EFT acceptance studies for $VHb\bar{b}$ analysis in $Z \to ee/\mu\mu$ channel

Peng Wang

University of Cambridge, UK

Supervisor: Aliya Nigamova, Rainer Mankel

September 9, 2021

Abstract

This report presents the analysis of Higgs Boson decaying into $b\bar{b}$ associating with the Z boson decaying into electron or muon pairs. The 2018 Run-2 CMS Monte Carlo(MC) samples were analysed in leading order Standard Model Effective Field Theory(SMEFT) together with Simplified Template Cross Sections(STXS) framework stage 1.2. Existing Monte Carlos samples were reweighted using EFT2Obs package to linearly parametrise the most relevant Wilson Coefficients in each STXS bin. Parametrisation was performed on both GEN-level and RECO-level MC samples, consistency was shown with some small deviations observed due to the acceptance effect.

Contents

1 Introduction

Since the discovery of the Higgs boson(H) in 2012, various analyses have been performed on Higg's decay and production processes to examine its properties and search for signs of possible new physics. At present, the full Large Hadron Collider(LHC) Run-2 data which was collected from 2015 to 2018 is now accessible for Higg's analysis.

This study is using the 2018 CMS Run-2 13TeV proton-proton collision MC samples to analyse the associated Vector Boson-Higgs Boson production and the Higgs to $b\bar{b}$ decay ($VHb\bar{b}$ process) with a focus on channel $Z \rightarrow ee/\mu\mu$. All measurements are processed under the Standard Model Effective Field Theory(SMEFT), which is looking for Beyond Standard Model(BSM) new physics at an energy scale of $\Lambda >> \nu$ where ν is the electroweak energy scale.

EFT effects will be studied within the Simplified Template Cross Sections(STXS) framework, which is adopted to extract the cross-section measurement in CMS and ATLAS. Parameters of Wilson Coefficients can be derived and compared to identify the acceptance effect embedded.

In this report, a brief introduction of SMEFT will be given in section 2. The use of STXS Framework and theories of production and decay parametrisation is summarised in section 3. A general description of the reweighting procedure is included in section 4, followed by results presented in section 5. Finally, a conclusion is given in section 6.

2 Standard Model Effective Field Theory (SMEFT)

In the theoretical framework of EFT, one needs to choose the Higgs' representation and the basis that they will be working in. SMEFT is a sub-division of EFT which represents the Higgs using Standard Model(SM) Higgs doublet. Its lagrangian can be written as an expansion of the SM lagrangian added by some higher dimension operators as shown in eq.1. Where $\mathcal{O}_i^n$ stands for the i^{th} operator with dimension n, N_{dn} is the total number of operators in dimension n, c_i are some coefficients that we want to study, and Λ stand for the energy scale which is set to equal 1TeV.

$$\mathcal{L}_{SMEFT} = \mathcal{L}_{SM} + \sum_i^{N_{d5}} \frac{c_i^{(5)}}{\Lambda}\mathcal{O}_i^{(5)} + \sum_i^{N_{d6}} \frac{c_i^{(6)}}{\Lambda^2}\mathcal{O}_i^{(6)} + \sum_i^{N_{d7}} \frac{c_i^{(7)}}{\Lambda^3}\mathcal{O}_i^{(7)} + \sum_i^{N_{d8}} \frac{c_i^{(8)}}{\Lambda^4}\mathcal{O}_i^{(8)} + ... \quad (1)$$

Dimension 5 and 7 operators violate the conservation of the lepton numbers, as a result, they will be ignored and we have the Leading Order(LO) to be dimension-6 operators and Next Leading Order(NLO) to be dimension-8 operators, the lagrangian therefore simplifies to eq.2 [1].

$$\mathcal{L}_{SMEFT} = \mathcal{L}_{SM} + \sum_i^{N_{d6}} \frac{c_i^{(6)}}{\Lambda^2}\mathcal{O}_i^{(6)} + \sum_i^{N_{d8}} \frac{c_i^{(8)}}{\Lambda^4}\mathcal{O}_i^{(8)} + ... \quad (2)$$

This analysis is carried out in LO, therefore only terms involving $\mathcal{O}_i^6$ are considered. Coefficients in front of the dimension-6 operators($c_i^{(6)}$), named Wilson Coefficients, will cause the SMEFT lagrangian to deviate from the SM lagrangian if they have non-zero values. In this study, the SMEFTsim package is used for SMEFT Leading Order generation in the Warsaw basis [2].

3 STXS Framework

Simplified Template Cross Sections(STXS) is a common framework used for providing cross-section measurements for Higgs data coming from the LHC experiment, it divides a large data set into several bins according to their kinematics. This framework is commonly adopted because it makes data more fine-grained for precise analysis and reduces the theoretical error folded inside.

Combining with the SMEFT, it is feasible to parametrise each of the STXS bins using the relevant Wilson Coefficients. Different Wilson Coefficients have various sensitivity in separate interaction channels. Therefore it is vital to consider parametrisation independently for the decay and production channel.

3.1 Production Parametrisation

The STXS production cross section can be written as a sum of the SM term, the BSM term and their interference term shown in eq.3 [1].

$$\sigma_{STXS} = \sigma_{SM} + \sigma_{int} + \sigma_{BSM} \tag{3}$$

The interference term σ_{int} has a single SMEFT dimension-6 operator embedded, giving a suppression in the scale of Λ^{-2}. Whereas σ_{BSM} is the pure SMEFT term, involving a product of two SMEFT operators, giving a total suppression of Λ^{-4}. They are usually expressed in the format shown in eq.4 as a cross-section ratio μ_i to the SM.

$$\mu_i(c_j) = \frac{\sigma^i_{STXS}}{\sigma^i_{SM}} = 1 + \sum_j A^i_j c_j + \sum_{jk} B^i_{jk} c_j c_k \tag{4}$$

The interference cross-section ratio to the SM ($\frac{\sigma^i_{int}}{\sigma^i_{SM}}$) is replaced with the linear summation of Wilson Coefficients $\sum_j A^i_j c_j$, the SMEFT term $\frac{\sigma^i_{BSM}}{\sigma^i_{SM}}$ is replaced with quadratic summation $\sum_{jk} B^i_{jk} c_j c_k$, where A^i_j and B^i_{jk} are constant parameters to be measured. As a result of the Λ suppression discussed above, the assumption is made that the quadratic term is negligible, therefore analysis can be carried out in the linear regime.

STXS bins used for analysing the $q\bar{q} \to ZH$ production process and their corresponding labels in this study are summarised in Table.1 [3], and Wilson Coefficients chosen to be focused on for the production process are $c^{(1)}_{Hq}$, $c^{(3)}_{Hq}$, c_{Hu} and c_{Hd}.

Label	STXS Bins (p_T)
400-401	FWD
401-402	0-75 GeV
402-403	75-150 GeV
403-404	150-250 GeV 0j
404-405	150-250 GeV ≥1j
405-406	>250 GeV

Table 1: STXS Bins for $q\bar{q} \to ZH$ production process

3.2 Decay Parametrisation

The use of SMEFT indicates there is negligible interference between the production and decay process due to the fact SM Higgs is a scalar particle with narrow width. Hence decay parametrisation can be factorised from production as shown in eq.5 [1] where $B^{H\to X}$ is the branching ratio for a particular decay process $H \to X$.

$$\begin{aligned} (\sigma \times B)^{i,H\to X} &= \sigma^i \times B^{H\to X} \\ &= (\sigma^i_{STXS}) \times \left(\frac{\Gamma^{H\to X}_{SM} + \Gamma^{H\to X}_{int} + \Gamma^{H\to X}_{BSM}}{\Gamma^H_{SM} + \Gamma^H_{int} + \Gamma^H_{BSM}} \right) \end{aligned} \tag{5}$$

By carrying out similar simplification as the production cross-section and only retain the linear terms, eq.6 can be derived. The significance of this equation is that the parametrisation of the production and decay processes can be carried out independently and re-combined via multiplication at the end.

$$(\sigma \times B)^{i,H\to X} = (\sigma \times B)^{i,H\to X}_{SM} \times \left(1 + \sigma_j A^{\sigma_i}_j c_j\right) \times \left(\frac{1 + \Sigma_k A^{\Gamma^H}_k c_k}{1 + \Sigma_k A^{\Gamma^{H\to X}}_k c_k} \right) \tag{6}$$

In comparison to the production STXS bins, where a large data-set is separated into more fine-grained regions according to kinematics of the additional jets, the decay process does not require additional bin splitting. According to eq.7 [4], it is clear that the decay width in dimension-6 has no dependence on p_T. Therefore parametrisation is completed on top of one merged bin that includes all kinematic regions as presented in Table.2. Wilson Coefficients chosen to be focused on for the production process are $c_{H\Box}$, c_{HDD}, c_{dH}, $c^{(3)}_{Hl}$ and c'_{ll}.

$$\Gamma^{(6,0)} = \frac{N_c m_H m_b^2}{4\pi} \left[C_{H\Box} - \frac{C_{HD}}{4} \left(1 - \frac{\hat{c}^2_w}{\hat{s}^2_w} \right) + \frac{\hat{c}_w}{\hat{s}_w} C_{HWB} - \frac{\hat{v}_T C_{bH}}{m_b \sqrt{2}} \right] \tag{7}$$

Where C_i are Wilson Coefficients sensitive for the $H \to b\bar{b}$ process, definitions of $\hat{v}_T$, $\hat{c}_T$ and $\hat{s}_T$ can be found in eq.8.

$$\begin{aligned} \hat{v}_T &\equiv \frac{2M_W \hat{s}_w}{e}, \\ \hat{c}_w^2 &\equiv \frac{M_W^2}{M_Z^2}, \\ \hat{s}_w^2 &\equiv 1 - \hat{c}_w^2, \end{aligned} \tag{8}$$

Label	STXS Bins (p_T)
400-406	All Kinematic regions

Table 2: STXS Bins for $H \to b\bar{b}$ decay process

4 Reweighting

MC samples are reweighted using the EFT2Obs package. For each of the decay and production processes, Feynmann diagrams using matrix elements containing different Wilson Coefficients are generated. The new weight is calculated from matrix elements and the original SM Monte Carlo weight. These are all computed up to LO via the reweighting formula summarised in eq.9 [5].

$$W_{new} = \frac{|M_{new}|^2}{|M_{orig}|^2} W_{orig} \tag{9}$$

The original MC weight is scaled by the ratio of the matrix elements squared to obtain the new weight. This reweighting is processed separately for the decay and production processes.

5 Results

5.1 Reweighted GEN-level and RECO-level production process

The 2018 Run-2 CMS GEN-level Monte Carlo(MC) samples have gone through selection to obtain the RECO-level samples. Figure.1 shows the cross-sections of each STXS bin when one of the Wilson Coefficients is set to equal 0.1 for both reweighted GEN-level and RECO-level MC samples. The black line labelled Nominal is the SM prediction. It is clear setting non-zero Wilson Coefficients results in deviations from the SM model, therefore one can conclude from these graphs that the Wilson Coefficients causing this deviation should have a non-zero parameter. The greatest deviation is seen for $c_{Hq}^{(3)} = 0.1$, hence it should have the largest parameter. The difference in histogram shape is a result of the additional selection cut applied for RECO-level samples.

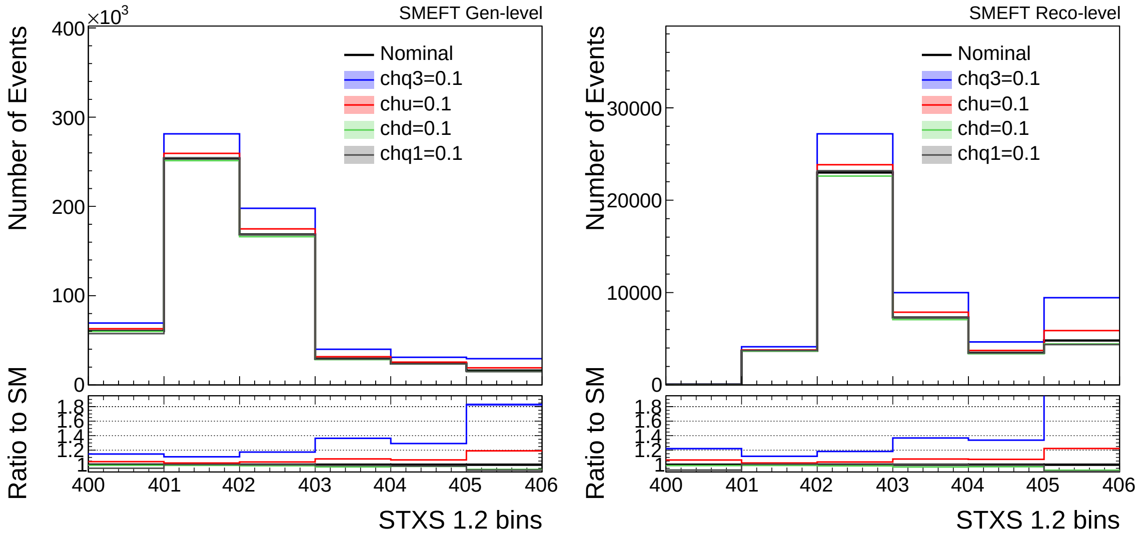

Figure 1: Graph showing the number of events in each STXS bin when $c_{Hq}^{(1)}$, $c_{Hq}^{(3)}$, c_{Hu} or c_{Hd} is set to be 0.1, for both GEN-level(left) and RECO-level(right)

5.2 Parametrisation

5.2.1 Production

The parametrisation is done by processing the reweighted MC samples using the EFT2Obs package. A summarised parameter table for each of the STXS bins using GEN-level and RECO-level can be found in Table.3 and Table.4. Comparison graphs for coefficients $c_{Hq}^{(1)}$, $c_{Hq}^{(3)}$, c_{Hu} and c_{Hd} are drawn in Figure.2. Generally, consistency was shown for both sample structures, however, small deviations are observed showing the effect of acceptance. Although the deviation is small, it is not negligible. Therefore it is preferable to use parametrisation from the RECO-level samples which include the acceptance effect.

STXS Bins	Parametrisation: GEN-level
400–401	$1 - 0.49\ c_{Hq}^{(1)} + 1.47\ c_{Hq}^{(3)} + 0.42\ c_{Hu} - 0.08\ c_{Hd} + 0.72\ c_{HW} + 0.33\ c_{HWB}$
401–402	$1 + 0.03\ c_{Hq}^{(1)} + 1.09\ c_{Hq}^{(3)} + 0.22\ c_{Hu} - 0.10\ c_{Hd} + 0.64\ c_{HW} + 0.30\ c_{HWB}$
402–403	$1 - 0.01\ c_{Hq}^{(1)} + 1.74\ c_{Hq}^{(3)} + 0.37\ c_{Hu} - 0.15\ c_{Hd} + 0.75\ c_{HW} + 0.34\ c_{HWB}$
403–404	$1 - 0.12\ c_{Hq}^{(1)} + 3.64\ c_{Hq}^{(3)} + 0.79\ c_{Hu} - 0.31\ c_{Hd} + 0.89\ c_{HW} + 0.38\ c_{HWB}$
404–405	$1 - 0.19\ c_{Hq}^{(1)} + 2.91\ c_{Hq}^{(3)} + 0.66\ c_{Hu} - 0.24\ c_{Hd} + 0.83\ c_{HW} + 0.36\ c_{HWB}$
405–406	$1 - 0.74\ c_{Hq}^{(1)} + 8.28\ c_{Hq}^{(3)} + 1.91\ c_{Hu} - 0.66\ c_{Hd} + 0.94\ c_{HW} + 0.39\ c_{HWB}$

Table 3: Production Parametrisation on GEN-level MC samples for each STXS Bin

STXS Bins	Parametrisation: RECO-level
400–401	$1 - 0.75\ c_{Hq}^{(1)} + 2.21\ c_{Hq}^{(3)} + 0.64\ c_{Hu} - 0.13\ c_{Hd} + 0.82\ c_{HW} + 0.37\ c_{HWB}$
401–402	$1 + 0.07\ c_{Hq}^{(1)} + 1.16\ c_{Hq}^{(3)} + 0.23\ c_{Hu} - 0.11\ c_{Hd} + 0.67\ c_{HW} + 0.32\ c_{HWB}$
402–403	$1 + 0.09\ c_{Hq}^{(1)} + 1.82\ c_{Hq}^{(3)} + 0.37\ c_{Hu} - 0.17\ c_{Hd} + 0.75\ c_{HW} + 0.34\ c_{HWB}$
403–404	$1 + 0.01\ c_{Hq}^{(1)} + 3.68\ c_{Hq}^{(3)} + 0.78\ c_{Hu} - 0.32\ c_{Hd} + 0.87\ c_{HW} + 0.38\ c_{HWB}$
404–405	$1 - 0.09\ c_{Hq}^{(1)} + 3.38\ c_{Hq}^{(3)} + 0.73\ c_{Hu} - 0.29\ c_{Hd} + 0.85\ c_{HW} + 0.37\ c_{HWB}$
405–406	$1 - 0.92\ c_{Hq}^{(1)} + 9.63\ c_{Hq}^{(3)} + 2.23\ c_{Hu} - 0.76\ c_{Hd} + 0.94\ c_{HW} + 0.40\ c_{HWB}$

Table 4: Production Parametrisation on RECO-level samples for each STXS Bin

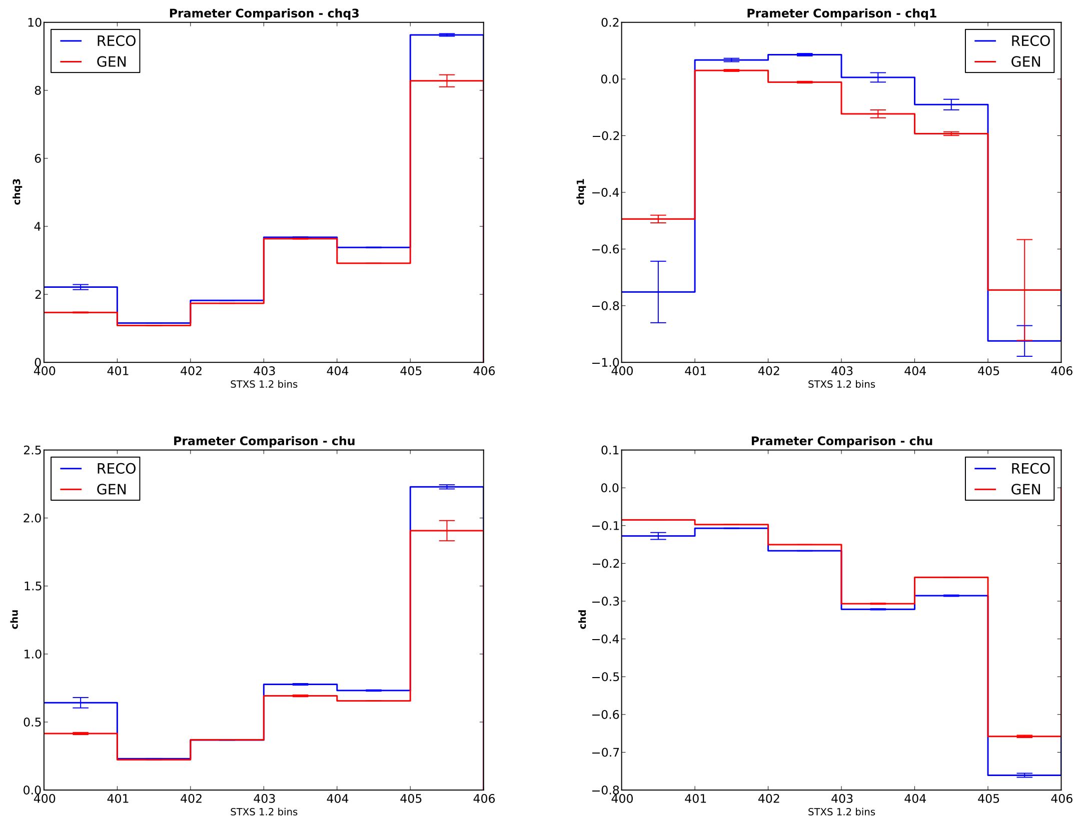

Figure 2: GEN-level and RECO-level parameter comparison for $c_{Hq}^{(3)}$(upper left), $c_{Hq}^{(1)}$(upper right), c_{Hu}(lower left) and c_{Hd}(lower right)

The larger deviation in $c_{Hq}^{(1)}$ graph is because low-valued parameters result in a larger error when sensitivity is limiting. The Wilson coefficient which has the largest contribution is $c_{Hq}^{(3)}$ as its parameters are of the highest magnitude. A validation graph was plotted for $c_{Hq}^{(3)}$ to compare RECO-level parametrisation obtained with ATLAS measurements [1] (Figure.3). Consistency was observed, the fact that two lines are not overlapping exactly can be explained by two reasons. The first reason could be ATLAS is using slightly different selection criteria which leads to different acceptances. Another reason is ATLAS has included more Wilson Coefficients when running the reweighting module, which will cause variations on other Wilson Coefficients, including the ones that are being measured and compared in this study.

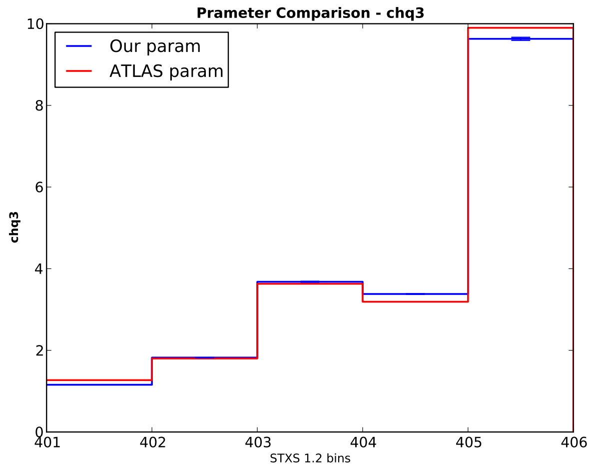

Figure 3: Comparison of RECO-level parametrisation for $c_{Hq}^{(3)}$ with ATLAS measurements

5.2.2 Decay

Separate reweighting was done for $H \to b\bar{b}$, parametrisation has been processed for $c_{H\Box}$, c_{HDD}, c_{dH}, $c_{Hl}^{(3)}$ and $c_{ll}^{'}$. Table.5 shows a summarisation of the values obtained at RECO-level. It is compared with ATLAS parametrisation of the same decay process in Table.6. As a result of merging STXS bins, there will be less deviation when compared with ATLAS parameters - perfect consistency is shown.

STXS Bins	**Parametrisation: RECO-level**
400–406	$1 + 0.12\ c_{H\Box} - 0.03\ c_{HDD} - 0.12\ c_{dH} - 0.12\ c_{Hl}^{(3)} + 0.06\ c_{ll}^{'}$

Table 5: Parametrisation for $H \to b\bar{b}$ with merged STXS bins

c_i	**ATLAS measurement [1]**	**RECO-level Measurement**
$c_{H\Box}$	0.12	0.12
c_{HDD}	-0.030	-0.030
c_{dH}	-0.121	-0.12
$c_{Hl}^{(3)}$	-0.121	-0.12
$c_{ll}^{'}$	0.061	0.061

Table 6: Parametrisation for $H \to b\bar{b}$ with merged STXS bins

6 Conclusion

This study presents the analysis of Higgs Boson decaying into $b\bar{b}$ associating with the Z boson decaying into electron or muon pairs using the 2018 Run-2 CMS MC samples. Calculations were performed in leading order Standard Model Effective Field Theory(SMEFT) together with the use of Simplified Template Cross Sections(STXS) framework stage 1.2. EFT2Obs package was adopted to reweight the existing MC samples and carry out parametrisation of Wilson Coefficients. Reweighting and parametrisation process is separate for production and decay processes.

Production parametrisation was done on both GEN-level and RECO-level samples, to examine the effect of acceptance. General consistency was observed with some small but non-negligible deviations, therefore it is preferable to use RECO-level parametrisation which includes the acceptance. Both decay and production RECO parameters are compared with ATLAS measurements. Production parametrisation showed slight deviations from ATLAS parameters possibly due to different acceptances being adopted at ATLAS. Variations may also be caused by having two analyses including different numbers of Wilson Coefficients throughout the reweighting procedure. Decay parametrisation was perfectly consistent with the ATLAS measurements. Better parametrisation would be expected to be obtained if the study expands to use the full Run-2 CMS MC samples.

References

[1] ATLAS Collaboration. Interpretations of the combined measurement of higgs boson production and decay. Oct 2020.

[2] Ilaria Brivio, Yun Jiang, and Michael Trott. The smeftsim package, theory and tools. *Journal of High Energy Physics*, 2017(12), Dec 2017.

[3] Nicolas Berger, Claudia Bertella, Thomas P. Calvet, Milene Calvetti, Valerio Dao, Marco Delmastro, Michael Duehrssen-Debling, Paolo Francavilla, Yacine Haddad, Oleh Kivernyk, Jonathon M. Langford, Changqiao Li, Giovanni Marchiori, Predrag Milenovic, Carlo E. Pandini, Edward Scott, Frank J. Tackmann, Kerstin Tackmann, Lorenzo Viliani, and Meng Xiao. Simplified template cross sections - stage 1.1, 2019.

[4] Jonathan M. Cullen, Benjamin D. Pecjak, and Darren J. Scott. Nlo corrections to h $\rightarrow$ b$\bar{b}$ decay in smeft. *Journal of High Energy Physics*, 2019(8), Aug 2019.

[5] Olivier Mattelaer. On the maximal use of monte carlo samples: re-weighting events at nlo accuracy. *The European Physical Journal C*, 76(12), Dec 2016.