

CMC-H & SM-PDF

Stefano Carrazza, Stefano Forte
Università di Milano & INFN

in collaboration with Z. Kassabov, J. I. Latorre and J. Rojo
Les Houches, June 4, 2015



MC2HESSIAN AND CMC-H

mc2hessian

An **unbiased** Hessian **representation** for MC PDFs [[arXiv:1505.06736](https://arxiv.org/abs/1505.06736)]

Basic Idea

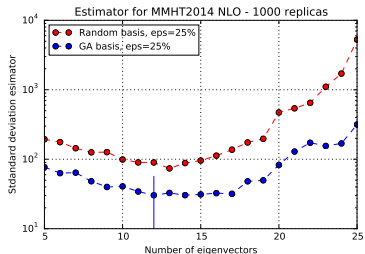
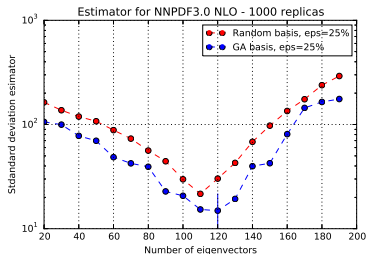
use **MC replicas** as a basis functions for Hessian representation

1. conversion quality based on the **relative uncertainty estimator**
2. symmetric eigenvectors determined by a **Genetic Algorithm**
3. There is **optimal** number of eigenvectors
 - technical details in SC talk @ PDF4LHC (April 15'), see wiki

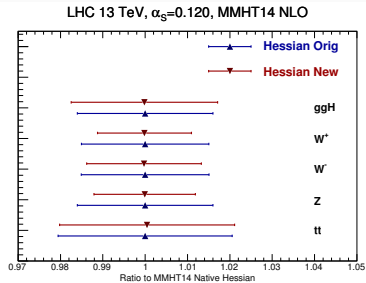
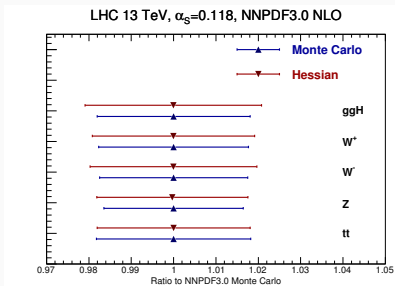


Examples:

- **NNPDF3.0:** native MC & **MMHT2014 MC:** native Hessian
 - optimal number of **eigenvectors** similar to the **compression algorithm**
 - $N_{\text{eig}} = 120$ for **NNPDF3.0** and $N_{\text{eig}} = 12$ for **MMHT2014 MC**

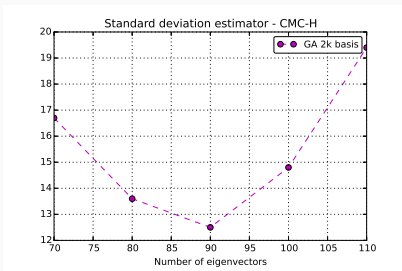


- **Good agreement** between representations at the level of:
 - PDFs, luminosities and **observables**.
- **Example:** inclusive cross-sections



CMC-H:

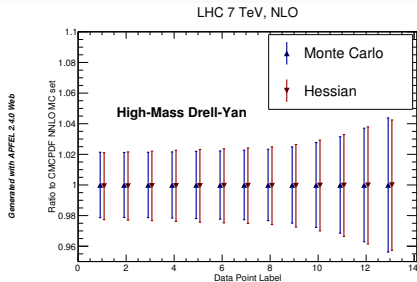
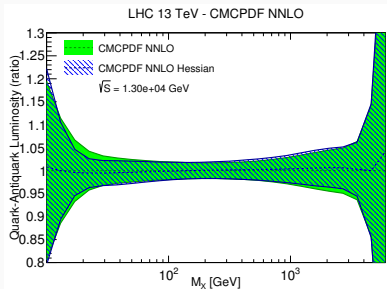
- apply **mc2hessian conversion** to the “combined MC set”
- **optimal number of eigenvectors** \Rightarrow **compression**



- new **Hessian CMC representation (CMC-H)**
- $N_{\text{eig}} = 90$ symmetric eigenvectors
- no **parametrization bias**



- Good agreement between representations, examples:



- Full details in [\[arXiv:1505.06736\]](https://arxiv.org/abs/1505.06736)
- Public tool at:
<https://github.com/scarrazza/mc2hessian>



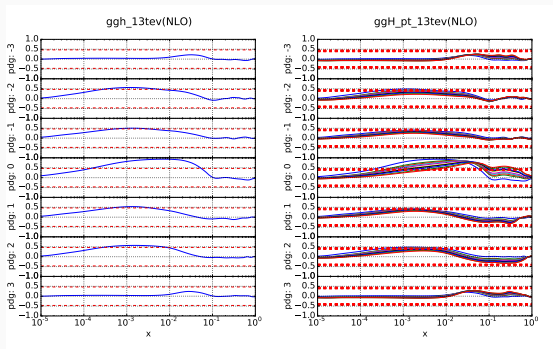
SPECIALIZED MINIMAL PDF

Specialized Minimal PDFs:

- **Idea:** perform a mc2hessian tool inspired on **Principal Component Analysis** (see Appendix of [\[arXiv:1505.06736\]](#)):
 - 1. compute the **covariance matrix** in a grid of x and flavors
 - 2. represent this matrix on a basis of replicas through covariance Single Value Decomposition (DVS)
- **Optimized for specific observable(s):**
 - 3. compute **correlations** between the basis of replicas and **observables**
 - 4. **remove** all points where **correlations** are smaller than a **threshold**
 - 5. sort **eigenvectors** by **eigenvalues** size
- **Output:**
 - **specialized** set of PDFs based on the **kinematics** of the processes with a **small number of eigenvectors**



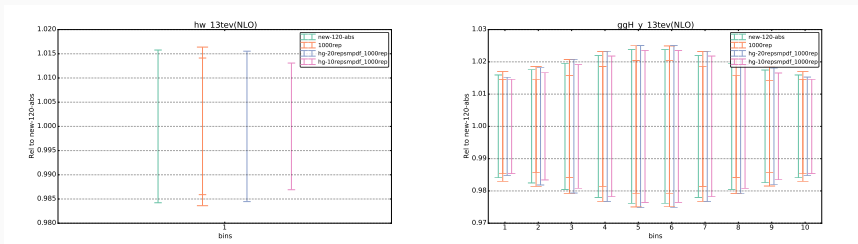
- **Example:** SM-PDF for Higgs, where we include:
 - $ggH, Ht\bar{t}, HZ, HW$ + differential distributions: $H p_T$ and y .



- Select the points in x , **larger** than a **threshold** (50% of maximum bin correlation), \Rightarrow e.g. from $N_x N_f = 700$ we keep $N_x N_f \sim 200$ points



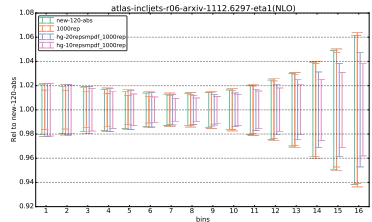
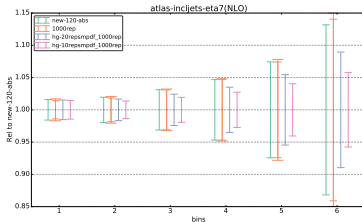
- Preliminary results for the **NNPDF3.0 NLO**:
 - differences between the MC and Hessian representations due to non-gaussian behavior ($68\% \text{ cl} \neq 1-\sigma$)



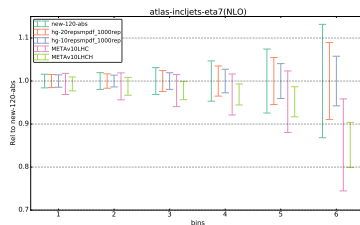
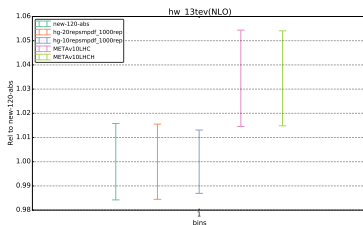
- For **Higgs processes** we need **~20 eigenvectors** to reproduce the best Hessian representation of **NNPDF3.0** (120 eigenvectors).



- The SM-PDFs are based on the **kinematics** of the process so we expect good predictions also for observables/processes not included in the procedure but with **similar kinematics**.
- **Example:** the Higgs SM-PDF is also able to reproduce some of the inclusive jet distributions



- SM-PDF and MetaPDF LHCH behave **similarly** for the input data...



- ... but for **jet predictions** MetaPDF LHCH deteriorates faster than SM-PDF.



SM-PDF delivery:

Public automatic tool for specialized grids for any process or combination of processes determined by the **user**:

- interface to **APPLgrid** (for LO and NLO)
- interface to **parton level MC** (useful for NNLO predictions)
- option to produce **comparison** of **observables** with multiple **PDFs**

<https://github.com/scarrazza/smpdf>



DISCUSSION

