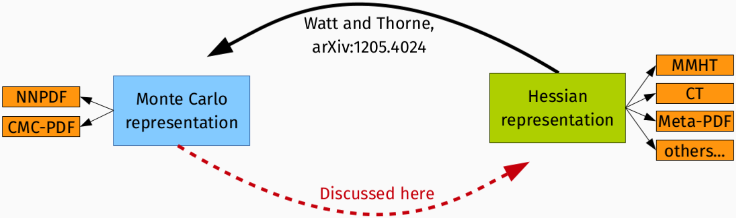
mc²hessian

An unbiased Hessian representation for Monte Carlo PDFs

Stefano Carrazza
in collaboration with S. Forte, Z. Kassabov, J.I. Latorre and J. Rojo

PDF4LHC, April 13, 2015, CERN

Panorama of PDF representations:



Problem addressed here:

⇒ Determine an **unbiased Hessian representation** for **MC** PDFs.

- Some **advantages** of each approach:

Monte Carlo approach:

- unbiased PDF parametrization
- no linear approximation to propagate uncertainties

Hessian approach:

- PDF probability density is Gaussian
- the use of PDF profiling

- Some **advantages** of each approach:

Monte Carlo approach:

- unbiased PDF parametrization
- no linear approximation to propagate uncertainties

Hessian approach:

- PDF probability density is Gaussian
- the use of PDF profiling

MC to Hessian conversion problem:

Avoid the definition of an **intermediate** functional form which introduces **bias** in the **PDF parametrization**

- Some **advantages** of each approach:

Monte Carlo approach:

- unbiased PDF parametrization
- no linear approximation to propagate uncertainties

Hessian approach:

- PDF probability density is Gaussian
- the use of PDF profiling

MC to Hessian conversion problem:

Avoid the definition of an **intermediate** functional form which introduces **bias** in the **PDF parametrization**

Our Strategy:

use MC replicas ***themselves*** *as the* ***basis*** *of the linear representation*

- **Methodology**
  - Definition and implementation
- **Validation and Benchmarking**
  - Hessian representation of MC PDFs
  - Self-closure of the algorithm
  - LHC phenomenology
- **Delivery**

THE `mc2hessian` METHODOLOGY

Given a Monte Carlo prior set of PDFs

$$\{f_\alpha^{(k)}\}_{k=1,\ldots,N_{\rm rep}}\,, \quad \alpha = \{g, u, d, s, \ldots\}\,,$$

use a subset of replicas as parameters of linear expansion:

$$f_\alpha^{(k)} \approx f_{H,\alpha}^{(k)} \equiv f_\alpha^{(0)} + \sum_{i=1}^{N_{\rm eig}} a_i^{(k)} (\eta_\alpha^{(i)} - f_\alpha^{(0)})\,, \quad k = 1, \ldots, N_{\rm rep}$$

Given a Monte Carlo prior set of PDFs

$$\{f_\alpha^{(k)}\}_{k=1,\ldots,N_{\mathrm{rep}}}\,, \quad \alpha = \{g, u, d, s, \ldots\}\,,$$

use a subset of replicas as parameters of linear expansion:

$$f_\alpha^{(k)} \approx f_{H,\alpha}^{(k)} \equiv f_\alpha^{(0)} + \sum_{i=1}^{N_{\mathrm{eig}}} a_i^{(k)} (\eta_\alpha^{(i)} - f_\alpha^{(0)})\,, \quad k = 1, \ldots, N_{\mathrm{rep}}$$

- approximate each replica of the original MC ensemble $f_\alpha^{(k)}$

Given a Monte Carlo prior set of PDFs

$$\{f_\alpha^{(k)}\}_{k=1,\dots,N_{\rm rep}}\,,\quad \alpha=\{g,u,d,s,\dots\}\,,$$

use a subset of replicas as parameters of linear expansion:

$$f_\alpha^{(k)} \approx f_{H,\alpha}^{(k)} \equiv f_\alpha^{(0)} + \sum_{i=1}^{N_{\rm eig}} a_i^{(k)}(\eta_\alpha^{(i)} - f_\alpha^{(0)})\,,\quad k=1,\dots,N_{\rm rep}$$

- approximate each replica of the original MC ensemble $f_\alpha^{(k)}$
- by the linear combination $f_{H,\alpha}^{(k)}$

Given a Monte Carlo prior set of PDFs

$$\{f_\alpha^{(k)}\}_{k=1,\dots,N_{\rm rep}}\,, \quad \alpha=\{g,u,d,s,\dots\}\,,$$

use a subset of replicas as parameters of linear expansion:

$$f_\alpha^{(k)} \approx f_{H,\alpha}^{(k)} \equiv f_\alpha^{(0)} + \sum_{i=1}^{N_{\rm eig}} a_i^{(k)} (\eta_\alpha^{(i)} - f_\alpha^{(0)})\,, \quad k=1,\dots,N_{\rm rep}$$

- approximate each replica of the original MC ensemble $f_\alpha^{(k)}$
- by the linear combination $f_{H,\alpha}^{(k)}$
- with coefficients $a_i^{(k)}$

Given a Monte Carlo prior set of PDFs

$$\{f_\alpha^{(k)}\}_{k=1,\ldots,N_{\rm rep}}\,, \quad \alpha = \{g, u, d, s, \ldots\}\,,$$

use a subset of replicas as parameters of linear expansion:

$$f_\alpha^{(k)} \approx f_{H,\alpha}^{(k)} \equiv f_\alpha^{(0)} + \sum_{i=1}^{N_{\rm eig}} a_i^{(k)} (\eta_\alpha^{(i)} - f_\alpha^{(0)})\,, \quad k = 1, \ldots, N_{\rm rep}$$

- approximate each replica of the original MC ensemble $f_\alpha^{(k)}$
- by the linear combination $f_{H,\alpha}^{(k)}$
- with coefficients $a_i^{(k)}$
- of deviations from the central value $f_\alpha^{(0)}$

Given a Monte Carlo prior set of PDFs

$$\{f_\alpha^{(k)}\}_{k=1,\dots,N_{\mathrm{rep}}}\,, \quad \alpha = \{g, u, d, s, \dots\}\,,$$

use a subset of replicas as parameters of linear expansion:

$$f_\alpha^{(k)} \approx f_{H,\alpha}^{(k)} \equiv f_\alpha^{(0)} + \sum_{i=1}^{N_{\mathrm{eig}}} a_i^{(k)} (\eta_\alpha^{(i)} - f_\alpha^{(0)})\,, \quad k = 1, \dots, N_{\mathrm{rep}}$$

- approximate each replica of the original MC ensemble $f_\alpha^{(k)}$
- by the linear combination $f_{H,\alpha}^{(k)}$
- with coefficients $a_i^{(k)}$
- of deviations from the central value $f_\alpha^{(0)}$
- expanded in the basis of a subset of replicas $\{\eta_\alpha^{(i)}\}_{i=1,\dots,N_{\mathrm{eig}}} \subset \{f_\alpha^{(k)}\}$

Given a Monte Carlo prior set of PDFs

$$\{f_\alpha^{(k)}\}_{k=1,\ldots,N_{\rm rep}}\,, \quad \alpha=\{g,u,d,s,\ldots\}\,,$$

use a subset of replicas as parameters of linear expansion:

$$f_\alpha^{(k)} \approx f_{H,\alpha}^{(k)} \equiv f_\alpha^{(0)} + \sum_{i=1}^{N_{\rm eig}} a_i^{(k)}(\eta_\alpha^{(i)} - f_\alpha^{(0)})\,, \quad k=1,\ldots,N_{\rm rep}$$

- approximate each replica of the original MC ensemble $f_\alpha^{(k)}$
- by the linear combination $f_{H,\alpha}^{(k)}$
- with coefficients $a_i^{(k)}$
- of deviations from the central value $f_\alpha^{(0)}$
- expanded in the basis of a subset of replicas $\{\eta_\alpha^{(i)}\}_{i=1,\ldots,N_{\rm eig}} \subset \{f_\alpha^{(k)}\}$

The methodology in 4 steps:

(a) We define the **covariance matrix** for the prior set

$$\mathrm{cov}^{\mathrm{pdf}}_{ij,\alpha\beta} \equiv \frac{N_{\mathrm{rep}}}{N_{\mathrm{rep}}-1}\left(\left\langle f_\alpha^{(k)}(x_i,Q_0^2)\cdot f_\beta^{(k)}(x_j,Q_0^2)\right\rangle_{\mathrm{rep}} - \left\langle f_\alpha^{(k)}(x_i,Q_0^2)\right\rangle_{\mathrm{rep}}\left\langle f_\beta^{(k)}(x_j,Q_0^2)\right\rangle_{\mathrm{rep}}\right)$$

The methodology in 4 steps:

(a) We define the **covariance matrix** for the prior set

$$\mathrm{cov}^{\mathrm{pdf}}_{ij,\alpha\beta} \equiv \frac{N_{\mathrm{rep}}}{N_{\mathrm{rep}}-1}\left(\left\langle f^{(k)}_{\alpha}(x_i,Q_0^2)\cdot f^{(k)}_{\beta}(x_j,Q_0^2)\right\rangle_{\mathrm{rep}} - \left\langle f^{(k)}_{\alpha}(x_i,Q_0^2)\right\rangle_{\mathrm{rep}}\left\langle f^{(k)}_{\beta}(x_j,Q_0^2)\right\rangle_{\mathrm{rep}}\right)$$

(b) Then we minimize the **figure of merit**

$$\chi^{2(k)}_{\mathrm{pdf}} \equiv \sum_{i,j=1}^{N_x}\sum_{\alpha,\beta=1}^{N_f}\left(\left[f^{(k)}_{H,\alpha}(x_i,Q_0^2) - f^{(k)}_{\alpha}(x_i,Q_0^2)\right]\cdot\left(\mathrm{cov}^{\mathrm{pdf}}\right)^{-1}_{ij,\alpha\beta}\cdot\left[f^{(k)}_{H,\beta}(x_j,Q_0^2) - f^{(k)}_{\beta}(x_j,Q_0^2)\right]\right)$$

- in a suitable sampling in x and flavors $\rightarrow (N_x, N_f)$
- at fixed Q_0^2 $\rightarrow$ higher values by DGLAP evolution

DESCRIPTION OF THE METHOD

The methodology in 4 steps:

(a) We define the **covariance matrix** for the prior set

$$\mathrm{cov}^{\mathrm{pdf}}_{ij,\alpha\beta} \equiv \frac{N_{\mathrm{rep}}}{N_{\mathrm{rep}}-1}\left(\left\langle f^{(k)}_{\alpha}(x_i,Q_0^2)\cdot f^{(k)}_{\beta}(x_j,Q_0^2)\right\rangle_{\mathrm{rep}} - \left\langle f^{(k)}_{\alpha}(x_i,Q_0^2)\right\rangle_{\mathrm{rep}}\left\langle f^{(k)}_{\beta}(x_j,Q_0^2)\right\rangle_{\mathrm{rep}}\right)$$

(b) Then we minimize the **figure of merit**

$$\chi^{2(k)}_{\mathrm{pdf}} \equiv \sum_{i,j=1}^{N_x}\sum_{\alpha,\beta=1}^{N_f}\left(\left[f^{(k)}_{H,\alpha}(x_i,Q_0^2) - f^{(k)}_{\alpha}(x_i,Q_0^2)\right]\cdot\left(\mathrm{cov}^{\mathrm{pdf}}\right)^{-1}_{ij,\alpha\beta}\cdot\left[f^{(k)}_{H,\beta}(x_j,Q_0^2) - f^{(k)}_{\beta}(x_j,Q_0^2)\right]\right)$$

- in a suitable sampling in x and flavors $\rightarrow (N_x, N_f)$
- at fixed Q_0^2 $\rightarrow$ higher values by DGLAP evolution

Minimization strategy

- impose $\chi^{2(k)}_{\mathrm{pdf}} \rightarrow 0$ for each replica of the prior (e.g. SVD)
- determination of the coefficients $\{a_i^{(k)}\}$ for each original replica k

(c) We construct the covariance matrix of $\{a_i^{(k)}\}$ coefficients

$$\mathrm{cov}_{ij}^{(\mathrm{a})} \equiv \frac{N_{\mathrm{rep}}}{N_{\mathrm{rep}}-1}\left(\left\langle a_i \cdot a_j \right\rangle_{\mathrm{rep}} - \left\langle a_i \right\rangle_{\mathrm{rep}} \left\langle a_j \right\rangle_{\mathrm{rep}}\right), \quad i,j = 1,\ldots,N_{\mathrm{eig}}$$

(d) We diagonalize the inverse of $\mathrm{cov}_{ij}^{(\mathrm{a})}$, the one-sigma uncertainty is

$$\sigma_{H,\alpha}^{\mathrm{PDF}}(x,Q^2) = \sqrt{\sum_{i=1}^{N_{\mathrm{eig}}}\left[\sum_{j=1}^{N_{\mathrm{eig}}} \frac{v_{ij}}{\sqrt{\lambda_i}}\left(\eta_{\alpha}^{(j)}(x,Q^2) - f_{\alpha}^{(0)}(x,Q^2)\right)\right]^2}$$

where v_{ij} is rotation matrix, and λ_i the set of eigenvalues.

The final symmetric Hessian eigenvectors

$$\widetilde{f}_{\alpha}^{(i)}(x, Q^2) = f_{\alpha}^{(0)}(x, Q^2) + \sum_{j=1}^{N_{\mathrm{eig}}} \frac{v_{ij}}{\sqrt{\lambda_i}} \left(\eta_{\alpha}^{(j)}(x, Q^2) - f_{\alpha}^{(0)}(x, Q^2) \right)$$

The one-sigma uncertainty band is then

$$\sigma_{H,\alpha}^{\mathrm{PDF}}(x, Q^2) = \sqrt{\sum_{i=1}^{N_{eig}} \left(\widetilde{f}_{\alpha}^{(i)}(x, Q^2) - f_{\alpha}^{(0)}(x, Q^2) \right)^2}$$

The final symmetric Hessian eigenvectors

$$\widetilde{f}_{\alpha}^{(i)}(x, Q^2) = f_{\alpha}^{(0)}(x, Q^2) + \sum_{j=1}^{N_{\rm eig}} \frac{v_{ij}}{\sqrt{\lambda_i}} \left(\eta_{\alpha}^{(j)}(x, Q^2) - f_{\alpha}^{(0)}(x, Q^2) \right)$$

The one-sigma uncertainty band is then

$$\sigma_{H,\alpha}^{\rm PDF}(x, Q^2) = \sqrt{\sum_{i=1}^{N_{eig}} \left(\widetilde{f}_{\alpha}^{(i)}(x, Q^2) - f_{\alpha}^{(0)}(x, Q^2) \right)^2}$$

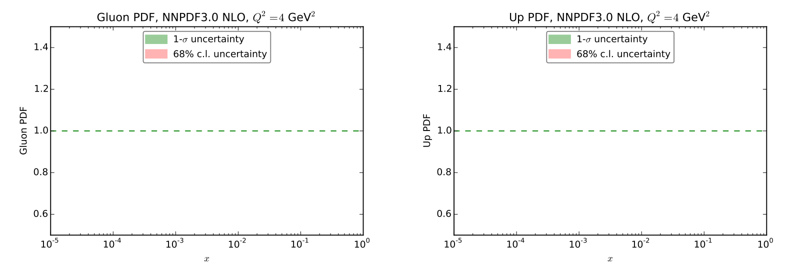
To be compared to the standard deviation of the prior MC set:

$$\sigma_{\alpha}^{\rm PDF}(x, Q^2) = \sqrt{\left\langle \left(f_{\alpha}^{(k)}(x, Q^2) \right)^2 \right\rangle_{\rm rep} - \left\langle f_{\alpha}^{(k)}(x, Q^2) \right\rangle_{\rm rep}^2}$$

THE `mc2hessian` NUMERICS

Practical implementation issues

1. the **grid of points in** x: 1-σ and the 68% confidence level intervals

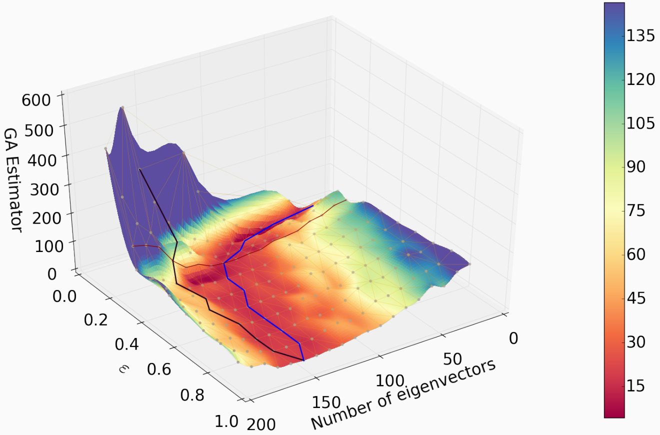


2. the **optimal basis** of replicas, the **optimal number** of symmetric eigenvectors for the Hessian representation

We define an **estimator** which measures the distance between the prior MC and its Hessian representation:

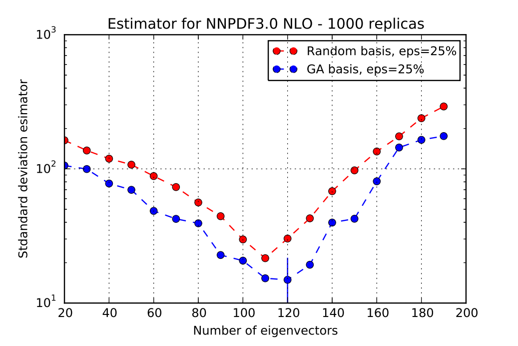
$$\mathrm{ERF}_\sigma = \sum_{i=1}^{N_x} \sum_{\alpha=1}^{N_f} \left| \frac{\sigma_{H,\alpha}^{\mathrm{PDF}}(x_i, Q_0^2) - \sigma_{\alpha}^{\mathrm{PDF}}(x_i, Q_0^2)}{\sigma_{\alpha}^{\mathrm{PDF}}(x_i, Q_0^2)} \right|$$

- We **introduce** an ϵ threshold for the **exclusion** of regions in x where the Gaussian approximation is no reliable.
- We **implement** a Genetic Algorithm in function of ϵ and N_{eig} which **minimizes** the **estimator**.



- **Surface:** GA minimum for estimator in function of ϵ and N_{eig}.
- **Blue curve:** surface minimum; **black curve:** estimator with large ϵ.

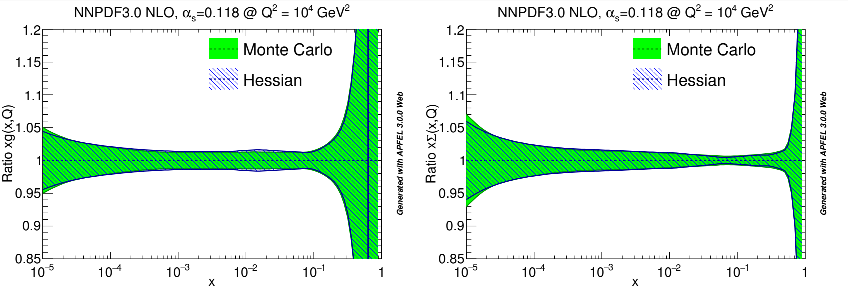
Estimator: Random vs. GA basis



- **NNPDF3.0 NLO** $\longrightarrow$ Hessian representation with $N_{\rm eig} = 120$
- We use $\epsilon = 25\%$ motivated by the previous slide.

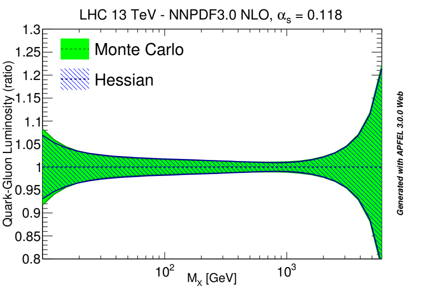
VALIDATION AND BENCHMARKING

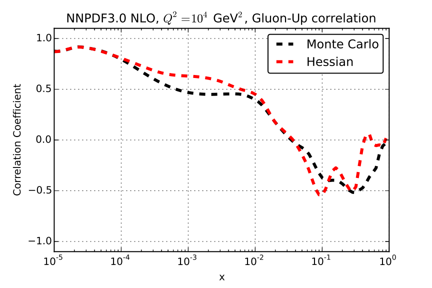
PDF comparison: MC vs. Hessian representations



- **Good agreement:** differences in the one-sigma PDF uncertainty bands of the order 5% at most between the two representations.

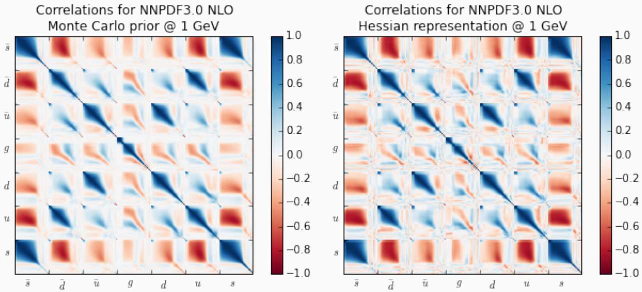
Luminosities and Correlations:





- **Reasonable agreement:** small differences due to the information loss when moving from the MC to the Hessian representation.

Full correlations matrix:



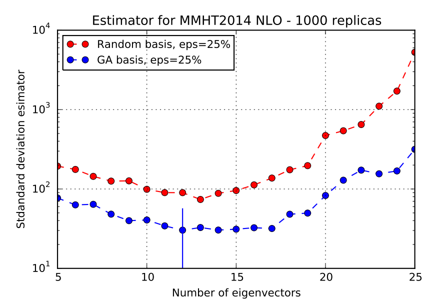
- **Good agreement:** small differences due to the information loss when moving from the MC to the Hessian representation.

SELF-CLOSURE TEST

Starting from the original Hessian MMHT14 NLO set:

1. Construct its **MC representation** (Watt & Thorne '12)
2. Run the `mc2hessian` algorithm

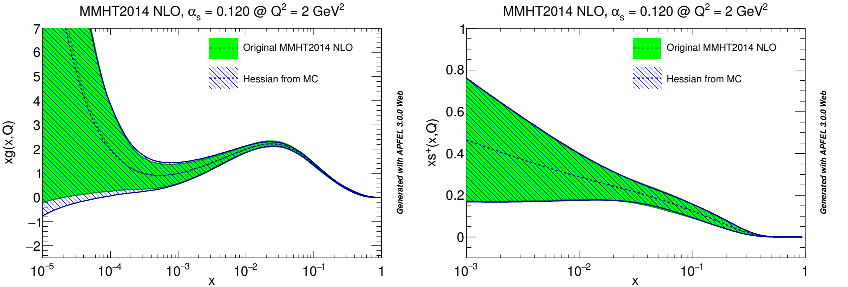
In this case, the estimator minimum is obtained with **12 symmetric eigenvectors**.



Self-closure Test Output:

`mc2hessian` successful ⇒ **original** and the **new** Hessian representations are **close to each other**

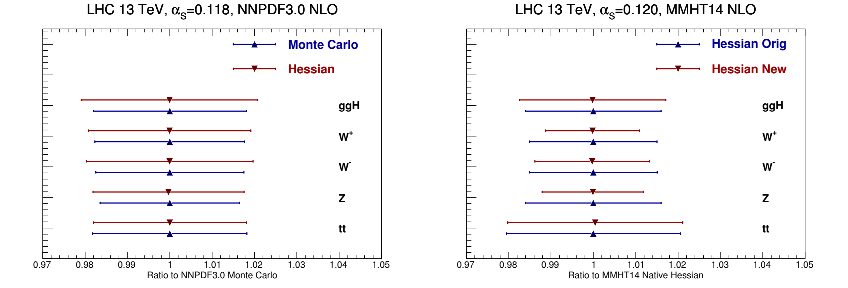
PDF comparison: Original MMHT2014 vs. Hessian from MC



- Agreements **better than 5%** of the uncertainty between the two Hessian representations.
- The `mc2hessian` is able to **compress information** of the native Hessian representations, **reducing the total number of eigenvectors**.

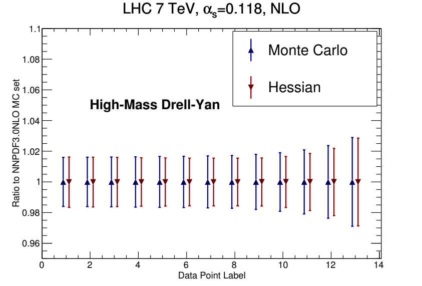
PHENOMENOLOGY

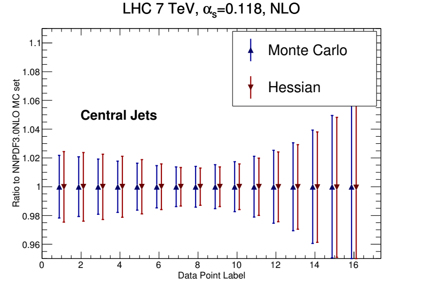
LHC inclusive cross-sections @ 13 TeV



- **Good agreement** for LHC inclusive cross-sections, discrepancies below 10%.

LHC differential distributions @ 7 TeV for NNPDF3.0 NLO





- **Very good agreement** for a large number of differential distributions at the LHC 7 TeV, differences always below 10%.

DELIVERY

- The `mc2hessian` program is public available at

  `github.com/scarrazza/mc2hessian`

- Further **optimizations** in progress before final release.
- NNPDF3.0 Hessian version available in `LHAPDF6` soon:
  - `NNPDF30_nlo_as_0118_hessian`
  - `NNPDF30_nnlo_as_0118_hessian`
- Any other MC set can be converted using directly the public code.

Questions?