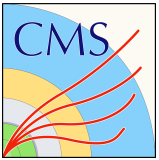


SMEFT and HEFT di-Higgs Event-Level Reweighting

Matheus Coelho[1], Pedro Galli Mercadante[1],
Alexandra Carvalho[2], Torben Lange[3]

1 - São Paulo Research and Analysis Center (SPRACE) - Federal University of ABC (BR)
2 - Peking University (CN)
3 - National Institute of Chemical Physics and Biophysics (EE)

Outline

- ☐ SMEFT and HEFT (ggF `HH`): A 5D problem for experimentalists
- ☐ Event-Level Reweighting: A tool for di-Higgs non-resonant searches
- ☐ Tests on generator-level variables
  - p_T^h raw distribution
  - Validation in realistic phase spaces: m_{hh} boosted selection
    - Run 2: $HH \to 4b$
    - Run 3: $HH \to 4b$
- ☐ Applications on experimental analyses
- ☐ Final remarks

SMEFT and HEFT (ggF HH): A 5D problem for experimentalists

Common EFT modeling for di-Higgs non-resonant searches

$$\mathcal{L}^{HH}_{HEFT}$$
$$= -m_t \left(\kappa_t \frac{h}{\nu} + c_2 \frac{h^2}{\nu^2}\right) \bar{t}t - \kappa_\lambda \frac{m_H^2}{2\nu} h^3 + \frac{1}{3}\frac{\alpha_s}{4\pi}\left(c_g \frac{h}{\nu} - \frac{1}{2} c_{2g} \frac{h^2}{\nu^2}\right) G^a_{\mu\nu} G^{a,\mu\nu}$$

Up to chiral dimension $d_\chi = 4$ (G. Buchalla et al., A. Carvalho et al.)

$$\mathcal{L}^{HH}_{\text{Warsaw}}$$
$$= \frac{C_{H\Box}}{\Lambda^2}(\phi^\dagger\phi)\Box(\phi^\dagger\phi) + \frac{C_{HD}}{\Lambda^2}(\phi^\dagger D_\mu \phi)^*(\phi^\dagger D^\mu \phi) + \frac{C_H}{\Lambda^2}(\phi^\dagger\phi)^3$$
$$+ \frac{C_{tH}}{\Lambda^2}\left(\phi^\dagger\phi \bar{q}_L \tilde{\phi} t_R + h.c.\right) + \frac{C_{HG}}{\Lambda^2}\phi^\dagger\phi G^a_{\mu\nu} G^{a,\mu\nu}$$

Up to dim6 operators and $\sigma \simeq \sigma_{(\text{SM+dim6})\times(\text{SM+dim6})}$
(G. Heinrich, J. Lang, L. Scyboz)

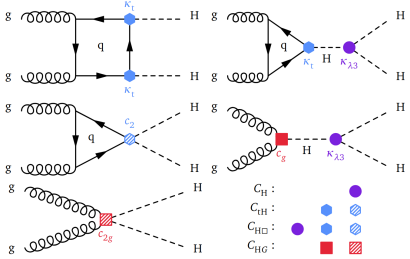


Figure: Main contributions to the process.

SMEFT and HEFT (ggF HH): A 5D problem for experimentalists

$\{\kappa_\lambda, \kappa_t, c_2, c_g, c_{2g}\}$, $\{C_{H\Box}, C_{HD}, C_H, C_{tH}, C_{HG}\}$

How to cover such a large parameter space in a real analysis?

- New physics may manifest as deviations hidden in an infinitude of differential shapes across such parameter space

SMEFT and HEFT (ggF HH): A 5D problem for experimentalists

$\{\kappa_\lambda, \kappa_t, c_2, c_g, c_{2g}\}$, $\{C_{H\Box}, C_{HD}, C_H, C_{tH}, C_{HG}\}$

How to cover such a large parameter space in a real analysis?

- New physics may manifest as deviations hidden in an infinitude of differential shapes across such parameter space
- Practical problems:
  - Experimental analyses have a huge demand for MC production
  - A naive parameter scan is unfeasible
    - `Powheg-Box` samples for HEFT and SMEFT at NLO QCD are expensive

SMEFT and HEFT (ggF HH): A 5D problem for experimentalists

$\{\kappa_\lambda, \kappa_t, c_2, c_g, c_{2g}\}$, $\{C_{H\Box}, C_{HD}, C_H, C_{tH}, C_{HG}\}$

How to cover such a large parameter space in a real analysis?

- New physics may manifest as deviations hidden in an infinitude of differential shapes across such parameter space
- Practical problems:
  - Experimental analyses have a huge demand for MC production
  - A naive parameter scan is unfeasible
    - `Powheg-Box` samples for HEFT and SMEFT at NLO QCD are expensive
  - Any reweighed signal however would still need to be processed, stored, statistically evaluated
    $\rightarrow$ More suited for benchmarks, less suited for actual scans
    - ATLAS and CMS typically use the so-called "matrix-inversion method" for the extraction of final results
    - This technique becomes very complicated in 5D spaces
      event-level reweighting can also help in this case

Differential parametrization at NLO QCD

$$
\begin{aligned}
d\sigma^{\text{NLO}}_{\text{HEFT}} &= Poly(\mathbf{c}_i, \mathbf{A}) \\
&= A_1 c_t^4 + A_2 c_{tt}^2 + (A_3 c_t^2 + A_4 c_{ggh}^2) c_{hhh}^2 + A_5 c_{gghh}^2 \\
&+ (A_6 c_{tt} + A_7 c_t c_{hhh}) c_t^2 + (A_8 c_t c_{hhh} + A_9 c_{ggh} c_{hhh}) c_{tt} \\
&+ A_{10} c_{tt} c_{gghh} + (A_{11} c_{ggh} c_{hhh} + A_{12} c_{gghh}) c_t^2 \\
&+ (A_{13} c_{hhh} c_{ggh} + A_{14} c_{gghh}) c_t c_{hhh} + A_{15} c_{ggh} c_{gghh} c_{hhh} \\
&+ A_{16} c_t^3 c_{ggh} + A_{17} c_t c_{tt} c_{ggh} + A_{18} c_t c_{ggh}^2 c_{hhh} + A_{19} c_t c_{ggh} c_{gghh} \\
&+ A_{20} c_t^2 c_{ggh}^2 + A_{21} c_{tt} c_{ggh}^2 + A_{22} c_{ggh}^3 c_{hhh} + A_{23} c_{ggh}^2 c_{gghh}
\end{aligned}
$$

$$
\begin{aligned}
d\sigma^{\text{NLO}}_{\text{SMEFT}} &= Poly(\mathbf{c}_i, \mathbf{A}) \\
&= A_1 + A_2 c_{H,\Box} + A_3 c_{HD} + A_4 c_H + A_5 c_{uH} + A_6 c_{HG} \\
&+ A_7 c_{H,\Box}^2 + A_8 c_{HD}^2 + A_9 c_H^2 + A_{10} c_{uH}^2 + A_{11} c_{HG}^2 \\
&+ A_{12} c_{H,\Box} c_{HD} + A_{13} c_{H,\Box} c_H + A_{14} c_{H,\Box} c_{uH} + A_{15} c_{H,\Box} c_{HG} \\
&+ A_{16} c_{HD} c_H + A_{17} c_{HD} c_{uH} + A_{18} c_{HD} c_{HG} + A_{19} c_H c_{uH} \\
&+ A_{20} c_H c_{HG} + A_{21} c_{uH} c_{HG}
\end{aligned}
$$

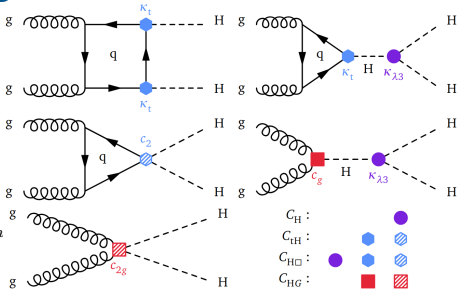


Figure: Main contributions to the process.

- $$\frac{d}{d\xi}\sigma^{\text{NLO}}_{HH}(\mathbf{c}_i) = \frac{Poly(\mathbf{c}_i, \mathbf{A}^j)}{Poly(\mathbf{c}_{\text{SM}}, \mathbf{A}^j)} \frac{d}{d\xi}\sigma^{\text{NLO}}_{HH}(\mathbf{c}_{\text{SM}})$$

Binning

- Preliminary binning
  - 1D binning in the high-mass region ($m_{hh} \geq 1000$ GeV)
  - m_{hh}= {250, 270, 290, 310, 330, 350, 370, 390, 410, 430, 450, 470, 490, 510, 530, 550, 570, 590, 610, 630, 650, 670, 690, 710, 730, 750, 800, 850 900, 950, 1000, 1050, 1150, 1200, 1300, 1400, 1500, 1600, 1800, 2000, 13000}
  - $|\cos\theta^*|$={0.0, 0.25, 0.5, 0.75, 1} (applied to $m_{hh} < 1000$ GeV)
  - p_T^{hh} = {0, 20, 40, 60, 80, 100, 140, 200, 290, 2500} (applied to $m_{hh} < 1000$ GeV)
    - Other variables were tested, but p_T^{hh} is more suitable when considering experimental cuts in the differential distributions

A-coefficients fit

HEFT

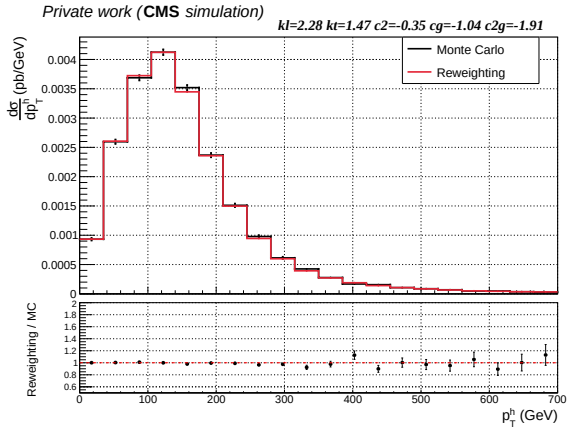
- 5000 (1000) samples were used at 13.6 TeV (13 TeV)
- All LO $A-$coefficients ($A_1 - A_{15}$) converged well for both $\sqrt{s} = 13$ and 13.6 TeV
- NLO coefficients ($A_{16} - A_{23}$) not so well converged as the LO ones

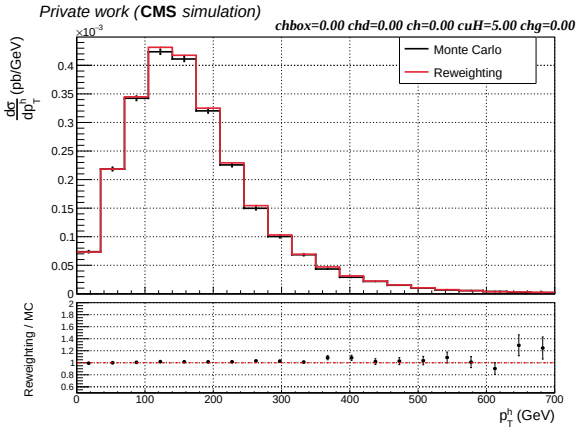
SMEFT

- We used $\sim$200 samples in both cases, 13 and 13.6 TeV
  - Most of the samples vary just one/two Wilson coefficients
  - For the best result in all corners of Wilson coefficients phase space, we need better constructed scans, but a first result is presented and show good closure

Tests on generator-level variables: p_T^h raw distribution

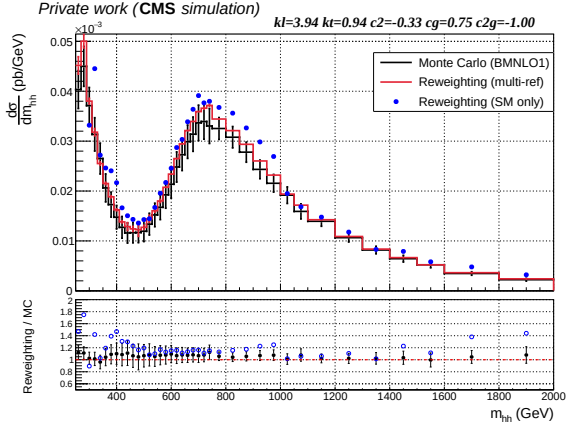
Tests on HEFT 13.6 TeV and SMEFT 13 TeV

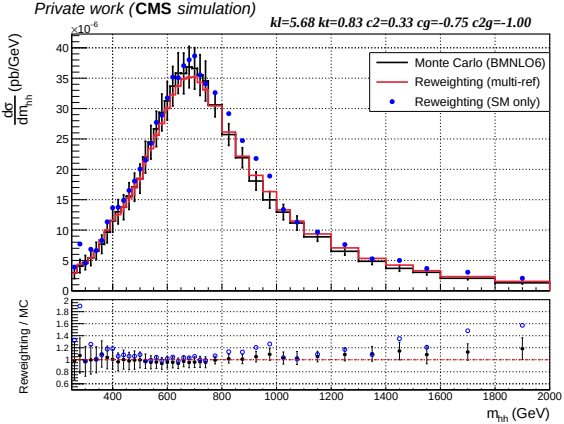




Tests on generator-level variables: Validation in realistic phase spaces

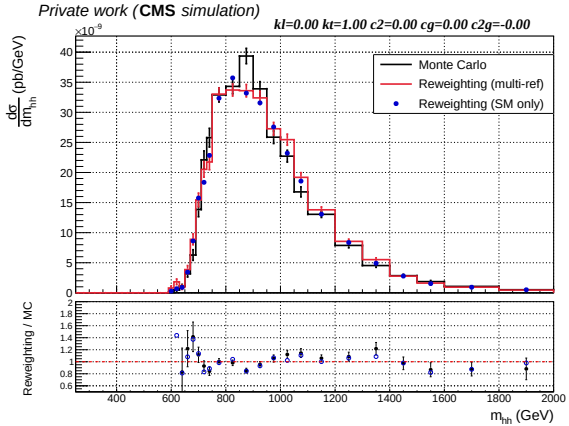
Tests on benchmarks defined at M. Capozi & G. Heinrich
(baseline cuts $HH \to 4b$ boosted selection, Run 2—HEFT)

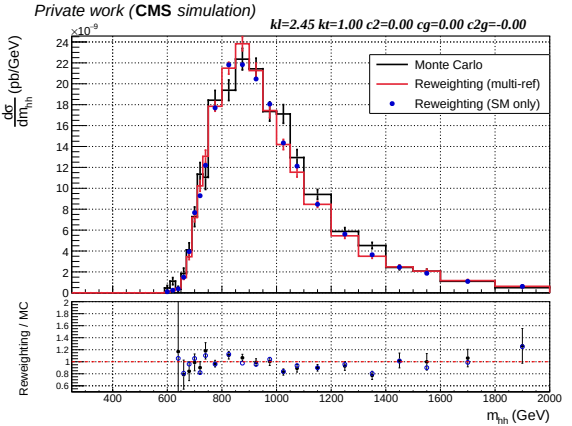




Tests on generator-level variables: Validation in realistic phase spaces

Tests on $\kappa_\lambda - \kappa_t$ benchmarks
(baseline cuts $HH \to 4b$ boosted selection, Run 3—HEFT)





Matrix-inversion method [1]

- In a nutshell:
  - The matrix-inversion method allows us to perform all cuts and statistical evaluation just in a subset of basis samples, and then extrapolate this to other point in the EFT space
  - Event-level reweighting allows a consistent scan of signal topologies to find the best input choices/basis choice

- CMS and ATLAS have used this approach in simpler parameter spaces
  - CMS: $\sigma^{\texttt{ggf}}_{HH}(\kappa_\lambda, \kappa_t | c_2 = c_g = c_{2g} = 0)$ (3 samples), $\sigma^{\texttt{ggf}}_{HH}(\kappa_\lambda, \kappa_t, c_2 | c_g = c_{2g} = 0)$ (6 samples)
  - ATLAS: $\sigma^{\texttt{ggf}}_{HH}(\kappa_\lambda, c_2, c_g | \kappa_t = 1, c_{2g} = 0)$ (10 samples)
  - How to find the full 5D 23-component basis for HEFT? $\sigma^{\texttt{ggf}}_{HH}(\kappa_\lambda, \kappa_t, c_2, c_g, c_{2g})$ (23 samples)

[1] Please see Torben's matrix inversion method for ggHH in the last LHCHWG meeting (here)

Final remarks

- ☐ Preliminary formulas for HEFT and SMEFT at 13 and 13.6 TeV: DONE
- ☐ First results look promising
- ☐ Significant improvements have been achieved in the 23-component HEFT basis

- ☐ Ongoing efforts:
  - Further improvements to the SMEFT coefficients are still needed
    - We're improving the SMEFT grid
    - Alternative parameterizations are being explored
  - Couplings present in HHH (κ_4) can be explored in di-Higgs via loop contributions
    - We're progressing towards formulas covering these coefficients (HEFT)

Final Remarks

- Paper is underway
- The final reweighting formulas should also appear in the upcoming LHC WG4 document, with recommendations concerning their usage:
  - i) Raw m_{hh} shapes are best described by the 1D m_{hh} formula
    - $\rightarrow$ To determine the HEFT 5D basis, use the 1D formula.

Final Remarks

- Paper is underway
- The final reweighting formulas should also appear in the upcoming LHC WG4 document, with recommendations concerning their usage:
  - i) Raw m_{hh} shapes are best described by the 1D m_{hh} formula
    - $\rightarrow$ To determine the HEFT 5D basis, use the 1D formula.
  - ii) p_T^h shapes are best described by 3D formulas
    - e.g., $\{m_{hh}, \cos\theta^*, p_T^{hh}\}$

Final Remarks

- Paper is underway
- The final reweighting formulas should also appear in the upcoming LHC WG4 document, with recommendations concerning their usage:
  - i) Raw m_{hh} shapes are best described by the 1D m_{hh} formula
    - $\rightarrow$ To determine the HEFT 5D basis, use the 1D formula.
  - ii) p_T^h shapes are best described by 3D formulas
    - e.g., $\{m_{hh}, \cos\theta^*, p_T^{hh}\}$
  - iii) For describing shapes with experimental cuts, the key variables affected by such cuts must be included as basis variables in the reweighting procedure
    - e.g., p_T^{hh} encodes cuts applied at the analysis level

Acknowledgments

- We thank:
  - Nicola D. Filippis for HEFT and SMEFT samples
  - Yuxiang Cao for SMEFT samples
  - Irene Dutta for the $HH \to 4b$ Run 2 samples with the baseline cuts
  - Stephane Cooperstein, Raffaele Gerosa and Soumya Mukherjee for the help in defining a proxy region of the boosted $HH \to 4b$ analysis and providing the samples (Run 3)

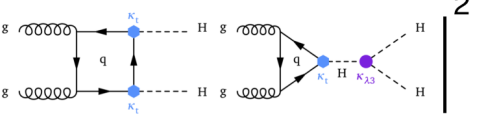
Thank you!

Backup

Backup: Matrix-inversion method

Example: 2D parameter space
3 samples needed for the matrix-inversion method

$$\sigma_{HH}^{ggF}(\kappa_\lambda, \kappa_t) \propto |A|^2 = \left| \right|^2$$



$$\kappa_\lambda^2 \cdot \kappa_t^2 \cdot |\triangle|^2 + \kappa_t^4 \cdot |\square|^2 + \kappa_t^3 \cdot \kappa_\lambda \cdot |\triangle^* \square + \square^* \triangle| = \mathrm{c}^T(\kappa_\lambda, \kappa_t) \cdot \mathrm{v} = \sigma_{HH}^{ggf}(\kappa_\lambda, \kappa_t)$$

$$\vec{\sigma} = \begin{pmatrix} \sigma^a \\ \sigma^b \\ \sigma^c \end{pmatrix} = \begin{pmatrix} c_1^a & c_2^a & c_3^a \\ c_1^b & c_2^b & c_3^b \\ c_1^c & c_2^c & c_3^c \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \mathbf{C} \cdot \mathbf{v}$$

$$\mathrm{v} = \mathrm{C}^{-1} \cdot \vec{\sigma}$$

$$\sigma_{HH}^{ggf}(\kappa_\lambda, \kappa_t) = \mathrm{c}^T(\kappa_\lambda, \kappa_t) \cdot \mathrm{v} = \mathrm{c}^T(\kappa_\lambda, \kappa_t) \cdot \mathrm{C}^{-1} \cdot \vec{\sigma}$$

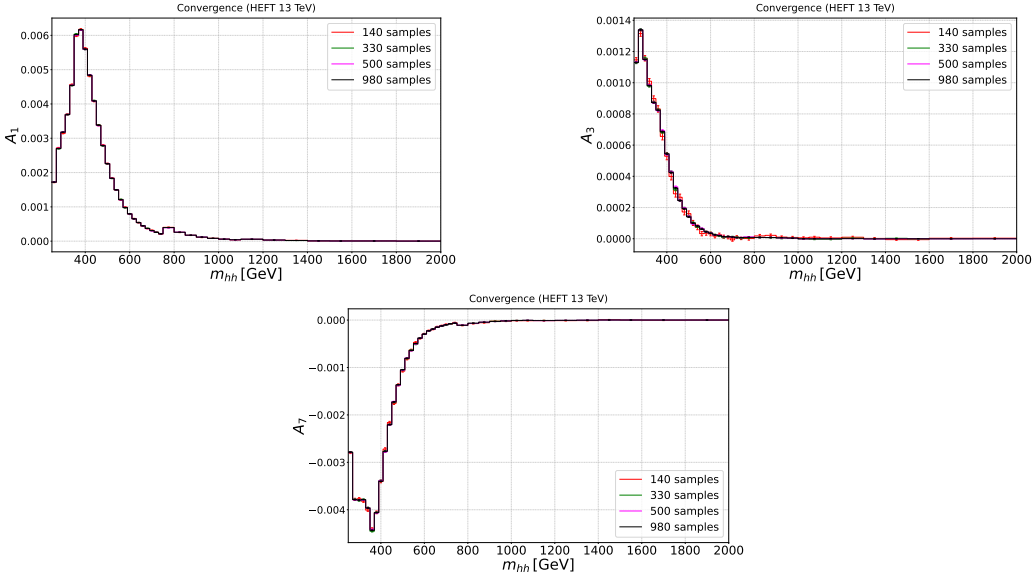


Figure: Convergence of the coefficients defining the SM-like subspace

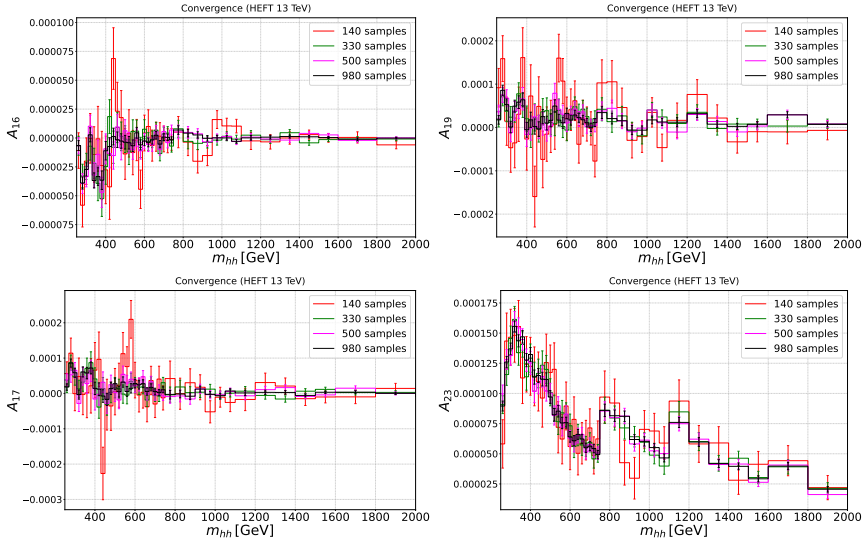


Figure: Convergence of four of the eight coefficients related to NLO QCD contributions in the differential cross section. As a general behavior, all NLO coefficients carry greater uncertainties than LO ones. (13 TeV HEFT)

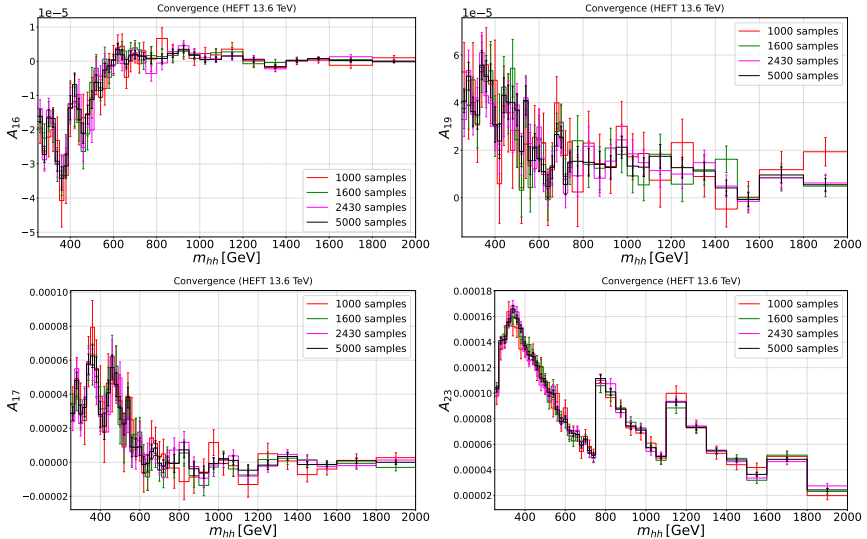
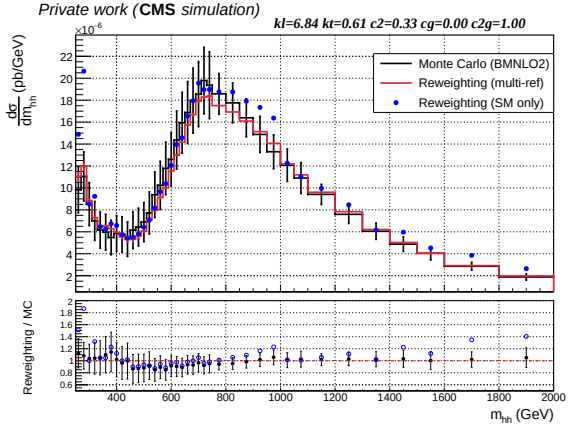
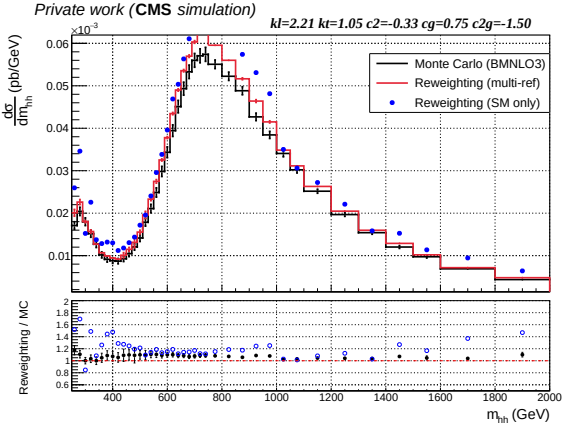


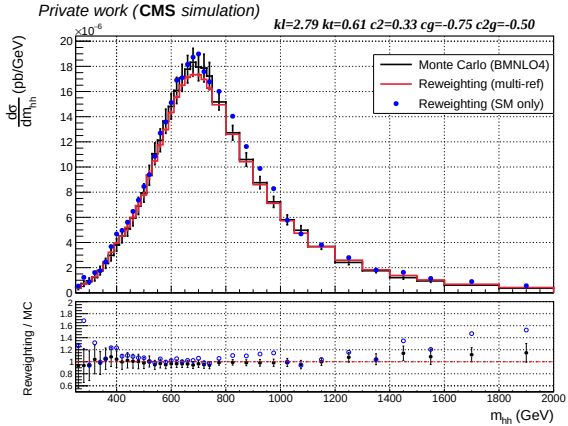
Figure: Convergence of four of the eight coefficients related to NLO QCD contributions in the differential cross section. As a general behavior, all NLO coefficients carry greater uncertainties than LO ones. (13.6 TeV HEFT)

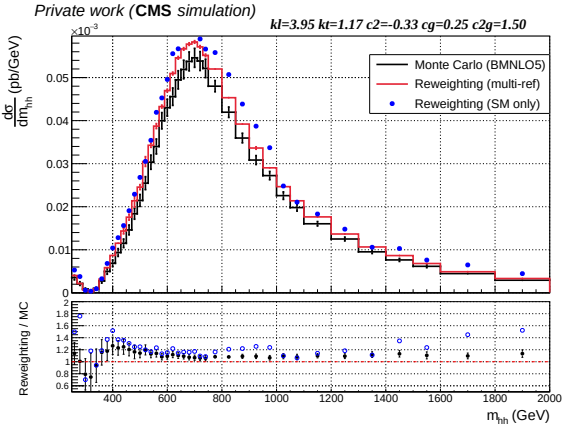
Tests on benchmarks defined at M. Capozi & G. Heinrich
(baseline cuts $HH \to 4b$ boosted selection, Run 2—HEFT)



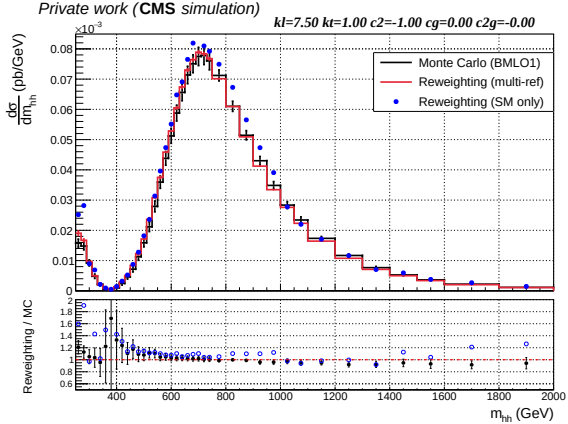


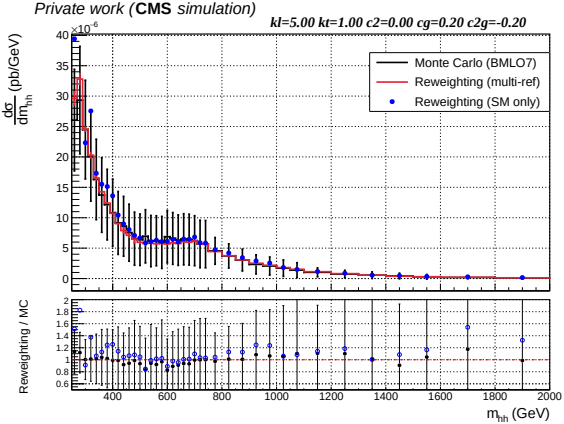
Tests on benchmarks defined at M. Capozi & G. Heinrich
(baseline cuts $HH \to 4b$ boosted selection, Run 2—HEFT)





Tests on benchmarks defined at A. Carvalho et al.
(baseline cuts $HH \to 4b$ boosted selection, Run 2—HEFT)





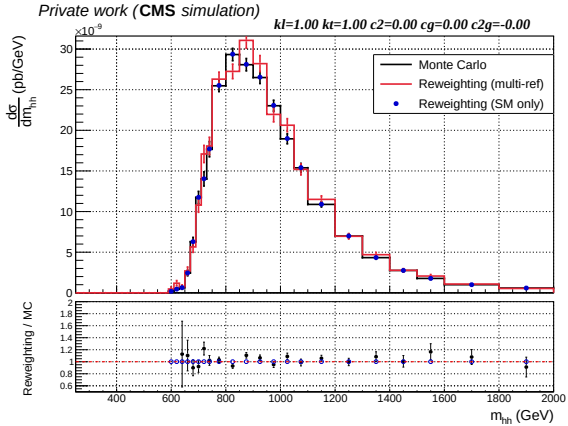
Backup—Validation in realistic phase spaces

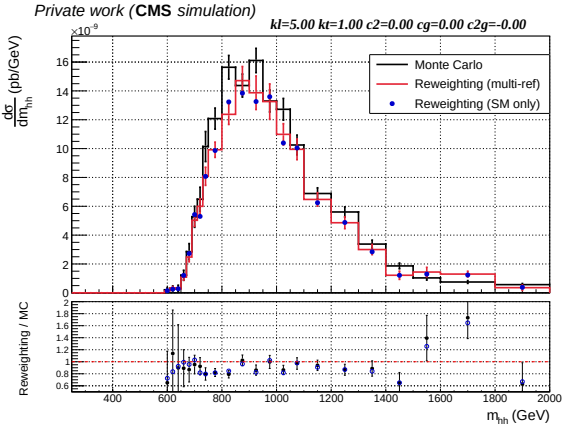
Benchmarks	Reweighted σ (fb)	MC σ (fb)	Central Deviation
BMNLO1	28.40 ± 0.07	26.8 ± 0.7	$+6\%$
BMNLO2	15.21 ± 0.02	15.3 ± 0.5	-1%
BMNLO3	42.95 ± 0.07	40.1 ± 1.0	$+7\%$
BMNLO4	9.05 ± 0.02	9.1 ± 0.2	-1%
BMNLO5	34.96 ± 0.06	31.7 ± 0.8	$+10\%$
BMNLO6	21.21 ± 0.04	20.9 ± 0.6	$+1\%$
BMNLO7	27.68 ± 0.04	26.0 ± 0.9	$+6\%$

Backup—Validation in realistic phase spaces

Benchmarks	Reweighted σ (fb)	MC σ (fb)	Central Deviation
BMLO1	36.49 ± 0.05	36.6 ± 0.6	$\sim 0\%$
BMLO2	5.919 ± 0.008	5.52 ± 0.03	$+7\%$
BMLO3	124.8 ± 0.2	126.8 ± 0.2	-2%
BMLO4	700 ± 2	723 ± 3	-3%
BMLO5	18.95 ± 0.03	17.76 ± 0.06	$+7\%$
BMLO6	3.221 ± 0.008	3.09 ± 0.06	$+4\%$
BMLO7	6.74 ± 0.04	6.7 ± 0.1	$+1\%$
BMLO8	7.17 ± 0.01	7.21 ± 0.05	-1%
BMLO9	28.83 ± 0.04	28.11 ± 0.09	$+3\%$
BMLO10	42.5 ± 0.1	45 ± 1	-6%
BMLO11	25.98 ± 0.06	24.2 ± 0.3	$+7\%$
BMLO12	147.1 ± 0.5	156 ± 2	-6%

Tests on $\kappa_\lambda - \kappa_t$ bencharmks
(baseline cuts $HH \to 4b$ boosted selection, Run 3—HEFT)





Benchmarks defined at M. Capozi & G. Heinrich

Benchmarks	κ_λ	κ_t	c_2	c_g	c_{2g}
1	3.94	0.94	$-\frac{1}{3}$	$\frac{3}{4}$	-1
2	6.84	0.61	$\frac{1}{3}$	0.0	1
3	2.21	1.05	$-\frac{1}{3}$	$\frac{3}{4}$	$-\frac{3}{2}$
4	2.79	0.61	$\frac{1}{3}$	$-\frac{3}{4}$	$-\frac{1}{2}$
5	3.95	1.17	$-\frac{1}{3}$	$\frac{1}{4}$	$\frac{3}{2}$
6	5.68	0.83	$\frac{1}{3}$	$-\frac{3}{4}$	-1
7	-0.10	0.94	1.0	$\frac{1}{4}$	$\frac{1}{2}$

Benchmarks defined at A. Carvalho et al.

Benchmarks	κ_λ	κ_t	c_2	c_g	c_{2g}
1	7.5	1.0	−1.0	0.0	0.0
2	1.0	1.0	0.5	−0.8	0.6
3	1.0	1.0	−1.5	0.0	−0.8
4	−3.5	1.5	−3.0	0.0	0.0
5	1.0	1.0	0.0	0.8	−1.0
6	2.4	1.0	0.0	0.2	−0.2
7	5.0	1.0	0.0	0.2	−0.2
8a	1.0	1.0	0.5	0.4	0.0
9	1.0	1.0	1.0	−0.6	0.6
10	10.0	1.5	−1.0	0.0	0.0
11	2.4	1.0	0.0	1.0	−1.0
12	15.0	1.0	1.0	0.0	0.0