Common format & toolchain for SMEFT parametrisations

Joint WG2/LHCEFTWG Area 2 activity

I. Brivio, A. Cueto, C. Knight, J. Langford, K. Mimasu, **E. Rossi**

Motivation

Idea: *Create a* ***common data format*** *to publish SMEFT parametrisations (Warsaw basis) & provide* ***tool*** *for generating them +* ***benchmarks*** *useful for own parameterisation*

- Theoretical parameterisations are key inputs to statistical models for interpretations of HEP data
- In the SMEFT, parameterisations are rather simple, e.g. at dim-6:

$$O = O_{\mathrm{SM}} + \sum_i \frac{C_i}{\Lambda^2} O_i^{\mathrm{int.}} + \sum_{i,j} \frac{C_i C_j}{\Lambda^4} O_{i,j}^{\mathrm{sq.}} \propto \left| \mathcal{A}_{\mathrm{SM}} + \sum_i \frac{C_i}{\Lambda^2} \mathcal{A}_i \right|^2$$

1 + n(n+3)/2 monomials for n operators

Previous talks in LHCHWG and LHC EFT WG

- Parameterisations can be quite challenging to obtain
- Improve efficiency & accuracy/validity of SMEFT interpretation workflow

Efficiency, accuracy & validity

- CMS, ATLAS & theorists often derive their **own parameterisations**
  - Duplication of efforts, reinventing the wheel
  - **Challenging** & **time-consuming**: potential for human error
  - Important to **validate calculations**
- A lot of exp./th. cross-talk already: technical expertise required
  - Support for existing tools `SMEFTsim` & `SMEFT@NLO` (e.g. **flavor assumptions)**
  - Developers spend time explaining how to do the same thing to different people
- More experiment/theory interactions → better SMEFT interpretations
  - From theory: state-of-the art models/analytical results, technical advices on input parameters/linear vs quadratic…
  - From Experiments: Acceptance corrections, tools like EFT2Obs

Common format

- **Stores** 1+n(n+3)/2 monomial coefficients
- **Includes** information about uncertainties
  - MC, PDF, scale, …
- **Sufficient metadata** to ensure that the results are reproducible
  - Settings, code versions, assumptions, … + cards/scripts and instructions to reproduce the parameterisation
- Important results can be reused by whole community
  - Subsequent experimental analyses or independent global interpretations
- Our exercise: **use STXS 1.2** as example implementation (+ decay)
  - Example of a “fixed” binning, agreed between LHC experiments

Example data format (`.json`)

- "metadata" field:
  - Information on shape of observable (number/list/matrix) & coeffs on which it depends
- "data" field:
  - contains monomial coefficients + errors
  - More than one error can be stored (MC, PDF, scale, …)

Not currently fixed. We would like feedback!

```
"metadata": {
 "coefficients": [   "chl3",   "chbox",   "chl1",   "chwb",   "chwbtil",
 "observable_shape": "(1,)",
 "observable_names": [   "ZZ"  ],
  ...
 },
"data": {
 "central": {
  "a_chl3": [        1.00                    ],
  ...
 "u_MC": {
  "a_chl3": [        0.01                    ],
  ...
```

"Main" prediction (a,b coefficients) →

e.g. MC error ("u_XY")→

More metadata fields → **reproducibility**


```
{
 "metadata": {
  "coefficients": [   "chl3",   "chbox",   "chl1",   "chwb",   "chwbtil",   "che",   "chdd",   "chb",   "chbtil",
  "observable_shape": "(1,)",
  "observable_names": [   "ZZ"  ],
  "description": "",
  "documentation": [   "https://github.com/MatthewDKnight/EFT2Obs/tree/Run2Legacy/cards"  ],
  "tool_version": "MG5_aMC_v2_6_7",
  "UFO": "SMEFTsim_topU3l_MwScheme_UFO",
  "flavor_scheme": "topU3l",
  "inputs": {
   "Lambda": 1000,
   "MW": 80.379,
   "MZ": 91.1876,
   "GF": 1.16638e-05,
   "aS": 0.1181,
   "MH": 125.0,
   "MB": 3.237,
   "MT": 173.2
  },
  "EW_input_scheme": "MW_MZ_GF",
  "EFT_order": "quadratic",
  "pertubative_order_QCD": "LO",
  "pertubative_order_QED/EW": "LO",
  "method": "reweighting"
 },
```


Not currently fixed. We would like feedback!

- Basis choice, EW input schemes, flavor assumptions,...
- Links to documentation
- Tools, versions,...
- MC settings (scale choice,...)
- Orders (EFT, perturbative,...)
- Method used
- Free-form fields
- Any other ideas? “Required” fields?

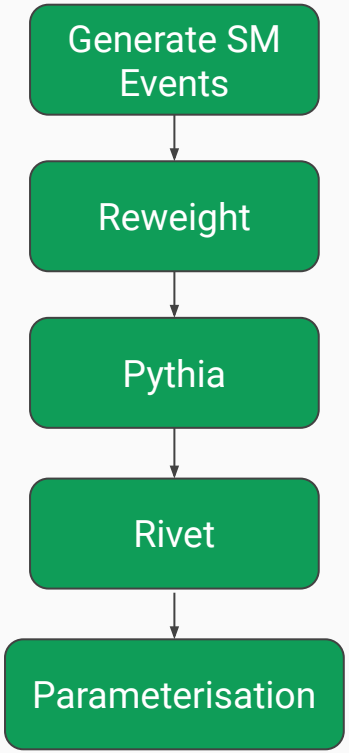
EFT2Obs toolchain

"A tool to automatically parametrize the effect of EFT coefficients on arbitrary observables." EFT2Obs

1. Generate events with MadGraph
2. Reweight those events during generation
   - c_i=0.5 and 1.0 for each c_i → A_i and B_ii terms
   - c_i=1, c_j=1 for each combination of c_i and c_j → B_ij i≠j terms
3. Events passed through Pythia
4. Classification into STXS performed by Rivet
5. Equations extracted from cross sections in each bin at each of the reweighting points

More details in the EFT2Obs README

Primarily reweighting based… but extensions have been developed for a generation based approach where necessary.



ATLAS/CMS validation exercise

- SMEFT parameterisation derived using EFT2Obs currently being validated by ATLAS
- Flavour assumptions used by the experiments:
  - ATLAS **top,** CMS **topU3l**
- Careful comparison of inputs (settings, EW parameters)
- Example of expected differences:
  - H->bb cbH coefficient: strictly depend on the b-mass choice which is different between the two experiments, e.g. pole mass or mass at the Higgs scale?
- Status: Higgs decays have been validated, production under way
- Numbers not yet public-> target: early 2024

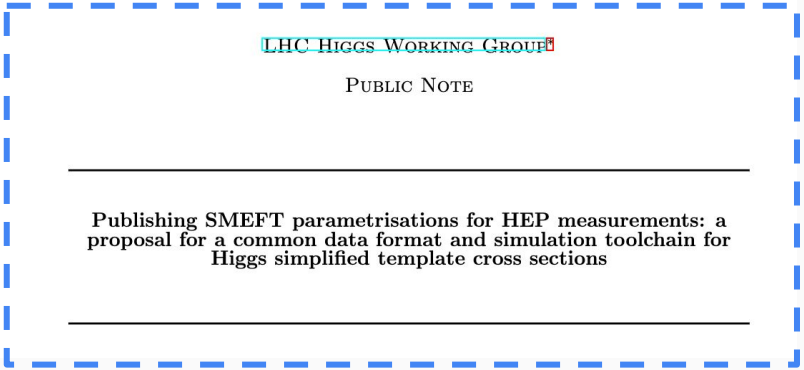
Settings & assumptions for STXS parameterisation

- Currently the parameterisation will be provided including common choices like mW scheme, topU3l flavour assumption, scale, STXS 1.2
- Analytical expressions for some channels
- How can this be further extended?
  - **Useful variants**: flavor schemes/propagator corrections/acceptance corrections/fixed vs dynamical scale
    - More relaxed flavor assumptions - benchmarks might be included to simplify checks
- Plan to publish: parameterisations + instructions and tools (EFT2Obs -> already public) to reproduce parameterisation
- Any requests/recommendations?

Plans, feedback? ideas?

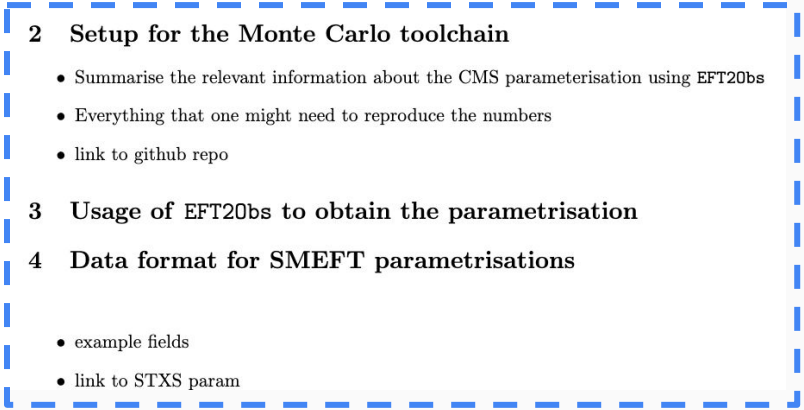
- Note in preparation
- Timeline for the publication: **Early 2024**

- Feedback on how to publish: currently we will put them on the EFT2Obs repo, zenodo, hepdata, some kind of DOI?
- E.g. the recent ATLAS STXS analysis is currently working on the HEPdata entry, would be nice to have a version of the parameterisation in this format


LHC Higgs Working Group

Public Note

Publishing SMEFT parametrisations for HEP measurements: a proposal for a common data format and simulation toolchain for Higgs simplified template cross sections


2 Setup for the Monte Carlo toolchain

- Summarise the relevant information about the CMS parameterisation using EFT2Obs
- Everything that one might need to reproduce the numbers
- link to github repo

3 Usage of EFT2Obs to obtain the parametrisation

4 Data format for SMEFT parametrisations

- example fields
- link to STXS param

THANK YOU!!