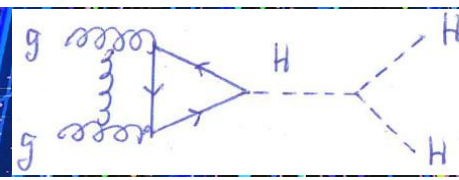
Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
ICSC Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing

HH XSECTION MODELING VIA ML TOOLS



Nicola De Filippis
Politecnico & INFN, Bari

Harrison Prosper, Kurtis Johnson, R. Starnes
Florida State University

Goals of this study

- ►Infer a detailed modeling of the HH xsection as a function of the HEFT/SMEFT model parameters by using ML tools
  - ►suitable for ICSC resources with CPU/GPUs

- ►Interpret HH results in terms of the 5-parameters HEFT and then SMEFT

- ►Use a likelihood-free approach to construct confidence interval → statistical interpretation (info in backup)

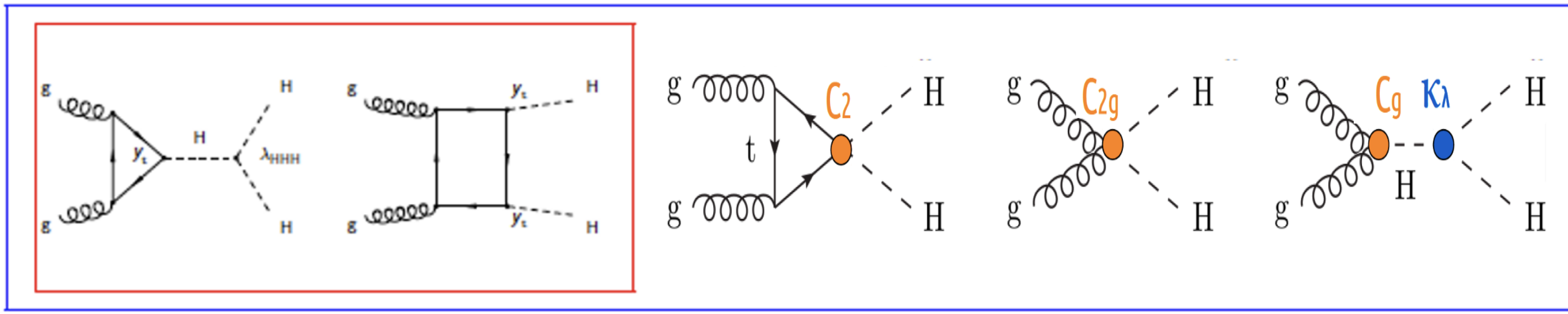
BSM physics with HH production: EFT

Effective field theories (EFTs) introduce higher dimensional operators of the SM Lagrangian → new physics exist at very large energy scales

HEFT ggF cross section modeling with three new contact interactions (couplings): ttHH (C_2), ggHH (C_{2g}) and ggH (C_g)

JHEP **11** (2019) 024

SM $(k_\lambda=1,\ k_t=1,\ c_{2g}=c_g=c_2=0)$ **BSM**

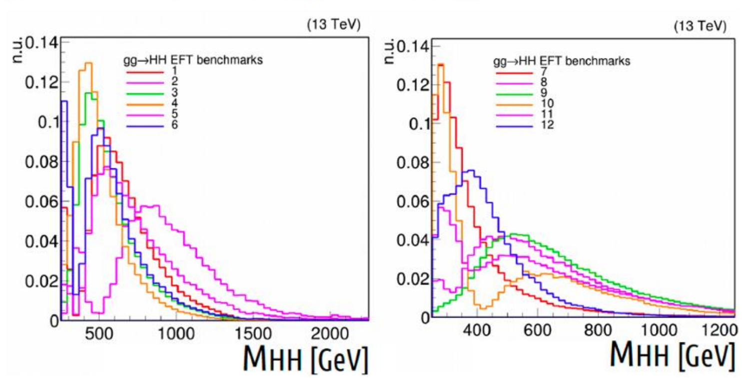


$$\mathcal{L}_h = \frac{1}{2}\partial_\mu h \partial^\mu h - \frac{1}{2} m_h^2 h^2 - k_\lambda \lambda_{SM} v h^3 - \frac{m_t}{v}\left(v + k_t h + \frac{c_2}{v} hh\right)(\bar{t}_L t_R + h.c.) + \frac{1}{4}\frac{\alpha_s}{3\pi v}\left(c_g h - \frac{c_{2g}}{2v} hh\right) G^{\mu\nu} G_{\mu\nu}$$

12 EFT benchmarks are defined for LHC searches in arXiv:1806.05162v3.

- they represent topologies of large regions of the 5D parameter space

7 EFT benchmarks from JHEP03(2020)091



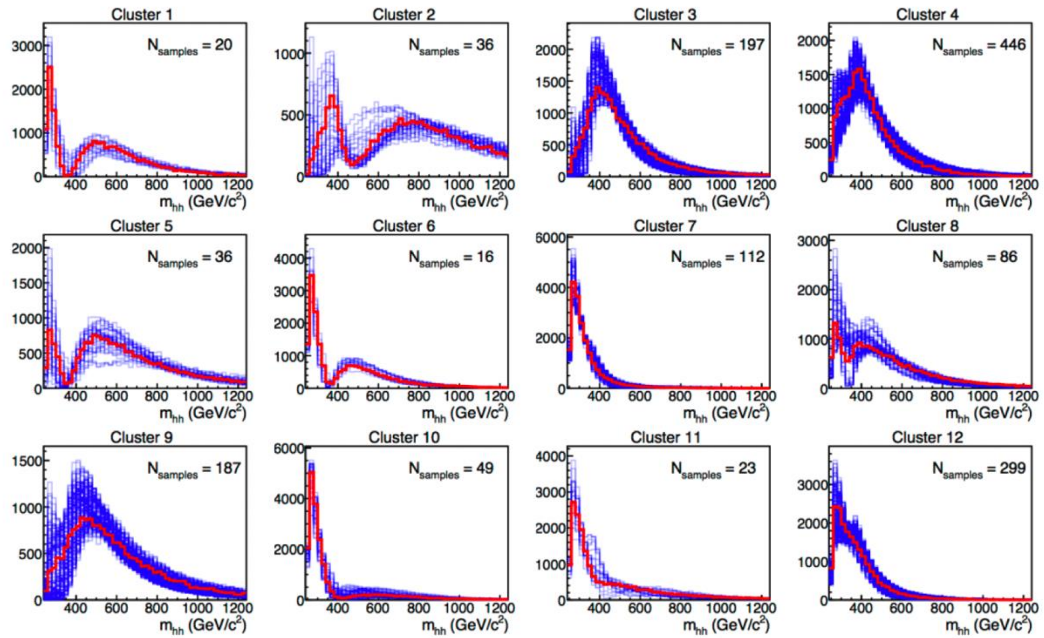
BSM physics with HH production: EFT

	1	2	3	4	5	6	7	8	9	10	11	12
κ_λ	7.5	1.0	1.0	-3.5	1.0	2.4	5.0	15.0	1.0	10.0	2.4	15.0
κ_t	1.0	1.0	1.0	1.5	1.0	1.0	1.0	1.0	1.0	1.5	1.0	1.0
C_2	-1.0	0.5	-1.5	-3.0	0	0	0	0	1.0	-1.0	0	1.0
C_g	0	-0.8	0	0	0.8	0.2	0.2	-1.0	-0.6	0	1.0	0
C_{2g}	0	0.6	-0.8	0	-1.0	-0.2	-0.2	1.0	0.6	0	-1.0	0

BSM benchmarks based on non-linear HEFT

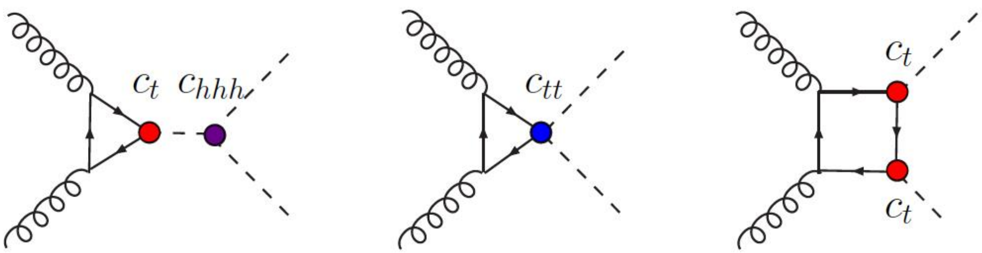
EWChL eq. (2.6)	Ref. [71]
c_{hhh}	κ_λ
c_t	κ_t
c_{tt}	c_2
c_{ggh}	$\frac{2}{3}c_g$
c_{gghh}	$-\frac{1}{3}c_{2g}$

Benchmark	c_{hhh}	c_t	c_{tt}	c_{ggh}	c_{gghh}
1	7.5	1.0	-1.0	0.0	0.0
2	1.0	1.0	0.5	$-\frac{1.6}{3}$	-0.2
3	1.0	1.0	-1.5	0.0	$\frac{0.8}{3}$
4	-3.5	1.5	-3.0	0.0	0.0
5	1.0	1.0	0.0	$\frac{1.6}{3}$	$\frac{1.0}{3}$
6	2.4	1.0	0.0	$\frac{0.4}{3}$	$\frac{0.2}{3}$
7	5.0	1.0	0.0	$\frac{0.4}{3}$	$\frac{0.2}{3}$
8a	1.0	1.0	0.5	$\frac{0.8}{3}$	0.0
9	1.0	1.0	1.0	-0.4	-0.2
10	10.0	1.5	-1.0	0.0	0.0
11	2.4	1.0	0.0	$\frac{2.0}{3}$	$\frac{1.0}{3}$
12	15.0	1.0	1.0	0.0	0.0
SM	1.0	1.0	0.0	0.0	0.0



BSM physics with HH production: EFT

ICSC
Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing



Lina Alasfar *et al.*, arXiv:2304.01968v1

$$\mathcal{M} = \mathcal{M}_{\rm SM} + \mathcal{M}_{\rm dim6} + \mathcal{M}_{\rm dim6^2}$$

$$\sigma \simeq \begin{cases} (a)\ \sigma_{\rm SM\times SM} + \sigma_{\rm SM\times dim6} \\ (b)\ \sigma_{\rm (SM+dim6)\times(SM+dim6)} \\ (c)\ \sigma_{\rm (SM+dim6)\times(SM+dim6)} + \sigma_{\rm SM\times dim6^2} \\ (d)\ \sigma_{\rm (SM+dim6+dim6^2)\times(SM+dim6+dim6^2)} \end{cases}$$

$$\sigma \sim |\mathcal{M}|^2$$

The HH production xsection via gluon fusion can be parameterised for any set of HEFT Wilson coefficients at NLO:

- A is a set of 23 coefficients determined from simulation
- c^T represents the vector of products of Wilson coefficients

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

EFT xsection modeling by using NN

Currently in the HH model the xsection is studies in terms of m_{hh} and $\cos(\theta)^*$ and parameterized as a function of C_2

The physics model of interest is the HEFT, defined by a 5D parameter space of Wilson coefficients $\theta = c_{hhh}, c_t, c_{tt}, c_{ggh}, c_{gghh}$

POWHEG ggHH package used for MC simulation (**G. Heinrich et al.).** Efficient workflow implemented in Bari ReCas computing center

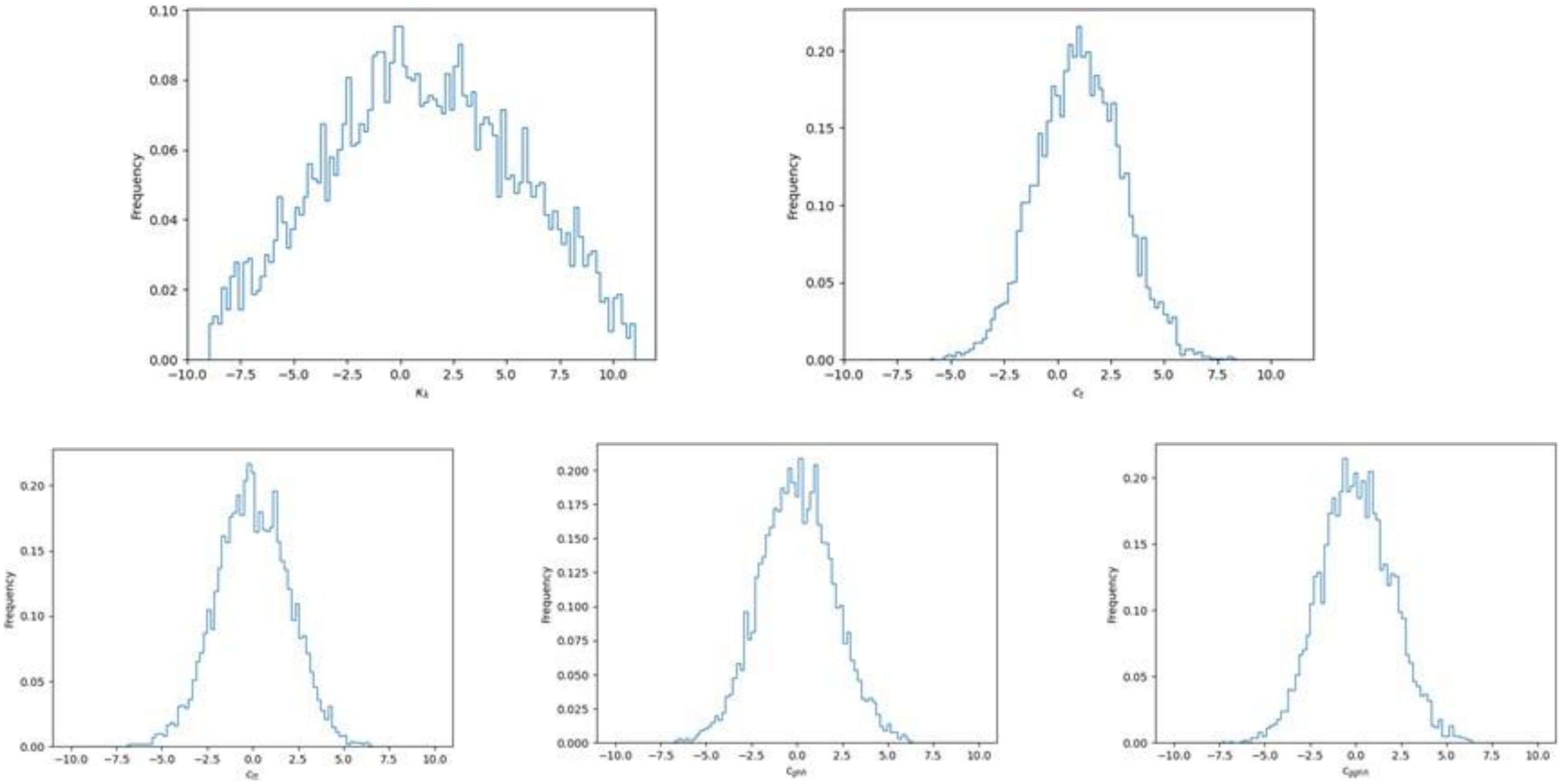
In principle we need to sample the HEFT parameter space sufficiently densely and, at every point, simulate data like the ones actually observed → computationally infeasible
→ a sparser sampling of the parameter space and reweight existing simulated data to mimic the sampling of data at any other point θ

To do this requires knowledge of the cross section per binned observables, and/or the differential cross section, as a function of both the *observables* (m_{hh}) and the *parameters* (θ) → one **first goal of the project is construct a parameterization of this function**.

HEFT modeling with 5 parameters by using NN

Scan of the 5D EFT parameter space:

- parameters distributed as a Gaussian distribution around the Standard Model value and randomly mixed



4000 samples generated (40k events each) + 1000 sample for SM configuration

EFT xsection modeling by using NN

As a first attempt, we restrict our attention to a single observable, the di-Higgs mass m_{hh} (23 functions).

$$\sigma(m_{hh}, \theta) = \boldsymbol{c}^T(\theta) \cdot \boldsymbol{a}(m_{hh}),$$

23 functions $a(m_{hh})$

$$\boldsymbol{c}^T(\theta) = (c_t^4, c_{tt}^2, c_t^2 c_{hhh}^2, c_{ggh}^2 c_{hhh}^2, c_{gghh}^2, c_{tt} c_t^2, c_{hhh} c_t^3, c_t c_{hhh} c_{tt}, c_{ggh} c_{hhh} c_{tt}, c_{tt} c_{gghh}, c_{ggh} c_{hhh} c_t^2, c_{gghh} c_t^2, c_{hhh}^2 c_{ggh} c_t, c_{gghh} c_t c_{hhh}, c_{ggh} c_{gghh} c_{hhh}, c_t^3 c_{ggh}, c_t c_{tt} c_{ggh}, c_t c_{ggh}^2 c_{hhh}, c_t c_{ggh} c_{gghh}, c_t^2 c_{ggh}^2, c_{tt} c_{ggh}^2, c_{ggh}^3 c_{hhh}, c_{ggh}^2 c_{gghh}),$$

We modeled the 23 functions $a(m_{hh})$ using a single DNN with one input, m_{hh} , and 23 outputs, one for each function, and a 23-parameter neural network

Our NN model, **HEFTNET**, models the mapping

$$f : m_{hh}, k_{lambda}, c_t, c_{tt}, c_{ggh}, c_{gghh} \rightarrow \sigma$$

using the known functional dependence of the cross section on the Wilson coefficients, $\theta = c_{hhh}, c_t, c_{tt}, c_{ggh}, c_{gghh}$, and the m_{hh}-dependent functions **a**.

The mapping $g : m_{hh} \rightarrow \boldsymbol{A}$ is modeled with a function of the form $A_j(m_{hh}) = P_j(m_{hh}) e^{Q_j(m_{hh})}$ where P is a neural network with one input and 23 outputs and $Q_j = w_j\, m_{hh}$ is a 23-parameter neural network with free parameters w_j.

Resources used for NN workflow

Different hyperparamter configurations run at:

- Jupyter Hub at RECAS Bari
- CERN SWAN interactive service
- ICSC resources:
  https://hub.131.154.98.51.myip.cloud.infn.it

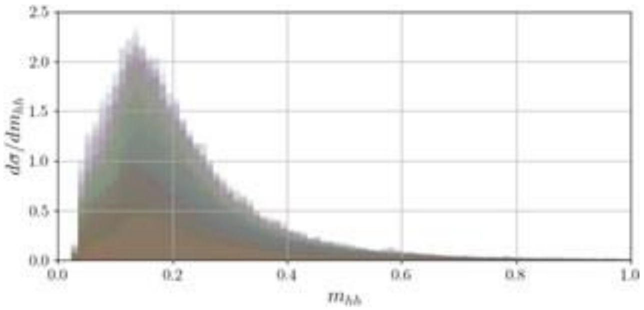
CPU vs GPUs also tested for the NN training:

- time consumption from 1 day to 1 hour !

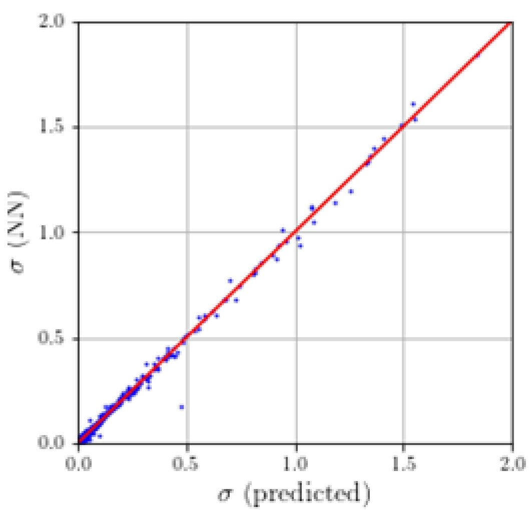
- Precise tests and measurements need to be done in a more systematic and coherent way → suggestions about benchmarks / metrics for ICSC official studies?

EFT xsection modeling by using NN: results

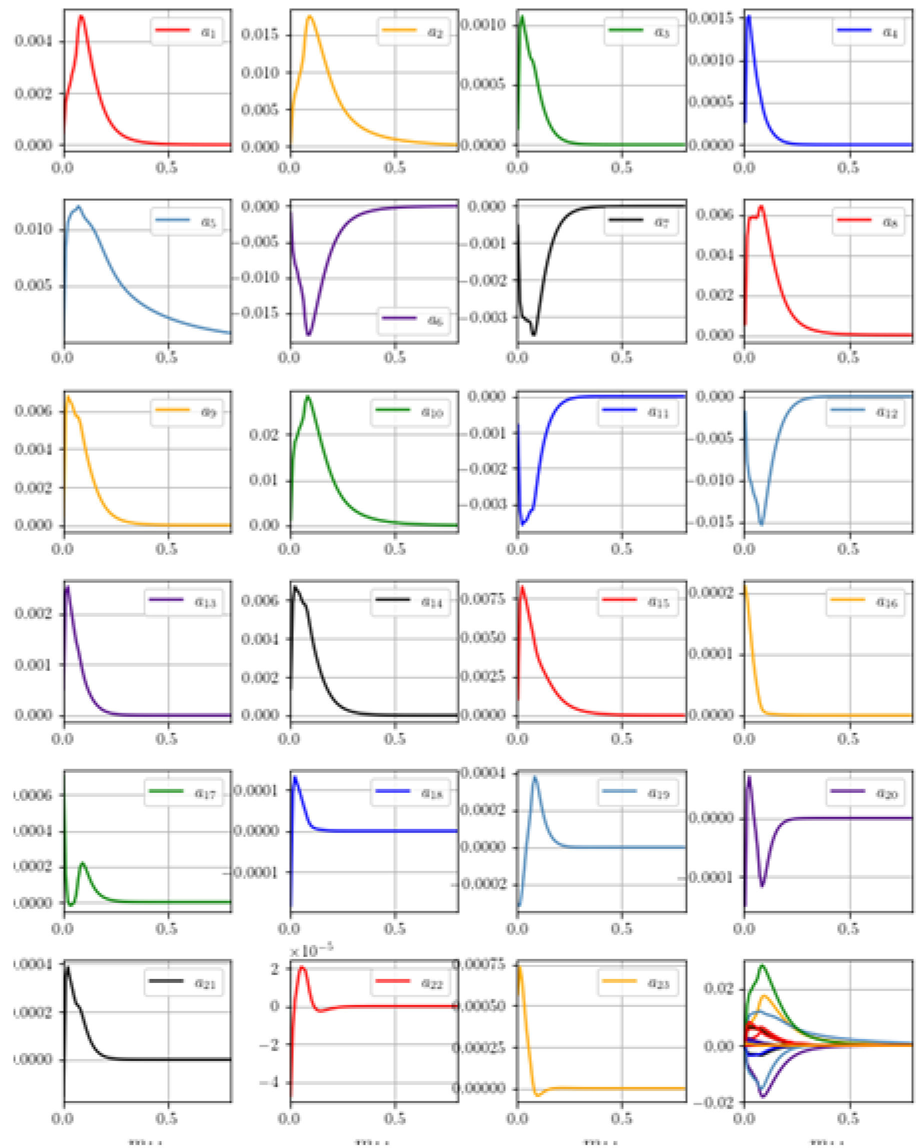
Di-Higgs mass m_{hh} sampled bin per bin and provided as an input to the HEFTNET model



NN prediction vs. predicted xsection

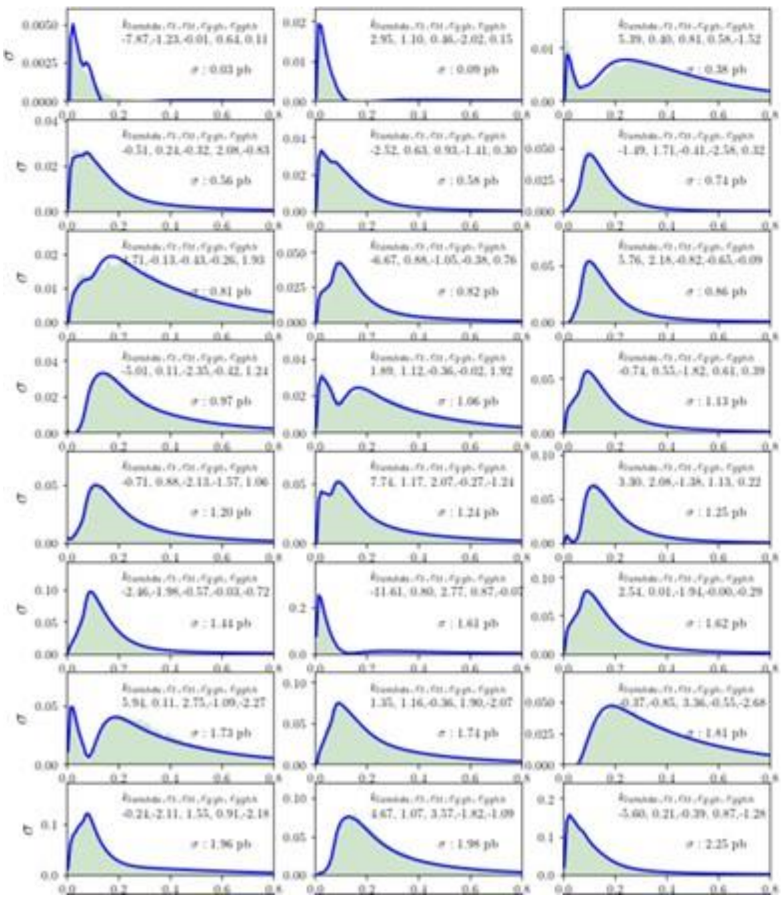


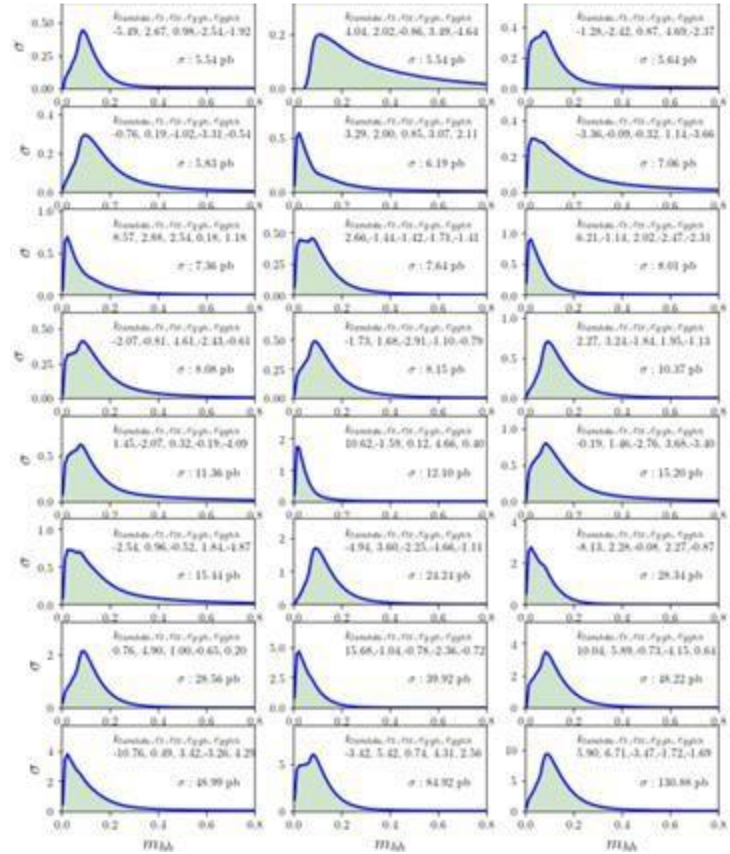
a(m_H) coefficients



EFT xsection modeling by using NN: results

Prediction in terms of differential and total xsection





Very good agreement so far!

Target: given the $a(m_{hh})$ coefficients and the known dependence of the xsection on the Wilson coefficients, we can compute the differential xsection for any combination of the Wilson coefficients → effective reweighting procedure based on NN → github package including the NN model and an example on how to use it

Summary and To Do List

- Infer a detailed modeling of the HH xsection as a function of the HEFT/SMEFT model parameters
- Tune the effectiveness of the xsection parameterization in the whole EFT parameter space → provide a parameterization
- **Testing the workflow with ICSC and other HPC resources → benchmarking the test and results to be done**
- Moving to the statistical treatment by using a likelihood-free approach → produce HH results in the EFT parameter space
- Generalizing the procedure in case of SMEFT

Backup

Likelihood-free approach

For the purpose of the statistical analysis and interpretation of HH results:

- we need to built the likelihood function λ and then the cumulative distribution function $\mathbb{P}(\lambda \leq \lambda_0 | \theta)$ if θ is the set of parameters of the HEFT/SMEFT model

- in principle, to approximate the cumulative distribution function $\mathbb{P}(\lambda \leq \lambda_0 | \theta)$ we need to sample the HEFT/SMEFT parameter space sufficiently densely and, at every point, simulate data like the ones actually observed → then the likelihood is known

- however, if it is computationally infeasible to sample the parameter space densely enough, the next option is to use a sparser sampling of the parameter space and reweight existing simulated data to mimic the sampling of data at any other point θ.

We propose to use a likelihood-free frequentist inference (LF2I) approach that makes it possible to construct confidence sets with the p-value function and to use the same function to check the coverage explicitly at any given parameter point.

LF2I

Ann Lee *et al.*, https://arxiv.org/abs/2107.03920
Ali Al Kadhim *et al.*, https://iopscience.iop.org/article/10.1088/2632-2153/ad218e

Likelihood-free approach

We shall use simulation-based inference to construct **confidence sets**, at confidence level (CL) τ, in the HEFT and later SMEFT parameter spaces.

This requires approximating $\mathbb{P}(\lambda \leq \lambda_0 | \theta) = \mathbb{E}(Z | \theta)$ via:

1) a machine learning (ML) model, where for a given hypothesis $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$, λ_0 is the observed value of a test statistic λ with the property that large values of the test statistic disfavor the hypothesis H_0.

2) the function $\lambda(m_{hh}, \theta)$ can be modeled directly by minimizing the average exponential loss, or indirectly by treating the problem as a classification. In this case, we minimize the average cross-entropy loss.

Likelihood-free approach

With a balanced dataset of HEFT and SM events, minimizing the average cross entropy yields a model that approximates

D is the outcome of the NN classificator

$$D = \frac{f(x,\theta)}{f(x,\theta) + f_{SM}(x,\theta)}$$

For the SM, $f_{SM}(x,\theta) = f_{SM}(x)\pi_\theta$ because the observables and the HEFT parameters are obviously unrelated. We can write

$$\ln\frac{1-D}{D} = \ln\frac{f_{SM}(x,\theta)}{f(x,\theta)}$$

$$= \ln\frac{f_{SM}(x)\pi_\theta}{f(x|\theta)\pi_\theta} = \lambda(x;\theta)$$

Given the cumulative distribution function (cdf), $\mathbb{P}(\lambda \leq \lambda_0 | \theta)$, a confidence set at CL τ is the set of θ values for which $\mathbb{P}(\lambda \leq \lambda_0 | \theta) \leq \tau$ → statistical interpretation

→ **To be completed**