Positivity of physical cross-sections in the NNPDF fits

The NNPDF Collaboration

Abstract

Note about the positivity of physical cross-sections in the NNPDF fits.

1 Introduction

It is a basic physical requirement that cross-section must be positive-definite, as they are probabilities. In the framework of perturbative QCD factorization in hadronic collisions, cross-sections and parton distributions are related by the following relation:

$$\sigma(p+p \to X) = \sum_{i,j=-5}^{5} \hat{\sigma}\left(q_i + q_j \to X\right) \otimes q_i(x, M_X) \otimes q_j(x, M_X)\,, \tag{1.1}$$

where the sum runs over all active partons in the proton (with $i = 0$ corresponding to the gluon), $\hat{\sigma}$ is the partonic cross-section, and M_X is the invariant mass of the produced final state. The parton distribution function of the proton $\{q_i(x, Q)\}$ cannot be computed from first principles and therefore need to be extracted from experimental data.

At leading order (LO), for standard processes (such as deep-inelastic scattering or Drell-Yan production) the partonic cross-section is just a number (generally diemensionful), so the hadronic cross-section is proportional to the PDF, which must therefore be a probability and therefore non-negative:

$$q_i(x, Q) \geq 0\,, \forall x\,, \forall i\,. \tag{1.2}$$

Beyond LO, the partonic cross-section is not even a function, but rather a distribution, and the factorization Eq. (1.1) depends on the factorization scheme, so while the hadronic cross-section must remain non-negative, the partonic cross-section and PDF can separately become negative.

In PDF sets based on fixed functional forms, such as CT or MMHT, the PDFs at large-x are often restricted to be non-negative by a particular choice of functional form, such as

$$q_i(x, Q^2) \propto (1-x)^{b_1}, \tag{1.3}$$

but this may introduce theoretical bias, especially in regions where there is limited experimental information, such as the large x region close to kinematic boundaries.

2 Positivity in NNPDF

In the NNPDF PDF determination no ad-hoc requirements on the shape of the PDFs are imposed, and PDFs are parametrized in a fully unbiased way by means of neural networks. Whereas in leading-order PDF sets (such as NNPDF3.0LO) non-negativity of each single replica is imposed and verified, in NLO and NNLO sets in order to avoid bias, no PDF positivity is imposed. Rather, positivity is required for a set of selected cross-sections by means of Lagrange multipliers (i.e., regardless of the availability of data for these processes). These cross-sections correspond either to physical processes, or to processes which do not exist in the standard model but which could be present either in an extension of the standard model or more in general assuming new interactions. They are selected in order to be effective in enforcing on PDFs the constraints of positivity of physical observables. Hence NNPDF PDF replicas are allowed to become negative, but provided they lead to non-negative cross-section. Henceforth, we well refer as "positivity of PDF" to the constraint that positivity of physical observables impose on that PDF: so for instance positivity of the gluon effectively means positivity of gluon-dominated processes.

Specifically, we impose positivity of

- Deep-inelastic structure functions:
  - F_2^u;
  - F_2^d;
  - F_2^s;
  - F_L^l, the light component of the longitudinal structure function, defined as the contribution to the structure function F_L when all quark electric charges are set to zero but those of the three lightest flavors.
- Flavor-tagged Drell-Yan rapidity distributions:
  - $d\sigma_{u\bar{u}}^{\mathrm{DY}}/dy$;
  - $d\sigma_{d\bar{d}}^{\mathrm{DY}}/dy$;
  - $d\sigma_{s\bar{s}}^{\mathrm{DY}}/dy$.
- Scalar "Higgs-like" production:
  - $d\sigma_{gg}^{H}/dy$, the cross-section for the production in gluon-gluon fusion of a Higgs-like scalar with mass $m_H^2 = 5$ GeV^2.

The combination of structure functions and Drell-Yan-like cross-sections are designed to impose positivity of individual quark flavors and antiflavors: hence the choice of having one constraint per flavor, in such a way that there are six constraints for six light quark and antiquark PDFs. The longitudinal structure function and scalar cross-sections are designed to constrain the gluon, both at small and large x. These positivity conditions are imposed at the low scale of $Q^2_{\mathrm{pos}} = 5$ GeV^2, and are then automatically propagated to higher scales when PDFs are evolved upwards in Q^2.

Despite the presence of these positivity constraints, violations of positivity may still occur for individual replicas. On the one hand, this could be simply due to the fact that in principle all physical observables must be non-negative, but there is no known way of imposing the most general positivity constraint on PDFs. Hence, it is always possible that some other physical observable might impose a more stringent constraint than that imposed by the above constraints.

More to the point, especially for final states with large invariant mass the cross-section is very small and often compatible with zero. The cross-section can then turn out to be negative but very small in absolute value due to loss of numerical accuracy.

A prescription is thus given in order to handle these cases.

3 The prescription

3.1 Positivity prescription

NNPDF PDF sets are delivered as PDF replica samples $\{q^k\}$ with $k = 1, \ldots, N_{\mathrm{rep}}$. The computation of any physical observable $\mathcal{F}$ is performed by computing the observable N_{rep} times, for each of the PDF replicas in the sample: $\mathcal{F}[\{q^{(k)}\}]$. Predictions are then obtained by performing statistics over the sample of replicas of the observable.

The prescription for handling cases in which $\mathcal{F}[\{q^{(k)}\}] < 0$ for some i, while physical constraints would impose $\mathcal{F} \geq 0$ (e.g. if $\mathcal{F}$ is a cross-section) is simply

- whenever $\mathcal{F}[\{q^{(k)}\}] < 0$, set $\mathcal{F}[\{q^{(k)}\}] = 0$.

This means that in such case one or more of the replicas of he observable will then give zero.

3.2 PDF uncertainties

When applying this prescription, it is a common situation that one may end up with a non-Gaussian distribution, because the probability distribution of replicas may have a large positive uncertainty, but is cut off at zero. In this case the standard way of computing PDF uncertainties as a standard deviation over the replica sample

$$\sigma_{\mathcal{F}} = \left(\frac{1}{N_{\rm rep} - 1} \sum_{k=1}^{N_{\rm rep}} \left(\mathcal{F}[\{q^{(k)}\}] - \langle \mathcal{F}[\{q\}] \rangle \right)^2 \right)^{1/2} . \quad (3.1)$$

may lead to misleading results, because the use of the standard deviation suggests a symmetric uncertainty about the mean. In such cases, it might be more appropriate to view the above PDF uncertainty as a 68% confidence level, and then more in general estimate the confidence level associated to the PDF uncertainty as an interquantile range.

Specifically, the X% confidence level can be estimated as the difference between the $\frac{1}{2}(100 - X)$ and $\frac{X}{2}$ percentiles. This can in turn be computed from the replica sample as follows:

1. Order the $N_{\rm rep}$ predictions for the cross-section $\mathcal{F}$ from the smallest to the largest, in a way that we have
$$\mathcal{F}_1 \leq \mathcal{F}_2 \leq \ldots \mathcal{F}_{N_{\rm rep}-1} \leq \mathcal{F}_{N_{\rm rep}} \quad (3.2)$$

2. Now remove (100-X)% of replicas ($\frac{1}{2}(1 - \frac{X}{100}) \times N_{\rm rep}$) with the highest and with the lowest values of the cross-section.

3. The resulting interval defines the X% confidence level interval for the cross section $\mathcal{F}$. For example, a 68% c.l. (corresponding to one σ for a Gaussian) is obtained by keeping the central 68% replicas, i.e.
$$\left[\mathcal{F}_{0.16 N_{\rm rep}}, \mathcal{F}_{0.84 N_{\rm rep}} \right] \quad (3.3)$$
so that a 68% c.l. PDF uncertainty on the cross-section is given by
$$\sigma_{\mathcal{F}} = \frac{\mathcal{F}_{0.84 N_{\rm rep}} - \mathcal{F}_{0.16 N_{\rm rep}}}{2} \quad (3.4)$$
Specifically, for $N_{rep} = 100$, the PDF error computed this way is
$$\sigma_{\mathcal{F}} = \frac{\mathcal{F}_{84} - \mathcal{F}_{0.16}}{2} \quad (3.5)$$
where the indices number cross-section replicas which have been ordered in ascending value.

We recommend against using one-σ uncertainties and using instead the central 68% confidence level whenever the positivity prescription leads to setting to zero a number of replicas which is large enough that the uncertainty band becomes asymmetric about the central value.

3.3 Central value

The standard way of computing the central value for a prediction is as the mean

$$\langle \mathcal{F} \rangle = \frac{1}{N_{\rm rep}} \sum_{k=1}^{N_{\rm rep}} \mathcal{F}[\{q^{(k)}\}]. \tag{3.6}$$

However, if the distribution of replicas is non-Gaussian and specifically asymmetric this can be misleading. In this case it is more appropriate to use as best prediction the median, namely

$$\mathcal{F}_{\rm median} = \mathcal{F}_{N_{\rm rep}/2}, \tag{3.7}$$

where again the index refers to replicas ordered by increasing size of the prediction.

In case a very large number of replicas has been set to zero by positivity this median value could be zero. Physically, this means that our best prediction for the observable is compatible with zero, and we can only give an upper bound on the observable, but not a meaningful central value.

3.4 Final recommendation

When applying the positivity prescription we recommend to use the median Eq. (3.7) as central prediction and the 68% confidence level Eq. (3.4) as PDF uncertainty band. Whenever the median is zero, this means that only an upper bound can be provided on the given observable. In this case, we recommend the users to adopt the prescription that they normally follow whenever only the upper limit for a quantity can be given.