Taming a leading theoretical uncertainty in HH measurements via accurate simulations for $\mathrm{b\bar{b}H}$ production

Stefano Manzoni,[a] **Elena Mazzeo,**[b] **Javier Mazzitelli,**[c] **Marius Wiesemann**[d] **and Marco Zaro**[b]

[a] *CERN, CH-1211 Geneva 23, Switzerland*

[b] *Università degli Studi di Milano & INFN, Sezione di Milano,*
Via Celoria 16, 20133 Milano, Italy

[c] *Paul Scherrer Institut, 5232 Villigen PSI, Switzerland*

[d] *Max-Planck-Institut für Physik, Föhringer Ring 6, 80805 München, Germany*

E-mail: stefano.manzoni@cern.ch, elena.mazzeo@cern.ch,
javier.mazzitelli@psi.ch, marius.wiesemann@mpp.mpg.de,
marco.zaro@mi.infn.it

ABSTRACT: We present a new simulation for Higgs boson production in association with bottom quarks ($b\bar{b}H$) at next-to-leading order (NLO) accuracy matched to parton showers in hadronic collisions. Both contributions, the standard one proportional to the bottom-quark Yukawa coupling and the loop-induced one proportional to the top-quark Yukawa coupling from the gluon-fusion process, are taken into account in a scheme with massive bottom quarks. Therefore, we provide the full simulation of the $b\bar{b}H$ final state in the Standard Model, which constitutes also a crucial background to measurements for Higgs-boson pair (HH) production at the Large Hadron Collider when at least one of the Higgs bosons decays to bottom quarks. So far, the modeling of the $b\bar{b}H$ final state induced one of the dominant theoretical uncertainties to HH measurements, as the gluon-fusion component was described only at the leading order (LO) with uncertainties of $\mathcal{O}(100\%)$. Including NLO corrections in its simulation allows us to reduce the scale dependence to $\mathcal{O}(50\%)$ so that it becomes subdominant with respect to other systematic uncertainties. As a case study, we provide an in-depth analysis of the $b\bar{b}H$ background to HH measurements with realistic selection cuts in the $2b2\gamma$ channel. We also compare our novel simulation with the currently-employed ones, discussing possible issues and shortcomings of a scheme with massless bottom quarks. Finally, we propagate the effect of the new $b\bar{b}H$ simulation to HH searches in the $2b2\gamma$ and $2b2\tau$ final states, and we find an improvement of up to 10% (20%) on the current (HL-LHC) limits on $\sigma_{\mathrm{SM}}^{HH}$.

KEYWORDS: Bottom Quarks, Higgs Production, Higher-Order Perturbative Calculations

ARXIV EPRINT: 2307.09992

Contents

1 Introduction

The measurement of the properties of the Higgs boson (H), discovered a decade ago [1, 2] at the Large Hadron Collider (LHC), is one of the major quests of the particle physics community. Being naturally the least explored sector of the Standard Model (SM), its characterization is of utmost importance in the search for new-physics phenomena. Current measurements of its coupling to top (t) and bottom (b) quarks, W and Z bosons, and tau leptons show a picture fully consistent with the SM expectation [3, 4]. With the continuous increase of data taken at the LHC, those measurements will progressively become more precise, improving their sensitivity to deviations with respect to the SM. On the other hand, further couplings will become accessible, which are currently restricted due to large statistical uncertainties. First and foremost, this will be the case for the self interaction of the Higgs boson.

Through the mechanism of spontaneous electroweak symmetry breaking, the Higgs field generates not only the masses of the other particles in the SM, but also its own mass. Crucial to this mechanism is the shape of the Higgs boson potential, which after symmetry breaking induces a self interaction for the physical Higgs boson. In the SM the strength of this self interaction, λ_{HHH}, is fully determined by the Higgs mass and its vacuum expectation value. An experimental determination of this coupling will therefore provide an essential verification whether this fundamental prediction of the SM is realized in nature, and it will be our first exploration of the Higgs boson potential. Indeed, several new-physics models suggest a modification of the Higgs boson potential and the strength of the self interaction, so that any deviation from the SM expectation would be a clear signal of new physics.

The main avenue for the extraction of the Higgs self interaction is the measurement of Higgs boson pair (HH) production (for a review, see ref. [5]). At the LHC, the main production mechanism is through gluon fusion, similarly to the case of single-Higgs production.
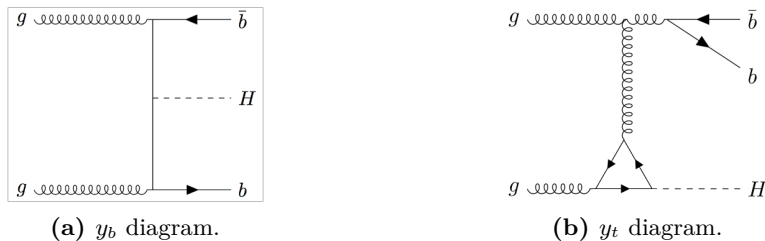
(a) y_b diagram.
    (b) y_t diagram.

Figure 1. Sample Feynman diagrams for $b\bar{b}H$ production proportional to y_b at tree-level and proportional to y_t induced by a top-quark loop.

Its measurement is extremely challenging, though, due to the very small production cross section of two Higgs bosons, which is three orders of magnitude smaller than for a single Higgs boson. For example, the most recent measurement from the ATLAS experiment [6] limits the production cross section to be below 2.4 times the SM prediction ($\sigma_{\rm SM}^{HH}$) in the combination of all HH search channels, and the self-coupling (by performing an exclusive λ_{HHH} variation) is bound to the range $-0.6 < \lambda_{HHH} < 6.6$, both at 95% confidence level. With the increased volume of data at the end of the High-Luminosity phase of the LHC (HL-LHC), the Higgs-pair production cross section is expected to be measured with a significance of 4.9σ (3.4σ) [7], when combining all search channels. The expected limits on the cross section will be reduced to $0.39(0.55) \times \sigma_{\rm SM}^{HH}$, and the determination of the Higgs self coupling is expected to be within $0.3(0.0) < \lambda_{HHH} < 1.9(2.5)$, again at 95% level, where the numbers shown in and outside the brackets include different assumptions on the systematic uncertainties [7]. Analogous results have been published by CMS [4, 8–11].

Accurate theoretical predictions are crucial for the prospects of measuring Higgs boson pair production, and, more importantly, for the subsequent extraction of limits on $\sigma_{\rm SM}^{HH}$ as well as λ_{HHH}. An impressive amount of effort has been devoted to the development of precise theoretical tools to predict the HH cross section (see refs. [12–30] and references therein), notably complicated by its loop-induced nature. A precise description of the signal, however, is not enough, and a good theoretical control over the relevant backgrounds of HH measurements is equally important, as it is necessary not to lose significance in the signal extraction.

Given the very small HH production rate, all of the phenomenologically relevant decay modes have (at least) one Higgs boson decaying into a pair of bottom quarks, since it has the highest branching ratio for a SM-like Higgs boson. As a result, Higgs boson production in association with a bottom quark-antiquark pair ($b\bar{b}H$) is one of the irreducible backgrounds to all of these HH searches at the LHC.

The $b\bar{b}H$ process receives contributions proportional to the bottom Yukawa coupling (y_b) where the Higgs couples to a bottom-quark line, see figure 1 (a), as well as contributions proportional to the top Yukawa coupling (y_t) where the Higgs boson couples to a closed top-quark loop, see figure 1 (b). The latter corresponds to the gluon-fusion process with an additional $b\bar{b}$ pair generated through QCD radiation. Both mechanisms have comparably large production rates (with the y_t^2 contribution actually being dominant, see ref. [31]). The

QCD perturbative corrections are particularly large for both the y_b^2 contributions and the y_t^2 contributions, so that leading-order (LO) results do not provide an adequate description. We would like to point out that the $b\bar{b}H$ final state can be produced via other production modes (VH associated production with $V \to b\bar{b}$ and $b-$associated vector-boson fusion), with an impact on the cross section at the level of some/several tenths of percents [32, 33]. However, accurate simulations for these production channels exist since long ago, and the inclusion of their contribution is thus trivial.

Since the bottom mass is neither particularly small nor particularly large compared to the typical scale of the $b\bar{b}H$ process, suitable predictions can be obtained either in a four-flavour scheme (4FS) with massive bottom quarks or in a five-flavour scheme (5FS) where the bottom quark is treated as being massless. In the 5FS, where the calculations are technically much simpler, significant progress has been made over the past years for the y_b^2 contribution [34–54] with the cross sections at the third order in the strong coupling being the most remarkable advancement. As far as 4FS calculations are concerned, there has been significantly less progress in higher-order calculations for the y_b^2 contribution [32, 48, 49, 55–61], given the much more involved structure of the LO process, with next-to-LO (NLO) QCD corrections (matched to parton showers) and combined with NLO electroweak ones, still being the state-of-the-art. By contrast, a consistent study of both $b\bar{b}H$ production modes (y_b^2 and y_t^2) has been presented only in the 4FS by including NLO QCD corrections in all relevant coupling structures (y_b^2, y_t^2 and $y_b y_t$ interference contributions) [31], and it was shown that the NLO corrections to the y_t^2 contribution are similarly large as the LO y_t^2 (and y_b^2) cross section. A substantial effort was also made in understanding differences between 4FS and 5FS cross sections, see e.g. refs. [62, 63], as well as in obtaining consistent combinations of the two schemes [64–68]. The advantages of either scheme in the context of simulating $b\bar{b}H$ production have been discussed in detail in ref. [48]. Throughout this paper we employ the 4FS, owing to its better description of differential observables related to final-state bottom quarks and the definition of bottom-flavoured jets.

In the present paper, we build upon the NLO calculation of ref. [31] and match both y_b^2 and y_t^2 contributions to parton showers (PS) using the MADGRAPH5_AMC@NLO framework. We perform a detailed study at NLO, including PS-matching uncertainties, of the $b\bar{b}H$ background to Higgs boson pair production and show that including NLO corrections to the y_t^2 cross section allows us to tame the corresponding theoretical uncertainty in HH measurements. To this end, we focus on the $2b2\gamma$ final state, and consider realistic selection cuts that are typically used in corresponding HH analyses. We compare our results to the size of the HH signal, as well as to the LO approximation of the y_t^2 terms based on the NNLOPS generator for inclusive Higgs boson production [69, 70], which is currently employed by the experimental analyses. Finally, we propagate our improved $b\bar{b}H$ modeling to HH searches performed by the ATLAS collaboration in the $2b2\gamma$ and $2b2\tau$ final states [71, 72], and we find an improvement of up to 10% (20%) on the current (HL-LHC) limits on $\sigma_{\mathrm{SM}}^{HH}$.

The paper is organized as follows: we describe our calculation in section 2. In section 3 phenomenological results for $b\bar{b}H$ production are discussed in the context of HH measurements. In particular, we start by describing the considered setup (section 3.1), then we

study predictions for integrated and fiducial cross sections (section 3.2), we analyze differential distributions (section 3.3), we compare our new $b\bar{b}H$ NLO+PS simulation in the 4FS with the NNLOPS prediction and describe how to consistently combine them (section 3.4), and finally we study the impact of the new $b\bar{b}H$ modeling on experimental HH searches (section 3.5). The paper is summarized in section 4.

2 Outline of the calculation

The calculation of the full NLO QCD corrections to $b\bar{b}H$ production has been discussed in detail in ref. [31], which we refer to at this point. Here, we briefly recall only the most relevant aspects.

The LO contribution to $b\bar{b}H$ production in the 4FS starts at $\mathcal{O}(\alpha_s^2)$ in QCD perturbation theory, and is mediated by the bottom-quark Yukawa coupling. Hence, the coupling structure of the LO process is $y_b^2 \alpha_s^2$. A sample Feynman diagram is shown in figure 1 (a). The squared gluon-fusion diagram in figure 1 (b), where the Higgs boson couples to a closed top-quark loop and which thus yields a contribution proportional to y_t^2, enters only at $\mathcal{O}(\alpha_s^4)$, i.e. it is formally only a next-to-NLO (NNLO) QCD correction to the LO y_b^2 contribution. However, since this subprocess is enhanced by y_t^2/y_b^2 with respect to the LO process proportional to y_b^2, it involves a (potentially enhanced) $g \to b\bar{b}$ splitting, and it receives large NLO QCD corrections (considerably larger than the y_b^2 one [31]), the y_t^2 contribution actually yields the dominant share to the $b\bar{b}H$ cross section in the SM. Interference contributions between the two processes exist in the 4FS, and they are formally part of the NLO QCD corrections to the y_b^2 induced LO process and thus enter at $\mathcal{O}(y_b y_t \alpha_s^3)$. However, when compared to the y_b^2 and y_t^2 contributions, the size of the interference is much smaller, with effects at the few-to-ten percent level.

Due to their different coupling structures, y_b^2, $y_b y_t$, and y_t^2 terms define three separately gauge-invariant contributions under higher-order QCD corrections. As a result, we can compute the NLO QCD corrections for each of the three contributions individually and decompose the $b\bar{b}H$ cross section up to $\mathcal{O}(\alpha_s^5)$, i.e. formally next-to-NNLO (N³LO) w.r.t. the LO process, as follows:

$$\mathrm{d}\sigma = y_b^2 \alpha_s^2 \left(\Delta_{y_b^2}^{(0)} + \alpha_s \Delta_{y_b^2}^{(1)} \right) + y_t y_b \alpha_s^3 \left(\Delta_{y_b y_t}^{(0)} + \alpha_s \Delta_{y_b y_t}^{(1)} \right) + y_t^2 \alpha_s^4 \left(\Delta_{y_t^2}^{(0)} + \alpha_s \Delta_{y_t^2}^{(1)} \right) . \quad (2.1)$$

Since the NLO corrections to the contributions proportional to y_t require the calculation of two-loop $2 \to 3$ amplitudes with internal massive fermion lines, which are beyond current technology, the approximation of a heavy top quark is employed to derive those contributions at NLO. The effective field theory resulting from integrating out the top quark is well known, it has been employed for several processes and it has been implemented in a model [73, 74] in the Universal Feynrules Output (UFO) format [75], so that it can be employed through automated Monte Carlo tools, such as MADGRAPH5_AMC@NLO [76, 77].

The validity of the heavy top approximation for $b\bar{b}H$ production has been studied and discussed at length in ref. [31] for the y_t^2 contribution. In fact, the case at hand corresponds essentially to Higgs+jet ($H+g$) production in gluon fusion with a $g \to b\bar{b}$ splitting either

in the initial or in the final state. For Higgs+jet production it has been shown in various publications, see for instance refs. [78–82], that the heavy top limit provides an excellent approximation to the exact cross section, especially for the computation of radiative corrections, and even well beyond the naively expected validity range, as ref. [82] has shown quite impressively. Also for the y_t^2 contribution to the $b\bar{b}H$ process, ref. [31] comes to the conclusion that at LO the heavy top-mass approximation works extremely well, with differences from the full computation below 10% as long as the probed momentum scales (Higgs or leading b-jet transverse momentum, or invariant mass of the b-jet pair) do not exceed 200 GeV. Given the relatively small impact, we refrain from performing the Born-improved (BI) approximation of the y_t^2 contribution here, which rescales the effective-field-theory result with the exact top-mass dependence at LO, since the mass effects are expected to be far below the perturbative uncertainties in the signal region under consideration [31].

We have computed all NLO QCD corrections corresponding to the decomposition in eq. (2.1) using the MADGRAPH5_AMC@NLO framework, including the matching to PS. Events for the y_b^2, $y_b\,y_t$, and y_t^2 contributions can be either generated simultaneously or separately for each contribution, with the latter option resulting in a better numerical convergence. Given that, as it has already been mentioned, the interference contribution is comparably small (about -5% to -10%), and given the relatively large NLO scale uncertainties of the y_b^2 and y_t^2 contributions, we deem the $y_b\,y_t$ contribution not relevant to the purpose of this paper and we refrain from including it here. We stress however that the $y_b\,y_t$ interference can be computed/turned on within our implementation of the $b\bar{b}H$ generator in MADGRAPH5_AMC@NLO.

3 Phenomenological results

3.1 Setup

We have used the same settings as in ref. [31], which allowed us to validate our new NLO+PS implementation against the previous NLO results. In particular, we consider proton-proton collisions with 13 TeV centre-of-mass energy and set following values for the on-shell masses:

$$m_b = 4.92\,\text{GeV}, \quad m_t = 172.5\,\text{GeV}, \quad m_H = 125\,\text{GeV}. \tag{3.1}$$

While also y_t is renormalized on-shell using the top-mass value above, for y_b we use the $\overline{\text{MS}}$ scheme, evaluated from an input value $m_b(m_b) = 4.18\,\text{GeV}$ and evolved to the respective central renormalization scale (μ_R^0) with a four-loop running. From that value $y_b(\mu_R^0)$ scale variations are determined with a two-loop evolution, consistent with the order of our calculation. This procedure follows the recommendation by the LHC Higgs cross section working group [83]. For the parton densities we use the NLO set of NNPDF3.1 [84] with 4 flavours and the corresponding two-loop running of α_S (LHAID=320500). The central factorization and renormalization scales are set to

$$\mu_R^0 = \mu_F^0 = \frac{H_T}{4} = \frac{1}{4}\sum_{i\,\in\,\text{final}}\sqrt{m^2(i) + p_T^2(i)}, \tag{3.2}$$

Contr.	LO	$\delta\mu_{R,F}$	NLO	$\delta\mu_{R,F}$
y_b^2	247	$^{+54\%}_{-33\%}$	374	$^{+18\%}_{-20\%}$
y_t^2	289	$^{+69\%}_{-38\%}$	689	$^{+61\%}_{-35\%}$
sum	536	$^{+62\%}_{-35\%}$	1064	$^{+46\%}_{-29\%}$

Table 1. Inclusive cross section for $pp \to b\bar{b}H$. Numbers are in fb.

where the sum runs over all final state particles and $m(i)$, $p_T(i)$ are their respective mass, transverse momentum. The associated scale uncertainties are determined through the customary 9-point envelope obtained by varying independently the two scales up and down by a factor 2. The shower scale is chosen with a lower reference value $Q_{\rm sh} = \frac{H_T}{4}$ than the default one in MadGraph5_aMC@NLO,[1] as suggested in ref. [48]. Here, we also study shower-scale variations around the central value by a factor of two up and down, which are quoted separately from those associated to the renormalization and factorization scales.

We have generated separately the contributions proportional to y_b^2 and y_t^2, while, as pointed out before, we neglect the $y_b y_t$ interference, which has a subleading numerical impact, well within the scale uncertainties. For reference, we quote the total inclusive cross section in table 1 separated by y_b^2, y_t^2 and their sum. As already observed in ref. [31], the NLO QCD corrections in all cases are substantial, and they lead to a reduction of the scale uncertainties. For the y_b^2 contribution the correction amounts to more than $+50\%$, while for the y_t^2 contributions we observe effects as large as $+140\%$ at NLO QCD, rendering LO predictions completely unreliable.[2]

As a representative case of HH searches, we consider here one of the most sensitive search channels, where one Higgs boson decays into bottom quarks and the other decays into photons, assuming a $H \to \gamma\gamma$ branching ratio of $\mathrm{BR}(H \to \gamma\gamma) = 0.227\%$ [85]. As far as the fiducial setup is concerned, throughout this paper, we consider selection cuts motivated by a recent HH search by the ATLAS collaboration [71]. In particular, we select anti-k_T [86] jets as implemented in FastJet [87] with $R = 0.4$ and define bottom-flavoured jets (short b-jets) as those containing at least one B hadron with the requirements

$$p_T(j) > 25\,\mathrm{GeV} \qquad \text{and} \qquad |\eta(j)| < 2.5\,, \tag{3.3}$$

assuming a b-tagging efficiency of 100% and without mistagging for our theoretical study. The HH signal region is defined by requiring (exactly) two b-jets and two photons (the QED shower is disabled in our simulations). The invariant mass of the b-jet pair is required to be within

$$80\,\mathrm{GeV} < m(b_1, b_2) < 140\,\mathrm{GeV}\,.[3] \tag{3.4}$$

---

[1]The setup employed in this paper corresponds to setting `shower_scale_factor=0.5` in the `run_card`.

[2]We note that, while the same setup as ref. [31] was employed in our work, the quoted numbers are different, because of an update of the 4-flavour NNPDF3.1 set [84], see also footnote 6 on page 11 in ref. [31].

[3]We have applied this cut instead of the full discriminant employed by the analysis of ref. [71], since this invariant-mass region provides the highest score for the HH signal.

Notice that for the distribution in the number of b-jets this requirement is lifted. The two photons must satisfy the following relations:

$$105\,\text{GeV} < m(\gamma_1, \gamma_2) < 160\,\text{GeV}, \quad |\eta(\gamma_i)| < 2.37, \quad \frac{p_T(\gamma_1)}{m(\gamma_1, \gamma_2)} > 0.35, \quad \frac{p_T(\gamma_2)}{m(\gamma_1, \gamma_2)} > 0.25\,. \tag{3.5}$$

In practice, since no detector effects are applied and no QED shower is included, we always have $m(\gamma_1, \gamma_2) - m_H = \mathcal{O}(\Gamma_H)$, so that the first requirement is trivially fulfilled.

Besides the above set of cuts, which we will refer to as *fiducial cuts*, we define the variables

$$m_{2b2\gamma} = m(b_1, b_2, \gamma_1, \gamma_2)\,, \tag{3.6}$$

and

$$m_{2b2\gamma}^{\star} = m_{2b2\gamma} - m(b_1, b_2) - m(\gamma_1, \gamma_2) + 2m_H\,, \tag{3.7}$$

and we consider three possible categories for cuts on $m_{2b2\gamma}^{\star}$:

$$m_{2b2\gamma}^{\star} < \infty, \qquad m_{2b2\gamma}^{\star} < 500\,\text{GeV}, \qquad m_{2b2\gamma}^{\star} < 350\,\text{GeV}\,. \tag{3.8}$$

Thus, the first scenario corresponds to the fiducial cuts, and the others apply increasingly stronger requirements on $m_{2b2\gamma}^{\star}$. These three regions provide complementary information on the Higgs potential. In particular, the region close to threshold has an enhanced sensitivity to the trilinear Higgs coupling.

In the presentation of our phenomenological results we will compare to reference predictions for both the $y_t^2\, b\bar{b}H$ background and the HH signal. The former is obtained from the NNLOPS generator for inclusive Higgs boson production [69, 70], which is currently used to model the $y_t^2\, b\bar{b}H$ background by the experiments, and it is formally LO+PS accurate for that contribution. To this end, we have followed closely the corresponding simulation employed by ATLAS [88]. The NNLOPS generator merges 0 and 1-jet multiplicities at NLO QCD using the MiNLO′ [69] method and then reaches NNLO QCD accuracy through reweighting to NNLO rapidity distribution. In addition, the sample is normalized to the reference gluon-fusion Higgs production cross section [83], which includes the N³LO corrections [89–92]. The renormalisation and factorisation scales in the NNLOPS calculation are set to $\mu_R = \mu_F = m_H/2$, the PDF4LHC15 [93] parton densities are used, and PYTHIA8 [94] is employed to perform the parton showering and to include the decay of the Higgs boson to two photons.

The HH signal is simulated at NLO QCD including the full top mass effects [95], using as reference value for the renormalisation and factorisation scales $\mu_R = \mu_F = m_{HH}/2$ and PDF4LHC15 as parton densities. The calculation is matched to PYTHIA8 to include parton showering and to include the decay of the two Higgs bosons to bottom quarks and photons. The HH signal cross section is normalized to the value $\sigma_{\text{ggF}}^{\text{SM}}(pp \to HH) = 31.0^{+2.1}_{-7.2}$ fb from ref. [21], which is obtained by combining the full NLO QCD calculation with the NNLO corrections obtained in the heavy-top limit and improved by including partial finite top mass effects via a suitable reweighting of the scattering amplitudes.

Cut	Contr.	Run	LO	NLO	$\delta\mu_{R,F}$	δQ_{sh}	NNLOPS (y_t^2 LO)	HH signal
No cut	y_b^2	PY8	561	849	+18% −20%	+0% +0%	4867	82.1
		PY8-Δ		848		+0% +0%		
		HW7	561	851		+0% +0%		
	y_t^2	PY8	655	1565	+61% −35%	+0% +0%	$g\not\to b\bar{b}$: 2140	
		PY8-Δ		1595		+0% +0%		
		HW7	655	1578		+0% +0%		
	sum	PY8	1217	2414	+46% −29%	+0% +0%		
		PY8-Δ		2443		+0% +0%		
		HW7	1216	2429		+0% +0%		
Fid. cuts	y_b^2	PY8	3.15	4.22	+15% −15%	+10% −4%	29.9	22.7
		PY8-Δ		4.75		+0% −2%		
		HW7	2.59	4.08		+8% −12%		
	y_t^2	PY8	8.24	18.1	+58% −34%	+10% −7%	$g\not\to b\bar{b}$: 17.2	
		PY8-Δ		19.2		+3% −1%		
		HW7	6.83	16.6		+4% −5%		
	sum	PY8	11.4	22.3	+50% −30%	+10% −6%		
		PY8-Δ		23.9		+2% −1%		
		HW7	9.42	20.7		+4% −6%		
Fid. cuts $+ m^{\star}_{2b2\gamma} < 500\,\mathrm{GeV}$	y_b^2	PY8	3.11	4.15	+15% −15%	+11% −4%	22.3	15.7
		PY8-Δ		4.69		+0% −2%		
		HW7	2.56	4.02		+8% −13%		
	y_t^2	PY8	5.33	12.3	+60% −34%	+12% −8%	$g\not\to b\bar{b}$: 13.3	
		PY8-Δ		12.8		+2% −1%		
		HW7	4.31	11.3		+5% −5%		
	sum	PY8	8.44	16.5	+49% −29%	+12% −7%		
		PY8-Δ		17.5		+1% −1%		
		HW7	6.86	15.3		+6% −7%		
Fid. cuts $+ m^{\star}_{2b2\gamma} < 350\,\mathrm{GeV}$	y_b^2	PY8	2.71	3.65	+15% −16%	+9% −4%	11.5	2.84
		PY8-Δ		4.11		+0% −2%		
		HW7	2.22	3.54		+8% −15%		
	y_t^2	PY8	2.32	5.78	+61% −34%	+13% −9%	$g\not\to b\bar{b}$: 6.82	
		PY8-Δ		6.05		+1% +0%		
		HW7	1.88	5.43		+5% −3%		
	sum	PY8	5.03	9.43	+44% −27%	+12% −7%		
		PY8-Δ		10.2		+0% +0%		
		HW7	4.10	8.97		+6% −8%		

Table 2. Rates for the process $pp \to b\bar{b}H$ with $H \to \gamma\gamma$ decay. Numbers are in ab. The NNLOPS predictions labeled with $g\not\to b\bar{b}$ are obtained vetoing $g \to b\bar{b}$ splittings in the parton shower.

3.2 Fiducial rates

In this section, we study in detail the relevance of the accurate modeling of the $b\bar{b}H$ production as a background to HH searches. We start by discussing total and fiducial rates for $b\bar{b}H$ production including the decay $H \to \gamma\gamma$ in table 2. In particular, we show LO and NLO QCD predictions for the y_b^2 and y_t^2 contributions as well as their sum for the fully inclusive cross section and in the three fiducial categories introduced in the previous section with different requirements on $m_{2b2\gamma}^\star$. All of these results are provided for both the matching to PYTHIA8 [94] and to HERWIG7 [96]. In the case of PYTHIA8, we also present results obtained with the recently developed MC@NLO-Δ matching procedure [97]. For our novel NLO+PS results we provide residual uncertainties due to variations of the factorization and renormalization scales as well as separately due to variations of the shower scale. For reference, we quote the cross section of the currently adopted approximation of the y_t^2 contribution $gg \to b\bar{b}H$ by the LHC experiments, which stems from the NNLOPS generator in refs. [69, 70] and is effectively described only at LO and for massless bottom quarks, i.e. in the 5FS, in the heavy top approximation (reweighted by the exact $gg \to gH$ amplitude [70]). In addition, we show the HH signal cross section in the different categories as a reference.

The main conclusions drawn from table 2 can be summarized as follows:

- The NLO QCD corrections are very large, and their inclusion becomes absolutely crucial in order to obtain a reliable estimate of the $b\bar{b}H$ background in any of the shown categories. They are significantly larger for the y_t^2 contributions, increasing the LO prediction by about $+150\%$, while the effect in the y_b^2 ones is roughly $+50\%$. The residual uncertainties from renormalization and factorization scale variations are at the level of $+50\%$ and -30% for the full $b\bar{b}H$ cross section (sum of y_b^2 and y_t^2 terms) in all fiducial categories.

- At NLO, the impact of using either PYTHIA8 (with or without the MC@NLO-Δ procedure) or HERWIG7 as a PS, as well as uncertainties related to the variation of the shower scale, are at the level of 10%. Thus, they can be considered as subleading with respect to the perturbative uncertainties. Only at LO, where the uncertainties of the predictions are generally much more significant, the difference of using the PYTHIA8 or HERWIG7 PS is more noticeable.

- The fiducial cuts have a substantial impact on the size of the $b\bar{b}H$ background. With respect to the fully inclusive cross section, the baseline cuts reduce the cross section by roughly a factor of 100. Further restrictions on $m_{2b2\gamma}^\star$ lead to a further decrease of the cross section by about 25% (55%) for $m_{2b2\gamma}^\star < 500\,\text{GeV}$ ($m_{2b2\gamma}^\star < 350\,\text{GeV}$). It is interesting to notice that the inclusive $b\bar{b}H$ cross section is almost two orders of magnitude larger than the HH signal, but the strong suppression induced by the fiducial selection cuts on the $b\bar{b}H$ background leads to similarly large cross sections in the three cut scenarios, with the exception of the $m_{2b2\gamma}^\star < 350\,\text{GeV}$ category, where the HH signal is more than a factor of two smaller than the $b\bar{b}H$ background. Regardless, the fact that the sum of y_b^2 and y_t^2 terms is at least as large as the

HH signal cross section in all categories, underlines the importance of an accurate modeling of the $b\bar{b}H$ background at NLO QCD.

- When we compare our results to the previously adopted NNLOPS simulation for the y_t^2 terms, which are effectively described only with LO accuracy in the matrix elements, we find large differences between our NLO+PS 4FS calculation and the 5FS NNLOPS one, where the latter yields a cross section that is about a factor of two larger. So far, in the ATLAS experiment analyses, the NNLOPS calculation is assigned a 100% uncertainty. Nevertheless, it is important to understand the origin of these significant differences to judge which prediction can be trusted. We have traced back a substantial effect in the NNLOPS results to $g \to b\bar{b}$ splittings in the PS, which make up almost half of the fiducial rates for the NNLOPS result, and even more inclusively. The corresponding value when turning off $g \to b\bar{b}$ splittings in the PS is given below the nominal NNLOPS prediction. We will discuss this point in more detail at the end of section 3.3, but we already point out our main conclusions: when the PS generates hard radiation, in particular bottom quarks through $g \to b\bar{b}$ splittings, it acts outside its validity range and the results are subject to very large uncertainties. Even worse, we have checked that in the case at hand, the enhancement of the cross section due to the PS originates almost entirely from events with exactly two bottom quarks coming from a gluon splitting, which thus creates two hard and separated b-tagged jets. Not only are such topologies not accurately described by a PS in general (due to its underlying soft/collinear approximation), but also the same kinematical region should already be described by the LO $gg \to b\bar{b}H$ matrix element. This suggests that by applying the PS to the inclusive $pp \to H$ NNLOPS sample, topologies with two hard b-jets may be double counted. Note that, of course, the NNLOPS matching is consistent up to relative α_s^2 accuracy (relative to the LO $gg \to H$ cross section), the same order where the perturbative series of the y_t^2-induced $b\bar{b}H$ cross section starts, so that such a double counting is formally beyond the nominal accuracy of that prediction. As a result, we conclude that the assigned 100% uncertainty (and possibly more) applied to the NNLOPS 5FS prediction by the ATLAS experiment so far is realistic, due to its insufficiencies in describing the bottom-quark/b-jet kinematics. By contrast, the 4FS calculation that describes these contributions at the level of the hard matrix elements up to NLO QCD provides a more accurate and realistic determination of the cross section with two bottom quarks/b-jets and their kinematics. Moreover, a clear reduction of the nominal uncertainties over the previously used NNLOPS can be achieved with the 4FS calculation.

Before turning to the description of differential distributions, we briefly comment on the amount of negative weights in our simulations. Negative weights are unavoidable in the MC@NLO matching [98] employed by MADGRAPH5_AMC@NLO. A large number of negative weights hampers the quality and the performance of the simulations, in the sense that larger event samples have to be generated in order to attain the same statistical uncertainties of a simulation with solely positive weights. The need to reduce negative weights in MADGRAPH5_AMC@NLO has motivated the development of new strategies,

Sample	$Q_{\rm sh}$	PY8	PY8-Δ	HW7
y_b^2	$\times 2$	39.1	33.8	36.8
	$\times 1$	38.5	33.9	36.4
	$\times 0.5$	38.1	33.8	36.3
y_t^2	$\times 2$	33.0	28.8	30.7
	$\times 1$	32.5	29.0	30.8
	$\times 0.5$	32.5	28.9	31.1

Table 3. Percentage of negative weights in the various samples generated.

such as the MC@NLO-Δ approach [97]. In table 3 we show the fractions of events with negative weights F in our samples and make the following observations: first, the fraction of negative weights has a mild dependence on the choice of shower-scale in general. Second, for both the y_b^2 and y_t^2 contributions the following patterns arise: $F_{\rm PY8} > F_{\rm HW7} > F_{\rm PY8-\Delta}$. Third, there is a noticeable improvement due to the usage of MC@NLO-Δ, which is particularly relevant for the y_b^2 term, where one has $F_{\rm PY8}^{y_b^2} - F_{\rm PY8-\Delta}^{y_b^2} \simeq 5\%$. While there is a slightly smaller reduction for the y_t^2 contribution, F is already smaller in that case using the standard MC@NLO matching. Although these numbers may look small, a reduction of 5% in the negative weight fraction can halve the computational cost of the simulation when $F \simeq 35\%$, as it is displayed in figure 1 of ref. [97] for instance.

3.3 Differential distributions

We now turn to studying differential distributions, which allow us to make further statements on the relevant kinematics of the $b\bar{b}H$ process. These results may be useful, for instance, to obtain a better understanding of which phase-space regions could be used in HH measurements to reduce the $b\bar{b}H$ background or, in general, what impact different cuts may have on the $b\bar{b}H$ process. Moreover, we will analyze further the different sources of uncertainties, including the differences observed using different PS or matching procedures.

The figures throughout this section are organized as follows: there is a main frame showing the absolute predictions as cross section per bin for the y_t^2 and y_b^2 contributions at NLO+PS using PYTHIA8 (orange, light blue) and using HERWIG7 (red, green). Then there are four ratio panels below. The first two show the envelope of the 9-point variations of renormalization and factorization scales for the y_t^2 and y_b^2 contributions at NLO+PS as a shaded band normalized to the central prediction, respectively. In these two frames one can also appreciate the size NLO corrections, as the respective LO+PS result is shown as a dashed curve. The last two ratio panels show, again for the y_t^2 and y_b^2 contributions respectively, shaded bands for the dependence on the shower scale for both PYTHIA8 and HERWIG7, using the same color scheme as the main inset. Besides these two set of curves, predictions obtained with PYTHIA8 in conjunction with the MC@NLO-Δ matching are also shown (solid curves in brown and cyan), as well as the NLO fixed-order predictions (dashed curves in magenta and dark blue). All results are normalized to the standard PYTHIA8 central prediction.
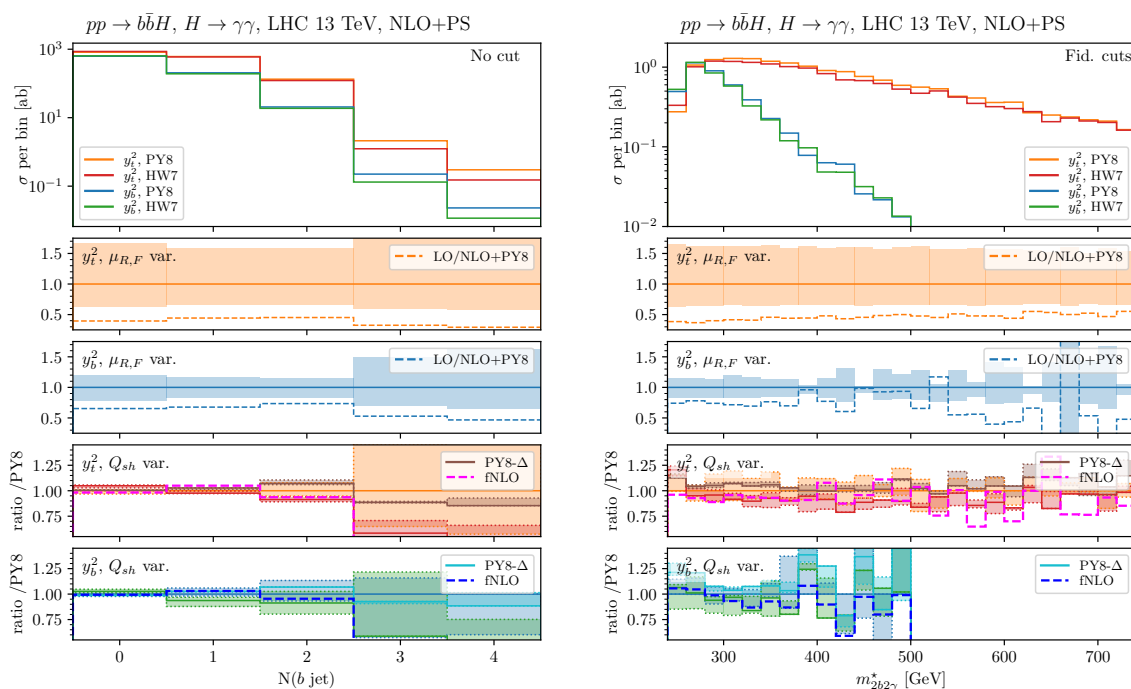
Figure 2. Distribution in the number of b-jets without further cuts (left) and in the $m^{\star}_{2b2\gamma}$ variable defined in eq. (3.7) with fiducial baseline selection cuts applied (right).

We start by discussing the exclusive cross section as a function of the number of b-jets ($N_{b\text{-jet}}$) in the fully inclusive phase space in figure 2 (left). From the main frame we can see that, while being similarly large for $N_{b\text{-jet}} = 0$, the y_t^2 cross section becomes increasingly larger compared to the y_b^2 one with $N_{b\text{-jet}}$. This has already been found in ref. [31], which concluded that requiring less (or softer) b-jets would increase the relative size of the y_b^2 contribution. As observed for the total inclusive cross section discussed in table 2, the y_t^2 contribution features even larger NLO QCD corrections than the y_b^2 one. This can be observed from the first two ratio panels by comparing the NLO+PS predictions with the LO+PS curves. As a result, the y_t^2 contribution is subject to a much larger uncertainty stemming from variations of the perturbative scales. This underlines again the crucial importance of NLO QCD corrections for $b\bar{b}H$ production, especially in the case of the component originating from the gluon-fusion process proportional to y_t^2. Looking at the last two ratio panels, it is clear that up to $N_{b\text{-jet}} = 2$, i.e. the multiplicities described at NLO+PS accuracy, shower-scale uncertainties are subleading. Only at higher multiplicities, which are effectively described only by the parton shower for $N_{b\text{-jet}} > 3$, the dependence on the shower scale becomes significantly larger, as expected. Comparing the different NLO-matched predictions in these two ratio panels, we observe differences that increase with the b-jet multiplicity, both for the y_b^2 and the y_t^2 contributions, with PYTHIA8 in conjunction with MC@NLO-Δ displaying the hardest spectra, while HERWIG7
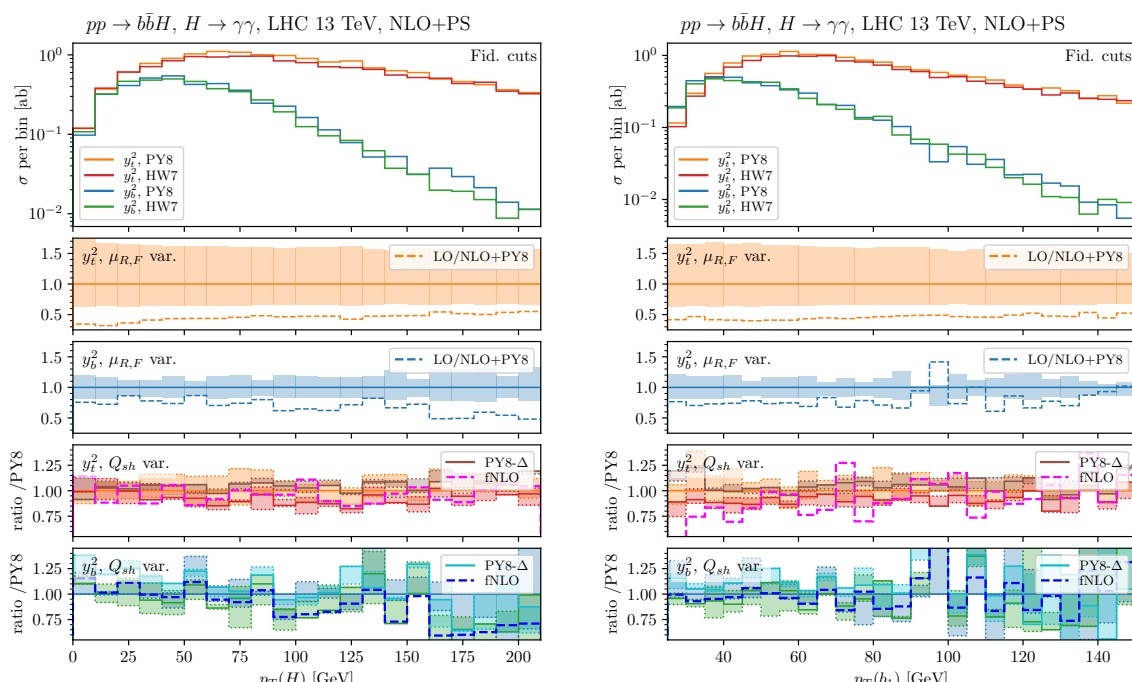
Figure 3. Distribution in the transverse momentum of the Higgs boson reconstructed from the two photons (left) and of the leading b-jet (right), with the fiducial baseline selection cuts applied.

shows the softest ones, with the latter also being closer to the fixed-order predictions. However, differences are rather mild: if we consider the bin with $N_{b\text{-jet}} = 2$, which is the one relevant for the fiducial cuts employed, they amount to about 20% for both the y_t^2 and the y_b^2 contribution (being slightly larger for the latter) when the most different predictions are considered. Indeed, the results in the fiducial phase space in table 2 display a very similar pattern. Similar results have been observed for other processes featuring bottom quarks and heavier objects [48, 99, 100]. Considering the relatively large perturbative scale variations, the predictions from PYTHIA8 (standard and MC@NLO-Δ) and HERWIG7 are compatible within their respective uncertainties, at least in the NLO+PS accurate multiplicities $N_{b\text{-jet}} \leq 2$.

Next, we consider the $m_{2b2\gamma}^{\star}$ distribution in figure 2 (right) in the fiducial phase space.[4] Also here we can appreciate from the main frame that the y_b^2 contribution is strongly suppressed with increasing $m_{2b2\gamma}^{\star}$, while it becomes similarly large as the y_t^2 one for $m_{2b2\gamma}^{\star} \lesssim$ 300 GeV. This behaviour is also confirmed by the rates quoted in table 2, where the relative size of the y_b^2 contribution is largest in the last category with $m_{2b2\gamma}^{\star} < 350$ GeV. This behaviour (reappearing in various kinematical distributions) originates from different features of the y_b^2 contribution: firstly, the two bottom quarks are in general less correlated; secondly,

---

[4]Because of the much steeper decrease of the differential cross section in the y_b^2 contribution, we show it only in a subset of the plotting range, up to $m_{2b2\gamma}^{\star} = 500$ GeV.

the bottom quarks are predominantly generated from initial-state splittings; thirdly, radiation off bottom quarks tends to be soft. Instead, for the y_t^2 contribution the bottom quarks are dominantly produced by a hard gluon which recoils against the Higgs boson. Indeed, we can observe the suppression of the y_b^2 contribution also at large transverse momenta of the Higgs boson $p_T(H)$ and of the hardest b-jet $p_T(b_1)$, shown in figure 3. As far as the scale variations and NLO QCD corrections visible in the first two ratio panels is concerned, we find very similar results as before for the number of b-jet also for $m_{2b2\gamma}^{\star}$, $p_T(H)$ and $p_T(b_1)$: the size of both the uncertainties and of the NLO corrections is larger for the y_t^2 contribution. We also notice that the relative behaviour of the various parton-shower predictions and the fixed order prediction is in general not the same for the y_b^2 and y_t^2 contributions.

3.4 Comparison and combination with the NNLOPS prediction

We now come back to the comparison of our 4FS NLO+PS predictions with results from the 5FS NNLOPS generator used so far to model the y_t^2-induced $b\bar{b}H$ background in HH measurements. In figure 4 we consider the $m_{2b2\gamma}^{\star}$ distribution, which is used to define different fiducial categories by the experiments, and show the NNLOPS prediction with (blue solid curve) and without $g \to b\bar{b}$ splittings in the PS (blue dotted curve) together with our LO+PS (orange dashed curve) and NLO+PS (orange solid curve with orange band) predictions in the 4FS. We immediately notice that the NNLOPS prediction is substantially larger than the 4FS prediction, especially at low $m_{2b2\gamma}^{\star}$ where it is even outside the orange uncertainty band. Moreover, the NNLOPS prediction reduces drastically and becomes more compatible with the NLO+PS 4FS prediction when secondary $g \to b\bar{b}$ splittings generated by the PS are turned off. This is in line without our findings for the fiducial cross section in table 2. Still, it is somewhat surprising that the LO-accurate 5FS result is as large as the NLO-accurate 4FS prediction (at low $m_{2b2\gamma}^{\star}$ even larger), when the NLO corrections are of the order of 100%. One should bear in mind, however, that the scale setting of the NNLOPS prediction is quite different. Since it is the typical scale for the inclusive production of a Higgs boson (not exclusive in the two bottom quarks), $\mu_R = \mu_F = m_H/2$ is used for the NNLO prediction. We also observe that the 5FS predictions tend more towards small $m_{2b2\gamma}^{\star}$, which is not unexpected due to the fact that the bottom quarks are massless in the matrix elements and only put on the mass shell through reshuffling of their momenta in the PS matching, which induces some arbitrariness in their kinematics.

To further investigate why $g \to b\bar{b}$ splittings in the PS lead to such a substantial and unexpected increase of the cross section with two hard b-jets, we will try to understand the type of topologies/kinematics that lead to this large contribution. First of all, the PS acts somewhat outside its validity range if it creates two hard b-jets, unless these originate from two hard jets each splitting into a collinear b-jet pair, thus giving rise to four bottom quarks in the event. We have checked that in the relevant phase space region the vast majority of events have exactly two bottom quarks in total. Therefore, the PS creates two b-jets at large angle out of a single $g \to b\bar{b}$ splitting, which in principle should barely happen and is very poorly described in the soft/collinear approximation of the PS. In this situation, there are only two possible configurations that could lead to two hard b-jets. In the first one, two soft and wide-angle bottom quarks are created, which the jet-clustering algorithm combines
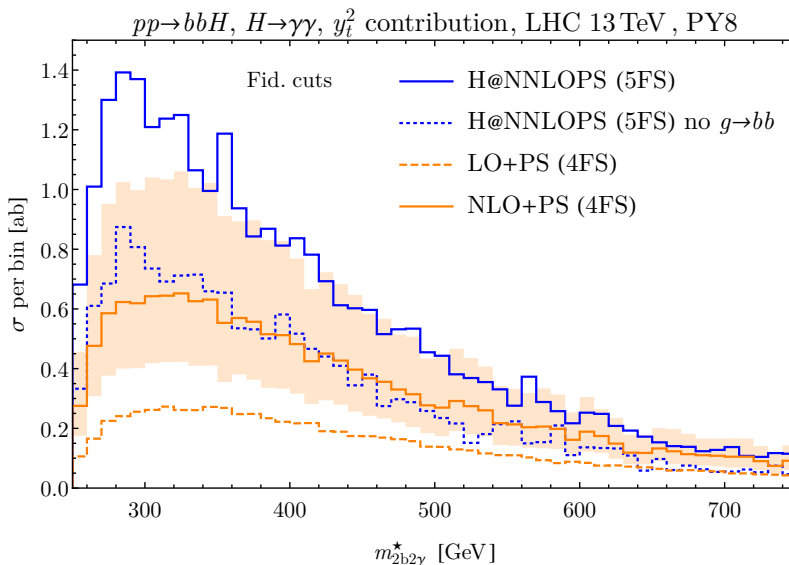
Figure 4. Distribution in $m^{\star}_{2b2\gamma}$ with fiducial selection cuts, for the y_t^2 contribution. Predictions at LO and NLO in the 4FS are shown, as well as 5FS results, the latter with and without contributions from $g \to b\bar{b}$ splittings in the shower.

with other hard partons to form b-jets. In other words, in these configurations, only a small fraction of the jet energy is carried by the bottom quarks. Such an effect could be missing in our 4FS $b\bar{b}H$ calculation at NLO, as we generally include two hard b-jets and (at most) one additional hard light jet at matrix-element level. All other jets are generated by the PS in our calculation and should be softer. The relevant hard matrix element for this configuration has two bottom quarks and two light jets in the final state ($pp \to b\bar{b}Hjj$), which enters beyond the accuracy of our calculation, but, in principle, could be included through multi-jet merging. In the second possible configuration, the PS creates two hard wide-angle bottom quarks, each of which creates a separate b-jet subsequently. In this case, the bottom quark will carry a rather large fraction of the b-jet energy. Should this configuration be the dominant one, then not only would the PS describe such kinematics very poorly, but also the relevant hard matrix-element for such configurations is simply $gg \to b\bar{b}H$ (and possibly further soft partons). This contribution is actually already accounted for in the LO matrix element of the 5FS NNLOPS calculation and should not be created again by the PS.

In order to asses which of the two aforementioned configurations is responsible for the observed large effect due to $g \to b\bar{b}$ splittings in the PS, we show two further distributions in figure 5, namely the invariant mass of the two hardest B hadrons in the signal region and the number of light jets, in both cases turning on and off $g \to b\bar{b}$ splittings in the PS. What we can see is that the PS creates two hard B hadrons at wide angle that form two separate b-jets. Their invariant mass is quite close to the window required for the two b-jets in eq. (3.4), displayed in the figure as a shaded region. Thus, we can rule out the first scenario. With the distribution in the number of light jets, we can see that a large fraction (more than one third) of the events has no extra light jets. All this strongly points to the fact

that the contribution due to secondary $g \to b\bar{b}$ splittings is counted twice (since it is already accounted for by the $gg \to b\bar{b}H$ hard matrix element), and that the PS is acting beyond its validity range and generates two hard and separated bottom quarks. Moreover, events with two or more light jets represent less than one third of the contribution from $g \to b\bar{b}$ splittings. Such a contribution may actually be missing in our 4FS calculation, however, it should be well covered by the $\sim +60\%$ higher-order uncertainties on the predictions.

Before concluding, we would like to point out that the NNLOPS calculation is used in experimental analyses not only as an estimate of the $b\bar{b}H$ y_t^2 background, as discussed at length in this work, but also to model the contribution originating from inclusive Higgs production with fake b-jets, i.e. where light jets are mistagged as b-jets. Therefore, even if our NLO+PS description of the $b\bar{b}H$ background in the 4FS is to be used, replacing the NNLOPS calculation for the y_t^2 contribution in this case, a consistent way of combining both simulations needs to be devised to be able to simulate the fake component as well. The most simple and naive option, which already provides a largely consistent prediction, can be achieved at the level of the events, without the need of regenerating the event samples: in the 5FS NNLOPS events, all contributions with final-state bottom quarks (irrespective whether they originate from the hard matrix element, from initial-state splittings of bottom quarks, or final-state $g \to b\bar{b}$ splittings generated by the shower) would have to be removed from the final result. This would cancel the $b\bar{b}H$ background entirely from the 5FS NNLOPS sample, and only the contributions from light jets (and potential fakes) would be kept. On top of this, the contribution from 4FS NLO+PS $b\bar{b}H$ events can be added incoherently. The same approach has been used in ref. [100] to estimate bottom mass effects in the Z-boson transverse momemtum. The only drawback of this approach is that formally a 5FS and a 4FS calculation are mixed. Thus, to be fully consistent, one would have to rerun the NNLOPS calculation with 4FS PDF sets and strong coupling, as well as setting the number of light flavours to $n_f = 4$ in the calculation. In this case, events with initial- or final-state bottom quarks (at the level of the hard matrix element) would be removed from the calculation by construction. This can be achieved by removing the appropriate flavour structures from the original 5FS implementation of the NNLOPS generator. Moreover, the $g \to b\bar{b}$ splittings will have to be turned off when showering the NNLOPS events, to avoid double counting with the 4FS NLO+PS $b\bar{b}H$ calculation. As a result, one obtains a consistent 4FS prediction including both NNLO QCD accuracy to inclusive Higgs production and NLO QCD accuracy for the $b\bar{b}H$ background, including their matching to the parton shower.

3.5 Impact of the new b$\bar{\text{b}}$H modeling for the HH searches

We have estimated the impact of the NLO QCD modeling of the y_t^2 contribution with respect to previously adopted NNLOPS prediction on the current limits for the HH cross section (and λ_{HHH}). The rates and the uncertainties from the new $b\bar{b}H$ prediction are propagated to the search for HH production in the $2b2\gamma$ final state performed by the ATLAS experiment by using the publicly available information in ref. [71].

The inclusive $gg \to H$ background involves both the $gg \to b\bar{b}H$ process and Higgs boson production with additional jets where at least one is mistagged as a b-jet (i.e. fake b-jets). The decomposition of the $gg \to H$ background in terms of the flavour of the two
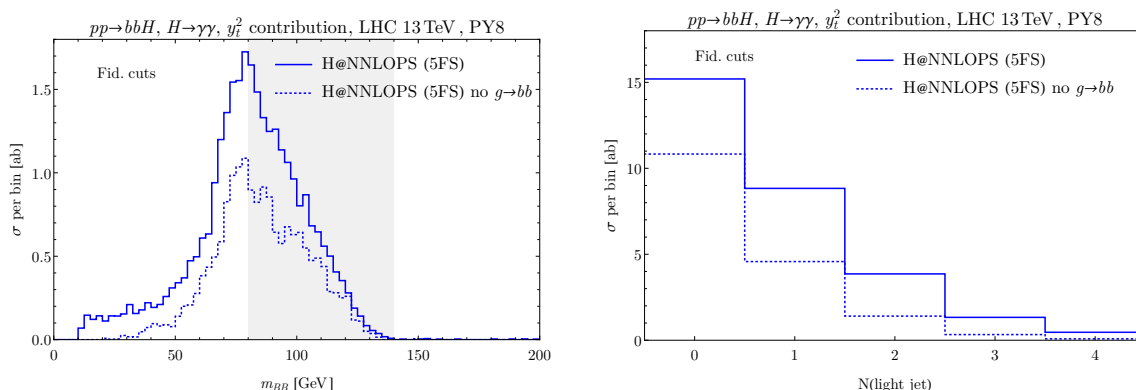
Figure 5. Distribution in the invariant mass of the two hardest B hadrons (left) and in the number of light jets (right) in the 5FS calculation, with and without the contributions from $g \to b\bar{b}$ splittings in the shower, in the fiducial region. The shaded area in the left-hand side plot corresponds to the invariant-mass requirement on the b-jets employed in our fiducial cuts.

additional jets is estimated by evaluating the rates of the $gg \to H$ plus b-jets, c-jets, or light jets contributions from the NNLOPS prediction, and multiplying each with the appropriate b-tagging (or mistagging) efficiency of the b-tagging algorithm employed in the $HH \to 2b2\gamma$ search [101]. This results in 80% (20%) of the inclusive $gg \to H$ background arising from events with true (fake) b-jets. To assess the impact of the new $b\bar{b}H$ calculation, we rescaled only the 80% of the inclusive $gg \to H$ background estimated by the analysis with the rates from the new y_t^2-induced $b\bar{b}H$ calculation quoted in table 2. Furthermore, the 100% uncertainty assigned to the $gg \to H$ process in the current analysis is replaced with the uncertainty arising from the scale variations of the y_t^2-induced $b\bar{b}H$ component reported in table 2.

Following the procedure described above, we have estimated that new $b\bar{b}H$ modeling improves the limit on the HH production cross section by a few percent. The relatively small improvement is mostly due to the fact that the $gg \to H$ process is not the main contribution to the background for the $HH \to 2b2\gamma$ search, which is dominated by the SM production of $\gamma\gamma jj$ final states. Moreover, the current analysis, based on Run 2 data, corresponding to a total integrated luminosity of $140\,\text{fb}^{-1}$, is largely limited by statistical uncertainties. Similarly, we have studied the effect of the new $b\bar{b}H$ modeling on HH searches during the High-Luminosity phase of the LHC by projecting these results to $3\,\text{ab}^{-1}$ integrated luminosity, as done in ref. [102]. At the HL-LHC, the systematic uncertainty on the $gg \to H$ background starts to be one of the limiting factors for the sensitivity of the analysis. In this case, the improvement coming from the new $b\bar{b}H$ modeling increases to around 5% for the limit on both the HH production cross section and λ_{HHH}, as well as for the HH discovery significance.

Finally, we have applied the same recipe to estimeate the effect of the new $b\bar{b}H$ modeling on the HH search in the $2b2\tau$ channel [72]. The same flavour decomposition found for the $2b2\gamma$ final state is assumed for the $2b2\tau$ channel, and so is the contribution of the $gg \to H$ process to the inclusive single Higgs background with respect to the other single Higgs production modes. In addition, the fiducial region defined by the $2b2\tau$ analysis was not

studied in detail. Instead, the same $b\bar{b}H$ rates showed in table 2, have been assumed. For the $2b2\tau$ final state, employing the new $b\bar{b}H$ prediction improves the limit on the HH production cross section by up to 10%, and, after extrapolating the analysis to the HL-LHC scenario, the improvement on both the discovery significance and the HH cross section limit increases up to 20%. The larger effect seen on the $2b2\tau$ final state is mainly due to the fact that the analysis is less dominated by the statistical uncertainty and that the single Higgs background is more relevant with respect to the $2b2\gamma$ channel.

4   Conclusions

We have presented a new study of $b\bar{b}H$ production in the SM, focusing on its impact in the Higgs-pair production signal region. The dominant contributions to the $b\bar{b}H$ process, namely those proportional to y_b^2 and y_t^2, are simulated at NLO in the 4FS and matched to parton showers by employing the MadGraph5_aMC@NLO framework. This is the first calculation of this kind for the latter contribution. The accurate modeling of the $b\bar{b}H$ background to Higgs-pair measurements is particularly relevant, since in all sensitive search channels at least one of the two Higgs bosons decays to bottom quarks, rendering the continuum production of the $b\bar{b}H$ final state an irreducible and very sizeable background. In fact, due to the low accuracy in the simulation of the y_t^2 contribution used in the current experimental searches, $b\bar{b}H$ production yields the dominant theoretical uncertainty in HH measurements.

In this paper, we have studied in detail the impact of including the full NLO QCD corrections to the $b\bar{b}H$ process in the context of HH production. We found that corrections of $\mathcal{O}(+100\%)$ and larger increase the y_t^2 contribution at NLO in QCD, which renders it vital to include these effects for a sufficiently accurate modeling of the $b\bar{b}H$ background in HH analyses. Although the total uncertainties (from factorization-scale, renormalization-scale, shower-scale and shower variations) are still relatively large (at the level of +50% and -30%), this constitutes a substantial improvement over the previously employed LO approximation for the y_t^2 contribution to the $b\bar{b}H$ background from the NNLOPS generator with $\mathcal{O}(100\%)$ uncertainties, not only in terms of accuracy, but also in terms of precision.

When comparing our results to the previously employed 5FS NNLOPS calculation, which describes the y_t^2 $b\bar{b}H$ background effectively with LO+PS accuracy, we notice that the NNLOPS prediction is almost a factor of two larger than our NLO+PS prediction, despite missing the $\mathcal{O}(+100\%)$ NLO correction. We have traced this back to two major sources: firstly, contributions with two hard bottom quarks generated from a single wide-angle $g \to b\bar{b}$ splitting in the parton shower yield about half of the 5FS NNLOPS prediction. Not only does the parton shower act vastly beyond its validity range in this case, but also a large part of these configurations have no extra light jets and should already be accounted for by the hard LO matrix element, which suggests that there is some kind of double counting in this kinematical regime. Secondly, the scale setting in the NNLOPS sample adopted for inclusive Higgs production, which is of the order of the Higgs mass, does not really reflect the hardness of the $b\bar{b}H$ final state. With this different scale setting the LO+PS 5FS result (without $g \to b\bar{b}$ splittings from the shower) is as large as our 4FS

NLO+PS prediction. We note that the issues with the 5FS description may affect also other processes including bottom quarks in the final state.

We have estimated the potential impact of accurate predictions for the $b\bar{b}H$ background for HH searches. We concluded that the current limits on $\sigma_{\rm SM}^{HH}$ could improve by few percent and up to 10% in the $2b2\gamma$ and $2b2\tau$ channels, respectively. At HL-LHC the improvements on the $\sigma_{\rm SM}^{HH}$ limits and on the HH discovery significance increases to 5% (20%) for the $2b2\gamma$ ($2b2\tau$) final state. Given that the current limit is several times the SM value of the HH cross section, this is not a particularly dramatic change. However, since with the end of the High-Luminosity phase of the LHC Higgs pair production is expected to be observed at almost five standard deviations, the accurate modeling of the $b\bar{b}H$ background, which we found to be as large as the HH cross section in the relevant fiducial phase space regions, will be indispensable for maximizing the sensitivity to the HH signal. Moreover, it will ensure that bounds/extraction of the HH cross section as well as the trilinear Higgs coupling will be fully reliable. In addition, an improved description of the $b\bar{b}H$ process can be beneficial not only in the context of HH physics, but also in analyses targeting single-Higgs production, such as measurements of Higgs boson production in association with a top quark pair, where $b\bar{b}H$ production also contributes as a background [103]. For all of these reasons, the $b\bar{b}H$ generator presented in this work will represent a very useful tool for future analyses performed by the LHC collaborations. In principle, also the inclusion of NNLO QCD corrections to this type of process is feasible, which is left for future work.

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 collaboration, *A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery*, Nature **607** (2022) 52 [*Erratum ibid.* **612** (2022) E24] [arXiv:2207.00092] [INSPIRE].

[4] CMS collaboration, *A portrait of the Higgs boson by the CMS experiment ten years after the discovery*, Nature **607** (2022) 60 [arXiv:2207.00043] [INSPIRE].

[5] J. Alison et al., *Higgs boson potential at colliders: status and perspectives*, Rev. Phys. **5** (2020) 100045 [arXiv:1910.00012] [INSPIRE].

[6] ATLAS collaboration, *Constraints on the Higgs boson self-coupling from single- and double-Higgs production with the ATLAS detector using pp collisions at $\sqrt{s} = 13$ TeV*, Phys. Lett. B **843** (2023) 137745 [arXiv:2211.01216] [INSPIRE].

[7] ATLAS collaboration, *HL-LHC prospects for the measurement of Higgs boson pair production in the $b\bar{b}b\bar{b}$ final state and combination with the $b\bar{b}\gamma\gamma$ and $b\bar{b}\tau^+\tau^-$ final states at the ATLAS experiment*, ATL-PHYS-PUB-2022-053, CERN, Geneva, Switzerland (2022).

[8] CMS collaboration, *Search for Higgs boson pair production in the four b quark final state in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Phys. Rev. Lett. **129** (2022) 081802 [arXiv:2202.09617] [INSPIRE].

[9] CMS collaboration, *Search for nonresonant pair production of highly energetic Higgs bosons decaying to bottom quarks*, Phys. Rev. Lett. **131** (2023) 041803 [arXiv:2205.06667] [INSPIRE].

[10] CMS collaboration, *Search for nonresonant Higgs boson pair production in final state with two bottom quarks and two tau leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Phys. Lett. B **842** (2023) 137531 [arXiv:2206.09401] [INSPIRE].

[11] CMS collaboration, *Search for Higgs boson pairs decaying to WW^*WW^*, $WW^*\tau\tau$, and $\tau\tau\tau\tau$ in proton-proton collisions at $\sqrt{s} = 13$ TeV*, JHEP **07** (2023) 095 [arXiv:2206.10268] [INSPIRE].

[12] D. de Florian and J. Mazzitelli, *Two-loop virtual corrections to Higgs pair production*, Phys. Lett. B **724** (2013) 306 [arXiv:1305.5206] [INSPIRE].

[13] D. de Florian and J. Mazzitelli, *Higgs boson pair production at next-to-next-to-leading order in QCD*, Phys. Rev. Lett. **111** (2013) 201801 [arXiv:1309.6594] [INSPIRE].

[14] D.Y. Shao, C.S. Li, H.T. Li and J. Wang, *Threshold resummation effects in Higgs boson pair production at the LHC*, JHEP **07** (2013) 169 [arXiv:1301.1245] [INSPIRE].

[15] R. Frederix et al., *Higgs pair production at the LHC with NLO and parton-shower effects*, Phys. Lett. B **732** (2014) 142 [arXiv:1401.7340] [INSPIRE].

[16] F. Maltoni, E. Vryonidou and M. Zaro, *Top-quark mass effects in double and triple Higgs production in gluon-gluon fusion at NLO*, JHEP **11** (2014) 079 [arXiv:1408.6542] [INSPIRE].

[17] J. Grigo, K. Melnikov and M. Steinhauser, *Virtual corrections to Higgs boson pair production in the large top quark mass limit*, Nucl. Phys. B **888** (2014) 17 [arXiv:1408.2422] [INSPIRE].

[18] D. de Florian and J. Mazzitelli, *Higgs pair production at next-to-next-to-leading logarithmic accuracy at the LHC*, JHEP **09** (2015) 053 [arXiv:1505.07122] [INSPIRE].

[19] S. Borowka et al., *Higgs boson pair production in gluon fusion at next-to-leading order with full top-quark mass dependence*, *Phys. Rev. Lett.* **117** (2016) 012001 [*Erratum ibid.* **117** (2016) 079901] [`arXiv:1604.06447`] [InSPIRE].

[20] J. Baglio et al., *Gluon fusion into Higgs pairs at NLO QCD and the top mass scheme*, *Eur. Phys. J. C* **79** (2019) 459 [`arXiv:1811.05692`] [InSPIRE].

[21] M. Grazzini et al., *Higgs boson pair production at NNLO with top quark mass effects*, *JHEP* **05** (2018) 059 [`arXiv:1803.02463`] [InSPIRE].

[22] D. De Florian and J. Mazzitelli, *Soft gluon resummation for Higgs boson pair production including finite M_t effects*, *JHEP* **08** (2018) 156 [`arXiv:1807.03704`] [InSPIRE].

[23] R. Bonciani, G. Degrassi, P.P. Giardino and R. Gröber, *Analytical method for next-to-leading-order QCD corrections to double-Higgs production*, *Phys. Rev. Lett.* **121** (2018) 162003 [`arXiv:1806.11564`] [InSPIRE].

[24] J. Davies, G. Mishima, M. Steinhauser and D. Wellmann, *Double Higgs boson production at NLO in the high-energy limit: complete analytic results*, *JHEP* **01** (2019) 176 [`arXiv:1811.05489`] [InSPIRE].

[25] L.-B. Chen, H.T. Li, H.-S. Shao and J. Wang, *The gluon-fusion production of Higgs boson pair: N^3LO QCD corrections and top-quark mass effects*, *JHEP* **03** (2020) 072 [`arXiv:1912.13001`] [InSPIRE].

[26] J. Davies and M. Steinhauser, *Three-loop form factors for Higgs boson pair production in the large top mass limit*, *JHEP* **10** (2019) 166 [`arXiv:1909.01361`] [InSPIRE].

[27] J. Baglio et al., *$gg \to HH$: combined uncertainties*, *Phys. Rev. D* **103** (2021) 056002 [`arXiv:2008.11626`] [InSPIRE].

[28] A.H. Ajjath and H.-S. Shao, *N^3LO+N^3LL QCD improved Higgs pair cross sections*, *JHEP* **02** (2023) 067 [`arXiv:2209.03914`] [InSPIRE].

[29] J. Davies et al., *Higgs boson contribution to the leading two-loop Yukawa corrections to $gg \to HH$*, *JHEP* **08** (2022) 259 [`arXiv:2207.02587`] [InSPIRE].

[30] M. Mühlleitner, J. Schlenk and M. Spira, *Top-Yukawa-induced corrections to Higgs pair production*, *JHEP* **10** (2022) 185 [`arXiv:2207.02524`] [InSPIRE].

[31] N. Deutschmann, F. Maltoni, M. Wiesemann and M. Zaro, *Top-Yukawa contributions to bbH production at the LHC*, *JHEP* **07** (2019) 054 [`arXiv:1808.01660`] [InSPIRE].

[32] D. Pagani, H.-S. Shao and M. Zaro, *RIP $Hb\bar{b}$: how other Higgs production modes conspire to kill a rare signal at the LHC*, *JHEP* **11** (2020) 036 [`arXiv:2005.10277`] [InSPIRE].

[33] C. Grojean, A. Paul and Z. Qian, *Resurrecting $b\bar{b}h$ with kinematic shapes*, *JHEP* **04** (2021) 139 [`arXiv:2011.13945`] [InSPIRE].

[34] D. Dicus, T. Stelzer, Z. Sullivan and S. Willenbrock, *Higgs boson production in association with bottom quarks at next-to-leading order*, *Phys. Rev. D* **59** (1999) 094016 [`hep-ph/9811492`] [InSPIRE].

[35] C. Balazs, H.-J. He and C.P. Yuan, *QCD corrections to scalar production via heavy quark fusion at hadron colliders*, *Phys. Rev. D* **60** (1999) 114001 [`hep-ph/9812263`] [InSPIRE].

[36] R.V. Harlander and W.B. Kilgore, *Higgs boson production in bottom quark fusion at next-to-next-to leading order*, *Phys. Rev. D* **68** (2003) 013001 [`hep-ph/0304035`] [InSPIRE].

[37] J.M. Campbell, R.K. Ellis, F. Maltoni and S. Willenbrock, *Higgs-boson production in association with a single bottom quark*, *Phys. Rev. D* **67** (2003) 095002 [hep-ph/0204093] [INSPIRE].

[38] R.V. Harlander, K.J. Ozeren and M. Wiesemann, *Higgs plus jet production in bottom quark annihilation at next-to-leading order*, *Phys. Lett. B* **693** (2010) 269 [arXiv:1007.5411] [INSPIRE].

[39] K.J. Ozeren, *Analytic results for Higgs production in bottom fusion*, *JHEP* **11** (2010) 084 [arXiv:1010.2977] [INSPIRE].

[40] R. Harlander and M. Wiesemann, *Jet-veto in bottom-quark induced Higgs production at next-to-next-to-leading order*, *JHEP* **04** (2012) 066 [arXiv:1111.2182] [INSPIRE].

[41] S. Bühler, F. Herzog, A. Lazopoulos and R. Müller, *The fully differential hadronic production of a Higgs boson via bottom quark fusion at NNLO*, *JHEP* **07** (2012) 115 [arXiv:1204.4415] [INSPIRE].

[42] A. Belyaev, P.M. Nadolsky and C.-P. Yuan, *Transverse momentum resummation for Higgs boson produced via b\bar{b} fusion at hadron colliders*, *JHEP* **04** (2006) 004 [hep-ph/0509100] [INSPIRE].

[43] R.V. Harlander, A. Tripathi and M. Wiesemann, *Higgs production in bottom quark annihilation: transverse momentum distribution at NNLO+NNLL*, *Phys. Rev. D* **90** (2014) 015017 [arXiv:1403.7196] [INSPIRE].

[44] T. Ahmed et al., *Two-loop QCD corrections to Higgs → b + \bar{b} + g amplitude*, *JHEP* **08** (2014) 075 [arXiv:1405.2324] [INSPIRE].

[45] T. Gehrmann and D. Kara, *The Hb\bar{b} form factor to three loops in QCD*, *JHEP* **09** (2014) 174 [arXiv:1407.8114] [INSPIRE].

[46] C. Duhr, F. Dulat and B. Mistlberger, *Higgs boson production in bottom-quark fusion to third order in the strong coupling*, *Phys. Rev. Lett.* **125** (2020) 051804 [arXiv:1904.09990] [INSPIRE].

[47] R. Mondini and C. Williams, *Bottom-induced contributions to Higgs plus jet at next-to-next-to-leading order*, *JHEP* **05** (2021) 045 [arXiv:2102.05487] [INSPIRE].

[48] M. Wiesemann et al., *Higgs production in association with bottom quarks*, *JHEP* **02** (2015) 132 [arXiv:1409.5301] [INSPIRE].

[49] F. Krauss, D. Napoletano and S. Schumann, *Simulating b-associated production of Z and Higgs bosons with the SHERPA event generator*, *Phys. Rev. D* **95** (2017) 036012 [arXiv:1612.04640] [INSPIRE].

[50] A.H. Ajjath et al., *NNLO QCD⊕QED corrections to Higgs production in bottom quark annihilation*, *Phys. Rev. D* **100** (2019) 114016 [arXiv:1906.09028] [INSPIRE].

[51] A.H. Ajjath et al., *Resummed prediction for Higgs boson production through b\bar{b} annihilation at N^3LL*, *JHEP* **11** (2019) 006 [arXiv:1905.03771] [INSPIRE].

[52] S. Forte, T. Giani and D. Napoletano, *Fitting the b-quark PDF as a massive-b scheme: Higgs production in bottom fusion*, *Eur. Phys. J. C* **79** (2019) 609 [arXiv:1905.02207] [INSPIRE].

[53] T. Ahmed, V. Ravindran, A. Sankar and S. Tiwari, *Two-loop amplitudes for di-Higgs and di-pseudo-Higgs productions through quark annihilation in QCD*, *JHEP* **01** (2022) 189 [arXiv:2110.11476] [INSPIRE].

[54] S. Badger, H.B. Hartanto, J. Kryś and S. Zoia, *Two-loop leading-colour QCD helicity amplitudes for Higgs boson production in association with a bottom-quark pair at the LHC*, *JHEP* **11** (2021) 012 [arXiv:2107.14733] [INSPIRE].

[55] S. Dittmaier, M. Krämer and M. Spira, *Higgs radiation off bottom quarks at the Tevatron and the CERN LHC*, *Phys. Rev. D* **70** (2004) 074010 [hep-ph/0309204] [INSPIRE].

[56] S. Dawson, C.B. Jackson, L. Reina and D. Wackeroth, *Exclusive Higgs boson production with bottom quarks at hadron colliders*, *Phys. Rev. D* **69** (2004) 074027 [hep-ph/0311067] [INSPIRE].

[57] S. Dawson, C.B. Jackson, L. Reina and D. Wackeroth, *Higgs production in association with bottom quarks at hadron colliders*, *Mod. Phys. Lett. A* **21** (2006) 89 [hep-ph/0508293] [INSPIRE].

[58] N. Liu, L. Wu, P. Wu and J.M. Yang, *Complete one-loop effects of SUSY QCD in $b\bar{b}h$ production at the LHC under current experimental constraints*, *JHEP* **01** (2013) 161 [arXiv:1208.3413] [INSPIRE].

[59] S. Dittmaier et al., *Neutral MSSM Higgs-boson production with heavy quarks: NLO supersymmetric QCD corrections*, *Phys. Rev. D* **90** (2014) 035010 [arXiv:1406.5307] [INSPIRE].

[60] Y. Zhang, *NLO electroweak effects on the Higgs boson production in association with a bottom quark pair at the LHC*, *Phys. Rev. D* **96** (2017) 113009 [arXiv:1708.08790] [INSPIRE].

[61] B. Jager, L. Reina and D. Wackeroth, *Higgs boson production in association with b jets in the POWHEG BOX*, *Phys. Rev. D* **93** (2016) 014030 [arXiv:1509.05843] [INSPIRE].

[62] F. Maltoni, G. Ridolfi and M. Ubiali, *b-initiated processes at the LHC: a reappraisal*, *JHEP* **07** (2012) 022 [*Erratum ibid.* **04** (2013) 095] [arXiv:1203.6393] [INSPIRE].

[63] M. Lim, F. Maltoni, G. Ridolfi and M. Ubiali, *Anatomy of double heavy-quark initiated processes*, *JHEP* **09** (2016) 132 [arXiv:1605.09411] [INSPIRE].

[64] S. Forte, D. Napoletano and M. Ubiali, *Higgs production in bottom-quark fusion in a matched scheme*, *Phys. Lett. B* **751** (2015) 331 [arXiv:1508.01529] [INSPIRE].

[65] S. Forte, D. Napoletano and M. Ubiali, *Higgs production in bottom-quark fusion: matching beyond leading order*, *Phys. Lett. B* **763** (2016) 190 [arXiv:1607.00389] [INSPIRE].

[66] M. Bonvini, A.S. Papanastasiou and F.J. Tackmann, *Resummation and matching of b-quark mass effects in $b\bar{b}H$ production*, *JHEP* **11** (2015) 196 [arXiv:1508.03288] [INSPIRE].

[67] M. Bonvini, A.S. Papanastasiou and F.J. Tackmann, *Matched predictions for the $b\bar{b}H$ cross section at the 13 TeV LHC*, *JHEP* **10** (2016) 053 [arXiv:1605.01733] [INSPIRE].

[68] C. Duhr, F. Dulat, V. Hirschi and B. Mistlberger, *Higgs production in bottom quark fusion: matching the 4- and 5-flavour schemes to third order in the strong coupling*, *JHEP* **08** (2020) 017 [arXiv:2004.04752] [INSPIRE].

[69] K. Hamilton, P. Nason, C. Oleari and G. Zanderighi, *Merging H/W/Z + 0 and 1 jet at NLO with no merging scale: a path to parton shower + NNLO matching*, *JHEP* **05** (2013) 082 [arXiv:1212.4504] [INSPIRE].

[70] K. Hamilton, P. Nason and G. Zanderighi, *Finite quark-mass effects in the NNLOPS POWHEG+MiNLO Higgs generator*, *JHEP* **05** (2015) 140 [arXiv:1501.04637] [INSPIRE].

[71] ATLAS collaboration, *Search for Higgs boson pair production in the two bottom quarks plus two photons final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Rev. D* **106** (2022) 052001 [arXiv:2112.11876] [INSPIRE].

[72] ATLAS collaboration, *Search for resonant and non-resonant Higgs boson pair production in the $b\bar{b}\tau^+\tau^-$ decay channel using 13 TeV pp collision data from the ATLAS detector*, *JHEP* **07** (2023) 040 [arXiv:2209.10910] [INSPIRE].

[73] P. Artoisenet et al., *A framework for Higgs characterisation*, *JHEP* **11** (2013) 043 [arXiv:1306.6464] [INSPIRE].

[74] F. Demartin et al., *Higgs characterisation at NLO in QCD: CP properties of the top-quark Yukawa interaction*, *Eur. Phys. J. C* **74** (2014) 3065 [arXiv:1407.5089] [INSPIRE].

[75] C. Degrande et al., *UFO — the Universal FeynRules Output*, *Comput. Phys. Commun.* **183** (2012) 1201 [arXiv:1108.2040] [INSPIRE].

[76] 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].

[77] R. Frederix et al., *The automation of next-to-leading order electroweak calculations*, *JHEP* **07** (2018) 185 [*Erratum ibid.* **11** (2021) 085] [arXiv:1804.10017] [INSPIRE].

[78] R.V. Harlander, T. Neumann, K.J. Ozeren and M. Wiesemann, *Top-mass effects in differential Higgs production through gluon fusion at order $\mathcal{O}(\alpha_s^4$*, *JHEP* **08** (2012) 139 [arXiv:1206.0157] [INSPIRE].

[79] T. Neumann and M. Wiesemann, *Finite top-mass effects in gluon-induced Higgs production with a jet-veto at NNLO*, *JHEP* **11** (2014) 150 [arXiv:1408.6836] [INSPIRE].

[80] T. Neumann, *NLO Higgs + jet production at large transverse momenta including top quark mass effects*, *J. Phys. Comm.* **2** (2018) 095017 [arXiv:1802.02981] [INSPIRE].

[81] S.P. Jones, M. Kerner and G. Luisoni, *Next-to-leading-order QCD corrections to Higgs boson plus jet production with full top-quark mass dependence*, *Phys. Rev. Lett.* **120** (2018) 162001 [*Erratum ibid.* **128** (2022) 059901] [arXiv:1802.00349] [INSPIRE].

[82] X. Chen et al., *Top-quark mass effects in H + jet and H + 2 jets production*, *JHEP* **03** (2022) 096 [arXiv:2110.06953] [INSPIRE].

[83] LHC HIGGS CROSS SECTION WORKING GROUP collaboration, *Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector*, arXiv:1610.07922 [DOI:10.23731/CYRM-2017-002] [INSPIRE].

[84] NNPDF collaboration, *Parton distributions from high-precision collider data*, *Eur. Phys. J. C* **77** (2017) 663 [arXiv:1706.00428] [INSPIRE].

[85] LHC HIGGS CROSS SECTION WORKING GROUP collaboration, *Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector*, arXiv:1610.07922 [DOI:10.23731/CYRM-2017-002] [INSPIRE].

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

[87] M. Cacciari, G.P. Salam and G. Soyez, *FastJet user manual*, *Eur. Phys. J. C* **72** (2012) 1896 [arXiv:1111.6097] [INSPIRE].

[88] ATLAS collaboration, *Combined measurements of Higgs boson production and decay using up to $80\,fb^{-1}$ of proton-proton collision data at $\sqrt{s} = 13\,TeV$ collected with the ATLAS experiment*, *Phys. Rev. D* **101** (2020) 012002 [arXiv:1909.02845] [INSPIRE].

[89] C. Anastasiou et al., *High precision determination of the gluon fusion Higgs boson cross-section at the LHC*, *JHEP* **05** (2016) 058 [arXiv:1602.00695] [INSPIRE].

[90] C. Anastasiou et al., *Higgs boson gluon-fusion production beyond threshold in N^3LO QCD*, *JHEP* **03** (2015) 091 [arXiv:1411.3584] [INSPIRE].

[91] C. Anastasiou et al., *Higgs boson gluon-fusion production at threshold in N^3LO QCD*, *Phys. Lett. B* **737** (2014) 325 [arXiv:1403.4616] [INSPIRE].

[92] C. Anastasiou et al., *Higgs boson gluon-fusion production in QCD at three loops*, *Phys. Rev. Lett.* **114** (2015) 212001 [arXiv:1503.06056] [INSPIRE].

[93] J. Butterworth et al., *PDF4LHC recommendations for LHC run II*, *J. Phys. G* **43** (2016) 023001 [arXiv:1510.03865] [INSPIRE].

[94] C. Bierlich et al., *A comprehensive guide to the physics and usage of PYTHIA 8.3*, arXiv:2203.11601 [DOI:10.21468/SciPostPhysCodeb.8] [INSPIRE].

[95] G. Heinrich et al., *NLO predictions for Higgs boson pair production with full top quark mass dependence matched to parton showers*, *JHEP* **08** (2017) 088 [arXiv:1703.09252] [INSPIRE].

[96] J. Bellm et al., *Herwig 7.2 release note*, *Eur. Phys. J. C* **80** (2020) 452 [arXiv:1912.06509] [INSPIRE].

[97] R. Frederix, S. Frixione, S. Prestel and P. Torrielli, *On the reduction of negative weights in MC@NLO-type matching procedures*, *JHEP* **07** (2020) 238 [arXiv:2002.12716] [INSPIRE].

[98] S. Frixione and B.R. Webber, *Matching NLO QCD computations and parton shower simulations*, *JHEP* **06** (2002) 029 [hep-ph/0204244] [INSPIRE].

[99] C. Degrande, M. Ubiali, M. Wiesemann and M. Zaro, *Heavy charged Higgs boson production at the LHC*, *JHEP* **10** (2015) 145 [arXiv:1507.02549] [INSPIRE].

[100] E. Bagnaschi, F. Maltoni, A. Vicini and M. Zaro, *Lepton-pair production in association with a $b\bar{b}$ pair and the determination of the W boson mass*, *JHEP* **07** (2018) 101 [arXiv:1803.04336] [INSPIRE].

[101] ATLAS collaboration, *ATLAS flavour-tagging algorithms for the LHC run 2 pp collision dataset*, *Eur. Phys. J. C* **83** (2023) 681 [arXiv:2211.16345] [INSPIRE].

[102] ATLAS collaboration, *Measurement prospects of Higgs boson pair production in the $b\bar{b}\gamma\gamma$ final state with the ATLAS experiment at the HL-LHC*, ATL-PHYS-PUB-2022-001, CERN, Geneva, Switzerland (2022).

[103] ATLAS collaboration, *Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector*, *Phys. Lett. B* **784** (2018) 173 [arXiv:1806.00425] [INSPIRE].