An NNPDF4.0 determination of α_s – status & ideas

Roy Stegeman
The University of Edinburgh

NNPDF collaboration meeting

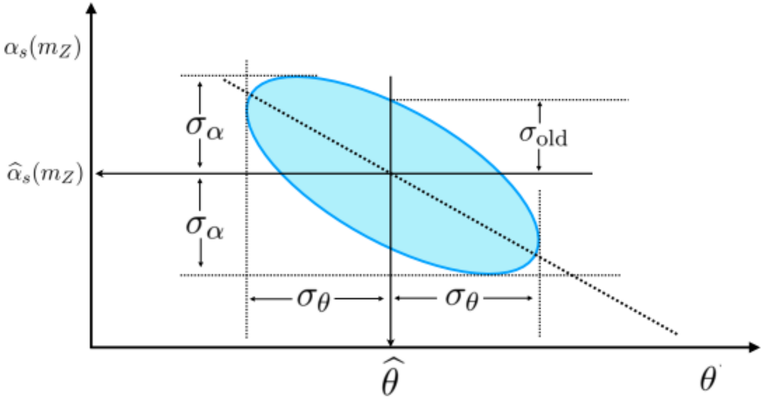
Gargnano, 26 September 2023

Three different methodologies

- Correlated replica method
- Theory covariance method based on [Ball & Pearson (2021); 2105.05114]
- SIMUnet [Iranipour & Ubiali (2022); 2201.07240]

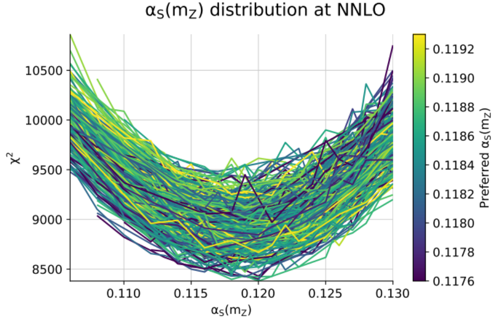
Correlated replica method – 1802.03398

Determining α_s with a fixed input PDF results in underestimated uncertainty

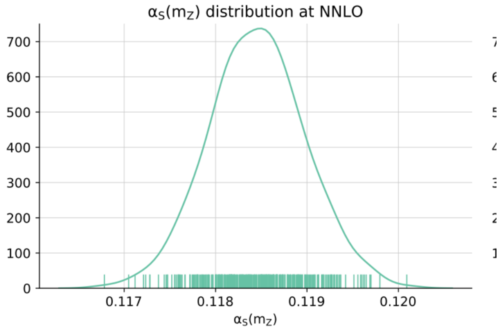


Correlated replica method – 1802.03398

1. Produce fits to a data replica, D^k at different values of α_s, such that we can determine $\alpha_s^{(k)} = \mathrm{argmin}\left[\chi^{2(k)}(\alpha_s)\right]$
2. Perform quadratic fit to χ^2 profiles



3. Analyze the resulting probability distribution



Correlated replica method – 1802.03398

NNPDF3.1 (arXiv:1802.03398): 0.1185 ± 0.0005

Data set	Methodological details	Corr. reps. meth. (updated), mean and std
NNPDF4.0	NNPDF4.0 repeated e.g. 230131-rs-nnpdf40-corr-01140-a	0.12149 ± 0.00055
NNPDF4.0	NNPDF4.0	0.12154 ± 0.00053
NNPDF3.1-like	NNPDF4.0	0.12012 ± 0.00076
NNPDF3.1-like	NNPDF4.0 (w/o nuc. uncs)	0.11940 ± 0.00055
NNPDF3.1-like	NNPDF4.0 (no NNPDF4.0 int./pos. constraints)	0.11967 ± 0.00228 (outliers!)
NNPDF3.1-like (no nuc. data)	NNPDF4.0	0.12080 ± 0.00080
NNPDF3.1-like	NNPDF4.0 (w/o nuc. uncs. for deut. data sets)	0.12009 ± 0.00086
NNPDF3.1-like	NNPDF4.0 (all nuc. data dw but CHORUS)	0.11959 ± 0.00070
NNPDF4.0 "conservative"	NNPDF4.0	0.12155 ± 0.00072

- Large value of α_s for NNPDF4.0 compared to world average
- $\sim 2\sigma$ disagreement between NNPDF3.1-like and NNPDF4.0 datasets
- What is the impact of MHOU?

Theory covariance method – 2105.05114

$$P(T \mid D) \propto \exp\left(-\frac{1}{2}(T-D)^T(C+S)^{-1}(T-D)\right),\ S = \beta\beta^T$$ [Ball, Nocera, Pearson (2019); 1812.09074]

Introduce a nuisance parameter λ, to model theory uncertainty as correlated shift $(T \to T + \lambda\beta)$ in the theory prediction:

$$P(T \mid D\lambda) \propto \exp\left(-\frac{1}{2}(T+\lambda\beta-D)^T C^{-1}(T+\lambda\beta-D)\right)$$

We can marginalize over λ to recover $P(T \mid D)$:

$$P(T \mid D) = \int d\lambda P(T \mid D\lambda) P(\lambda)$$

Which is a Gaussian integral for the choice $P(\lambda) = \exp\left(-\frac{1}{2}\lambda^2\right)$, reproducing the result on top of the slide:

$$P(T \mid D) \propto \int d\lambda \exp\left(-\frac{1}{2}Z^{-1}(\lambda-\bar{\lambda})^2 - \frac{1}{2}\chi^2\right) \propto \exp\left(-\frac{1}{2}\chi^2\right)$$

With $Z = \left(1+\beta^T C^{-1}\beta\right)^{-1} = 1 - \beta^T(C+S)^{-1}\beta$ independent of T and D,
and $\bar{\lambda}(T,D) = Z\beta^T C^{-1}(D-T) = \beta^T(C+S)^{-1}(D-T)$.

Theory covariance method – 2105.05114

Previously we recovered the known result

$$P(T \mid D) \propto \int d\lambda \exp\left(-\frac{1}{2}Z^{-1}(\lambda-\bar{\lambda})^2 - \frac{1}{2}\chi^2\right) \propto \exp\left(-\frac{1}{2}\chi^2\right).$$

However, we can also calculate the posterior of the nuisance parameter

$$P(\lambda \mid TD) = \frac{P(T \mid D\lambda)P(\lambda)}{P(T \mid D)} \propto \exp\left(-\frac{1}{2}Z^{-1}(\lambda-\bar{\lambda}(T,D))^2\right)$$

where the expected value is given by $\bar{\lambda}$ and the uncertainty by Z.

Same idea can, with some work, be extended to the scenario of a PDF fit. This has been done by Richard and Rosalyn in their paper.

Theory covariance method (TCM) requires only a single fit!

Compare CRM and TCM

`https://www.wiki.ed.ac.uk/display/nnpdfwiki/Correlated+replica+method`

NNPDF3.1 (arXiv:1802.03398): 0.1185 ± 0.0005

Data set	Methodological details	Corr. reps. meth. (updated), mean and std	[0.116; 0.118; 0.120]	[0.114; 0.118; 0.122]	[0.114; 0.118; 0.122] (ite)	[0.116; 0.119; 0.122] (ite)
NNPDF4.0	NNPDF4.0 repeated e.g. 230131-rs-nnpdf40-corr-01140-a	0.12149 ± 0.00055				
NNPDF4.0	NNPDF4.0	0.12154 ± 0.00053	0.11990 ± 0.00038	0.12025 ± 0.00037	0.12032 ± 0.00039	0.12005 ± 0.00036
NNPDF3.1-like	NNPDF4.0	0.12012 ± 0.00076	0.11837 ± 0.00056	0.11842 ± 0.00060	0.11831 ± 0.00062	0.11835 ± 0.00058
NNPDF3.1-like	NNPDF4.0 (w/o nuc. uncs)	0.11940 ± 0.00055	0.11775 ± 0.00059	0.11786 ± 0.00064	0.11779 ± 0.00061	0.11769 ± 0.00058
NNPDF3.1-like	NNPDF4.0 (no NNPDF4.0 int./pos. constraints)	0.11967 ± 0.00228 (outliers!)	0.11775 ± 0.00059	0.11787 ± 0.00067	0.11777 ± 0.00064	0.11768 ± 0.00058
NNPDF3.1-like (no nuc. data)	NNPDF4.0	0.12080 ± 0.00080	0.11922 ± 0.00058	0.11928 ± 0.00067	0.11908 ± 0.00068	0.11908 ± 0.00066
NNPDF3.1-like	NNPDF4.0 (w/o nuc. uncs. for deut. data sets)	0.12009 ± 0.00086	0.11839 ± 0.00062	0.11843 ± 0.00060	0.11837 ± 0.00058	0.11839 ± 0.00060
NNPDF3.1-like	NNPDF4.0 (all nuc. data dw but CHORUS)	0.11959 ± 0.00070	0.11781 ± 0.00062	0.11795 ± 0.00063	0.11788 ± 0.00062	0.11778 ± 0.00057
NNPDF4.0 "conservative"	NNPDF4.0	0.12155 ± 0.00072	0.12045 ± 0.00057	0.12059 ± 0.00055	0.12049 ± 0.00053	0.12043 ± 0.00049

- Consistent shift of $\sim$0.0017 between CRM and TCM
- TCM stable upon changes to prior

Compare CRM and TCM

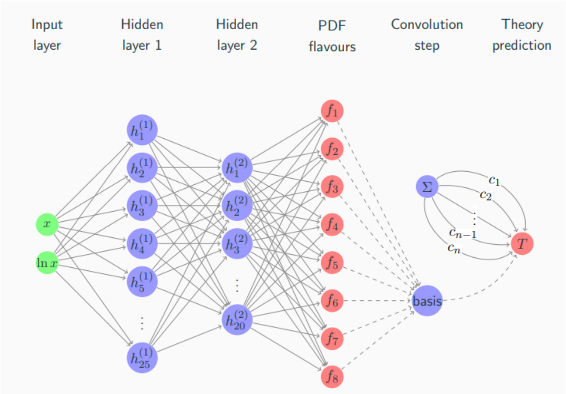
CRM and TCM disagree, but there exist differences between the methodologies. We have attempted to understand these differences and obtain the same (though not necessarily the correct) result by trying the following:

- t0 PDF varies between α_s in CRM $\rightarrow$ fix t0 PDF at all values of α_s
- Non-Gaussianity of CRM $\rightarrow$ minimal effect in most cases but keep in mind
- Theory covariance deweights data with large theory uncertainty in TCM $\rightarrow$ include theory covmat in CRM
- TCM assumes fitting and sampling covmat are the same $\rightarrow$ use the t0 covmat for both sampling and fitting in CRM and TCM

However, this results in shifts that differ per dataset and the pattern of the consistent shifts is lost

To be tested: what is the impact of expanding chi2 beyond quadratic in the nuisance parameter?

SIMUnet – the basic idea



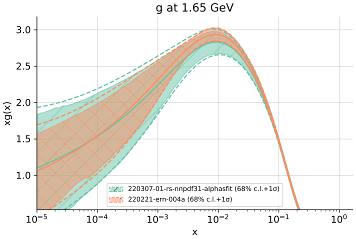
Can we use this to determine α_s by interpolating between FKTables?
Note that this involves loading many different theories and is thus very memory-expensive

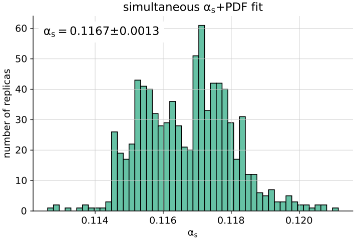
SIMUnet

NNPDF3.1-like dataset w/o nucl. uncertainties

High training-validation losses
$\rightarrow$ problem with optimization (possible to solve by careful tuning and initial freezing)

Two peaks in the α_s distribution
$\rightarrow$ peak at lower α_s corresponds to worse χ^2

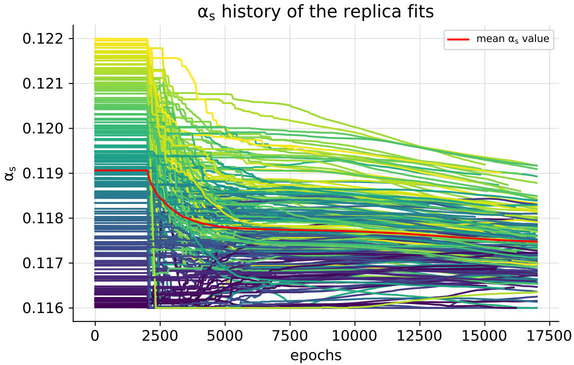




Idea: freeze α_s during the initial training to improve stability and then small learning rate

SIMUnet

Problem: α_s doesn't stabilize and depends on initial value



Way forward?

- Beyond quadratic expansion for TCM
- MHOU
- How to improve the training in SIMUnet?
- ...?