is thus crucial to investigate whether proposed future colliders could offer a handle to *directly* test processes for which the LHC has a poor sensitivity. Direct probes of such decays are in fact essential both to complement indirect limits on FCNC interactions arising from low-energy flavour observable measurements and in case of deviations from the SM expectations being observed to be able to pin down their origin. Correspondingly, it is the aim of this paper to study the potential of an e^+e^- collider in probing exotic decay modes of the Higgs boson in a bottom quark and a light quark. In doing so, we will focus on the planned International Linear Collider (ILC), adopting the proposed centre-of-mass energy, beam polarisation and integrated luminosity values of this machine [17, 18] and showing that the sensitivity on NP couplings responsible for $h \to bj$ processes competes with, and for some specific analysis surpasses, the one that can be obtained by low-energy flavour measurements in the effective field theory (EFT) approach.

The paper is organised as follows. In section 2 we describe the effective parametrisation that we use for our study while in section 3 we review the Higgs physics program at the ILC. In section 4 we describe the details of our numerical analysis and present our findings. We then conclude in section 5.

2   Parametrisation

We chose to parametrise the non-SM interactions of the Higgs boson with a pair of quarks using the EFT language adopting as a basis the one of the Warsaw convention [19]. At dimension-six level, the operators that couple the Higgs field with a pair of down-type quarks are

$$\mathcal{Q}_{dH} = (\bar{q}_L^i H^\dagger d_R^j)(H^\dagger H), \tag{2.1a}$$

$$\mathcal{Q}_{Hd} = (\bar{d}_R^i \gamma^\mu d_R^j)(H^\dagger i\overleftrightarrow{D}_\mu H), \tag{2.1b}$$

where q_L and d_R indicate the left-handed doublet and right-handed singlet of SM quarks with generation indices i, j, $H = (\phi^+, v + h)^\intercal/\sqrt{2}$ is the Higgs field with $v \simeq 246\,\text{GeV}$, h the physical Higgs state and $H^\dagger i\overleftrightarrow{D}_\mu H = H^\dagger iD_\mu H - (iD_\mu H^\dagger)H$. However, after electroweak symmetry breaking (EWSB) only the operator of eq. (2.1a) gives rise to interactions of the physical Higgs with a pair of SM quarks of the form

$$\mathcal{L}_{\text{EFT}} = -y_{ij} h d_L^i d_R^j + \text{h.c.} \tag{2.2}$$

Observable	Coupling	Constraint
B_d^0 oscillations	$|y_{db}|^2$, $|y_{bd}|^2$	$< 2.3 \cdot 10^{-8}$
	$|y_{db}y_{bd}|$	$< 3.3 \cdot 10^{-9}$
B_s^0 oscillations	$|y_{sb}|^2$, $|y_{bs}|^2$	$< 1.8 \cdot 10^{-6}$
	$|y_{sb}y_{bs}|$	$< 2.5 \cdot 10^{-7}$
K^0 oscillations	$\mathrm{Re}\, y_{ds}^2$, $\mathrm{Re}\, y_{sd}^2$	$[-5.9 \ldots 5.6] \cdot 10^{-10}$
	$\mathrm{Im}\, y_{ds}^2$, $\mathrm{Im}\, y_{sd}^2$	$[-2.9 \ldots 1.6] \cdot 10^{-12}$
	$\mathrm{Re}\, y_{ds}^* y_{sd}$	$[-5.6 \ldots 5.6] \cdot 10^{-11}$
	$\mathrm{Im}\, y_{ds}^* y_{sd}$	$[-1.4 \ldots 2.8] \cdot 10^{-13}$

Table 1. Current $95\,\%$ confidence level (CL) limits on the Wilson coefficient of eq. (2.2) from meson oscillations experiments. The reported bounds are taken from [20].

These interactions have numerous potentially observable effects. Firstly, they introduce additional decay modes for the Higgs boson, with a leading order partial width in the limit $m_{d^i}, m_{d^j} \ll m_h$ given by

$$\Gamma(h \to d^i d^j) = \frac{N_c m_h}{8\pi}\left(|y_{ij}|^2 + |y_{ji}|^2\right).\tag{2.3}$$

Secondly, they generate FCNC at tree-level, which are testable in K^0 and $B_{s,d}^0$ meson oscillation experiments. Flavour measurements impose stringent limits on the values of the Wilson coefficient of eq. (2.2), which are reported in table 1. These bounds in turn translate into upper limits on the exotic Higgs branching ratios. For example, assuming non-zero real Wilson coefficients only for the bs sector one gets $\mathcal{B}(h \to bs) < 1.8 \cdot 10^{-3}$ for $y_{bs} = y_{sb}$ and $\mathcal{B}(h \to bs) < 6.8 \cdot 10^{-3}$ with $y_{bs} \neq 0$ and $y_{sb} = 0$ or vice versa. However, given the unspecified nature of the ultraviolet structure of the model captured by the EFT framework, there could exist in principle additional effects giving rise to cancellations in the Feynman diagrams leading to mesons oscillations, still allowing for a sizable $h \to bs$ decay rate, see e.g. the recent [9]. For this reason we will consider the exotic Higgs branching ratio as a free parameters in the following.

3 Higgs physics at the ILC

One of the ILC's major physics goals is the precise measurement of the Higgs bosons' properties (for a review, see e.g. [21]). Thus, not only a deeper insight into the mechanism of EWSB will be acquired, but also NP that couples to the Higgs may be detected directly or indirectly. In particular, search channels suffering from large QCD backgrounds which are hard or impossible to see at the LHC may become easily accessible in the clean collider environment of the ILC. This is for example the case of the searches for the Higgs boson decaying into a pair of bottom quarks, for which only recently the ATLAS and CMS collaborations have reported a direct evidence of approximately $3.5\,\sigma$ [22, 23]. Needless to say, Higgs decays into a bottom and a light quark are even more challenging and possibly beyond the LHC reach.

In order to have a sufficient number of Higgs bosons to be investigated, all ILC operating scenarios include runs at a centre-of-mass energy of 250 GeV where the Higgs-strahlung production cross section, $e^+e^- \to Zh$, reaches its maximum of around 300 fb.[2] Hence, different event signatures are possible depending on the actual Z decay taking place. Of course, all of them come with their own advantages and drawbacks.

On the one hand, due to the large branching ratio $\mathcal{B}(Z \to \bar{q}q) \simeq 70\,\%$,[3] events with a hadronically decaying Z are the most abundant ones. Unfortunately, they are also the least clean ones. In particular, as in our case the signal process involves the Higgs decaying into quarks as well, the arising final state contains four jets. This signature comes along with a large SM background both due to processes where a Higgs boson decays into a pair of same flavour jets of which at least one flavour is mis-identified, and to four-jet production which, despite presenting a different final state kinematics, have a large production cross section. Still, we will investigate the potential of the hadronic search channel in section 4.2.

On the other hand, the Z may also decay into electron or muon pairs. These particles can be reconstructed and identified with good efficiencies. Thereby, it is possible to very reliably tag the produced Higgs via the recoil mass technique which allows to precisely measure the mass of the Higgs boson [18] and to reject backgrounds arising from processes without its resonant propagation. However, the corresponding Z branching ratio is roughly an order of magnitude smaller than in the hadronic case, $\mathcal{B}(Z \to \ell^+\ell^-) = \mathcal{B}(Z \to e^+e^-) + \mathcal{B}(Z \to \mu^+\mu^-) \simeq 6.7\,\%$, thus possibly limiting the statistical significance that can be reached with this search mode. The analysis of this channel is presented in section 4.3.1.

Finally, the Z can decay invisibly into neutrinos with a branching ratio of $\mathcal{B}(Z \to \bar{\nu}\nu) \simeq 20\,\%$. The signal signature then contains missing transverse energy[4] ($\not\!\!E_T$) that can be used to select the event along with the information from the Higgs decay products. However, the presence of missing energy implies that more background channels have to be taken into account since the $bj + \not\!\!E_T$ signature can also arise, e.g., through $\ell\nu cs$ final states where the lepton remains undetected and the charm quark is mis-identified. In any case the neutrino channel turns out to be a good compromise between event rate and cleanliness and will be covered in section 4.3.2.

All other Higgs production processes only give minor contributions at $\sqrt{s} = 250$ GeV. In contrast, at a centre-of-mass energy of 500 GeV most Higgs bosons are produced in W-fusion with a cross-section of approximately 160 fb. Obviously, this results in a signature very similar to the one from Higgs-strahlung with subsequent invisible Z decays, the cross-section of which is around 20 fb at this energy. We will investigate the reach of the neutrino search channel during the 500 GeV run in section 4.3.2.

_____

[2]All cross-sections quoted in this section are valid assuming polarised beams with $\mathcal{P}(e^-, e^+) = (-0.8, +0.3)$, see e.g. [18].

[3]For our analysis we adopt the Z branching ratios reported in [24].

[4]Given the precise knowledge about the colliding particles' energies, we could employ the total missing momentum instead of $\not\!\!E_T$, the latter being widely used in hadron collider studies. We expect, however, that due to the uncertainties resulting from a non-trivial beam energy spectrum, beamstrahlung and initial state radiation, this will not improve upon the results obtained.

4 Analysis and results

4.1 Analysis details

In order to perform a detector-level analysis, we first employed the `FeynRules` package [25] to implement the effective vertices of eq. (2.2) in the `UFO` format [26]. Events were then generated using `Whizard 2.5.0` [27] supplemented by the `O'Mega` matrix element generator [28]. In doing so, polarised beams and a non-trivial beam energy spectrum due to initial state radiation and beamstrahlung were taken into account. Results without the inclusion of the aforementioned effects were eventually cross checked with `MadGraph5_aMC@NLO` [29]. Parton showering and hadronisation were performed through `PYTHIA8` [30], while a fast detector simulation was carried out using `Delphes 3` [31] supplemented by the `DSiD` detector card suitable for performing analyses at e^+e^- linear colliders [32]. Jets were reconstructed with `FastJet` [33] via the `anti-kT` algorithm [34] with a cone radius of 0.4. Events have been finally analysed with `MadAnalysis5` [35].

For the purpose of accurately modeling the flavour identification of b and c jets at the ILC, we assumed the following tagging efficiencies

$$\mathcal{P}(b\text{-tag}\,|\,b) = 0.80\,, \quad \mathcal{P}(c\text{-tag}\,|\,c) = 0.70\,, \tag{4.1a}$$

and mis-identification rates

$$\begin{aligned} \mathcal{P}(b\text{-tag}\,|\,c) &= 0.08\,, \quad \mathcal{P}(b\text{-tag}\,|\,j) = 0.01\,, \\ \mathcal{P}(c\text{-tag}\,|\,b) &= 0.17\,, \quad \mathcal{P}(c\text{-tag}\,|\,j) = 0.10\,, \end{aligned} \tag{4.1b}$$

which are based on the expected performance of the `LCFIPlus` software [36]. Note that at the ILC it will not be possible to differentiate between jets from s and d quarks, rendering the Higgs decays to bs and bd effectively indistinguishable. Hence, our analysis is sensitive to the quantity $\mathcal{B}(h \to bj) := \mathcal{B}(h \to bs) + \mathcal{B}(h \to bd)$ rather than to the individual branching ratios.

One of the major advantages of a linear collider is the ability to polarise its particle beams. For the ILC, beam polarisations of $\mathcal{P}(e^-) = \pm 80\,\%$ and $\mathcal{P}(e^+) = \pm 30\,\%$ are expected. During our analyses we will, on the one hand, compare the performance of the four possible combinations and, on the other hand, we will consider two potential ways of how to realistically split the full integrated luminosity among the combinations in an actual run, namely [37]

$$\begin{aligned} \textbf{Scenario 1:} \quad & \mathcal{P}_{-+,+-,++,--} = [67.5\,\%, 22.5\,\%, 5\,\%, 5\,\%]\,, \\ \textbf{Scenario 2:} \quad & \mathcal{P}_{-+,+-,++,--} = [40\,\%, 40\,\%, 10\,\%, 10\,\%]\,, \end{aligned}$$

where the first (second) sign corresponds to the polarisation of the electron (positron) beam. Regarding the amount of accumulated data, we adopt the specifications of the H-20 scenario, i.e. we assume an integrated luminosity of $2\,\text{ab}^{-1}$ ($4\,\text{ab}^{-1}$) at a centre-of-mass energy of $250\,\text{GeV}$ ($500\,\text{GeV}$), with an initial phase at both energies collecting $500\,\text{fb}^{-1}$ each [37].

For later reference, we define the signal significance as

$$z = \frac{S}{\sqrt{S + B + \epsilon_{\mathrm{syst}}^2 B^2}} \, , \tag{4.2}$$

where S and B indicate the number of signal and background events, respectively, and ϵ_{syst} is the systematic uncertainty on the background determination, i.e. $\epsilon_{\mathrm{syst}} = \Delta B / B$. For given B and ϵ_{syst}, the 95 % CL upper limit on the number of signal events is then defined as the value $S_{95\%}$, for which eq. (4.2) gives $z = \Phi^{-1}(0.95) \approx 1.64$, where Φ^{-1} is the probit function.

The remaining sections will be divided in line with our discussion in section 3, according to signal signatures and centre-of-mass energies. Whenever an explicit rate for the signal process is of relevance, we choose as benchmark point $|y_{bs}| = |y_{sb}| = 10^{-3}$, which gives an exotic branching fraction of $\mathcal{B}(h \to bs) \simeq 0.73\,\%$ according to eq. (2.3).

4.2 Hadronic channel at 250 GeV

The SM background for the hadronic channel arises from the inclusive production of four jets, $e^+ e^- \to kkkk$, where k indicates gluons or any quark excluding the top. We have split this background into the following two contributions:

Non-resonant processes where the Higgs boson does not propagate in the Feynman diagrams (hereafter referred to as $4k$).

Resonant Higgs processes where the Higgs is produced on-shell in association with two quarks and then further decays into a pair of jets ($\bar{b}b, \bar{c}c, gg$). This class includes both Higgs-strahlung type processes with a hadronically decaying Z and topologies with quark pair production, where one quark radiates off a Higgs boson.

The above splitting has been performed in order to be able to rescale the final number of events in order to properly take into account the SM Higgs branching ratios as predicted by the Higgs cross section working group for a 125 GeV Higgs boson [38]. Note that this procedure removes the contribution to the background given by the interference amongst the two classes of processes which, however, we found to be negligible.

Signal events are selected with the following requirements. We ask for exactly four jets, with at least one b-tag jet and one light-flavour jet. All jets are supposed to satisfy $p_T > 10\,\mathrm{GeV}$ and $|\eta| < 1.8$. We further require that a b-tag jet and a light jet reconstruct the Higgs mass within $\Delta_{mh} := |m_{bj} - m_h| < 5\,\mathrm{GeV}$ and that two same-flavour jets, different from the ones identifying the Higgs mass, reconstruct the Z boson mass with $\Delta_{mZ} := |m_{kk} - m_Z| < 10\,\mathrm{GeV}$. In case that more than one pair falls into the chosen invariant mass window, we chose the pair closest to the true mass. We finally optimise the signal selection by asking the b-jet arising from the Higgs boson decay to have a transverse momentum greater than 50 GeV and a not too large angular separation from the light jet with which it reconstructs the Higgs mass: $\Delta R_{bj} < 2.6$ and $\Delta \phi_{bj} < 2.5$. These last two selection cuts have been imposed to further suppress the combinatorial $4k$ background. In figure 1, we
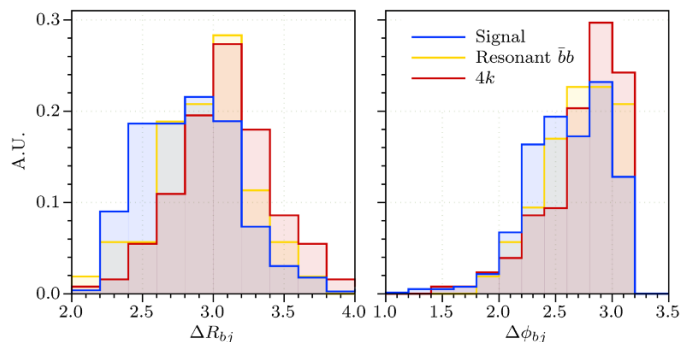
Figure 1. Normalised distributions of the ΔR and $\Delta \phi$ separation between the b jet and the light jet reconstructing the Higgs mass after the invariant mass cuts for $\sqrt{s} = 250\,\text{GeV}$ with polarisation $\mathcal{P}_{-+}$ for both the signal and the two main background processes.

	Signal	$4k$	$\bar{b}b$	$\bar{c}c$	gg
Exp.	2681	$2.8 \cdot 10^7$	$1.9 \cdot 10^5$	9490	$2.8 \cdot 10^4$
Jet tag.	1080	$1.9 \cdot 10^6$	$6.9 \cdot 10^4$	1150	3470
Δ_{mh}	147	$1.4 \cdot 10^5$	6730	87	223
Δ_{mZ}	21	6946	98	5	2
p_T^b	14	3503	81	3	1
$\Delta\phi_{bj}$	5	419	15	1	1
ΔR_{bj}	4	165	12	1	0

Table 2. Cutflow table for the signal and backgrounds in the hadronic channel at $250\,\text{GeV}$ with polarisation sharing scenario 1 and $2\,\text{ab}^{-1}$ of integrated luminosity. For the signal we have assumed $\mathcal{B}(h \to bj) \simeq 0.73\,\%$.

show the differential distributions for the two variables normalised to a unitary area after the invariant mass cuts which motivates our angular requirements.

For the above choices of selection cuts the event yields for the signal and background processes for polarisation scenario 1 and with an integrated luminosity of $2\,\text{ab}^{-1}$ are reported in table 2 for our benchmark point with $\mathcal{B}(h \to bs) \simeq 0.73\,\%$. Assuming a systematic background error of $1\,\%$, we then show in figure 2 the $95\,\%$ CL and $5\,\sigma$ contours in the luminosity-branching ratio plane for all the polarisation configurations individually as well as for the two sharing scenarios mentioned before. The hadronic channel will thus be able to exclude at $95\,\%$ CL branching ratios $\mathcal{B}(h \to bj)$ of around $5\,\%$ with $2\,\text{ab}^{-1}$ of accumulated data for both realistic polarisation sharing scenarios. A bound lower by a factor of two could in principle be obtained with the ILC running in the $\mathcal{P}_{+-}$ configuration up to the full targeted integrated luminosity at $\sqrt{s} = 250\,\text{GeV}$.

4.3 Leptonic channels

The previous section demonstrated that the reach of the hadronic channel is limited due to the large combinatorial four-jet background. If we want the signal signature to contain only a pair of jets there are two possibilities (cf. section 3). Either the final state involves a pair
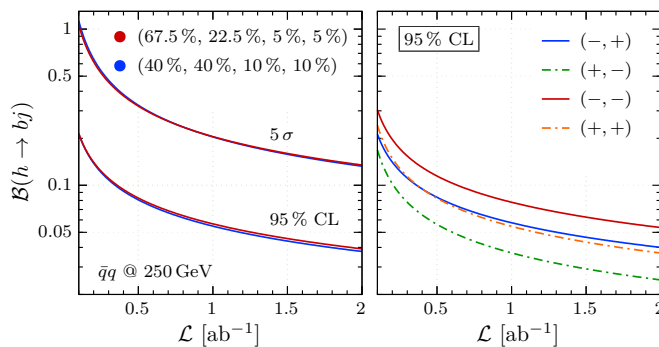
Figure 2. *Left*: expected 95 % CL exclusion and $5\,\sigma$ discovery reaches on $\mathcal{B}(h \to bj)$ in the hadronic channel at $\sqrt{s} = 250\,\text{GeV}$ as a function of the total integrated luminosity for the two polarisation scenarios described in the text. *Right*: expected exclusion reaches for the four possible polarisation configurations individually. Both plots assume $\epsilon_{\text{syst}} = 1\,\%$.

Process X	$\mathcal{B}(h \to X)$ [%]	ϵ_{tag} [%]	relative weight r
bj	0.73	71.2	1.0
bb	58.09	4.8	5.4
cc	2.884	3.52	0.20
gg	8.180	1.78	0.28

Table 3. Qualitative estimate of the irreducible backgrounds' relative importance for the leptonic search channels. Higgs branching ratios $\mathcal{B}$ are taken from [38], while tagging efficiencies ϵ_{tag} are calculated based on eq. (4.1).

of opposite-sign charged leptons from Z decays (*charged lepton channel*), or there is missing energy coming from a pair of neutrinos (*neutrino channel*). Clearly, both signatures come with different kinds of backgrounds, that we discuss in detail in the following subsections.

At this point, we only mention that the dominant SM background for both leptonic channels comes from on-shell Higgs production with its subsequent decay into a pair of same-flavour quarks. Due to flavour tagging imperfections, these processes can mimic the signal, while, in particular, exhibiting its identical kinematics. They are thus *de facto* irreducible and are only suppressed by flavour-tagging requirements. It is therefore simple, yet very instructive,[5] to estimate the relative importance of the individual background processes based only on the SM Higgs branching ratios and the flavour tagging specifications from eq. (4.1). The results are compiled in table 3, where the relative weight measure r is defined as

$$r_X = \frac{\mathcal{B} \cdot \epsilon_{\text{tag}}|_X}{\mathcal{B} \cdot \epsilon_{\text{tag}}|_{bj}}$$

with ϵ_{tag} being the probability that the given quark pair is tagged as bj. Consequently, the resonant $\bar{b}b$ process is expected to be by far the most severe background.

---

[5]For the hadronic channel such an estimate is less useful due to the large amount of combinatorial background.

Based on the above considerations, we can even obtain a qualitative estimate of the reachable statistical significance z given the expected number of signal events S and the sum of all background relative weights r_B, namely $z \approx \sqrt{S/(1 + r_B)}$. For instance, plugging in the numbers from table 3 yields $\sqrt{S}/2.6$. Clearly this is just a very rough upper bound for z due to several reasons including the fact that we took into account only the resonant background modes and ignored systematic errors. However, it helps us stress the importance of good flavour tagging for a successful analysis once more. Let us, e.g., assume the same specifications for b tagging as before, but a degraded c tagging efficiency of only $60\,\%$. Even though the mis-tag rates will then decrease accordingly (subject to the LCFIPlus performance [36]), redoing the above calculation shows a further enhancement of the $\bar{b}b$ background, $r_{bb} = 13.5$, with an expected worsened significance $z \approx \sqrt{S}/3.9$.

Note that the above estimates were obtained for a given fixed branching ratio $\mathcal{B}(h \to bj) = 0.73\,\%$. Alternatively, we can ask for the minimally achievable $95\,\%$ CL upper bound on $\mathcal{B}$. Again taking into account only resonant backgrounds and neglecting systematic errors, we obtain from eq. (4.2)

$$\mathcal{B}_{95\%} \gtrsim 1.3 \cdot \mathcal{B}_0/S_0 \left(1 + \sqrt{1 + 0.063 \cdot S_0/\mathcal{B}_0} \right), \tag{4.3}$$

where we used the SM Higgs branching ratios and tagging efficiencies of table 3. Notice that $\mathcal{B}_0$ is some reference signal branching fraction while S_0 is the corresponding number of observable events. Hence, the ratio $S_0/\mathcal{B}_0$ encodes the performance of a given analysis setup and is not sensitive to any BSM parameter. In section 4.3.2, we will compare the results of our actual analysis with the aforementioned estimate.

4.3.1 Charged lepton channel at $250\,\mathrm{GeV}$

The main SM background modes to the relevant signal process $e^+e^- \to Zh \to \ell^+\ell^- bs$ arise from the processes $e^+e^- \to \ell^+\ell^-\bar{b}b, \ell^+\ell^-\bar{c}c$ and $\ell^+\ell^- jj$ where the flavour of at least one final-state jet is mis-tagged and a bj combination is identified. Two classes of processes contribute to this background, namely

Resonant Higgs processes where the Higgs is produced in association with a leptonically decaying Z boson, through the same topology as the signal.

Non-resonant processes where the Higgs boson does not propagate in the Feynman diagrams, such as ZZ, $\gamma\gamma$ or $Z\gamma$ production.

As mentioned before, resonant backgrounds are the most problematic ones since they exhibit the same kinematics as the signal process. Hence, they are only suppressed by flavour-tagging requirements. Apart from the backgrounds listed above, charged-current processes like $e^+e^- \to \nu_\ell \ell sc$ could in principle also contribute, provided a second lepton that can arise for example from hadronic activity is detected. However, defining suitable isolation criteria[6] for the charged leptons and requiring them to have sufficiently large transverse momenta allows to reject this background very efficiently.

---

[6] We use a variant of the DSiD card that already implements electron, muon and photon isolation criteria. For details, see [32].

|	Signal	$\ell\ell\bar{b}b$	$\ell\ell\bar{c}c$	$\ell\ell jj$	$\nu_\ell \ell sc$	$\nu_\ell \ell du$
Exp.	225	$1.2 \cdot 10^5$	$1.4 \cdot 10^5$	$4.0 \cdot 10^5$	$5.2 \cdot 10^6$	$5.2 \cdot 10^6$
Jet tag.	82	4320	1904	2960	305	152
p_T^ℓ	79	3616	1552	2419	0	0
Δ_{mZ}	77	2784	1051	1746	0	0
$\Delta_{\rm recoil}$	71	1272	325	553	0	0
Δ_{mh}	51	247	34	66	0	0

Table 4. Cutflow table for the signal and relevant backgrounds in the charged lepton channel at 250 GeV with polarisation sharing scenario 1 and $2\,{\rm ab}^{-1}$ of integrated luminosity. For the signal we have assumed $\mathcal{B}(h \to bs) \simeq 0.73\,\%$.
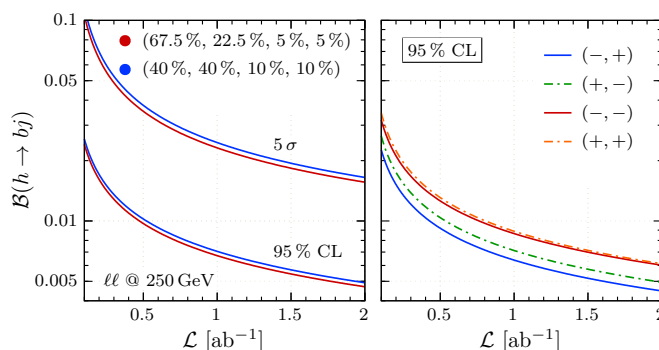
Figure 3. *Left*: expected $95\,\%$ CL exclusion and $5\,\sigma$ discovery reaches on $\mathcal{B}(h \to bj)$ in the charged lepton channel at $\sqrt{s} = 250\,{\rm GeV}$ as a function of the total integrated luminosity for the two polarisation scenarios described in the text. *Right*: expected exclusion reaches for the four possible polarisation configurations individually. Both plots assume $\epsilon_{\rm syst} = 1\,\%$.

In order to discriminate the signal from the remaining backgrounds, we adopt the following set of cuts. We ask for exactly one b jet and one light-flavour jet j which are supposed to both satisfy $p_T > 10\,{\rm GeV}$ and $|\eta| < 1.8$. The bj pair is required to have an invariant mass that fulfills $\Delta_{mh} < 30\,{\rm GeV}$. We further ask for exactly one pair of same-flavour and opposite-sign isolated leptons (electrons or muons) with $p_T > 10\,{\rm GeV}$ and $|\eta| < 2.44$ each. The lepton pair is additionally supposed to reconstruct the Z boson mass within $\Delta_{mZ} < 20\,{\rm GeV}$. Finally, we ask the recoil mass, defined via $m_{\rm rec}^2 := s - 2\sqrt{s}E_{\ell\ell} + m_{\ell\ell}^2$, to be within a $20\,{\rm GeV}$ mass window around the Higgs mass.

For these choices of selection cuts, the event yields for the charged lepton channel at $\sqrt{s} = 250\,{\rm GeV}$ are reported in table 4 assuming $2\,{\rm ab}^{-1}$ of accumulated data, $\mathcal{B}(h \to bs) \simeq 0.73\,\%$ for the signal and polarisation sharing of scenario 1. The exclusion and discovery reach on $\mathcal{B}(h \to bj)$ as a function of the integrated luminosity with an assumed $1\,\%$ systematic background error are shown in figure 3. The charged lepton channel will thus be able to exclude at $95\,\%$ CL values of $\mathcal{B}(h \to bj)$ down to around $0.5\,\%$, roughly one order of magnitude smaller than what can be obtained in the hadronic channel, with the ideal beam polarisation now being $\mathcal{P}_{-+}$.
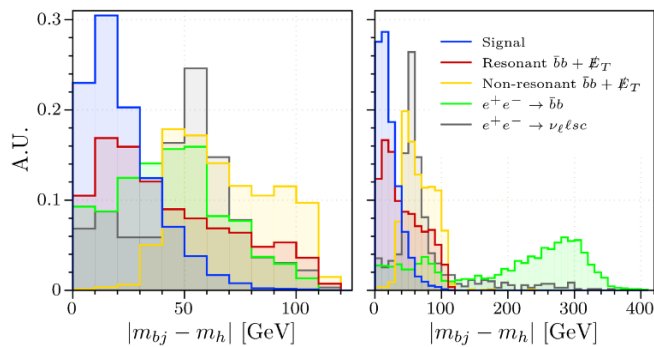
Figure 4. Normalised distributions of Δ_{mh} after the $\not{E}_T$ for the neutrino channel at $250\,\text{GeV}$ (left) and $500\,\text{GeV}$ (right) with polarisation $\mathcal{P}_{-+}$ for both the signal and background processes.

4.3.2 Neutrino channel at 250 GeV and 500 GeV

As outlined in section 3, the neutrino channel with final state $e^+e^- \to \bar{\nu}\nu bs$ is interesting at both the $250\,\text{GeV}$ and the $500\,\text{GeV}$ runs. The common search signature consists of exactly two jets plus missing transverse energy. Of the two jets, precisely one has to be b-tagged, while the other one must not receive any flavour tag. Since we found that the relevant SM backgrounds for this signature as well as the most important cuts to isolate the signal are very similar for both energies, we discuss the two cases largely in parallel.

Four classes of SM processes contribute significantly to the background and are listed in the following.

Resonant Higgs processes were already described in the beginning of section 4.3 and make up the largest contribution because they are only suppressed by flavour-tagging requirements, but are otherwise irreducible.

Non-resonant processes have the same final state as the resonant ones, but do not involve a Higgs decay. They may, for instance, originate from double-Z production.

Quark pair production ($e^+e^- \to \bar{b}b$, etc.) Uncertainties in the determination and/or reconstruction of the jets' four-momenta or other detector imperfections can lead to artificial missing energy. Although this background can be very efficiently reduced by a number of cuts (see below), it still contributes due to its large cross-section.

Charged-current processes ($e^+e^- \to \nu_\ell \ell sc$ etc.)
arise e.g. from double-W production and contribute when the charged lepton is missed in the detector.

All but the resonant processes do not involve Higgs decays, such that they can be efficiently rejected by the two-jet invariant mass cut, Δ_{mh}, see figure 4. Furthermore, requiring a significant amount of missing transverse energy helps to reduce the background from events with less than two neutrinos. We select signal events requiring exactly one b-tag jet and one light-flavour jet with $p_T > 10\,\text{GeV}$ and $|\eta| < 1.8$. We then veto on the presence of reconstructed isolated charged leptons. Further we require $\Delta_{mh} < 30\,\text{GeV}$ and

	$\sqrt{s} = 250\,\mathrm{GeV}$						$\sqrt{s} = 500\,\mathrm{GeV}$					
	Signal	$(\bar{\nu}\nu)\bar{b}b$	$(\bar{\nu}\nu)\bar{c}c$	$(\bar{\nu}\nu)jj$	$\nu_\ell\ell sc$	$\nu_\ell\ell du$	Signal	$(\bar{\nu}\nu)\bar{b}b$	$(\bar{\nu}\nu)\bar{c}c$	$(\bar{\nu}\nu)jj$	$\nu_\ell\ell sc$	$\nu_\ell\ell du$
Exp.	879	$2.4\cdot10^7$	$2.4\cdot10^7$	$7.2\cdot10^7$	$5.2\cdot10^6$	$5.2\cdot10^6$	2490	$7.5\cdot10^6$	$8.2\cdot10^6$	$2.3\cdot10^7$	$3.6\cdot10^6$	$3.6\cdot10^6$
Jet tag.	435	$1.0\cdot10^6$	$5.2\cdot10^5$	$8.6\cdot10^5$	$2.4\cdot10^5$	$66\,720$	1147	$1.6\cdot10^5$	$98\,378$	$1.6\cdot10^5$	$1.3\cdot10^5$	$36\,040$
No ℓ	429	$9.7\cdot10^5$	$5.1\cdot10^5$	$8.5\cdot10^5$	7316	2430	1137	$1.6\cdot10^5$	$97\,545$	$1.6\cdot10^5$	4092	1539
$\not{E}_T$	301	21\,151	11\,828	13\,108	4400	1208	1040	37\,247	20\,224	23\,285	3450	1217
Δ_{mh}	222	4506	2420	2898	935	402	695	4617	568	813	288	116
$\Delta\phi_{bj}$	201	1335	234	202	899	375	655	3946	299	522	270	107

Table 5. Cutflow table for the signal and backgrounds in the neutrino channel at $250\,\mathrm{GeV}$ with $2\,\mathrm{ab}^{-1}$ and polarisation sharing of scenario 1 and $500\,\mathrm{GeV}$ with $4\,\mathrm{ab}^{-1}$ and polarisation sharing of scenario 2. For the signal we have assumed $\mathcal{B}(h\to bs)\simeq 0.73\,\%$.

$\not{E}_T > 35\,\mathrm{GeV}$ for $\sqrt{s}=250\,\mathrm{GeV}$ and $\Delta_{mh} < 25\,\mathrm{GeV}$ and $\not{E}_T > 25\,\mathrm{GeV}$ for $\sqrt{s}=500\,\mathrm{GeV}$. We finally require that the azimuthal distance between the two jets satisfies $\Delta\phi_{bj} < 3.0$, which helps to reduce the background contribution from quark-pair final states for which the jets are mainly back-to-back. Note that the $\not{E}_T$ cut slightly differ for the two considered centre-of-mass energies. This is primarily caused by the differences in available energy as well as by the different kinematics of the two main production channels.

For these choices of selection cuts the event yields for the signal, with $\mathcal{B}(h\to bs)\simeq 0.73\,\%$, and the SM backgrounds are reported in table 5. For the $250\,\mathrm{GeV}$ run, polarisation sharing scenario 1 and $2\,\mathrm{ab}^{-1}$ of accumulated data were assumed, whereas we adopted scenario 2 and an integrated luminosity of $4\,\mathrm{ab}^{-1}$ for $\sqrt{s}=500\,\mathrm{GeV}$. The corresponding exclusion and discovery reach on $\mathcal{B}(h\to bj)$ as a function of the integrated luminosity for a $1\,\%$ systematic background error are displayed in figure 5. The results show that the neutrino channel at $\sqrt{s}=250\,\mathrm{GeV}$ is able to set a limit on $\mathcal{B}(h\to bj)$ slightly better than that obtained from the charged lepton channel at the same centre-of-mass energy. We finally see, that the best sensitivity of all search modes under consideration is offered by the neutrino channel at the $500\,\mathrm{GeV}$ run. To be precise, a $95\,\%$ CL upper limit on $\mathcal{B}(h\to bj)$ of approximately $0.2\,\%$ can be set at the end of the ILC operations. The optimal beam polarisation for the neutrino channels is found to be $\mathcal{P}_{-+}$.

Let us finally compare the obtained result of $\mathcal{B}_{95\%}\approx 2\cdot10^{-3}$ to the estimate in eq. (4.3). In a benchmark scenario with $\mathcal{B}_0 = 0.73\,\%$ and at a full $500\,\mathrm{GeV}$ run with only $\mathcal{P}_{-+}$ beams, we expect a total of 5256 signal events to be produced. Factoring in the assumed tagging efficiency from table 3, this number reduces to $S_0 = 3742$, giving a maximally achievable sensitivity of $\mathcal{B}_{95\%}\gtrsim 5\cdot10^{-4}$. Additionally taking into account particle reconstruction efficiencies, detector acceptance and a more realistic polarization sharing (but no further kinematical cuts) one ends up with $S_0 = 1137$ and $\mathcal{B}_{95\%}\gtrsim 9\cdot10^{-4}$. Note that in a realistic analysis, additional cuts would have to be applied in order to reduce non-resonant background contributions. The above estimate is already quite close to our full result thus showing that the performance of the neutrino channel is indeed limited by the resonant background modes and flavour tagging capabilities.
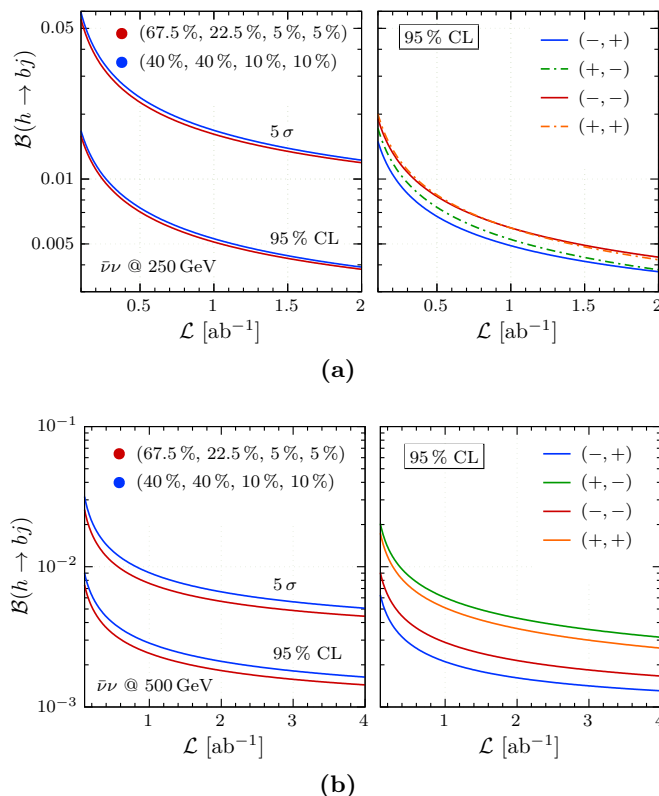
Figure 5. *(a) Left*: expected 95 % CL exclusion and 5 σ discovery reaches on $\mathcal{B}(h \to bj)$ in the neutrino channel at $\sqrt{s} = 250\,\text{GeV}$ as a function of the total integrated luminosity for the two polarisation scenarios described in the text. *(a) Right*: expected exclusion reaches for the four possible polarisation configurations individually. Both plots assume $\epsilon_{\text{syst}} = 1\,\%$. *(b)*: same as *(a)* but for $\sqrt{s} = 500\,\text{GeV}$.

5 Summary and conclusions

In this paper we have discussed the potential of a future e^+e^- collider in directly searching for quark flavour-violating Higgs decays involving a bottom quark and a light-flavour quark. As a benchmark machine, we chose the planned International Linear Collider (ILC). Focusing on the two main Higgs production modes at the ILC – the Higgs-strahlung and the W-fusion processes – we identified the most promising analysis channels for the 250 GeV run to be the ones involving a pair of neutrinos or charged leptons in the final state. Adopting standard ILC beam polarisation scenarios, we found that a dedicated analysis is expected to exclude at 95 % CL exotic Higgs branching ratios of the order of 0.5 % with $2000\,\text{fb}^{-1}$ of integrated luminosity, while a 5σ significance can be attained for order 1 % decay rates. In contrast, the hadronic final state will have a sensitivity reduced by a factor of ten due to the large combinatorial four-jet background.

At a centre-of-mass energy of 500 GeV with $4000\,\text{fb}^{-1}$ of accumulated data and by exploiting the final state with a pair of neutrinos, an upper limit on $\mathcal{B}(h \to bj)$ of approximately 0.2 % is expected, while a discovery will be possible if $\mathcal{B}(h \to bj)$ is greater than 0.5 %. Even the initial run with only $500\,\text{fb}^{-1}$ of integrated luminosity will be able to
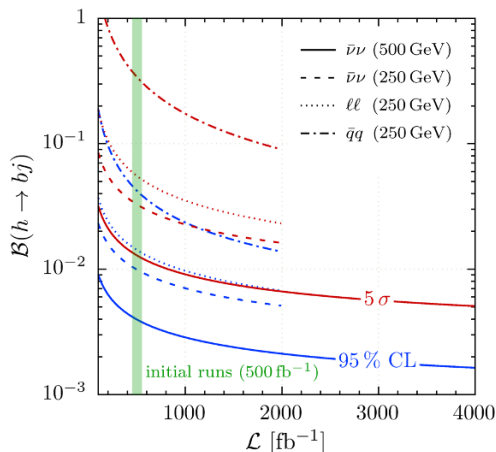
Figure 6. Comparison between the expected 95 % CL exclusion (blue) and 5 σ discovery (red) reaches on $\mathcal{B}(h \to bj)$ for the various channels as a function of total integrated luminosity specific for the given centre-of-mass energy. The polarisation sharing is scenario 1 and 2 for the channels at $\sqrt{s} = 250\,\mathrm{GeV}$ and $500\,\mathrm{GeV}$, respectively. The plot assumes $\epsilon_{\mathrm{syst}} = 1\,\%$.

test, at 95 % CL, branching ratios as small as 0.4 %. In particular, we demonstrated that our results are close to the fundamental bound on the sensitivity, which is determined by resonant background rates and flavour tagging capabilities only.

The outcome of our study is then summarised in figure 6 where both the 95 % CL exclusion (blue) and the 5 σ discovery (red) reaches on $\mathcal{B}(h \to bj)$ are illustrated for the four channels analysed in the paper showing only the polarisation scenario and integrated luminosity relevant for a given centre-of-mass energy.

In conclusion, our analysis has shown that the prospects for directly testing exotic Higgs decays at the ILC are indeed promising at both of its planned centre-of-mass energies, and that the projected limits from the full 500 GeV run are in the same ballpark as the ones that can be obtained from low-energy flavour measurements of the Wilson coefficient of the operators mediating flavour-changing neutral currents in the Higgs sector. These experiments provide, however, indirect probes of quark flavour-violating couplings to which direct tests at colliders are complementary.

Acknowledgments

References

[1] ATLAS collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett.* **B 716** (2012) 1 [arXiv:1207.7214] [INSPIRE].

[2] CMS collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys. Lett.* **B 716** (2012) 30 [arXiv:1207.7235] [INSPIRE].

[3] ATLAS and CMS collaborations, *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV* ATLAS-CONF-2015-044, (2015).

[4] S. Bejar, F. Dilme, J. Guasch and J. Solà, *Higgs boson flavor changing neutral decays into bottom quarks in supersymmetry*, *JHEP* **08** (2004) 018 [hep-ph/0402188] [INSPIRE].

[5] J.L. Diaz-Cruz and J.J. Toscano, *Lepton flavor violating decays of Higgs bosons beyond the standard model*, *Phys. Rev.* **D 62** (2000) 116005 [hep-ph/9910233] [INSPIRE].

[6] G.C. Branco, P.M. Ferreira, L. Lavoura, M.N. Rebelo, M. Sher and J.P. Silva, *Theory and phenomenology of two-Higgs-doublet models*, *Phys. Rept.* **516** (2012) 1 [arXiv:1106.0034] [INSPIRE].

[7] A. Crivellin, A. Kokulu and C. Greub, *Flavor-phenomenology of two-Higgs-doublet models with generic Yukawa structure*, *Phys. Rev.* **D 87** (2013) 094031 [arXiv:1303.5877] [INSPIRE].

[8] F.J. Botella, G.C. Branco, M. Nebot and M.N. Rebelo, *Flavour Changing Higgs Couplings in a Class of Two Higgs Doublet Models*, *Eur. Phys. J.* **C 76** (2016) 161 [arXiv:1508.05101] [INSPIRE].

[9] A. Crivellin, J. Heeck and D. Mueller, *Large $h \to bs$ in generic two-Higgs-doublet models*, arXiv:1710.04663 [INSPIRE].

[10] A. Arhrib, Y. Cheng and O.C.W. Kong, *Comprehensive analysis on lepton flavor violating Higgs boson to $\mu^{\mp}\tau^{\pm}$ decay in supersymmetry without R parity*, *Phys. Rev.* **D 87** (2013) 015025 [arXiv:1210.8241] [INSPIRE].

[11] K. Agashe and R. Contino, *Composite Higgs-Mediated FCNC*, *Phys. Rev.* **D 80** (2009) 075016 [arXiv:0906.1542] [INSPIRE].

[12] A. Azatov, M. Toharia and L. Zhu, *Higgs Mediated FCNC's in Warped Extra Dimensions*, *Phys. Rev.* **D 80** (2009) 035016 [arXiv:0906.1990] [INSPIRE].

[13] CMS collaboration, *Search for Lepton-Flavour-Violating Decays of the Higgs Boson*, *Phys. Lett.* **B 749** (2015) 337 [arXiv:1502.07400] [INSPIRE].

[14] CMS collaboration, *Search for lepton flavour violating decays of the Higgs boson to $\mu\tau$ and $e\tau$ in proton-proton collisions at $\sqrt{s} = 13$ TeV*, CMS-PAS-HIG-17-001 [INSPIRE].

[15] ATLAS collaboration, *Search for flavour-changing neutral current top quark decays $t \to Hq$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, *JHEP* **12** (2015) 061 [arXiv:1509.06047] [INSPIRE].

[16] CMS collaboration, *Search for top quark decays via Higgs-boson-mediated flavor-changing neutral currents in pp collisions at $\sqrt{s} = 8$ TeV*, *JHEP* **02** (2017) 079 [arXiv:1610.04857] [INSPIRE].

[17] T. Behnke et al., *The International Linear Collider Technical Design Report - Volume 1: Executive Summary*, `arXiv:1306.6327` [INSPIRE].

[18] H. Baer et al., *The International Linear Collider Technical Design Report - Volume 2: Physics*, `arXiv:1306.6352` [INSPIRE].

[19] B. Grzadkowski, M. Iskrzynski, M. Misiak and J. Rosiek, *Dimension-Six Terms in the Standard Model Lagrangian*, *JHEP* **10** (2010) 085 [`arXiv:1008.4884`] [INSPIRE].

[20] R. Harnik, J. Kopp and J. Zupan, *Flavor Violating Higgs Decays*, *JHEP* **03** (2013) 026 [`arXiv:1209.1397`] [INSPIRE].

[21] D.M. Asner et al., *ILC Higgs White Paper*, in *Proceedings, Community Summer Study 2013: Snowmass on the Mississippi (CSS2013): Minneapolis, MN, U.S.A., July 29 – August 6, 2013*, `arXiv:1310.0763` [INSPIRE].

[22] CMS collaboration, *Evidence for the decay of the Higgs Boson to Bottom Quarks*, CMS-PAS-HIG-16-044 [INSPIRE].

[23] ATLAS collaboration, *Evidence for the $H \to b\bar{b}$ decay with the ATLAS detector*, *JHEP* **12** (2017) 024 [`arXiv:1708.03299`] [INSPIRE].

[24] PARTICLE DATA GROUP collaboration, C. Patrignani et al., *Review of Particle Physics*, *Chin. Phys.* **C 40** (2016) 100001 [INSPIRE].

[25] 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].

[26] 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].

[27] W. Kilian, T. Ohl and J. Reuter, *WHIZARD: Simulating Multi-Particle Processes at LHC and ILC*, *Eur. Phys. J.* **C 71** (2011) 1742 [`arXiv:0708.4233`] [INSPIRE].

[28] M. Moretti, T. Ohl and J. Reuter, *O'Mega: An optimizing matrix element generator*, `hep-ph/0102195` [INSPIRE].

[29] 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].

[30] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159 [`arXiv:1410.3012`] [INSPIRE].

[31] 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].

[32] C.T. Potter, *DSiD: a Delphes Detector for ILC Physics Studies*, in *Proceedings, International Workshop on Future Linear Colliders (LCWS15): Whistler, B.C., Canada, November 2–6, 2015*, `arXiv:1602.07748` [INSPIRE].

[33] M. Cacciari, G.P. Salam and G. Soyez, *FastJet User Manual*, *Eur. Phys. J.* **C 72** (2012) 1896 [`arXiv:1111.6097`] [INSPIRE].

[34] M. Cacciari, G.P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, *JHEP* **04** (2008) 063 [`arXiv:0802.1189`] [INSPIRE].

[35] E. Conte, B. Fuks and G. Serret, *MadAnalysis 5, A User-Friendly Framework for Collider Phenomenology*, [*Comput. Phys. Commun.* **184** (2013) 222](https://doi.org/) [[arXiv:1206.1599](https://arxiv.org/abs/1206.1599)] [[INSPIRE]].

[36] T. Suehara and T. Tanabe, *LCFIPlus: A Framework for Jet Analysis in Linear Collider Studies*, [*Nucl. Instrum. Meth.* **A 808** (2016) 109](https://doi.org/) [[arXiv:1506.08371](https://arxiv.org/abs/1506.08371)] [[INSPIRE]].

[37] T. Barklow et al., *ILC Operating Scenarios*, [arXiv:1506.07830](https://arxiv.org/abs/1506.07830) [[INSPIRE]].

[38] LHC Higgs Cross section Working Group collaboration, D. de Florian et al., *Handbook of LHC Higgs Cross sections: 4. Deciphering the Nature of the Higgs Sector*, [arXiv:1610.07922](https://arxiv.org/abs/1610.07922) [[INSPIRE]].