Eur. Phys. J. C (2024) 84:697
https://doi.org/10.1140/epjc/s10052-024-12972-7

THE EUROPEAN PHYSICAL JOURNAL C

Regular Article - Theoretical Physics

Yadism: yet another deep-inelastic scattering module

Alessandro Candido[1,2], **Felix Hekhorn**[1,3,4], **Giacomo Magni**[5,6], **Tanjona R. Rabemananjara**[5,6], **Roy Stegeman**[7,a]

[1] Tif Lab, Dipartimento di Fisica, Università di Milano and INFN, Sezione di Milano, Via Celoria 16, 20133 Milan, Italy
[2] CERN, Theoretical Physics Department, 1211 Geneva 23, Switzerland
[3] Department of Physics, University of Jyvaskyla, P.O. Box 35, 40014 Jyvaskyla, Finland
[4] Helsinki Institute of Physics, University of Helsinki, P.O. Box 64, 00014 Helsinki, Finland
[5] Department of Physics and Astronomy, Vrije Universiteit, 1081 HV Amsterdam, The Netherlands
[6] Nikhef Theory Group, Science Park 105, 1098, XG Amsterdam, The Netherlands
[7] The Higgs Centre for Theoretical Physics, University of Edinburgh, JCMB, KB, Mayfield Rd, Edinburgh EH9 3JZ, Scotland

Received: 2 February 2024 / Accepted: 1 June 2024

Abstract We present yadism, a library for the evaluation of both polarized and unpolarized deep-inelastic scattering (DIS) structure functions and cross sections up to N^{3}LO in perturbative QCD. The package provides computations of observables in fixed-flavor and zero-mass variable flavor number schemes. The implementation of the general mass variable flavor number schemes is supported through the high virtuality limits for the heavy flavor coefficients. In addition, yadism provides a set of tools for the generation of interpolation grids in the PDF-independent PineAPPL format, allowing to test the PDF dependence on any DIS observable without needing to rerun the computation. This work is part of an ongoing effort to standardize the format of theory predictions in high-energy physics within the pineline framework. The code is open source, written in Python and documented to facilitate usage, integrations, and further extensions. Finally, the code has been benchmarked against the widely used APFEL++ and QCDNUM libraries.

Contents

1 Introduction

Deep-inelastic scattering (DIS) experiments provide strong constraints on the structure of hadrons, and around half of the experimental data points, used in the most recent PDF determinations [1–3], corresponds to charged current (CC) or neutral current (NC) DIS processes. These include both relatively recent data, such as the ones collected at HERA [4,5], and earlier data such as from the BCDMS [6,7] experiment. While in recent years the focus of particle physics phenomenology has shifted away from DIS in favor of the description of LHC data, the upcoming Electron-Ion Collider (EIC) projects in the US [8] and China [9] have renewed interest in DIS. Thus, an accurate description of these processes is required to optimally utilize their future data. Reliable predictions for DIS are furthermore relevant for the interpretation of neutrino scattering data [10] from neutrino telescopes such as IceCube [11] and experiments such as the FPF [12,13], the LHeC [14–16] or FASERν [17] and SND@LHC [18,19] at the LHC.

In this paper we present yadism, a new software library developed for the calculation of DIS observables with the requirements of the particle physics community of the current age in mind. Yadism differs from other QCD codes such as APFEL [20], APFEL++ [21], HOPPET [22], and QCDNUM

[a] e-mail: r.stegeman@ed.ac.uk (corresponding author)

Published online: 15 July 2024

[23] in various ways, which we briefly highlight in the following.

`Yadism` includes most of the currently available results in literature, specifically it allows for the computation of polarized [24] and unpolarized structure functions up to next-to-next-to-next-to-leading order (N^3LO) [25] in QCD. Thanks to its modular design, the library can be easily extended as the results of new computations become available. The coefficients, whenever possible, have been benchmarked against `APFEL++` and `QCDNUM`.

`Yadism` provides both renormalization and factorization scale variations consistently [26] and both can be implemented at any order. The currently implemented coefficients allow to perform renormalization scale variations up to N^3LO and factorization scale variations up to NNLO. Instead, N^3LO factorization scale variations can be included through the `EKO` evolution code [27].

`Yadism` can, together with `EKO`, be used to construct general-mass variable flavor number schemes (GM-VFNS) using coexisting PDFs with different numbers of active flavors. This can avoid [28] the perturbative expansion of the evolution kernel as is currently done in the construction of the FONLL scheme [29].

`Yadism` has a uniform treatment of all heavy quarks, i.e., all features that are available for charm are also available for bottom and top. This strategy opens up the possibility for computations with an intrinsic bottom quark [30,31]. We provide both the fixed-flavor number scheme (FFNS) and zero-mass variable-flavor number scheme (ZM-VFNS) calculation, as well as the asymptotic limit, $Q^2 \gg m^2$, of the FFNS (FFN0), which is required in the construction of the FONLL scheme [29].

`Yadism` is interfaced to `PineAPPL` [32,33], a library providing fast interpolation grids in a unique format and exposing an Application Programming Interface (API) for the programming languages `C`, `C++`, `Fortran`, `Rust`, and `Python` making it portable and easy to use. Fast interpolation grids provide a representation of predictions independent of PDFs and the strong coupling, and therefore do not require rerunning the entire toolchain of theory codes if one whishes to assess the impact of the PDF on a theory prediction. This feature can be useful both for PDF fitting and also for the determination of standard model parameters [34]. Fast interpolation grids were pioneered by FastNLO [35] and are a vital part in the toolchain for phenomenology of perturbative QCD, as such the grid technique has been adopted in various programs including APPLgrid [36] and APPLfast [37]. The `PineAPPL` grid output format allows `yadism` to be integrated into the `xFitter` framework [38–40] and the `pineline` framework [41]. Specifically, the latter consists of various codes with the aim to automate and efficiently compute theory predictions for collider physics processes. Through this toolchain, one can define a collection of consistent theory parameters and observables of interest for which both the partonic coefficients along with the DGLAP evolution kernels (through the `EKO` package [27]) can conveniently be calculated to produce fast interpolation grids.

`Yadism` is written in the `Python` programming language, which is known for its ease of use, and thus reduces the threshold for potential new contributors. For these reasons, development of new functionality can be quick to, e.g., rapidly adopt new computations. Proposed changes to the `yadism` code are reviewed thoroughly and are subjected to automated checks as part of a Continuous Integration (CI) policy.

So far `yadism` has already been used in various papers. Specifically, it has been used for the evaluation of neutrino structure functions in Refs. [10,42], and for the computation of polarized structure functions in Ref. [24]. Furthermore, `yadism` has been adopted by the NNPDF collaboration, who has used it in their most recent papers [25,26,43].

The `yadism` code is open source and free to use under a GPL−3.0 license and can be found in its Github repository:
https://github.com/NNPDF/yadism
along with a user-friendly and up-to-date documentation:
https://yadism.readthedocs.io/en/latest/.

In this paper, we aim to provide an overview of some of the functionalities provided by `yadism` while we refer the reader to the documentation for an extensive overview of all the available features.

The rest of the paper proceeds as follows: in Sect. 2, we briefly summarize the theory underlying deep-inelastic scattering and provide details on the `yadism` implementation. In Sect. 3, we discuss representative benchmarks in comparison to other available libraries. Finally, we conclude in Sect. 4 and provide a brief description on possible extensions. In addition, we include two appendices where we briefly comment on the calculation of a new set of formerly unknown coefficient functions in Appendix A and we give an explicit example on how to run `yadism` in Appendix B.

2 The `yadism` library

In this section we introduce `yadism` and provide an overview of its most important features. In the following we assume the standard notation on DIS calculations as can be found, e.g., in any textbook [44] or in the PDG review [45].

DIS structure functions can be evaluated within the framework of perturbative Quantum Chromodynamics (pQCD) using collinear factorization [46] by convoluting the PDFs with the relevant coefficient functions

$$F\left(x, Q^2\right) = \sum_j \int_x^1 \frac{dz}{z} C_j\left(z, \alpha_s\left(Q^2\right)\right) f_j\left(\frac{x}{z}, Q^2\right), \tag{2.1}$$

where j runs over all possible partons in the initial state, f_j denotes the PDF of flavor j and C_j are the coefficient function that can be calculated perturbatively as an expansion in the strong coupling α_s

$$C_j\left(z, \alpha_s\left(Q^2\right)\right) = \sum_{k=0}^{\infty}\left(\frac{\alpha_s}{4\pi}\right)^k C_j^{(k)}(z). \tag{2.2}$$

Here and in the following we will assume that the LO coefficient function are of $\mathcal{O}(\alpha_s^0)$ irrespective of the first non-zero order.

All structure functions depend on two kinematic variables: the Bjorken-x and the virtuality Q^2. Equation (2.1) highlights how the DIS structure functions, describing the lepton-hadron interaction, depend on a linear combination of PDFs. Thus, it is clear why DIS processes provide important constraints on flavor separation and are therefore fundamental for the PDFs determination from experimental data.

However, while the factorization formula, Eq. (2.1), is conceptually simple, if one wishes to actually compute a structure function one needs to define a number of theory parameters and parameters of the experimental setup. Such input settings are passed to `yadism` through *runcards* in YAML format,* and they are divided into two parts: an *observable runcard* describing the experimental setup (such as scattering particles, kinematic bins, or helicity settings) and a *theory runcard* describing the parameters of the theory framework (such as coupling strength, perturbative orders, or quark masses). While observable runcards are usually tailored to a given experiment, theory parameters are usually shared by multiple runs. An example of such runcards is reported in Sect. B, where we show explicitly how to run `yadism`.

For completeness, the current settings of DIS datasets used in the NNPDF framework are collected in the repository *pinecards*:

https://github.com/NNPDF/pinecards.

where we specify the setup for a number of measurements at NMC [47,48], SLAC [49,50], BCDMS [6,7], CHORUS [51], NuTeV [52], EMC [53], and HERA [4,5]. There, we also provide settings for several pseudo-measurements, which are used as theoretical constraints in NNPDF [3].

Below, we describe the most important options for the configuration of the observables (Sect. 2.1) and theories (Sect. 2.2) that can be defined in the respective runcard. Section 2.3 overviews the partonic coefficient functions implementation and technical details on the computation thereof.

* https://yaml.org/.

2.1 Observable configuration options

Projectile. `Yadism` supports computations of DIS coefficients with massless charged leptons and their associated neutrinos as projectiles in the scattering process. Specifically, to describe, e.g., the HERA data one needs both electrons and positrons and, e.g., for the CHORUS data both neutrinos and anti-neutrinos are needed. Charged leptons can interact both electromagnetically and weakly with the scattered nuclei, whereas neutrinos only carry weak charges. Recently, together with a machine-learning parametrization of experimental data, CC neutrino DIS predictions computed with `yadism` have been used to extend predictions for neutrino structure functions [10].

Target. `Yadism` supports computations with nuclei with mass number A and Z protons as targets in the scattering process. By acting on the coefficients associated to up and down partons `yadism` implements the isospin symmetry of the form:

$$\begin{pmatrix} c'_u \\ c'_d \end{pmatrix} = \frac{1}{A}\begin{pmatrix} Z & A-Z \\ A-Z & Z \end{pmatrix}\begin{pmatrix} c_u \\ c_d \end{pmatrix} \tag{2.3}$$

where c'_i and c_i are the effective and the proton coefficient associated with the parton i. This rotation is particularly useful in the context of proton PDF fitting where it can be used to relate neutron, deuteron, and heavier nuclear structure functions to the proton ones. In this way, isospin is used as a first approximation of nuclear correction by just swapping up and down contribution for the amount specified by the target nuclei. In particular for:

proton targets ($A = 1, Z = 1$): up and down are kept as they are.

neutron targets ($A = 1, Z = 0$): up and down components are fully swapped, such that the up coefficient function is matched to the down PDF and conversely.

isoscalar targets, i.e. deuteron ($A = 2, Z = 1$): the effective coefficient functions will be mixed such that c'_u will be half the original c_u and half the original c_d.

`Yadism` is completely general with respect to the nuclear target allowing a user to provide values for A and Z as input to the computation. Alternatively, for a number of targets, the name itself can also be provided as input. The readily available targets are: iron ($A = 49.618, Z = 23.403$), used to describe NuTeV data; lead ($A = 208, Z = 82$), used to describe CHORUS data; neon and marble ($CaCO_3$) with both $A = 20, Z = 10$, used to describe respectively the BEBCWA59 [54] and CHARM [55] data.

Exchanged electroweak gauge boson(s). DIS can be categorized into three different processes defined by the gauge boson mediating the interaction:

Electromagnetic current (EM) corresponds to a process where the exchanged boson is a photon.

Neutral current (NC) corresponds to cases where the exchanged boson does not carry any electric charge. Thus, it is a superset of the EM process where also the exchange of the Z boson is allowed. Since for NC two bosons are allowed, interference terms must be included. Because the Z boson has an axial coupling to the incoming lepton it introduces further complications related to γ_5 [56]. Note that at small virtualities the Z contributions are suppressed by a factor Q^2/M_Z^2.

Charged current (CC) corresponds to processes where a $W^\pm$ boson is exchanged. The CC process is a flavor-changing current where the CKM-matrix encodes the probabilities to transition between different quark flavors [57,58].

Cross sections `Yadism` supports the computation of both structure functions and (reduced) cross sections. In particular, for the unpolarized scattering, we implement the structure functions:

$$F_2,\ F_L,\ xF_3, \tag{2.4}$$

and their polarized counter parts:

$$g_4,\ g_L,\ 2xg_1, \tag{2.5}$$

where the normalization is chosen such that at LO, all the structure functions are proportional to different PDF combinations of the form $xf(x)$.

While structure functions may only depend on two variables, cross sections may also depend on the inelasticity y. Generally, we can write the (reduced) cross sections for a DIS process in terms of the structure functions as

$$\begin{aligned}\sigma(x, Q^2, y) &= N\Big(F_2(x, Q^2) - d_L F_L(x, Q^2) \\ &\quad + d_3 x F_3(x, Q^2)\Big),\end{aligned} \tag{2.6}$$

where N, d_L, and d_3 may depend on the experimental setup or the scattered lepton. The different reduced cross sections implemented in `yadism`, and their definitions in terms of N, d_L, and d_3 can be found in the online documentation.[†] The implemented definitions can be used to describe data from HERA, CHORUS, NuTeV, CDHSW [59], and FPF [42].

Finally, we provide the linearly dependent structure functions:

$$2xF_1 = F_2 - F_L, \qquad 2xg_5 = g_4 - g_L. \tag{2.7}$$

Flavor tagging. In general, any total DIS structure function F can be decomposed in three different components, according to the type of quark coupling to the exchanged EW boson:

$$F = F^{(l)} + F^{(h)} + F^{(hl)}, \tag{2.8}$$

where $F^{(l)}$ denotes the contribution coming from diagrams where all the fermion lines are massless, $F^{(h)}$ is the contribution due to heavy quarks coupling to the EW boson and $F^{(hl)}$ originates from higher order diagrams where a light quark is coupling to the boson, but heavy quarks lines are present.

Given Eq. (2.8), we support the calculation of fully inclusive (total) observables, where only the lepton is observed in the final state, and flavor tagged final state, where we require a specific heavy quark (charm, bottom, or top) to couple with the mediating boson. This definition coincides with $F^{(h)}$ and it is an infrared-safe definition [29]. For example, the charm structure function $F^{(c)}$ can be obtained by assuming the coupling of any quark other than charm and anti-charm to be zero. Instead, a naive definition of $F^{(c)}$ by the heavy final state tag would not be infrared safe. For completeness, also light structure functions $F^{(l)}$ are available, in isolation, although they do not corresponds to any physical observable.

2.2 Theory configuration options

Flavor number schemes. Flavor number schemes provide a prescription to resolve the ambiguous treatment of heavy quark masses. Generally, to achieve a faithful description of experimental data at scales roughly around the heavy quarks mass $Q \sim m$, quarks should be treated fully massive. However, in the region where $Q \gg m$, quarks should be considered massless. In `yadism` we allow for 3 different schemes. Only one single heavy quark is allowed at each time.

Fixed flavor number scheme (FFNS). The FFNS, is defined as a configuration with a fixed number of flavors at all scales, i.e. all quark masses are fixed to be either light, heavy or decoupled. The FFNS retains all power-like heavy quark corrections m^2/Q^2 and a finite number of logarithmic corrections $\ln(Q^2/m^2)$. This finite number of logarithms, as opposed to a full resummation, limits the perturbative stable region.

Zero mass-variable flavor number scheme (ZM-VFNS). In the ZM-VFNS all quark masses in the calculations are either light or decoupled. The number of light quarks n_f is not fixed, but instead varies with the number of active flavors depending on the scale of the process, i.e. $n_f(Q^2)$. Specifically, $n_f = 3$ below m_c and this increases as the heavy quark thresholds are crossed, i.e. $Q > m_h$, after which the corresponding heavy quark is treated to be light. The ZM-VFNS resums all logarithmic corrections as they are provided by DGLAP evolution. However, the ZM-VFNS does not contain any power-like heavy quark corrections m^2/Q^2 which may be phenomenolog-

[†] https://yadism.readthedocs.io/en/latest/theory/intro.html#cross-sections.

ical important in certain regions of the kinematic phase space.

Asymtotic fixed flavor number scheme (FFN0). The FFN0 is similar to the FFNS, but retains only the logarithmic corrections, i.e. it does not contain any power-like heavy quark corrections m^2/Q^2. The FFN0 is constructed as the overlap between FFNS and ZM-VFNS and can be used to construct a GM-VFNS flavor number schemes. A GM-VFNS can be constructed to overcome the limitation due to potentially large missing corrections of FFNS and ZM-VFNS. One possible scheme is the FONLL scheme [29], which is defined through a linear combination of the FFNS and ZM-VFNS while taking care of possible double counting through the FFN0. A detailed discussion on how to construct the FONLL scheme is given in Ref. [28]. `Yadism` does not provide it explicitly, but all the necessary ingredients FFNS, FFN0, and ZM-VFNS are available.

Renormalization and factorization scale variations. In perturbative QCD the coefficients C_j of Eq. (2.1), are expanded in powers of α_s. The estimate of the error introduced by the truncation of such series is quite relevant in multiple precision applications. Some information about the missing higher orders, and the related uncertainty (MHOU), can be extracted from the Callan-Symanzyk equations violation. In this sense, a practical approach to obtain a numerical estimate consists in varying the relevant unphysical scales.

In DIS, the two involved unphysical scales are the *renormalization scale*, arising from the subtraction of ultraviolet divergences, and the *factorization scale*, from the subtraction of collinear logarithms in the PDF definition.

The explicit expressions of the C_i expansion upon scale variations can be found, e.g., in Sect. 2 of Ref. [60]. Generally, these depend, order by order in perturbation theory, on the derivatives of α_s and the PDFs with respect to the mentioned scales. The former are the β-function coefficients and the latter the splitting functions. In `yadism`, necessary β-function coefficients are taken from the `EKO` package, while the x-space splitting functions are directly implemented.

At the level of structure function, scale variations can be cast into an additional convolution with a kernel K:

$$F(x, \mu \neq Q) = (K \otimes C \otimes f)(x) \tag{2.9}$$

It can be shown that the transformation can be applied a-posteriori to an already computed interpolation grid.

Target mass corrections. While Eq. (2.1) is usually derived for the scattering of two massless particles, it is possible to account for the finite mass of the scattering target through target mass corrections [24,61,62]. These corrections become relevant for either small virtualities or large Bjorken-x. They can be implemented as an additional convolution and we provide several approximations (corresponding to higher twist expansions) following Ref. [61].

2.3 Implemented partonic coefficients and computation details

Quark mass corrections. We can differentiate quarks into three different types: light ($m = 0$), heavy (m finite) and decoupled ($m = \infty$). Thus, each coefficient function of Eq. (2.1) can be categorized by the appearance of heavy quark lines in various parts of the diagrams:

Light does not contain massive corrections in any part, i.e. all quarks are either light or decoupled.

Heavy contains heavy quarks in the output. Note that while some calculation of coefficients with two mass scales are available [63], in `yadism` we currently only provide support for coefficients depending on a single heavy quark mass scale since the impact of the missing corrections are small.

Intrinsic contains contributions where the incoming parton is a heavy quark and which thus allows for intrinsic heavy quarks as opposed to radiatively produced heavy quarks.

Asymptotic is the $Q \gg m$ limit of either the heavy or intrinsic coefficient. The asymptotic contributions are used in the construction of general mass variable flavor number schemes.

This classification is not exclusive and it is useful to only distinguish coefficient functions but it does not correspond to a unique trivial mapping at the level of structure functions (see Eq. (2.8)). In fact, depending on the chosen variable flavor number scheme, the same coefficients can be reshuffled differently inside each of the components $F^{(l)}$, $F^{(h)}$, and $F^{(hl)}$, or might even not be present at all. For example in the FONLL scheme [29] $F^{(c)}$ is computed with massive quarks (using *heavy*), with massless quarks (using *light*), and in the asymptotic mass limit (using *asymptotic*).

Partonic coefficient functions.

`Yadism` implements both unpolarized and polarized coefficient functions up to N^{3}LO in fixed-order QCD. In Tables 1 and 2 we collect a summary of the coefficient functions as currently implemented. For each perturbative order[‡] and process, we distinguish contributions from light, heavy, asymptotic and intrinsic coefficients. In the unpolarized case, Table 1, the light, heavy, and asymptotic contributions are available up to N^{3}LO, except for the CC. The intrinsic components are also available only up to NLO, with the CC part computed very recently, see also Sect. A. The NNLO corrections to CC are known only through K-factors [73] and their implementation into `yadism` is currently work in progress. Instead, for the heavy NC N^{3}LO coefficients, a full analytical

[‡] Recall that we adopt an *absolute* terminology of perturbative order, i.e., LO $= O(a_s^0)$ irrespective of the first non-zero order, e.g. for F_L or $F_2^{(c)}$.

Table 1 Overview of the unpolarized DIS coefficients currently implemented in the `yadism` library at the corresponding order in perturbative QCD. In the columns we distinguish between light, heavy, intrinsic, and asymptotic. We mark in green coefficient function that are implemented in `yadism`, in red the ones which are not yet known and in yellow the ones which are not yet implemented in `yadism`, but available in literature

NLO $O(a_s)$	light	heavy	intrinsic	asymptotic
NC	✓[64–66]	✓[67]	✓[68]	✓[69]
CC	✓[70, 71]	✓[72]	✓see Appendix A	✓[69]
NNLO $O(a_s^2)$				
NC	✓[64–66]	✓[67]	✗	✓[69]
CC	✓[70, 71]	[73]*	✗	[69]
N^{3}LO $O(a_s^3)$				
NC	✓[64–66]	✓[74–76]	✗	✓[76–82]
CC	✓[70, 71]	✗	✗	✗

*Available as K-factors

Table 2 Same as Table 1 for NC polarized coefficients

	light	heavy	intrinsic	asymptotic
NLO $O(a_s)$	✓[83–85]	✓[67]	✗	✓[86]
NNLO $O(a_s^2)$	✓[83, 87]	✓[67]	✗	✓[86]
N^{3}LO $O(a_s^3)$	[88]*	✗	✗	[89–94]*

* Only for the g_1 structure function

result is not known, although approximations can be constructed using information already available through resummations and high virtuality limits [74,75]. These coefficient are then implemented together with an uncertainty which can be propagated to the final result.

Similarly, Table 2 provides an overview of the polarized NC coefficient functions currently implemented. As opposed to the unpolarized counterpart, intrinsic polarized coefficient functions are not yet known. At N^{3}LO, only the light coefficient functions for the structure function g_1 have been computed [88] together with the heavy asymptotic limit $Q^2 \gg m_h^2$ [89–94]. Their implementation in `yadism` is left for future updates.

Analytic structure of coefficient functions. The coefficient functions are not restricted to being regular functions, but they might also correspond to a Dirac delta function or singular distributions. In particular, the latter occur in the light coefficient functions because of massless quarks: whenever the mass of the quark does not prevent IR divergences, it generates *plus distributions* upon subtractions.

The presence of such distributions does not cause any issue at the analytical level, since the coefficients have to be convoluted with a PDF (or an interpolation polynomial as in Eq. (2.11)), and thus they always act within the scope of an integral. Instead, it does require a dedicated treatment at numerical level, since a distribution cannot be just evaluated (sampled) at given points, and integrated with some approximation, as it is done for regular functions. As common in literature, our x-space coefficient function implementation follows the so called *regular, singular, local* formalism, first described in [95].

Grid formalism. It is common for DIS calculations to provide coefficient functions that are directly convoluted with a given PDF, thus returning the predicted value for the requested DIS cross section. This is, however, not necessarily the most practical approach. Instead one may wish to store the computation of the DIS coefficient in an interpolation grid format, thus factorizing the PDF dependence. This is useful in situations where predictions for the same observable have to be computed for different PDFs. The case where this is clearly most relevant is in the context of PDF fits, where at each step of the fitting procedure new comparisons to data are required. In order to unify the treatment inside a PDF fit we follow the `pineline` framework [41] and provide interpolation grids, which are more beneficial in a PDF fitting environment.

We introduce an interpolation for the PDF $f(x)$ using the nodes x_k and its associated basis of interpolation polynomials $p_k(x)$ and write

$$f(x) = \sum_k f(x_k) p_k(x) = f^k p_k(x) \tag{2.10}$$

where we defined $f^k \equiv f(x_k)$ and, as usual, sum over repeated indices. While the choice of the nodes is left up to the user, the interpolation basis is fixed to piece-wise Lagrange polynomials, and provided by `EKO`. The PDF values can then be evaluated directly on the nodes from the original parametrization, or (re-)interpolated from distributed PDF grids (such as provided by LHAPDF [96]).

To compute an observable $\sigma(x)$ for a given Bjorken-x, we can then write

$$\sigma(x) = (C \otimes f)(x) = f^k (C \otimes p_k)(x) = f^k C_k(x) \tag{2.11}$$

and identify $C_k(x) = (C \otimes p_k)(x)$ as the sought-after interpolation grid. Note that in Eq. (2.11) we suppressed for the sake of readability the flavor dependency, the scale dependency, and the dependency on the strong coupling. In practice, however, we need to keep track of all of them and the `PineAPPL` format [32,33] supports such a full breakdown.

3 Benchmarking and validation

Having described the yadism library and its available features, we will now provide various benchmark analyses. First, we benchmark yadism with some of the most widely used libraries for the computation of DIS observables, namely APFEL++ and QCDNUM. Then we provide representative comparisons on the different prescriptions used to treat heavy quark masses in order to underline their relevance in the different kinematic regions. In all the subsequent comparisons, we adopt a fixed boundary condition defined as a PDF set at a given scale $Q = Q_0$. Evolution of the boundary condition, including changing of the number of active flavors, is performed using EKO.

3.1 Benchmarking

Let us start discussing the benchmarks of yadism in comparison to other available libraries for a set of representative structure functions using various flavor number schemes and different perturbative orders. In particular, we show benchmarks for both the unpolarised structure function F_2 and its polarised counterpart $2xg_1$ using ZM-VFNS[§] to highlight the accuracy of the massless calculation. Then we compare $F_2^{(c)}$ with FFNS with $n_f = 3$ light flavors to highlight heavy quark mass effects. As benchmark tools, we rely on two main programs:

APFEL++ [21] which provides DIS observables up to N^{3}LO for massless, unpolarized structure functions and up to NNLO for massless, polarized structure functions. It extends the functionalities of the previous Fortran code APFEL [20] and has an explicit dependence on the PDF, which can be interfaced via LHAPDF [96].

QCDNUM [23] which computes DIS structure functions up to NNLO for unpolarized parton densities and up to NLO for polarized parton densities. The program implements both FFNS and ZM-VFNS and uses polynomial spline interpolation to compute the structure function from a given PDF.

The results reported below show the agreement between yadism and other tools.

Massless coefficient functions.

In the ZM-VFNS only massless coefficient functions are involved, thus we expect to reach good agreement with different tools for a broad range of kinematics. Here we select $x \in [10^{-4}, 1]$ and $Q^2 \in [4.0, 10^4]\,\text{GeV}^2$, covering the relevant ranges for DIS phenomenology studies. For simplicity, we focus on NC structure functions, but analogous results hold also for CC DIS. First, in Fig. 1 (Fig. 2) we show the relative difference on F_2 (F_L) between yadism and QCDNUM computed at NLO (left) and NNLO (right) accuracy for different kinematics ranges. The overall agreement is within 0.05% with the largest discrepancies visible in the small-x corner.

The analogous comparison with APFEL++ is displayed in Fig. 3 for the polarized structure function $2xg_1$ and in Fig. 4 (Fig. 5) again for F_2 (F_L) but now at NLO (upper left), NNLO (upper right) and N^{3}LO (bottom) accuracy. Also here the agreement between the different codes is always within 0.05%. An exception is found for $2xg_1$ at NNLO, where the differences are around 0.5%. This larger difference is a result of the different implementation of the nonsinglet (NS) coefficient function – while yadism exploits the exact symmetry of $\Delta C^{(2)}_{1,\text{NS}} = C^{(2)}_{3,\text{NS}}$ [87], APFEL++ implements the analytical calculation from [83].

From the examples discussed, it is clear that the accuracy of the results does not depend on the perturbative order, i.e. the pattern is not affected by the complexity of the calculation.

Heavy quark mass effects. Benchmarks of massive calculations are more involved because massive effects are subdominant in most of the kinematic regions, and can be affected by different approximation of the massive coefficient functions [67,97].

In order to verify the accuracy of our implementation, we report the comparison for the EM charm-tagged structure function $F^{(c)}$, computed in FFNS with three light flavors. We adopt the same kinematic range in Q^2 as in the previous part, but we select $x \in [10^{-4}, 10^{-1}]$ excluding the large-x region where massive structure functions become small and relative uncertainties large. Moreover, a sufficiently large-x corresponds to an energy that is below the threshold to produce a heavy quark pair ($s < 4m^2$).

Figure 6 (Fig. 7) displays the relative difference between APFEL++ and yadism for a NLO and NNLO computations of $F_2^{(c)}$ ($F_L^{(c)}$). In this case, the agreement is around 0.02% at NLO for most of the kinematics and around 0.05% at NNLO.

The analogue comparison to QCDNUM is shown in Fig. 8 at NLO accuracy only demonstrating again a good level of agreement. Here, we cannot perform the comparison at NNLO as QCDNUM does not follow the infrared safe definition of $F^{(c)}$ (as discussed in Sect. 2), but instead includes diagrams with a light quark coupling to the boson into their respective $F^{(h)}$ result which start contributing at NNLO.

3.2 Flavour number schemes

As discussed in Sect. 2.2, yadism implements various FNSs with the aim of reducing the impact of missing logarithmic or power-like corrections that become large in certain regions of phase space. In this section we investigate their relevance.

[§] While ZM-VFNS allows for a variable number of active flavors, i.e. $n_f(Q^2)$, here, and in the rest of this section, we keep n_f fixed to simplify the discussion.

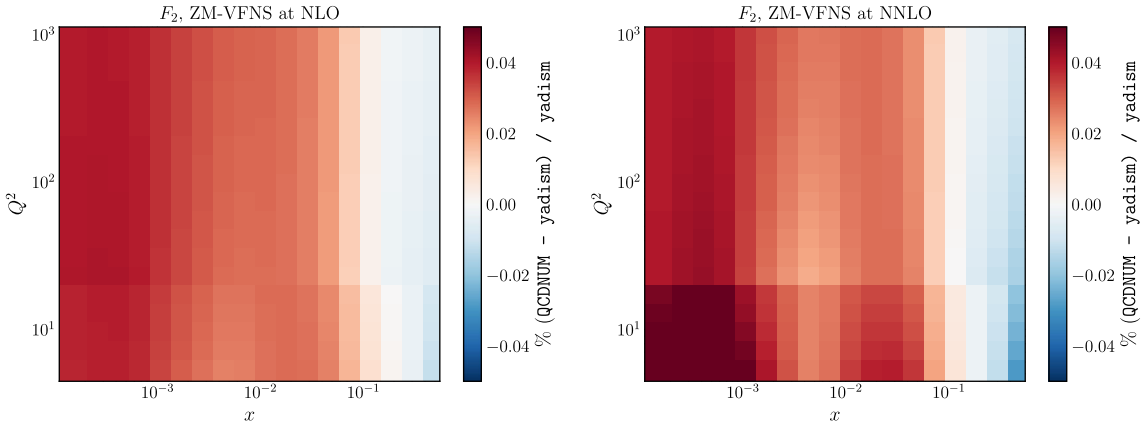

Fig. 1 Relative difference between `yadism` and `QCDNUM` for the structure function F_2 using ZM-VFNS as function of x and Q^2 at NLO (left) and NNLO (right) accuracy

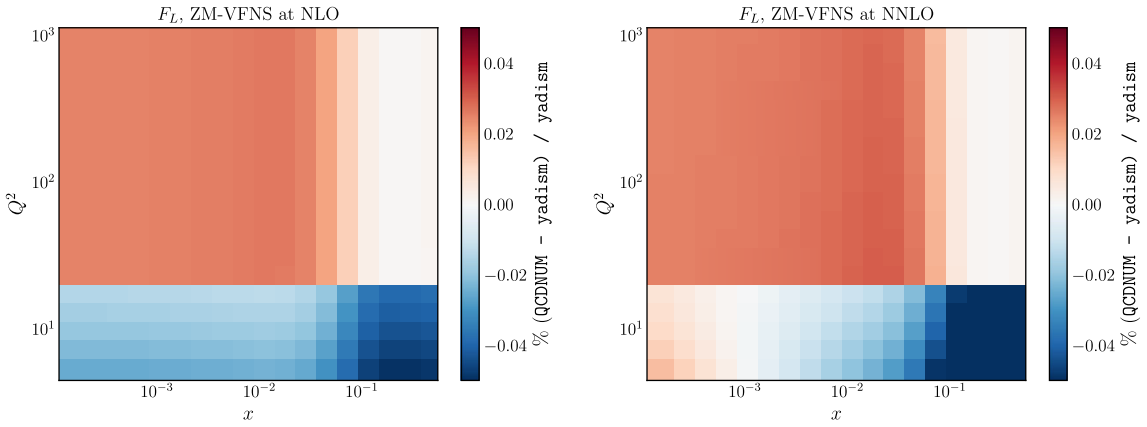

Fig. 2 Same as Fig. 1, but now comparing the structure function F_L

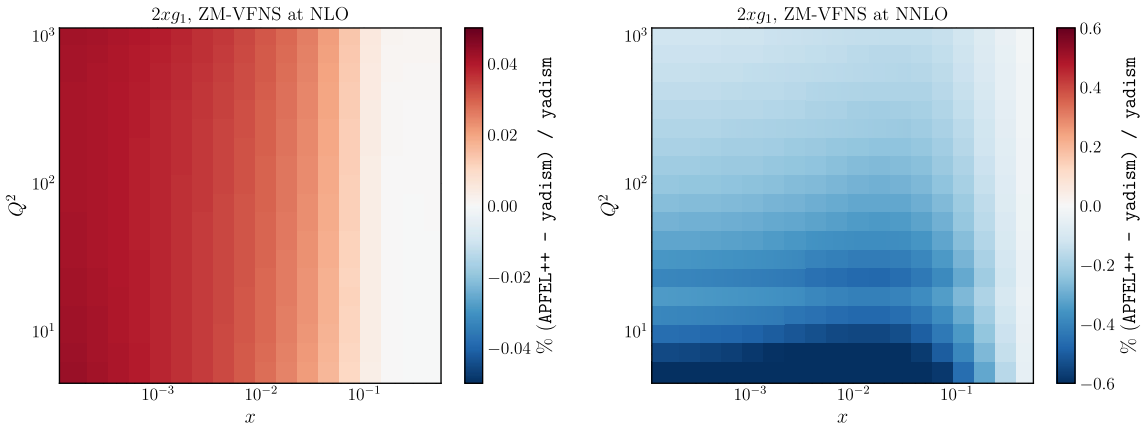

Fig. 3 Same as Fig. 1, but now comparing the structure function $2xg_1$ computed with `yadism` and `APFEL++`

First, in Fig. 9, we compare the ZM-VFNS and FFNS coefficient functions as a function of Q^2. Recall that the ZM-VFNS is defined by assuming all (active) quarks to be massless and the FFNS by considering a single heavy quark with a finite mass and the remaining quarks massless. We expect both calculations to differ more in the low-Q^2 region and progressively reach better agreement towards the large-Q^2 region. However, while ZM-VFNS fully resums all (collinear) logarithms $\log(m^2/Q^2)$, FFNS is a fixed order calculation which only collects a finite number of (collinear) logarithms and hence a finite difference between the two calculations remains. We indeed observe for both structure functions F_2 and $F_2^{(c)}$ this expected pattern, thus confirming a consistent implementation.

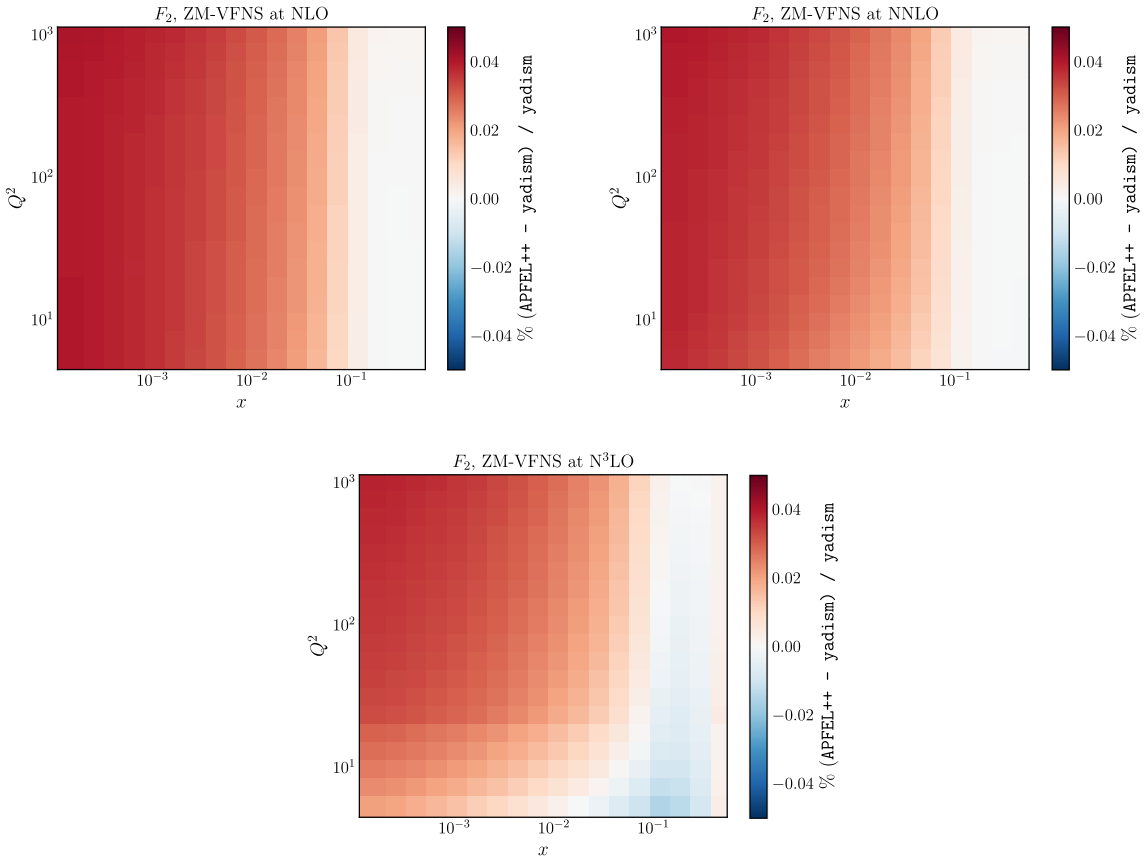

Fig. 4 Same as Fig. 1, but now comparing the structure function F_2 computed with `yadism` and `APFEL++` at NLO (upper left), NNLO (upper right) and N^{3}LO (bottom) accuracy

Next, in Fig. 10, we compare FFNS and FFN0 coefficient functions as a function of Q^2. Recall that FFN0 is derived from FFNS by only keeping the finite number of collinear logs and, hence, we expect both calculations to converge in the large-Q^2 region where any power-like mass corrections vanish. While we can indeed observe this convergence at large-Q^2, we also find a relevant region at mid to low Q^2 where mass effects can grow up to 25%. This latter region can reach up to $O(100)$ times the heavy quark mass and clearly demonstrates the need for GM-VFNSs to improve the accuracy of the prediction.

4 Conclusions and outlook

In this paper we presented `yadism`, a new software package to compute cross section and structure functions in deep-inelastic scattering. In Sect. 2 we reviewed some core features that are relevant in specifying the exact theoretical and experimental setup for which `yadism` is able to provides calculations. `Yadism` has been developed with much care to ensure the results are in agreement with the widely used packages `QCDNUM` and `APFEL++` when they should be, and to understand any differences where they do appear. The success of this effort has been shown in the benchmarking exercise presented in Sect. 3.

While `yadism` is able to reproduce results also available in `QCDNUM` and `APFEL++`, it provides value in following a modular design, allowing it to be easily extended with new coefficient functions, new observables or new DIS-like theories [98,99]. At the time of writing, `yadism` has been used for calculations of the photon PDF [43], where the photon PDF is computed from DIS cross sections in the LuxQED procedure [100–104], the study of heavy quark mass effects in polarised DIS scatterings [24], the study of neutrino-ion interactions at the Forward Physics Facility (FPF) [42] and the determination of low-energy neutrino structure functions [10]. `Yadism` is adopted by the NNPDF collaboration to perform PDF fits where it provides the computations for all fully inclusive DIS measurements by merit of its interface to `PineAPPL`, and so far this has resulted in the work presented in Refs. [26,43].

In future it will be possible to adjust the package structure to exploit synergies with a new software package dedicated to the computation of cross sections in semi-inclusive annihilation (SIA). Indeed, DIS and SIA are related by a crossing relation of Feynman diagrams, which makes the mathematical structure of convoluting a collinear distribution, in this

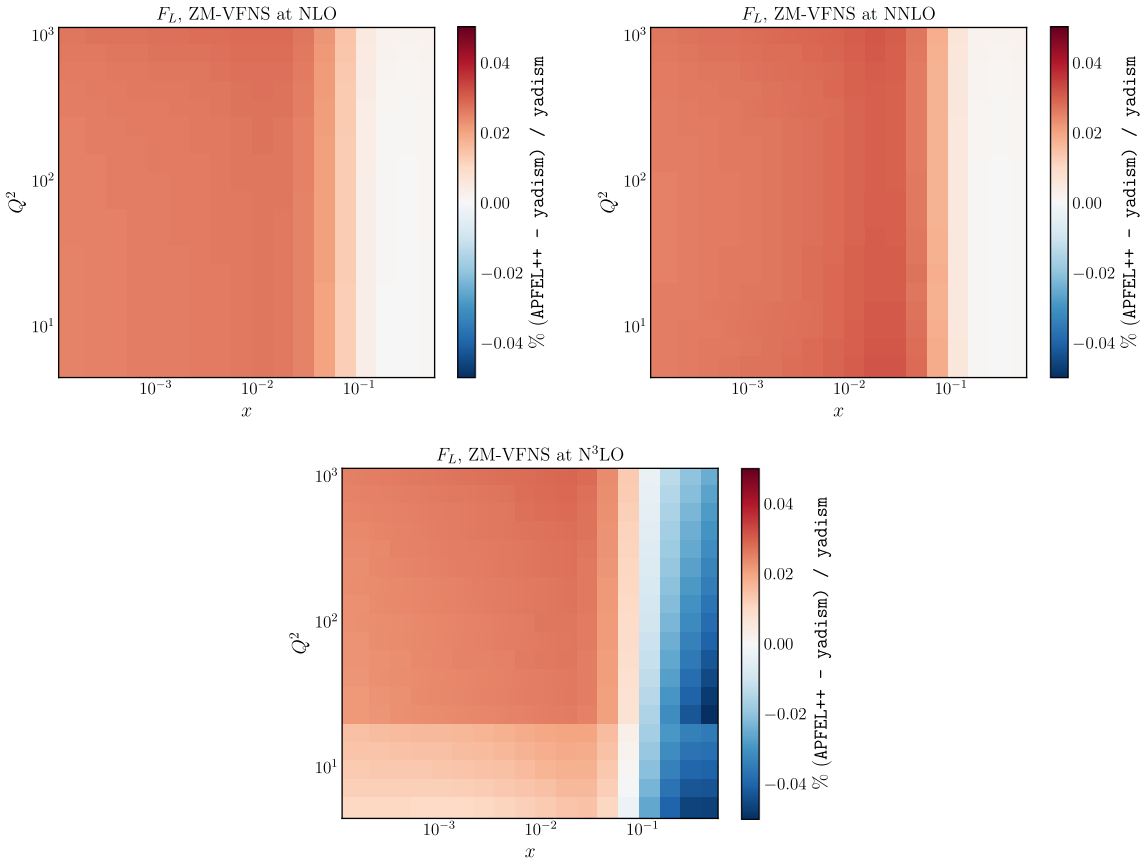

Fig. 5 Same as Fig. 4, but now comparing the structure function F_L

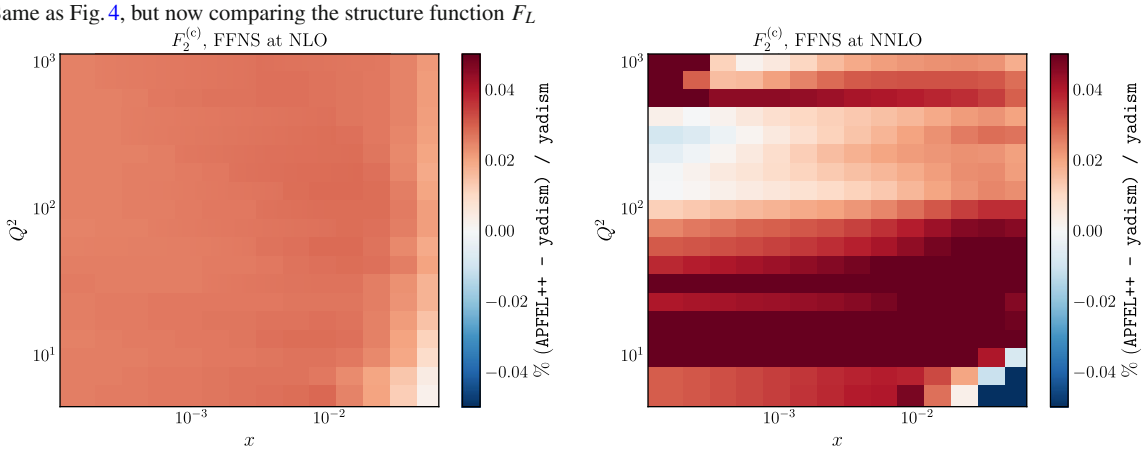

Fig. 6 Relative difference between `yadism` and `APFEL++` for the structure function $F_2^{(c)}$ using FFNS, $n_f = 3$, as function of x and Q^2 at NLO (left) and NNLO (right) accuracy

case a fragmentation function (FF), with a coefficient function very similar.

Acknowledgements We thank all the members of the NNPDF collaboration for numerous discussions during the development of this work. We thank in particular Stefano Forte for guidance during the early stages of the project. We thank Stefano Forte, Juan Rojo, and Christopher Schwan for a careful reading of the manuscript. We are grateful to Valerio Bertone for discussions about `APFEL` and `APFEL++`, and to Kirill Kudashkin for sharing the results of his NLO CC calculation with heavy initial states. F. H. is supported by the Academy of Finland project 358090 and is funded as a part of the Center of Excellence in Quark Matter of the Academy of Finland, project 346326. R. S. is supported by the U.K. Science and Technology Facility Council (STFC) grant ST/T000600/1. G. M. is supported by NWO, the Dutch Research Council. T. R. is partially supported by NWO and by the Netherlands eScience Center.

Data Availability Statement This manuscript has no associated data or the data will not be deposited. [Authors' comment: All data can be generated with the public code.]

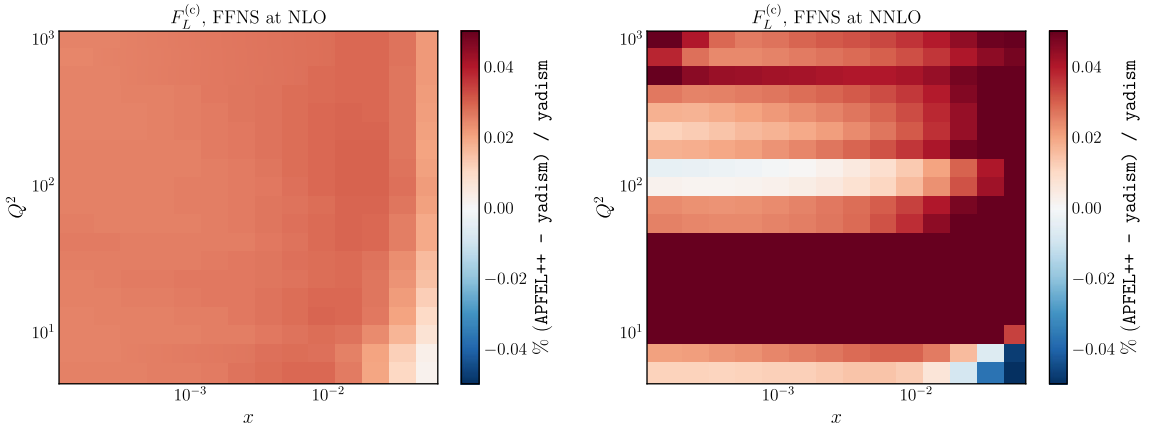

Fig. 7 Same as Fig. 6, but now comparing the structure function $F_L^{(c)}$

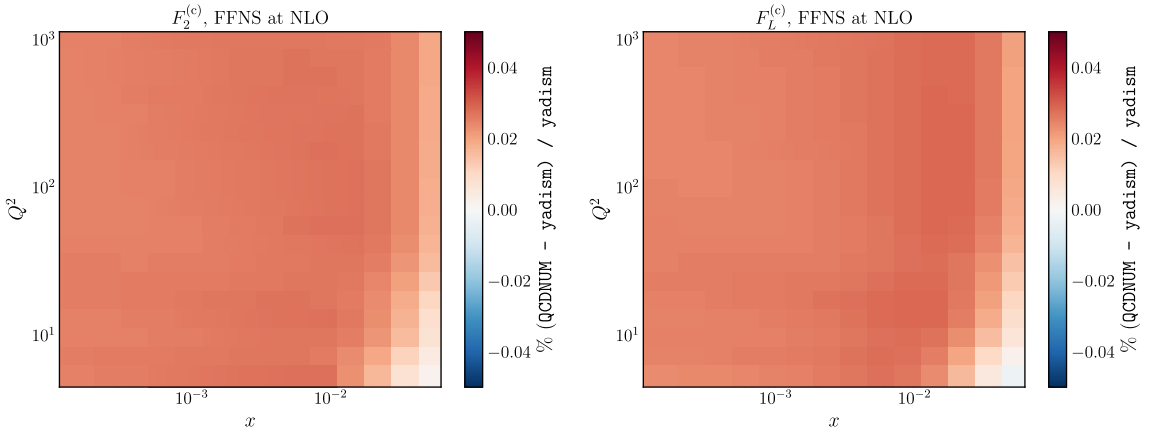

Fig. 8 Similar as Figs. 6 and 7, but now comparing structure function $F_2^{(c)}$ (left) and $F_L^{(c)}$ (right) computed with `yadism` and `QCDNUM`. Note that we can only compare at NLO due to the different conventions adopted by the programs - see text for an explanation

Code Availability Statement This manuscript has no associated code/-software. [Authors' comment: The code is publicly available from https://github.com/NNPDF/yadism.]

Funded by SCOAP[3].

A NLO CC heavy-to-light coefficient functions

In this appendix we sketch the setup used in the computation of NLO heavy-to-light CC coefficient functions. For the NC counterpart, the problem has been solved in [68]. There, the authors study the process

$$Q_1(k_1) + B^*(q) \to Q'(k_2) + X \tag{A.1}$$

with heavy quarks Q and Q' and their respective masses $k_1^2 = m_1^2$ and $k_2^2 = m_2^2$ and the scattered boson B with virtuality $Q^2 = -q^2$. Eventually they compute the coefficient functions up to NLO, which agrees with results that were already available (see Ref. [68, Sect. 2.2.1]).

In our case, for CC DIS, this calculation (beyond the LO) cannot be used directly, since we are interested in the scattering

$$Q_1(k_1) + W^*(q) \to q'(k_2) + X \tag{A.2}$$

now with $k_2^2 = 0$. Indeed, in Eq. (A.1), the explicit dependency on $m_2^2 > 0$ shields the calculation from additional infrared singularities, but this does not happen anymore in Eq. (A.2).

Thus, the CC coefficient functions, with massive initial states, require a dedicated calculation and we report the structure of the obtained results. More details will be discussed in a forthcoming publication [105].

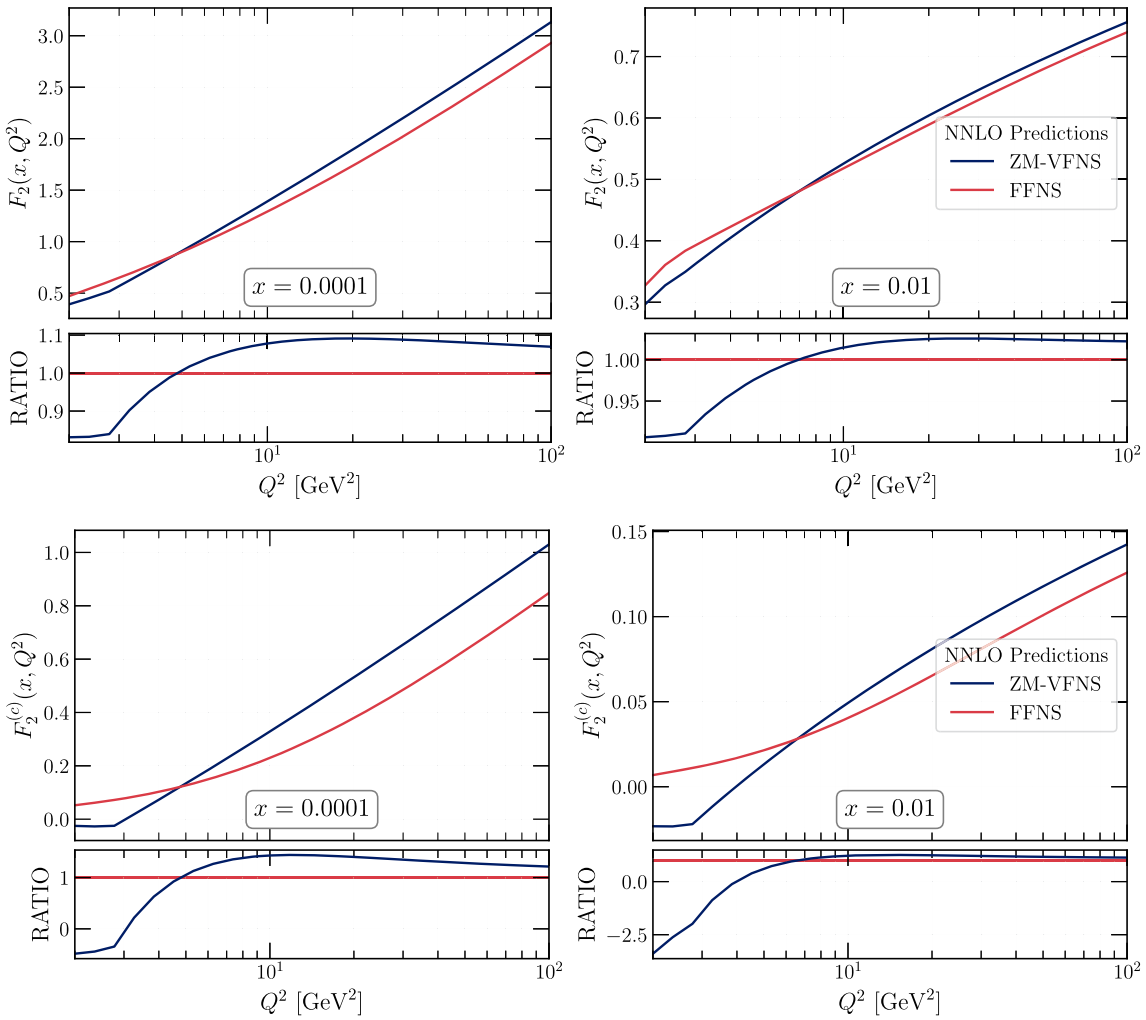

Fig. 9 Comparison of the structure functions F_2 (top) and $F_2^{(c)}$ (bottom) using FFNS and ZM-VFNS at NNLO accuracy. The top panels show the absolute comparisons while the bottom ones show the ratio w.r.t. ZM-VFNS

In particular, for the intrinsic charm contributions to all three unpolarized structure functions we have:

$$2F_1(x, Q^2) \supset \int_x^1 \frac{d\zeta}{\zeta} C_c^1(\zeta, \alpha_s(Q^2), Q^2/m_c^2) f_c(x/\zeta) \tag{A.3}$$

$$\frac{1}{x} F_2(x, Q^2) \supset \int_x^1 \frac{d\zeta}{\zeta} C_c^2(\zeta, \alpha_s(Q^2), Q^2/m_c^2) f_c(x/\zeta) \tag{A.4}$$

$$F_3(x, Q^2) \supset \int_x^1 \frac{d\zeta}{\zeta} C_c^3(\zeta, \alpha_s(Q^2), Q^2/m_c^2) f_c(x/\zeta) \tag{A.5}$$

where the intrinsic charm PDF $f_c(z)$ is now scale independent. The respective coefficient functions are expanded in powers of α_s as

$$C_c^1(\zeta, \alpha_s(Q^2), Q^2/m_c^2) = e^2 \Bigg(\delta(1-\zeta) + \frac{\alpha_s(Q^2) C_F}{2\pi} C_c^{1,(1)}(\zeta, y) \Bigg) \tag{A.6}$$

$$C_c^2(\zeta, \alpha_s(Q^2), Q^2/m_c^2) = e^2 \frac{(y-1)}{y} \Bigg(\delta(1-\zeta) + \frac{\alpha_s(Q^2) C_F}{2\pi} C_c^{2,(1)}(\zeta, y) \Bigg) \tag{A.7}$$

$$C_c^3(\zeta, \alpha_s(Q^2), Q^2/m_c^2) = e^2 \Bigg(\delta(1-\zeta) + \frac{\alpha_s(Q^2) C_F}{2\pi} C_c^{3,(1)}(\zeta, y) \Bigg) \tag{A.8}$$

with C_F the second Casimir constant of the fundamental color representation and $y = -Q^2/m_c^2$. The lengthy expressions of the actual NLO coefficient functions $C_c^{k,(1)}$ are provided as ancillary Mathematica files attached to the arXiv version of this publication. The reader can note the explicit appearance of soft-collinear contributions, which manifest themselves as $\left(\frac{\ln(1-\zeta)}{1-\zeta}\right)_+$ and which are not present in [68].

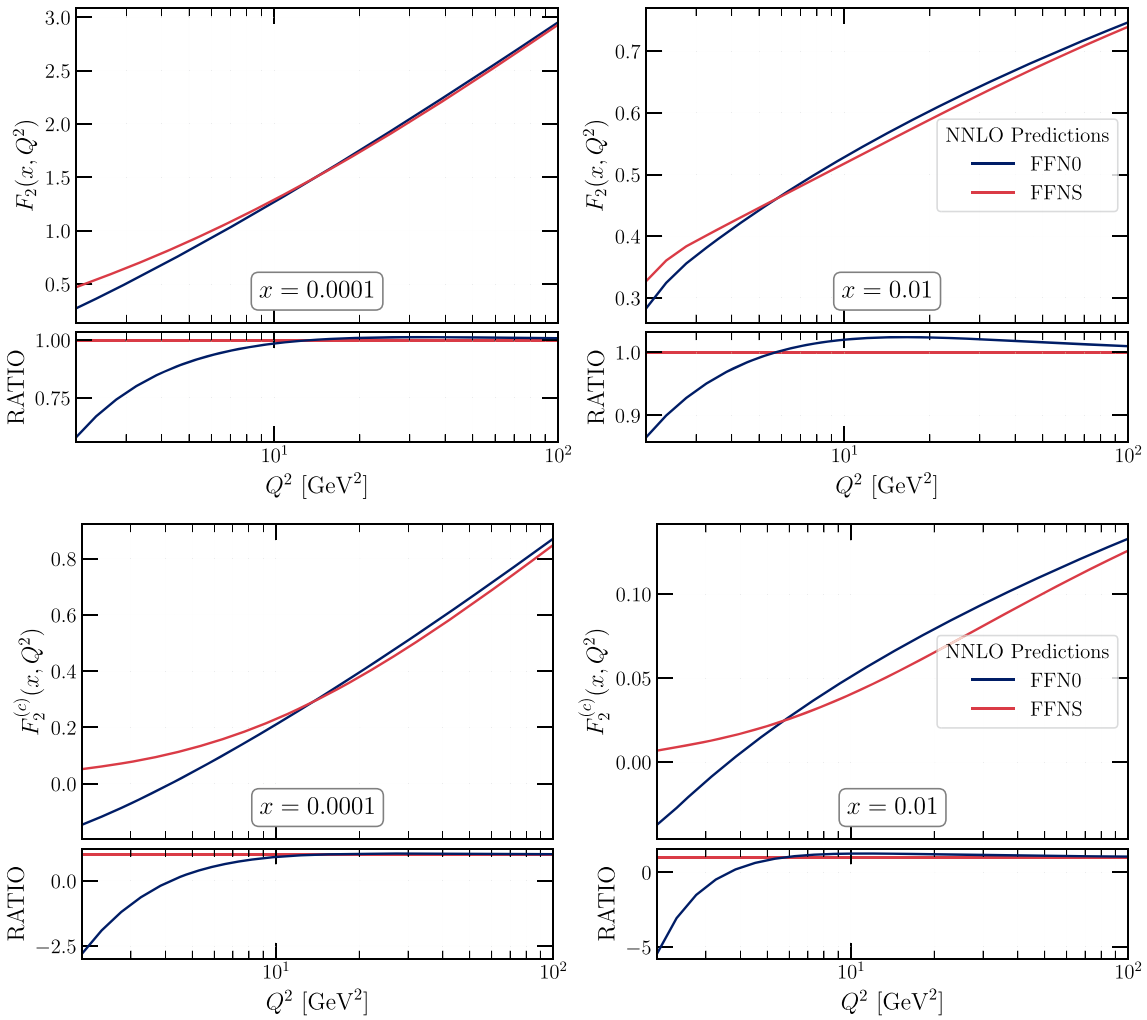

Fig. 10 Same as Fig. 9 but now comparing FFNS and FFN0

B User manual

The purpose of this appendix is to summarize the steps required to generate an interpolation grid using `yadism` by providing an explicit example. For more details, we refer to the online documentation containing, among other things, instructions on how to install and use `yadism`:

https://yadism.readthedocs.io/en/latest/index.html

B.1 Installation

The requirements to run and install `yadism` are a system with `Python3`,[¶] along with the python package installer `pip` which enables the installation of packages from the Python Package Index.

With these requirements met, `yadism`, specifically v0.13.2 discussed in this paper, can be installed by simply running

```
python -m pip install yadism==0.13.2
```

B.2 Generating results with `yadism`

Once `yadism` is installed, results can be produced by running a single function that takes as input two dictionaries: one with instructions on the observable to be computed, and one containing the theory parameters of the calculation. Once the computation has been run, the output can be saved as a `PineAPPL` interpolation grid, or convoluted directly to any PDF in LHAPDF format and thereby obtain the requested predictions.

At a first glance, the required number of theory parameters might seems redundant, but this is actually a design choice. In fact, for most of these parameters, `yadism` does not provide

[¶] The range of minor versions of `Python` supported by a given version of `yadism` can be found on the Python Package Index: https://pypi.org/project/yadism.

a default value, as we want the user to be fully aware of all the settings entering in the calculation.

The code snippet below provides a simple example of a script that can be used to compute the reduced charm HERA NC cross section.

```
import lhapdf
import yadism
from yadbox.export import dump_pineappl_to_file

# OBSERVABLE RUNCARD:
observablecard = {
  # Process type: "EM", "NC", "CC"
  "prDIS": "NC",
  # Projectile: "electron", "positron", "neutrino", "antineutrino"
  "ProjectileDIS": "electron",
  # Scattering target: "proton", "neutron", "isoscalar", "lead", "iron", "neon" or "marble"
  "TargetDIS": "proton",
  # Interpolation: if True use log interpolation
  "interpolation_is_log": True,
  # Interpolation: polynomial degree, 1 = linear, ...
  "interpolation_polynomial_degree": 4,
  # Interpolation: xgrid values
  # Note: for illustrative purposes the grid is chosen very small here
  "interpolation_xgrid": [1e-7, 1e-6, 1e-5, 1e-4, 1e-3, 1e-2, 1e-1, 1.0],
  # Observables configurations
  "observables": {
    "XSHERANCAVG_charm": [
      {
          "y": 0.8240707777909629,
          "x": 3e-05,
          "Q2": 2.5,
      },
      {
          "y": 0.3531731904818413,
          "x": 7e-05,
          "Q2": 2.5,
      },
      # Add here the kinematics of other datapoints
    ],
    # Potentially include observables other than XSHERANCAVG_charm,
    # each of them has to be: TYPE_heaviness, where heaviness can take:
    # "charm", "bottom", "top", "total" or "light".
  },
  # Projectile polarization faction, float from 0 to 1.
  "PolarizationDIS": 0.0,
  # Exchanged boson propagator correction
  "PropagatorCorrection": 0.0,
  # Restrict boson coupling to a single parton ? Monte Carlo PID or None for all partons
  "NCPositivityCharge": None,
}
```

```
# THEORY RUNCARD:
theorycard = {
  "CKM": "0.97428 0.22530 0.003470 0.22520 0.97345 0.041000 0.00862 0.04030 0.999152", # CKM matrix elements
  "FNS": "FFNS", # Flavour Number Scheme, options: "FFNS", "FFN0", "ZM-VFNS"
  "GF": 1.1663787e-05, # [GeV^-2] Fermi coupling constant
  "IC": 1, # 0 = perturbative charm only, 1 = intrinsic charm allowed
  "MP": 0.938, # [GeV] proton mass
  "MW": 80.398, # [GeV] W boson mass
  "MZ": 91.1876, # [GeV] Z boson mass
  "NfFF": 4, # (fixed) number of running flavors, only for FFNS or FFN0 schemes
  "PTO": 2, # perturbative order in alpha_s: 0 = LO (alpha_s^0), 1 = NLO (alpha_s^1) ...
  "Q0": 1.65, # [GeV] reference scale for the flavor patch determination
  "nf0": 4, # number of active flavors at the Q0 reference scale
  "Qref": 91.2, # [GeV] reference scale for the alphas value
  "nfref": 5, # number of active flavors at the reference scale Qref
  "alphas": 0.118, # alphas value at the reference scale
  "TMC": 1, # include target mass corrections: 0 = disabled, 1 = leading twist, 2 = higher twist approximated, 3 = higher twist exact
  "XIF": 1.0, # ratio of factorization scale over the hard scattering scale
  "XIR": 1.0, # ratio of renormalization scale over the hard scattering scale
  "alphaqed": 0.007496252, # alpha_em value
  "kcThr": 1.0, # ratio of the charm matching scale over the charm mass
  "kbThr": 1.0, # ratio of the bottom matching scale over the bottom mass
  "ktThr": 1.0, # ratio of the top matching scale over the top mass
  "mc": 1.51, # [GeV] charm mass
  "mb": 4.92, # [GeV] bottom mass
  "mt": 172.5, # [GeV] top mass
  "n3lo_cf_variation": 0, # N3LO coefficient functions variation: -1 = lower bound, 0 = central , 1 = upper bound
  "QED": 0, # QED correction to running of strong coupling: 0 = disabled, 1 = allowed
  "MaxNfAs": 5, # maximum number of flavors in running of strong coupling
  "HQ": "POLE", # heavy quark mass scheme (not yet implemented in yadism)
  "MaxNfPdf": 5, # maximum number of flavors in running of PDFs (ignored by yadism)
  "ModEv": "EXA", # evolution solver for PDFs (ignored by yadism)
}

out = yadism.run_yadism(theorycard, observablecard)

# After running yadism one can either store the result in a pineappl grid
dump_pineappl_to_file(out,"outputgrid.pineappl.lz4","XSHERANCAVG_charm")

# ... or apply a chosen PDF to compute observable predictions direcly
pdf = lhapdf.mkPDF("NNPDF40_nnlo_as_01180")
values = out.apply_pdf(pdf)
```

References

1. T.-J. Hou et al., New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC. Phys. Rev. D **103**, 014013 (2021). https://doi.org/10.1103/PhysRevD.103.014013. arXiv:1912.10053
2. S. Bailey, T. Cridge, L.A. Harland-Lang, A.D. Martin, R.S. Thorne, Parton distributions from LHC, HERA, Tevatron and fixed target data: MSHT20 PDFs. Eur. Phys. J. C **81**, 341 (2021). https://doi.org/10.1140/epjc/s10052-021-09057-0. arXiv:2012.04684
3. NNPDF collaboration, The path to proton structure at 1% accuracy. Eur. Phys. J. C **82**, 428 (2022). https://doi.org/10.1140/epjc/s10052-022-10328-7. arXiv:2109.02653
4. H1, ZEUS collaboration, Combined Measurement and QCD Analysis of the Inclusive e+- p Scattering Cross Sections at HERA. JHEP **01**, 109 (2010). https://doi.org/10.1007/JHEP01(2010)109. arXiv:0911.0884
5. H1, ZEUS collaboration, Combination and QCD analysis of charm and beauty production cross-section measurements in deep inelastic *ep* scattering at HERA. Eur. Phys. J. C **78**, 473 (2018). https://doi.org/10.1140/epjc/s10052-018-5848-3. arXiv:1804.01019
6. BCDMS collaboration, A High Statistics Measurement of the Proton Structure Functions $F_2(x, Q^2)$ and R from Deep Inelastic Muon Scattering at High Q^2. Phys. Lett. B **223**, 485 (1989). https://doi.org/10.1016/0370-2693(89)91637-7
7. BCDMS collaboration, A High Statistics Measurement of the Deuteron Structure Functions $F_2(x, Q^2)$ and R from Deep Inelastic Muon Scattering at High Q^2. Phys. Lett. B **237**, 592 (1990). https://doi.org/10.1016/0370-2693(90)91231-Y
8. R. Abdul Khalek et al. Science Requirements and Detector Concepts for the Electron-Ion Collider: EIC Yellow Report. Nucl. Phys. A **1026**, 122447 (2022). https://doi.org/10.1016/j.nuclphysa.2022.122447. arXiv:2103.05419
9. D.P. Anderle et al., Electron-ion collider in China. Front. Phys. (Beijing) **16**, 64701 (2021). https://doi.org/10.1007/s11467-021-1062-0. arXiv:2102.09222
10. A. Candido, A. Garcia, G. Magni, T. Rabemananjara, J. Rojo, R. Stegeman, Neutrino Structure Functions from GeV to EeV Energies. JHEP **05**, 149 (2023). https://doi.org/10.1007/JHEP05(2023)149. arXiv:2302.08527
11. IceCube collaboration, Observation of high-energy astrophysical neutrinos in three years of IceCube data. Phys. Rev. Lett. **113**, 101101 (2014). https://doi.org/10.1103/PhysRevLett.113.101101. arXiv:1405.5303
12. L.A. Anchordoqui et al., The Forward Physics Facility: Sites, experiments, and physics potential. Phys. Rept. **968**, 1 (2022). https://doi.org/10.1016/j.physrep.2022.04.004. arXiv:2109.10905
13. J.L. Feng et al., The Forward Physics Facility at the High-Luminosity LHC. J. Phys. G **50**, 030501 (2023). https://doi.org/10.1088/1361-6471/ac865e. arXiv:2203.05090
14. R. Abdul Khalek, S. Bailey, J. Gao, L. Harland-Lang, J. Rojo, Probing Proton Structure at the Large Hadron electron Collider. SciPost Phys. **7**, 051 (2019). https://doi.org/10.21468/SciPostPhys.7.4.051. arXiv:1906.10127
15. LHeC, FCC-he Study Group collaboration, *The Large Hadron–Electron Collider at the HL-LHC. J. Phys. G* **48**, 110501 (2021). https://doi.org/10.1088/1361-6471/abf3ba. arXiv:2007.14491
16. LHeC Study Group collaboration, A Large Hadron Electron Collider at CERN: Report on the Physics and Design Concepts for Machine and Detector. J. Phys. G **39**, 075001 (2012). https://doi.org/10.1088/0954-3899/39/7/075001. arXiv:1206.2913
17. FASER collaboration, Detecting and Studying High-Energy Collider Neutrinos with FASER at the LHC. Eur. Phys. J. C **80**, 61 (2020). https://doi.org/10.1140/epjc/s10052-020-7631-5. arXiv:1908.02310
18. SHiP collaboration, *SND@LHC*, arXiv:2002.08722
19. SND@LHC collaboration, SND@LHC: The Scattering and Neutrino Detector at the LHC. arXiv:2210.02784
20. APFEL collaboration, APFEL: A PDF Evolution Library with QED corrections. Comput. Phys. Commun. **185**, 1647 (2014). https://doi.org/10.1016/j.cpc.2014.03.007. arXiv:1310.1394
21. V. Bertone, APFEL++: A new PDF evolution library in C++. *PoS* DIS2017 201 (2018). https://doi.org/10.22323/1.297.0201. arXiv:1708.00911
22. G.P. Salam, J. Rojo, A Higher Order Perturbative Parton Evolution Toolkit (HOPPET). Comput. Phys. Commun. **180**, 120 (2009). https://doi.org/10.1016/j.cpc.2008.08.010. arXiv:0804.3755
23. M. Botje, QCDNUM: Fast QCD Evolution and Convolution. Comput. Phys. Commun. **182**, 490 (2011). https://doi.org/10.1016/j.cpc.2010.10.020. arXiv:1005.1481
24. F. Hekhorn, G. Magni, E.R. Nocera, T.R. Rabemananjara, J. Rojo, A. Schaus et al. *Heavy Quarks in Polarised Deep-Inelastic Scattering at the Electron-Ion Collider*. arXiv:2401.10127
25. NNPDF collaboration, *The Path to* $\mathrm{N^3LO}$ *Parton Distributions*. arXiv:2402.18635
26. NNPDF collaboration, Determination of the theory uncertainties from missing higher orders on NNLO parton distributions with percent accuracy. Eur. Phys. J. C **84**(5), 517 (2024). https://doi.org/10.1140/epjc/s10052-024-12772-z. arXiv:2401.10319
27. A. Candido, F. Hekhorn, G. Magni, EKO: evolution kernel operators. Eur. Phys. J. C **82**, 976 (2022). https://doi.org/10.1140/epjc/s10052-022-10878-w. arXiv:2202.02338
28. A. Barontini, A. Candido, F. Hekhorn, G. Magni, R. Stegeman, An FONLL prescription with coexisting flavor number PDFs. in preparation (2024)
29. S. Forte, E. Laenen, P. Nason, J. Rojo, Heavy quarks in deep-inelastic scattering. Nucl. Phys. B **834**, 116 (2010). https://doi.org/10.1016/j.nuclphysb.2010.03.014. arXiv:1001.2312
30. S.J. Brodsky, A. Kusina, F. Lyonnet, I. Schienbein, H. Spiesberger, R. Vogt, A review of the intrinsic heavy quark content of the nucleon. Adv. High Energy Phys. **2015**, 231547 (2015). https://doi.org/10.1155/2015/231547. arXiv:1504.06287
31. Y.N. Lima, V.P. Goncalves, A.V. Giannini, B-meson production at forward rapidities in pp collisions at the LHC: Estimating the intrinsic bottom contribution. arXiv:2403.04619
32. S. Carrazza, E.R. Nocera, C. Schwan, M. Zaro, PineAPPL: combining EW and QCD corrections for fast evaluation of LHC processes. JHEP **12**, 108 (2020). https://doi.org/10.1007/JHEP12(2020)108. arXiv:2008.12789
33. C. Schwan, A. Candido, F. Hekhorn, S. Carrazza, A. Barontini, *Nnpdf/pineappl: v0.6.2*, Oct., 2023. https://doi.org/10.5281/zenodo.8422248
34. T. Cridge, M.A. Lim, Constraining the top-quark mass within the global MSHT PDF fit. Eur. Phys. J. C **83**, 805 (2023). https://doi.org/10.1140/epjc/s10052-023-11961-6. arXiv:2306.14885
35. T. Kluge, K. Rabbertz, M. Wobisch, *FastNLO: Fast pQCD calculations for PDF fits*. In *14th International Workshop on Deep Inelastic Scattering*, pp. 483–486, 9, (2006). https://doi.org/10.1142/9789812706706_0110. arXiv:hep-ph/0609285
36. T. Carli, D. Clements, A. Cooper-Sarkar, C. Gwenlan, G.P. Salam, F. Siegert et al., A posteriori inclusion of parton density functions in NLO QCD final-state calculations at hadron colliders: The APPLGRID Project. Eur. Phys. J. C **66**, 503 (2010). https://doi.org/10.1140/epjc/s10052-010-1255-0. arXiv:0911.2985
37. D. Britzger et al., NNLO interpolation grids for jet production at the LHC. Eur. Phys. J. C **82**, 930 (2022). https://doi.org/10.1140/epjc/s10052-022-10880-2. arXiv:2207.13735

38. S. Alekhin et al., HERAFitter. Eur. Phys. J. C **75**, 304 (2015). https://doi.org/10.1140/epjc/s10052-015-3480-z. arXiv:1410.4412
39. xFitter Developers' Team collaboration, xFitter 2.0.0: An Open Source QCD Fit Framework. PoS **DIS2017**, 203 (2018). https://doi.org/10.22323/1.297.0203. arXiv:1709.01151
40. xFitter collaboration, xFitter: An Open Source QCD Analysis Framework. A resource and reference document for the Snowmass study. **6** (2022). arXiv:2206.12465
41. A. Barontini, A. Candido, J.M. Cruz-Martinez, F. Hekhorn, C. Schwan, Pineline: Industrialization of high-energy theory predictions. Comput. Phys. Commun. **297**, 109061 (2024). https://doi.org/10.1016/j.cpc.2023.109061. arXiv:2302.12124
42. J.M. Cruz-Martinez, M. Fieg, T. Giani, P. Krack, T. Mäkelä, T. Rabemananjara et al. The LHC as a Neutrino-Ion Collider. arXiv:2309.09581
43. NNPDF Collaboration., Ball, R.D., Barontini, A. et al. Photons in the proton: implications for the LHC. Eur. Phys. J. C **84**, 540 (2024). https://doi.org/10.1140/epjc/s10052-024-12731-8. arXiv:2401.08749
44. R. Ellis, W. Stirling, B. Webber, *QCD and collider physics*, vol. 8, Cambridge University Press (2011)
45. Particle Data Group collaboration, *Review of Particle Physics*, *PTEP* 2022 083C01 (2022). https://doi.org/10.1093/ptep/ptac097
46. J.C. Collins, D.E. Soper, G.F. Sterman, Factorization of Hard Processes in QCD. Adv. Ser. Direct. High Energy Phys. **5**, 1 (1989). https://doi.org/10.1142/9789814503266_0001. arXiv:hep-ph/0409313
47. New Muon collaboration, Accurate measurement of F2(d) / F2(p) and R**d - R**p. Nucl. Phys. B **487**, 3 (1997). https://doi.org/10.1016/S0550-3213(96)00673-6. arXiv:hep-ex/9611022
48. New Muon collaboration, Measurement of the proton and deuteron structure functions, F2(p) and F2(d), and of the ratio sigma-L / sigma-T. Nucl. Phys. B **483**, 3 (1997). https://doi.org/10.1016/S0550-3213(96)00538-X. arXiv:hep-ph/9610231
49. L.W. Whitlow, S. Rock, A. Bodek, E.M. Riordan, S. Dasu, A Precise extraction of R = sigma-L / sigma-T from a global analysis of the SLAC deep inelastic e p and e d scattering cross-sections. Phys. Lett. B **250**, 193 (1990). https://doi.org/10.1016/0370-2693(90)91176-C
50. L.W. Whitlow, E.M. Riordan, S. Dasu, S. Rock, A. Bodek, Precise measurements of the proton and deuteron structure functions from a global analysis of the SLAC deep inelastic electron scattering cross-sections. Phys. Lett. B **282**, 475 (1992). https://doi.org/10.1016/0370-2693(92)90672-Q
51. CHORUS collaboration, Measurement of nucleon structure functions in neutrino scattering. Phys. Lett. B **632**, 65 (2006). https://doi.org/10.1016/j.physletb.2005.10.062
52. NuTeV collaboration, Precise measurement of dimuon production cross-sections in ν_μFe and $\bar{\nu}_\mu$Fe deep inelastic scattering at the Tevatron. Phys. Rev. D **64**, 112006 (2001). https://doi.org/10.1103/PhysRevD.64.112006. arXiv:hep-ex/0102049
53. European Muon collaboration, Production of charmed particles in 250-GeV μ^+ - iron interactions. Nucl. Phys. B **213**, 31 (1983). https://doi.org/10.1016/0550-3213(83)90174-8
54. BEBC WA59 collaboration, Measurement of the Structure Functions F2 and Xf3 and Comparison With QCD Predictions Including Kinematical and Dynamical Higher Twist Effects. Z. Phys. C **36**, 1 (1987). https://doi.org/10.1007/BF01556159
55. CHARM collaboration, Experimental study of the nucleon longitudinal structure function in charged current neutrino and antineutrinos interactions. Phys. Lett. B **141**, 129 (1984). https://doi.org/10.1016/0370-2693(84)90575-6
56. C. Gnendiger et al., To *d*, or not to *d*: recent developments and comparisons of regularization schemes. Eur. Phys. J. C **77**, 471 (2017). https://doi.org/10.1140/epjc/s10052-017-5023-2. arXiv:1705.01827
57. N. Cabibbo, Unitary Symmetry and Leptonic Decays. Phys. Rev. Lett. **10**, 531 (1963). https://doi.org/10.1103/PhysRevLett.10.531
58. M. Kobayashi, T. Maskawa, CP Violation in the Renormalizable Theory of Weak Interaction. Prog. Theor. Phys. **49**, 652 (1973). https://doi.org/10.1143/PTP.49.652
59. J.P. Berge et al., A Measurement of Differential Cross-Sections and Nucleon Structure Functions in Charged Current Neutrino Interactions on Iron. Z. Phys. C **49**, 187 (1991). https://doi.org/10.1007/BF01555493
60. W.L. van Neerven, A. Vogt, NNLO evolution of deep inelastic structure functions: The Singlet case. Nucl. Phys. B **588**, 345 (2000). https://doi.org/10.1016/S0550-3213(00)00480-6. arXiv:hep-ph/0006154
61. I. Schienbein et al., A Review of Target Mass Corrections. J. Phys. G **35**, 053101 (2008). https://doi.org/10.1088/0954-3899/35/5/053101. arXiv:0709.1775
62. M. Goharipour, S. Rostami, Implementation of target mass corrections and higher-twist effects in the xFitter framework. Phys. Rev. D **101**, 074015 (2020). https://doi.org/10.1103/PhysRevD.101.074015. arXiv:2004.03403
63. J. Ablinger, J. Blümlein, A. De Freitas, A. Goedicke, C. Schneider, K. Schönwald, Two-mass three-loop effects in deep-inelastic scattering. PoS **LL2018**, 051 (2018). https://doi.org/10.22323/1.303.0051. arXiv:1807.07855
64. J.A.M. Vermaseren, A. Vogt, S. Moch, The Third-order QCD corrections to deep-inelastic scattering by photon exchange. Nucl. Phys. B **724**, 3 (2005). https://doi.org/10.1016/j.nuclphysb.2005.06.020. arXiv:hep-ph/0504242
65. S. Moch, J. Vermaseren, A. Vogt, The longitudinal structure function at the third order, Phys. Lett. B **606**, 123–129 (2005). https://doi.org/10.1016/j.physletb.2004.11.063. arXiv:hep-ph/0411112
66. S. Moch, J. Vermaseren, Deep inelastic structure functions at two loops. Nucl. Phys. B **573**, 853 (2000). https://doi.org/10.1016/S0550-3213(00)00045-6. arXiv:hep-ph/9912355
67. F. Hekhorn, Next-to-Leading Order QCD Corrections to Heavy-Flavour Production in Neutral Current DIS. Ph.D. thesis, Tubingen U., Math. Inst. (2019). https://doi.org/10.15496/publikation-34811. arXiv:1910.01536
68. S. Kretzer, I. Schienbein, Heavy quark initiated contributions to deep inelastic structure functions. Phys. Rev. D **58**, 094035 (1998). https://doi.org/10.1103/PhysRevD.58.094035. arXiv:hep-ph/9805233
69. M. Buza, Y. Matiounine, J. Smith, W.L. van Neerven, Charm electroproduction viewed in the variable flavor number scheme versus fixed order perturbation theory. Eur. Phys. J. C **1**, 301 (1998). https://doi.org/10.1007/BF01245820. arXiv:hep-ph/9612398
70. S. Moch, M. Rogal, A. Vogt, Differences between charged-current coefficient functions. Nucl. Phys. B **790**, 317 (2008). https://doi.org/10.1016/j.nuclphysb.2007.09.022. arXiv:0708.3731
71. S. Moch, J. Vermaseren, A. Vogt, Third-order QCD corrections to the charged-current structure function F(3). Nucl. Phys. B **813**, 220 (2009). https://doi.org/10.1016/j.nuclphysb.2009.01.001. arXiv:0812.4168
72. M. Gluck, S. Kretzer, E. Reya, The Strange sea density and charm production in deep inelastic charged current processes. Phys. Lett. B **380**, 171 (1996). https://doi.org/10.1016/0370-2693(96)00456-X. arXiv:hep-ph/9603304
73. J. Gao, Massive charged-current coefficient functions in deep-inelastic scattering at NNLO and impact on strange-quark distributions. JHEP **02**, 026 (2018). https://doi.org/10.1007/JHEP02(2018)026. arXiv:1710.04258
74. H. Kawamura, N.A. Lo Presti, S. Moch, A. Vogt, On the next-to-next-to-leading order QCD corrections to heavy-quark production

in deep-inelastic scattering. Nucl. Phys. B **864**, 399 (2012). https://doi.org/10.1016/j.nuclphysb.2012.07.001. arXiv:1205.5727
75. N. Laurenti, Construction of a next-to-next-to-next-to-leading order approximation for heavy flavour production in deep inelastic scattering with quark masses. arXiv:2401.12139
76. A. Barontini, M. Bonvini, N. Laurenti, *Implementation of DIS at N3LO for PDF determinations* in preparation (2024)
77. I. Bierenbaum, J. Blumlein, S. Klein, C. Schneider, Two-Loop Massive Operator Matrix Elements for Unpolarized Heavy Flavor Production to O(epsilon). Nucl. Phys. B **803**, 1 (2008). https://doi.org/10.1016/j.nuclphysb.2008.05.016. arXiv:0803.0273
78. I. Bierenbaum, J. Blümlein, S. Klein, Mellin Moments of the $O(\alpha_s^3)$ Heavy Flavor Contributions to unpolarized Deep-Inelastic Scattering at $Q^2 \gg m^2$ and Anomalous Dimensions. Nucl. Phys. B **820**, 417 (2009). https://doi.org/10.1016/j.nuclphysb.2009.06.005. arXiv:0904.3563
79. J. Ablinger, J. Blümlein, S. Klein, C. Schneider, F. Wissbrock, The $O(\alpha_s^3)$ Massive Operator Matrix Elements of $O(n_f)$ for the Structure Function $F_2(x, Q^2)$ and Transversity. Nucl. Phys. B **844**, 26 (2011). https://doi.org/10.1016/j.nuclphysb.2010.10.021. arXiv:1008.3347
80. J. Ablinger, A. Behring, J. Blümlein, A. De Freitas, A. Hasselhuhn, A. von Manteuffel et al., The 3-Loop Non-Singlet Heavy Flavor Contributions and Anomalous Dimensions for the Structure Function $F_2(x, Q^2)$ and Transversity. Nucl. Phys. B **886**, 733 (2014). https://doi.org/10.1016/j.nuclphysb.2014.07.010. arXiv:1406.4654
81. J. Ablinger, A. Behring, J. Blümlein, A. De Freitas, A. von Manteuffel et al., *The 3-Loop Pure Singlet Heavy Flavor Contributions to the Structure Function $F_2(x, Q^2)$ and the Anomalous Dimension*. arXiv:1409.1135
82. A. Behring, I. Bierenbaum, J. Blümlein, A. De Freitas, S. Klein, F. Wißbrock, The logarithmic contributions to the $O(\alpha_s^3)$ asymptotic massive Wilson coefficients and operator matrix elements in deeply inelastic scattering. Eur. Phys. J. C **74**, 3033 (2014). https://doi.org/10.1140/epjc/s10052-014-3033-x. arXiv:1403.6356
83. E.B. Zijlstra, W.L. van Neerven, Order-α_s^2 corrections to the polarized structure function $g_1(x, Q^2)$. Nucl. Phys. B **417**, 61 (1994). https://doi.org/10.1016/0550-3213(94)90538-X
84. D. de Florian, R. Sassot, O (alpha-s) spin dependent weak structure functions. Phys. Rev. D **51**, 6052 (1995). https://doi.org/10.1103/PhysRevD.51.6052. arXiv:hep-ph/9412255
85. M. Anselmino, P. Gambino, J. Kalinowski, New proton polarized structure functions in charged current processes at HERA. Phys. Rev. D **55**, 5841 (1997). https://doi.org/10.1103/PhysRevD.55.5841. arXiv:hep-ph/9607427
86. M. Buza, Y. Matiounine, J. Smith, W.L. van Neerven, O (alpha-s**2) corrections to polarized heavy flavor production at Q**2 is > > m**2. Nucl. Phys. B **485**, 420 (1997). https://doi.org/10.1016/S0550-3213(96)00606-2. arXiv:hep-ph/9608342
87. I. Borsa, D. de Florian, I. Pedron, The full set of polarized deep inelastic scattering structure functions at NNLO accuracy. Eur. Phys. J. C **82**, 1167 (2022). https://doi.org/10.1140/epjc/s10052-022-11140-z. arXiv:2210.12014
88. J. Blümlein, P. Marquard, C. Schneider, K. Schönwald, The massless three-loop Wilson coefficients for the deep-inelastic structure functions F_2, F_L, xF_3 and g_1. JHEP **11**, 156 (2022). https://doi.org/10.1007/JHEP11(2022)156. arXiv:2208.14325
89. A. Behring, J. Blümlein, A. De Freitas, A. von Manteuffel, C. Schneider, The 3-Loop Non-Singlet Heavy Flavor Contributions to the Structure Function $g_1(x, Q^2)$ at Large Momentum Transfer. Nucl. Phys. B **897**, 612 (2015). https://doi.org/10.1016/j.nuclphysb.2015.06.007. arXiv:1504.08217
90. J. Ablinger, A. Behring, J. Blümlein, A. De Freitas, A. von Manteuffel, C. Schneider et al., The three-loop single mass polarized pure singlet operator matrix element. Nucl. Phys. B **953**, 114945 (2020). https://doi.org/10.1016/j.nuclphysb.2020.114945. arXiv:1912.02536
91. A. Behring, J. Blümlein, A. De Freitas, A. von Manteuffel, K. Schönwald, C. Schneider, The polarized transition matrix element $A_{gq}(N)$ of the variable flavor number scheme at $O(\alpha_s^3)$. Nucl. Phys. B **964**, 115331 (2021). https://doi.org/10.1016/j.nuclphysb.2021.115331. arXiv:2101.05733
92. J. Blümlein, A. De Freitas, M. Saragnese, C. Schneider, K. Schönwald, Logarithmic contributions to the polarized O(αs3) asymptotic massive Wilson coefficients and operator matrix elements in deeply inelastic scattering. Phys. Rev. D **104**, 034030 (2021). https://doi.org/10.1103/PhysRevD.104.034030. arXiv:2105.09572
93. I. Bierenbaum, J. Blümlein, A. De Freitas, A. Goedicke, S. Klein, K. Schönwald, $O(\alpha_s^2)$ polarized heavy flavor corrections to deep-inelastic scattering at $Q^2 \gg m^2$. Nucl. Phys. B **988**, 116114 (2023). https://doi.org/10.1016/j.nuclphysb.2023.116114. arXiv:2211.15337
94. J. Ablinger, A. Behring, J. Blümlein, A. De Freitas, A. von Manteuffel, C. Schneider et al., The first-order factorizable contributions to the three-loop massive operator matrix elements $A_{Qg}^{(3)}$ and $\Delta A_{Qg}^{(3)}$. arXiv:2311.00644
95. E.G. Floratos, C. Kounnas, R. Lacaze, Higher Order QCD Effects in Inclusive Annihilation and Deep Inelastic Scattering. Nucl. Phys. B **192**, 417 (1981). https://doi.org/10.1016/0550-3213(81)90434-X
96. A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht et al., LHAPDF6: parton density access in the LHC precision era. Eur. Phys. J. C **75**, 132 (2015). https://doi.org/10.1140/epjc/s10052-015-3318-8. arXiv:1412.7420
97. W. Beenakker, H. Kuijf, W.L. van Neerven, J. Smith, QCD Corrections to Heavy Quark Production in p anti-p Collisions. Phys. Rev. D **40**, 54 (1989). https://doi.org/10.1103/PhysRevD.40.54
98. C. Bissolotti, R. Boughezal, K. Simsek, SMEFT probes in future precision DIS experiments. Phys. Rev. D **108**, 075007 (2023). https://doi.org/10.1103/PhysRevD.108.075007. arXiv:2306.05564
99. M. McCullough, J. Moore, M. Ubiali, The dark side of the proton. JHEP **08**, 019 (2022). https://doi.org/10.1007/JHEP08(2022)019. arXiv:2203.12628
100. A. Manohar, P. Nason, G.P. Salam, G. Zanderighi, How bright is the proton? A precise determination of the photon parton distribution function. Phys. Rev. Lett. **117**, 242002 (2016). https://doi.org/10.1103/PhysRevLett.117.242002. arXiv:1607.04266
101. A.V. Manohar, P. Nason, G.P. Salam, G. Zanderighi, The Photon Content of the Proton. JHEP **12**, 046 (2017). https://doi.org/10.1007/JHEP12(2017)046. arXiv:1708.01256
102. NNPDF collaboration, *Illuminating the photon content of the proton within a global PDF analysis*. SciPost Phys. **5**, 008 (2018). https://doi.org/10.21468/SciPostPhys.5.1.008. arXiv:1712.07053
103. T. Cridge, L.A. Harland-Lang, A.D. Martin, R.S. Thorne, QED parton distribution functions in the MSHT20 fit. Eur. Phys. J. C **82**, 90 (2022). https://doi.org/10.1140/epjc/s10052-022-10028-2. arXiv:2111.05357
104. K. Xie, B. Zhou, T.J. Hobbs, *The Photon Content of the Neutron*. arXiv:2305.10497
105. K. Kudashkin. in preparation, (2024)